Front. Phys. 12(6), 128905 (2017)
DOI 10.1007/s11467-017-0661-2

REVIEW ARTICLE

# New approaches in agent-based modeling of complex financial systems

**Ting-Ting Chen**[1,2], **Bo Zheng**[1,2,†], **Yan Li**[1,2], **Xiong-Fei Jiang**[1,2]

[1] *Department of Physics, Zhejiang University, Hangzhou 310027, China*
[2] *Collaborative Innovation Center of Advanced Microstructures, Nanjing 210093, China*
*Corresponding author. E-mail:* [†] *zhengbo@zju.edu.cn*
*Received November 9, 2016; accepted January 6, 2017*

Agent-based modeling is a powerful simulation technique to understand the collective behavior and microscopic interaction in complex financial systems. Recently, the concept for determining the key parameters of agent-based models from empirical data instead of setting them artificially was suggested. We first review several agent-based models and the new approaches to determine the key model parameters from historical market data. Based on the agents' behaviors with heterogeneous personal preferences and interactions, these models are successful in explaining the microscopic origination of the temporal and spatial correlations of financial markets. We then present a novel paradigm combining big-data analysis with agent-based modeling. Specifically, from internet query and stock market data, we extract the information driving forces and develop an agent-based model to simulate the dynamic behaviors of complex financial systems.

**Keywords** econophysics, complex systems

**PACS numbers** 89.65.Gh, 89.75.-k

Contents



## 1 Introduction

Complex financial systems typically have many-body interactions. The interactions of multiple agents induce various collective phenomena, such as abnormal distributions, temporal correlations, and sector structures [1–9]. Complex financial systems are also substantially influenced by external information that may, for example, drive the systems to non-stationary states, larger fluctuations, or extreme events [10–17].

Complex financial systems are important examples of open complex systems. Standard finance supposes that investors have complete rationality, but the progress of behavioral and experimental finance shows that investors

*Special Topic: Soft-Matter Physics and Complex Systems (Ed. Zhi-Gang Zheng). arXiv: 1703.06840.

© Higher Education Press and Springer-Verlag Berlin Heidelberg 2017

Frontiers of Physics

in real life have behavioral and emotional differences [18, 19]. To be more specific, agents who are not fully rational may have different personal preferences and interact with each other differently in financial markets [9, 20–26].

Information is a leading factor in complex financial systems. However, our understanding of external information and its controlling effect in agent-based modeling is rather limited [27–30]. In recent years, exploring the scientific impact of online big-data has attracted much attention of researchers from different fields. The massive new data sources resulting from human interactions with the internet offer a better understanding of the profound influence of external information on complex financial systems [31–39].

Agent-based modeling is a powerful simulation technique to understand the collective behavior in complex financial systems [40–45]. More recently, the concept of determining the key parameters of agent-based models from empirical data instead of setting them artificially was suggested [20]. A similar concept has also been applied to order-driven models, which were first proposed by Mike and Farmer [46] and improved by Gu and Zhou [47–50]. In this family of order-driven models, the parameters of order submissions and order cancellations are determined using real-order book data. For comparison, agent-based models focus more on the behaviors of agents [40–45], while order-driven models are mainly intended to explore the dynamics of order flows [46–50]. In Section 2, we review several agent-based models that are based on agents' behaviors with heterogeneous personal preferences and interactions. These models explore the microscopic origination of the temporal and spatial correlations of financial markets [9, 21, 24]. In Section 3, we present a novel paradigm combining big-data analysis with agent-based modeling [51].

## 2 New approaches in agent-based models

From the viewpoint of physicists, the dynamic behavior and community structure of complex financial systems can be characterized by temporal and spatial correlation functions. Recently, several agent-based models are proposed to explore the microscopic generation mechanisms of temporal and spatial correlations [9, 21, 24]. These models are microscopic herding models, in which the agents are linked with each other and trade in groups, and, in particular, contain new approaches to multi-agent interactions.

### 2.1 Basics of agent-based models

The stock price on day $t$ is denoted as $Y(t)$, and the logarithmic price return is $R(t) = \ln[Y(t)/Y(t-1)]$. For comparison of different time series of returns, the normalized return $r(t)$ is introduced as

$$r(t) = [R(t) - \langle R(t)\rangle]/\sigma, \tag{1}$$

where $\langle\cdots\rangle$ represents the average over time $t$, and $\sigma = \sqrt{\langle R^2(t)\rangle - \langle R(t)\rangle^2}$ is the standard deviation of $R(t)$. In stock markets, the information for investors is highly incomplete; therefore, an agent's decision of *buy*, *sell*, or *hold* is assumed to be random. In these models, there is only one stock and $N$ agents, and each agent operates one share every day. On day $t$, each agent $i$ makes a trading decision $\phi_i(t)$ so that

$$\phi_i(t) = \begin{cases} 1, & \text{buy}, \\ -1, & \text{sell}, \\ 0, & \text{hold}, \end{cases} \tag{2}$$

and the probabilities of *buy*, *sell*, or *hold* decisions are denoted as $P_{buy}(t)$, $P_{sell}(t)$, and $P_{hold}(t)$, respectively. The price return $R(t)$ is defined by the difference of the demand and supply of the stock as

$$R(t) = \sum_{i=1}^{N} \phi_i(t). \tag{3}$$

For simplicity, the volatility is defined as the absolute return $|R(t)|$. Other definitions yield similar results.

The investment horizon is introduced since agents' decisions are based on the previous stock performance of different time horizons [9, 21, 24]. It has been found that the relative portion $\gamma_i$ of agents with an $i$-days investment horizon follows a power-law decay, $\gamma_i \propto i^{-\eta}$ with $\eta = 1.12$ [20]. The maximum investment horizon is denoted as $M$. To describe the integrated investment basis of all agents, a weighted average return $R'(t)$ is introduced as

$$R'(t) = k \cdot \sum_{i=1}^{M} \left[ \gamma_i \sum_{j=0}^{i-1} R(t-j) \right], \tag{4}$$

where $k$ is a proportional coefficient. According to Ref. [52], the investment horizons of investors range from a few days to several months. For $M$ between 50 and 500, the results from simulations remain robust.

In complex financial systems, herding is one of the collective behaviors, which arises when investors imitate the decision of others rather than follow their own beliefs and judgment. In other words, investors cluster into groups when making decisions [53, 54]. Here, a herding degree, $D(t)$, is introduced to quantify the clustering degree of the herding behavior as

$$D(t) = n_A(t)/N, \tag{5}$$

where $n_A(t)$ is the average number of agents in each cluster on day $t$.

### 2.2 Agent-based model with asymmetric trading and herding

The negative and positive return-volatility correlations, that is, the so-called leverage and anti-leverage effects, are particularly important for the understanding of price dynamics [1, 6, 55, 56]. Although various macroscopic models have been proposed to describe the return-volatility correlation, it is very important to understand the correlations from the microscopic level. To study the microscopic origination of the return-volatility correlation in financial markets, two novel microscopic mechanisms, that is, investors' asymmetric trading and herding in bull and bear markets, are recently introduced in agent-based modeling [21].

#### *2.2.1 Two important behaviors of investors*

(i) Asymmetric trading. An investor's willingness to trade is affected by the previous price returns, leading the trading probability to be distinct in bull and bear markets. The model thus assumes dynamic probabilities for buying and selling, but with $P_{buy}(t) = P_{sell}(t)$. As the trading probability $P_{trade}(t) = P_{buy}(t) + P_{sell}(t)$, its average over time is set to be $\langle P_{trade}(t)\rangle = 2p$. We adopt the value of $p$ estimated in Ref. [20], where $p = 0.0154$. The market performance is defined to be bullish if $R'(t) > 0$ and bearish if $R'(t) < 0$. The investors' asymmetric trading in bull and bear markets gives rise to the distinction between $P_{trade}(t+1)|_{R'(t)>0}$ and $P_{trade}(t+1)|_{R'(t)<0}$. Thus, $P_{trade}(t+1)$ should take the following form:

$$P_{trade}(t+1) = \begin{cases} 2p\cdot\alpha, & R'(t) > 0 \\ 2p, & R'(t) = 0 \\ 2p\cdot\beta, & R'(t) < 0 \end{cases}. \tag{6}$$

Here, $\alpha$ and $\beta$ are asymmetric factors, and $\langle P_{trade}(t)\rangle = 2p$ requires $\alpha + \beta = 2$; that is, $\alpha$ and $\beta$ are not independent.

(ii) Asymmetric herding. Herding, as a collective behavior in financial markets, describes the fact that investors form clusters when making decisions, and these clusters can be large [53, 54]. However, the herding behavior in bull markets is not the same as that in bear ones [57, 58].

In general, herding should be related to previous volatilities [59, 60], and we set the average number of agents in each cluster $n_A(t+1) = |R'(t)|$. Thus, the herding degree on day $t+1$ is

$$D(t+1) = |R'(t)|/N. \tag{7}$$

This herding degree is symmetric for $R'(t) > 0$ and $R'(t) < 0$. However, investors' herding behaviors in bull and bear markets are asymmetric. Thus, $D(t+1)$ should be redefined as

$$D(t+1) = |R'(t) - \Delta R|/N. \tag{8}$$

Here, $\Delta R$ is the degree of asymmetry. Every day, agents in a cluster make the same trading decision, that is, decision to *buy*, *sell*, or *hold* with the same probability $P_{buy}$, $P_{sell}$, or $P_{hold}$.

#### *2.2.2 Determination of $\alpha$ and $\Delta R$*

Data from six representative stock-market indices are collected, that is, daily data for the S&P 500 Index, Shanghai Index, Nikkei 225 Index, FTSE 100 Index, HKSE Index, and DAX Index.

We assume that the trading probability is proportional to the trading volume. Thus, the ratio of average trading volumes for the bull and bear markets is

$$V_+/V_- = \frac{P_{trade}(t+1)|_{R'(t)>0}}{P_{trade}(t+1)|_{R'(t)<0}} = \alpha/\beta. \tag{9}$$

Together with the condition that $\alpha + \beta = 2$, $\alpha$ is determined from $V_+/V_-$ for the six representative stock market indices, as shown in Table 1.

From empirical analysis, the herding degrees of bull and bear stock markets are not equal, that is, $d_{bull} \neq d_{bear}$. To quantize this asymmetry, a shifting $\Delta r$ is introduced such that $d_{bull}[r'(t)] = d_{bear}[r'(t)]$ with $r'(t) = r(t) + \Delta r$. From this definition, $\Delta r$ is derived to be

$$\Delta r = \frac{1}{2}[d_{bear}(r(t)) - d_{bull}(r(t))]. \tag{10}$$

Here, the herding degrees of bull markets ($r(t) > 0$) and bear markets ($r(t) < 0$) are defined as the average $|r(t)|$ with weight $V(t)$, that is,

$$\begin{cases} d_{bull}[r(t)] = \sum\limits_{t,r(t)>0} V(t)\cdot r(t) / \sum\limits_{t,r(t)>0} V(t) \\ d_{bear}[r(t)] = \sum\limits_{t,r(t)<0} V(t)\cdot |r(t)| / \sum\limits_{t,r(t)<0} V(t) \end{cases}. \tag{11}$$

Then, shifting to the time series $R(t)$, which equalizes the herding degree, $D(t+1) = |R'(t) - \Delta R|/N$ in bull markets ($R'(t) > 0$) and bear markets ($R'(t) < 0$) is computed in a similar manner. Table 1 shows the values of $\Delta r$ and $\Delta R$ for different indices.

#### *2.2.3 Simulation results*

With $\alpha$ and $\Delta R$ determined for each index, the model produces the time series of returns $R(t)$. To describe how past returns affect future volatilities, the return-volatility correlation function $L(t)$ is defined as

$$L(t) = \langle r(t')\cdot |r(t'+t)|^2\rangle / Z, \tag{12}$$

**Table 1** The values of $\alpha$, $\Delta r$ and $\Delta R$ for the six indices. $\Delta R$ is computed from the linear relation between $\Delta r$ and $\Delta R$ for all these indices.

| Index | $\alpha$ | $\Delta r$ | $\Delta R$ |
|---|---|---|---|
| S&P 500 | $1.01 \pm 0.01$ | $0.067 \pm 0.007$ | 3 |
| Shanghai | $1.09 \pm 0.01$ | $-0.043 \pm 0.005$ | $-2$ |
| Nikkei 225 | $1.01 \pm 0.01$ | $0.039 \pm 0.005$ | 2 |
| FTSE 100 | $0.99 \pm 0.01$ | $0.028 \pm 0.003$ | 2 |
| Hangseng | $1.02 \pm 0.02$ | $0.032 \pm 0.003$ | 2 |
| DAX | $0.98 \pm 0.02$ | $0.013 \pm 0.002$ | 1 |

with $Z = \langle |r(t')|^2 \rangle^2$ [61]. Here, $\langle \cdots \rangle$ represents the average over time $t'$. As displayed in Fig. 1, $L(t)$ calculated with the empirical data of the S&P 500 Index shows negative values up to at least 15 days, and this is the well-known leverage effect [1, 6, 61]. On the other hand, $L(t)$ for the Shanghai Index remains positive for about 10 days. This is the so-called anti-leverage effect [6, 55]. The return-volatility correlation function produced in the model is in agreement with that calculated from empirical data on amplitude and duration for both the S&P 500 and Shanghai indices. This is the first time that the leverage and anti-leverage effects are simulated with a microscopic model. As displayed in Fig. 2, $L(t)$ for the simulations is also in agreement with that for the Nikkei, FTSE 100, HKSE, and DAX indices.

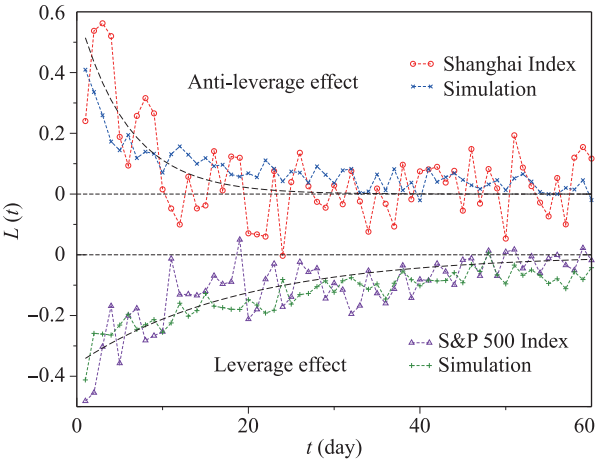


**Fig. 1** The return-volatility correlation functions for the S&P 500 and Shanghai indices, and for the corresponding simulations. The S&P 500 and Shanghai indices are simulated with $(\alpha, \Delta R) = (1.0, 3)$ and $(\alpha, \Delta R) = (1.1, -2)$, respectively. Dashed lines show the exponential fits $L(t) = c \cdot \exp(-t/\tau)$.

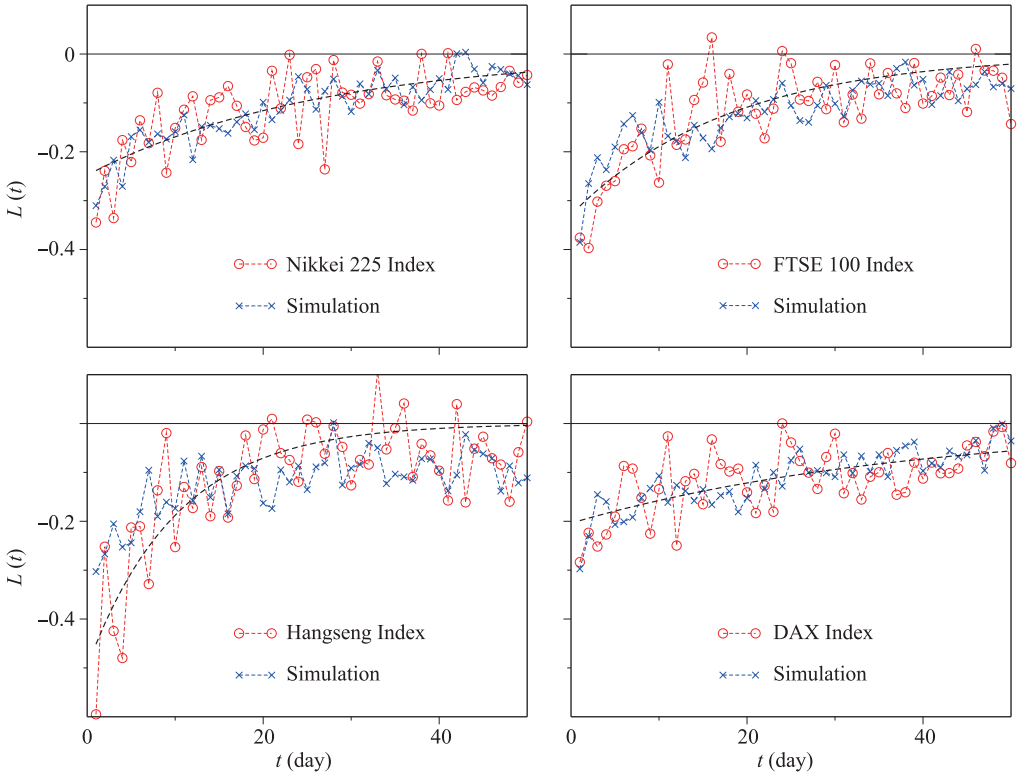


**Fig. 2** The return-volatility correlation functions for the four indices and the corresponding simulations. The Nikkei 225, FTSE 100, Hangseng, and DAX indices are simulated with $(\alpha, \Delta R) = (1.0, 2)$, $(1.0, 2)$, $(1.0, 2)$ and $(1.0, 1)$, respectively. Dashed lines show the exponential fits $L(t) = c \cdot \exp(-t/\tau)$.

As shown in Fig. 4 and Fig. 5 of Ref. [21], the model also produces volatility clustering and fat-tail distribution of returns [21]. The Hurst exponent of $A(t)$ is calculated to be 0.79, which also indicates the long-range correlation of volatilities [62]. The auto-correlation function of returns fluctuates around zero. The power-law exponent of the simulated returns is estimated to be 2.96, close to the so-called inverse cubic law [63–66].

### 2.3 Agent-based model with asymmetric trading preference

The problem of whether and how volatilities affect the price movement draws much attention. However, the usual volatility-return correlation function, which is local in time, typically fluctuates around zero. Recently, a dynamic observable nonlocal in time was constructed to explore the volatility-return correlation [9]. Strikingly, the correlation is found to be non-zero with an amplitude of a few percent and duration of over two weeks. This result provides compelling evidence that past volatilities nonlocal in time affect future returns. Alternatively, this phenomenon could be also understood as the non-stationary dynamic effect of complex financial systems.

To study the microscopic origin of the nonlocal volatility-return correlation, an agent-based model is constructed [9], in which a novel mechanism, that is, the asymmetric trading preference in volatile and stable markets, is introduced.

In financial markets, the market behaviors of buying and selling are not always in balance [67]. Hence, $P_{buy}$ and $P_{sell}$ are not always equal to each other. They are affected by previous volatilities, and the more volatile the market is, the more $P_{buy}$ differs from $P_{sell}$.

For an agent with an $i$-days investment horizon, the average volatility over previous $i$ days is taken into account, which is defined as

$$v_i(t) = \frac{1}{i}\sum_{j=1}^{i} v(t-j+1). \tag{13}$$

The background volatility is considered to be $v_M(t)$, with $M$ being the maximum investment horizon. On day $t$, the agent with an $i$-days investment horizon estimates the volatility of the market by comparing $v_i(t)$ with $v_M(t)$. Therefore, the integrated perspective of all agents on the recent market volatility is defined as

$$\xi(t) = \frac{1}{v_M(t)}\sum_{i=1}^{M}\gamma_i v_i(t). \tag{14}$$

Thus, the probabilities of buying and selling are assumed to be

$$\begin{cases} P_{buy}(t+1) = p[c\cdot\xi(t) + (1-c)], \\ P_{sell}(t+1) = 2p - P_{buy}(t+1). \end{cases} \tag{15}$$

Here, the parameter $c$ measures the degree of agents' asymmetric trading preference in volatile and stable markets. Compared with the model reviewed in Section 2.2, $c$ is the only additional parameter. In principle, $c$ could be determined from the trade and quote data of stock markets. Unfortunately, the data are currently not available to us. Thus, the question of how to determine $c$ from historical market data remains unanswered. Nevertheless, with this model, it is possible to simulate the non-zero volatility-return correlation nonlocal in time [9].

### 2.4 Agent-based model with multi-level herding

The spatial structure of stock markets is explored through the cross-correlations of individual stocks. With the random matrix theory (RMT), for example, communities can be identified, which are usually associated with business sectors in stock markets [7, 15, 68–72]. To simulate the sector structure with the agent-based model, we newly introduce the multi-level herding mechanism [24].

#### *2.4.1 Multi-level herding*

In the model, there are $N$ agents, $n$ stocks, and $n_{sec}$ sectors. Each sector contains $n/n_{sec}$ stocks. Every agent holds only one stock, which is randomly chosen from the $n$ stocks. The logarithmic price return of the $k$-th stock on day $t$ is denoted by $R_k(t)$. We assume that the agents' herding behavior comprises the herding at stock, sector, and market levels. The schematic diagram of the multi-level herding is displayed in Fig. 3(a).

(i) Herding at the stock level. The agents in each individual stock first cluster into groups, which are called $I$-groups. The herding degree $D^I$ quantifies the herding behavior at this level. On day $t$, the herding degree for the $k$-th stock is

$$D_k^I(t) = |R_k'(t-1)|/N_k. \tag{16}$$

In the $k$-th stock, the number of $I$-groups is $1/D_k^I(t)$, and the agents randomly join one of the $I$-groups. After the herding at the stock level, the number of $I$-groups in the $j$-th sector and the whole market are denoted, respectively, by $N_j^I(t)$ and $N_M^I(t)$ as follows:

$$\begin{cases} N_j^I(t) = \sum_{k\in j}[1/D_k^I(t)], \\ N_M^I(t) = \sum_{k}[1/D_k^I(t)]. \end{cases} \tag{17}$$

Here, $k \in j$ represents stock $k$ in sector $j$.

(ii) Herding at the sector level. Stocks in the same sector share the characteristics of the sector. At this level, agents' herding behavior is driven by the price co-movement of the sector; that is, the prices of stocks in a sector tend to rise and fall simultaneously. Thus, the

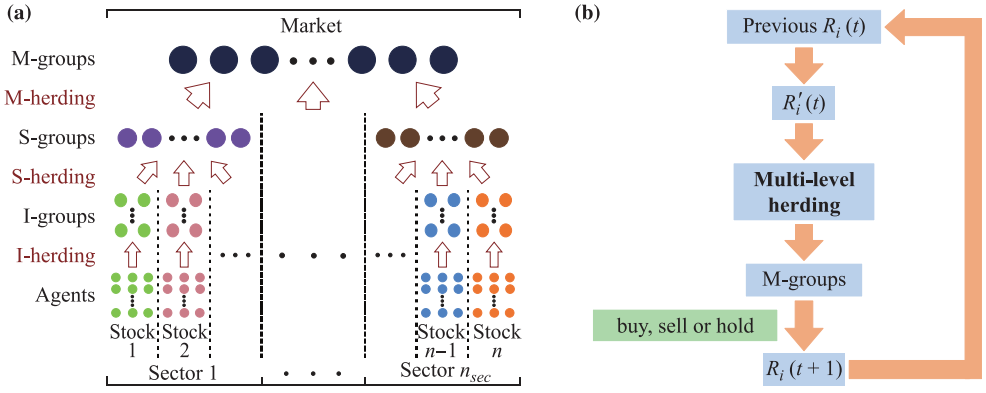


**Fig. 3** The schematic diagram of **(a)** the multi-level herding; **(b)** the procedure of simulation.

$I$-groups in the same sector would further form larger groups, which are called $S$-groups. $H_M$ and $H_j$ characterize the price co-movement degrees for stocks in the whole market and in sector $j$, respectively. For the $j$-th sector, the average number of $I$-groups in each $S$-group is set to be $n \cdot (H_j - H_M)$, which represents the pure price co-movement of the sector. Therefore, the herding degree is

$$D_j^S(t) = n \cdot (H_j - H_M)/N_j^I(t). \tag{18}$$

In sector $j$, the number of $S$-groups is $1/D_j^S(t)$, and each $I$-group joins one of the $S$-groups.

(iii) Herding at the market level. Agents' herding behavior at this level is driven by the price co-movement of the entire market. The $S$-groups in different sectors share common features of the whole market, and thus they cluster into larger groups. These groups are called $M$-groups. For the $S$-groups in sector $j$, the herding degree at market level is

$$D_j^M(t) = n \cdot H_M/N_j^M(t), \tag{19}$$

and the number of $M$-groups is $1/D_j^M(t)$. The total number of $M$-groups in the market is the maximum of $1/D_j^M(t)$ for different $j$. With all $M$-groups numbered, an $S$-group in sector $j$ joins one of the first $1/D_j^M(t)$ $M$-groups.

In the formation of $S$-groups, the $I$-groups in the same stock tend not to join the same $S$-group, otherwise these $I$-groups would have gathered together during the herding at the stock level. Similarly, in the formation of $M$-groups, the $S$-groups in the same sector tend not to join the same $M$-group.

After the herding for the three levels, all agents cluster into $M$-groups. The agents in the same $M$-group make the same trading decision $\phi_i(t)$ with the same probability. Similar to the previous models [20, 21], the buying and selling probabilities are equal, that is, $P_{buy} = P_{sell} = P$, thus $P_{hold} = 1 - 2P$. Here, $P$ is the buying or selling probability of an $M$-group, which can be calculated from the daily trading probability $p$ for each agent and the average number of agents in an $M$-group [24]. The return of the $k$-th stock is defined as $R_k(t) = \sum_{i \in k} \phi_i(t)$. Here, $i \in k$ represents agent $i$ in stock $k$.

### 2.4.2 *Determination of $H_M$ and $H_j$*

On each day $t$, according to the sign of $r_k(t)$, the stocks are grouped into two market trends, that is, rising and falling. The amplitudes of the rising and falling trends on day $t$ are defined as $v^+(t)$ and $v^-(t)$, respectively, as follows:

$$\begin{cases} v^+(t) = \sum\limits_{i, r_i(t)>0} r_i^2(t)/n_s, \\ v^-(t) = \sum\limits_{i, r_i(t)<0} r_i^2(t)/n_s. \end{cases} \tag{20}$$

Here, $n_s$ is the number of stocks in a sector, and $n_s = n$ in the calculation of $H_M$. The amplitude $v^d(t)$ of the dominating trend and the amplitude $v^n(t)$ of the non-dominating trend are

$$\begin{cases} v^d(t) = \max\{v^+(t), v^-(t)\}, \\ v^n(t) = \min\{v^+(t), v^-(t)\}. \end{cases} \tag{21}$$

Stocks grouped into the dominating trend are denoted as "dominating stocks".

To characterize the price co-movement degrees for stocks in the whole market and in sector $j$, the co-movement degrees $H_M$ and $H_j$ are computed as

$$\begin{cases} H_M = \langle \zeta(t) \rangle \cdot \langle v^d(t) - v^n(t) \rangle \big|_{\text{market}}, \\ H_j = \langle \zeta(t) \rangle \cdot \langle v^d(t) - v^n(t) \rangle \big|_{\text{sector } j}. \end{cases} \tag{22}$$

Here, $|_{\text{market}}$ and $|_{\text{sector } j}$ represent the stocks in the whole market and in the $j$-th sector, respectively. $\zeta(t)$ represents the similarity in the signs of the returns for different

**Table 2** The values of parameters $H_M$ and $H_j$ for the NYSE and HKSE.

| | $H_M$ | $H_1$ | $H_2$ | $H_3$ | $H_4$ | $H_5$ |
|---|---|---|---|---|---|---|
| NYSE | 0.363 | 0.491 | 0.414 | 0.438 | 0.431 | 0.546 |
| HKSE | 0.306 | 0.426 | 0.406 | 0.364 | 0.361 | 0.340 |

stocks, and it is defined as the percentage of the dominating stocks; that is, $\zeta(t) = n^d(t)/n_s$. $\langle v^d(t) - v^n(t)\rangle$ is the average total amplitude of the dominating stocks.

The co-movement degrees $H_M$ and $H_j$ for the NYSE and HKSE indices are shown in Table 2.

### *2.4.3 Simulation results*

Estimated from historical market data and investment reports, the buying or selling probability is $P = 0.363$ for the NYSE index and $P = 0.317$ for the HKSE index [24]. With $H_M$ and $H_j$ determined for the NYSE and HKSE indices, respectively, the model produces the time series $R_k(t)$ of each stock. The schematic diagram of the simulation procedure is displayed in Fig. 4(b).

To characterize the spatial structure, one may compute the equal-time cross-correlation matrix $C_{ij} = \langle r_i(t) r_j(t)\rangle$ [65, 71], where $\langle\cdots\rangle$ represents the average over time $t$, and $C_{ij}$ measures the correlation between the returns of the $i$-th and $j$-th stocks. The distribution of the eigenvalues is displayed for the NYSE and HKSE indices in Fig. 4, and the bulk of the distribution and three largest eigenvalues from the simulation are in agreement with those from the empirical data.

The first, second, and third largest eigenvalues of $C$ are denoted by $\lambda_0$, $\lambda_1$, and $\lambda_2$, respectively. $\lambda_0$ represents the market mode, that is, the price co-movement of the entire market, and the components of the corresponding eigenvector is rather uniform for all stocks. Other large eigenvalues stand for the sector modes, and the eigenvector of these eigenvalues is dominated by the stocks in a certain sector.

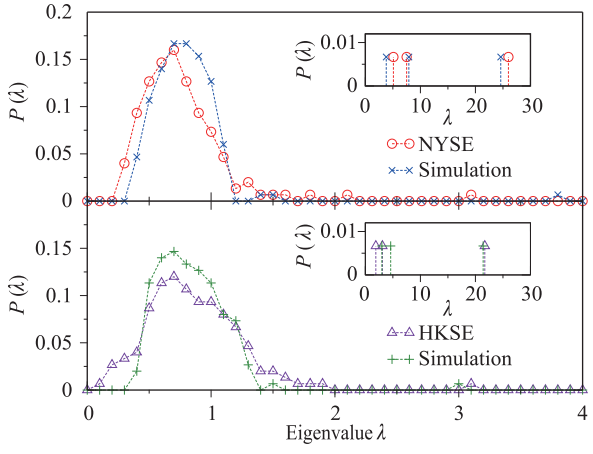


**Fig. 4** The probability distribution of the eigenvalues of the cross-correlation matrix $C$ for the NYSE and HKSE, and for the corresponding simulations. The inset shows the three largest eigenvalue for the NYSE and HKSE, and for the corresponding simulations.

The empirical result of the NYSE index is displayed in Fig. 5(a). The eigenvectors of $\lambda_1$ and $\lambda_2$ are dominated by sector (5) and sector (1), respectively, with the components significantly larger than those in other sectors. These features are reproduced in our simulation, and the results are shown in Fig. 5(b). For the HKSE index, the eigenvectors of $\lambda_1$ and $\lambda_2$ are respectively dominated by sector (1) and sector (2), and these features are also obtained [24]. From the simulated returns, we also observe volatility clustering.

## 3 Big data and agent-based modeling

Information is a leading factor in complex financial systems. In the past years, however, it has been difficult to quantify the effect of external information on financial systems due to lack of data. Our understanding of external information and its controlling effect in agent-based modeling is rather limited [27–30]. Fortunately, massive new data sources have resulted from human interactions with the internet in recent years. Therefore, we propose a novel paradigm by combining big-data analysis with agent-based modeling [51].

### 3.1 Information driving forces

Internet query data cannot only reflect the arrival of news but also provide a proxy measurement of the information gathering process of traders before their trading decisions. We collect weekly Google search volumes and corresponding historical market data for 108 components of the S&P 500. In this section, we define an information driving force and analyze how it drives the complex financial system.

The states of external information, that is, Google search volume $G_k(t)$ for the $k$-th stock, may be complicated [35, 39]. As a first approach, we simplify the information states to two states as follows:

$$S_k(t) = \begin{cases} 1, & G_k(t) > \bar{G}_k, \\ 0, & G_k(t) \le \bar{G}_k. \end{cases} \tag{23}$$

Here, $\bar{G}_k$ is the mean value of $G_k(t)$. Traders are more influenced by the external information at $S_k(t) = 1$ and less at $S_k(t) = 0$. The auto-correlation function of the time series $r(t')$ is defined as

$$A(t) = [\langle |r(t')||r(t+t')|\rangle - \langle |r(t')|\rangle^2]/A_0, \tag{24}$$

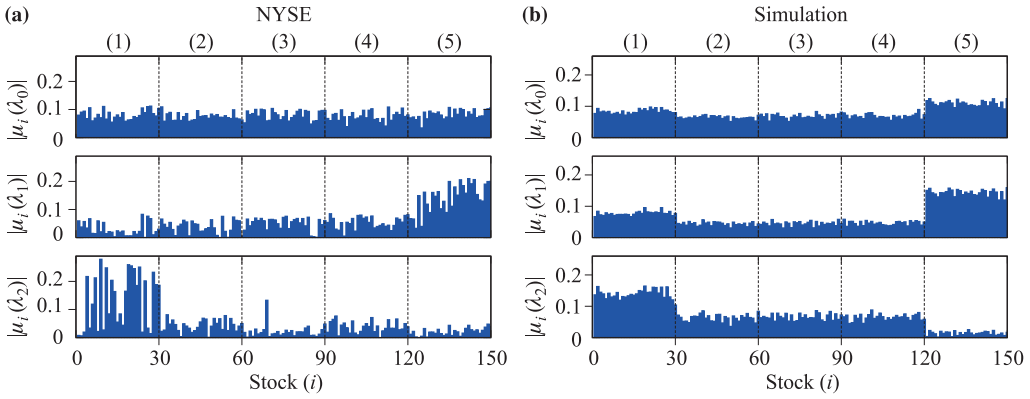


**Fig. 5** The absolute values of the eigenvector components $u_i(\lambda)$ corresponding to the three largest eigenvalues for the cross-correlation matrix $C$ calculated from **(a)** the empirical data in the NYSE; **(b)** the simulated returns for the NYSE. Stocks are arranged according to business sectors separated by dashed lines. (1) Basic Materials; (2) Consumer Goods; (3) Industrial Goods; (4) Services; (5) Utility.

where $A_0 = \langle |r(t')|^2 \rangle - \langle |r(t')| \rangle^2$ [21]. For each stock, the auto-correlation functions of $G_k(t')$, $S_k(t')$, and $V_k(t')$ are computed and averaged over $k$. As displayed in Fig. 6(a), the average auto-correlation functions of $G_k(t')$ and $S_k(t')$ exhibit a power-law-like behavior in a certain period of time. This is similar to that of $V_k(t')$. On the other hand, all three curves start deviating from the power law at about $t = 26$ weeks, which could be considered as the correlating time $\tau$.

To study how external information influences the trading behavior of traders, we calculate the moving time averages of the trading volumes in different information states. Here, we adopt the correlating time $\tau$ of the Google search volumes as the length of the moving time window. Denoting the moving time averages of the trading volumes at $S_k(t') = 1$ and $S_k(t') = 0$ by $V_k^1(t)$ and $V_k^0(t)$, respectively, one can simply compute

$$\begin{aligned} V_k^1(t) &= \langle V_k(t') \rangle_\tau \,|_{S_k(t)=1}, \\ V_k^0(t) &= \langle V_k(t') \rangle_\tau \,|_{S_k(t)=0}, \end{aligned} \tag{25}$$

where $\langle \cdots \rangle_\tau$ represents the average over the time window $t' \in (t, t+\tau)$. We then empirically define the information driving force for the $k$-th stock on time $t$ as follows:

$$\tilde{F}_k(t) = V_k^1(t)/V_k^0(t) - 1. \tag{26}$$

If $\tilde{F}_k(t) > 0$, that is, $V_k^1(t) > V_k^0(t)$, the traders trade more frequently at state $S_k(t) = 1$, and the external information does drive the market to be more active. The positive information driving forces reflect the information gathering process of the traders before their trading decisions. If $\tilde{F}_k(t) < 0$, that is, $V_k^1(t) < V_k^0(t)$, the traders trade less frequently at state $S_k(t) = 1$, and the market is not driven to be more active. The negative information driving forces may be related to ambiguous or uncertain information that does not play a key role in trading behavior. As displayed in Fig. 6(b), the probability distribution of $\tilde{F}_k(t)$ is obviously asymmetric with a heavier positive tail. This result indicates that external information usually drives the market to be more active, which is consistent with the previous empirical findings for internet query data or news [32, 35].

To study the information driving forces in different market states, we compute the average information driving forces, $\tilde{F}^{bear}$ and $\tilde{F}^{bull}$, for the bull and bear markets, respectively. Thus, their difference is defined as

$$\Delta\tilde{F} = (\tilde{F}^{bear} - \tilde{F}^{bull})/\langle \tilde{F} \rangle, \tag{27}$$

where $\langle \tilde{F} \rangle$ is the mean value of $\tilde{F}_k(t)$ for all different $t$ and $i$. The result is $\Delta\tilde{F} = 0.4$; that is, the information driving forces in the bear market are stronger than those in the bull market. The asymmetric information driving forces in the bull and bear markets indicate that traders are more sensitive in the bear market.

## 3.2 Agent-based model driven by information driving forces

### *3.2.1 Model framework*

As an application, we propose an agent-based model driven by the information driving force. We consider a stock market composed of $N$ agents, in which there is only one stock, and each agent operates one share every day. On day $t$, after all the agents have made their trading decision $\phi_i(t)$ according to Eq. (2), we can calculate the price return according to Eq. (4). We still assume $P_i^{buy}(t) = P_i^{sell}(t)$, but the trading probability of the $i$-th agent, $P_i(t) = P_i^{buy}(t) + P_i^{sell}(t)$, evolves with time.

The information driving force of the $i$-th agent in this section is denoted by $F_i(t)$, which is distinguished from the empirically-defined information driving force $\tilde{F}_k(t)$

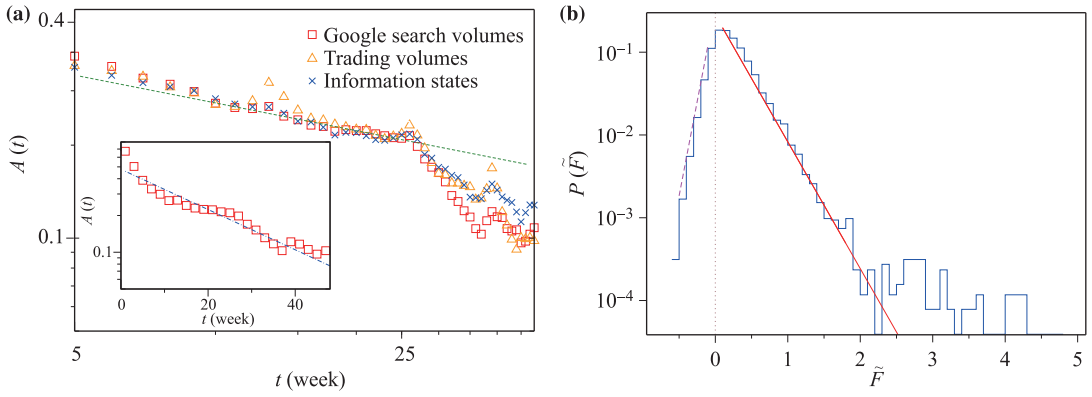


**Fig. 6** **(a)** The average auto-correlation functions of the Google search volumes, trading volumes and the information states of the S&P 500 components. A power-law fit is given by the dashed line. As shown in the inset, the curve is fitted with an exponential law $A(t) = c\exp(-t/\tau)$ with $\tau = 26$. **(b)** The probability distribution of the information driving forces for the S&P 500 components. An exponential fit $P(\tilde{F}_k) = a_1\exp(-b_1\tilde{F}_k)$ with $b_1 = 3.5$ for $\tilde{F}_k(t) > 0$ is displayed with the solid line. The dashed line corresponds to an exponential fit for $\tilde{F}_k(t) < 0$, i.e., $P(\tilde{F}_k) = a_2\exp(b_2\tilde{F}_k)$ with $b_2 = 10.5$.

of the $k$-th stock in Section 3.1. We assume that $F_i(t)$ induces a dynamic fluctuation of the trading probability,

$$P_i(t) = E(1 + F_i(t))P(0), \tag{28}$$

where $P(0)$ is the initial value of $P(t)$, and $E$ is the identity matrix. In our model, we set $P(0) = 2p/(1+\bar{F})$ to ensure the time average of the trading probabilities for the $i$-th agent $\langle P_i(t)\rangle = 2p$, where $p = 0.0154$ [20]. Here, $\bar{F}$ is the mean value of information driving forces $F_i(t)$.

### 3.2.2 Information states

As stated in Section 3.1, there are two information states for the market, that is, $S(t) = 1$ and $S(t) = 0$, and the information driving force plays an important role only at state $S(t) = 1$. Here, we omit the subscript $k$ of $S_k(t)$, as the difference between different stocks is not discussed in our model. The initial information state is randomly set to be $S(t) = 1$ or $S(t) = 0$. Then, the information state will flip between $S(t) = 1$ and $S(t) = 0$, with an average transition probability $p_t$. On average, an information state will persist for $1/p_t$. Then, we assume $p_t = 1/\tau$, where $\tau$ is the correlating time of the Google search volume for the S&P 500 components.

For simplicity, we only consider the positive information driving forces, since the negative ones are not dominant. In Section 3.1, the probability distribution of the empirically-defined information driving forces is fitted with the exponential function $Prob(\tilde{F}) \sim \exp(-b_1\tilde{F})$, with $b_1 = 3.5$. We suppose that $F_i(t)$ of different $i$ obeys the same distribution. Therefore, the simplest form of $F_i(t)$ should be

$$F_i(t) = s_i(t)y(t), \tag{29}$$

where the stochastic variable $y(t)$ obeys the distribution $Prob(y)$, and $s_i(t)$ is the state of the $i$-th agent. For each time $t$, we set the states for a dominating percentage of the agents to be $s_i(t) = S(t)$ and those for others to be $s_i(t) = 1 - S(t)$.

To describe the asymmetric trading behavior of agents in the bull and bear markets, we complete the form of $F_i(t)$ as

$$F_i(t) = s_i(t)y(t)[1 + a\cdot\mathrm{sgn}(R'(t))], \tag{30}$$

where $a$ is the asymmetric coefficient, and $R'(t)$ is the weighted return defined in Eq. (4). We assume that the asymmetric coefficient $a = \Delta\tilde{F}/2$. Here, $\Delta\tilde{F}$ is the difference between the empirically-defined information driving forces in the bull and bear markets, which is computed from Eq. (27) in Section 3.1.

The herding behavior can be explained by the information dispersion [53, 73]. The agents behave similarly because they are exposed to the same information. Here, we assume that the agents with positive information driving forces $F_i(t)$ are divided into clusters. The average number of agents in each cluster $n(t)$ should be related to the information driving force, and we set $n(t) = p_t^{-1}\Sigma_{i=1}^{N}F_i(t)/N$.

### 3.2.3 Simulation results

With the number of agents set to be $N = 10^4$, we perform the numerical simulation and obtain the return time series $R(t)$.

Our model reproduces the statistical features of real stock markets. For instance, the simulation is compared with the daily price returns of the S&P 500 components. To reduce the fluctuations, the calculations for the empirical data are averaged over all stocks. The probability distribution functions $P(|r(t)|)$ of the absolute values of

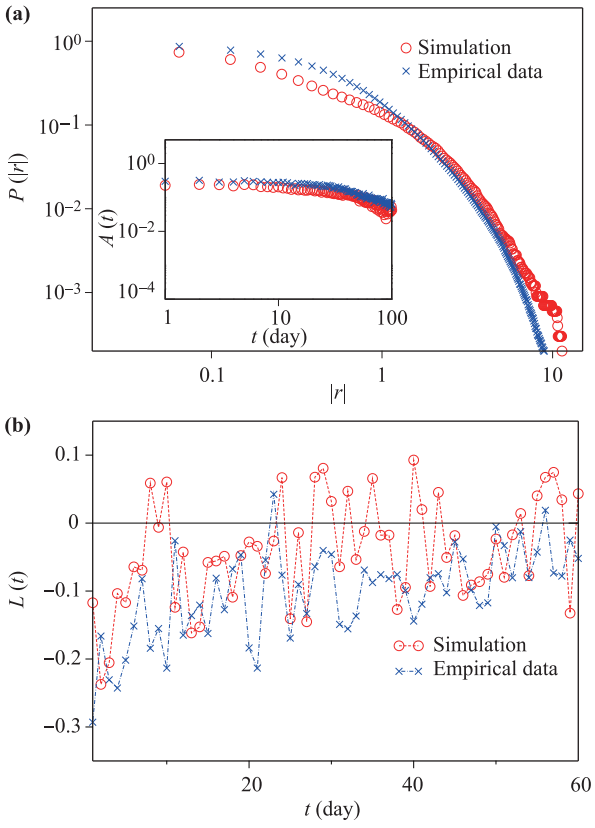


**Fig. 7** Comparison of the S&P 500 components and the simulations: **(a)** The probability distribution functions of the absolute values of returns. The auto-correlation functions of volatilities are displayed in the inset. **(b)** The return-volatility correlation functions.

returns are displayed in Fig. 7(a), and empirical fat tails are observed. The volatility clustering is characterized by the auto-correlation function of volatilities [3], which is defined in Eq. (24). As shown in the inset of figure Fig. 7(a), $A(t)$ from the simulation is in agreement with that from the empirical data.

To describe how past returns affect future volatilities, we compute the return-volatility correlation function $L(t)$ defined in Eq. (12). As displayed in Fig. 7(b), $L(t)$ from our simulation is consistent with that from empirical data.

## 4 Summary

We first review several agent-based models and new approaches to determine the key model parameters from historical market data. Based on the agents' behaviors with heterogeneous personal preferences and interactions, these models are successful in explaining the microscopic origination of the temporal and spatial correlations of financial markets. More specifically, asymmetric trading and asymmetric herding are introduced to agent-based modeling to understand the leverage and anti-leverage effects. The asymmetric trading preference in volatile and stable markets is proposed to explain the nonlocal return-volatility correlation. Finally, an agent-based model with multi-level herding is constructed to simulate the sector structure.

We then present a novel paradigm combining big-data analysis with agent-based modeling. From internet query and stock market data, we extract the information driving forces and develop an agent-based model to simulate the dynamic behaviors of complex financial systems. The key parameters of the model are determined from the statistical properties of the information driving forces. Our results provide a better understanding of the controlling effect of the information driving forces on complex financial systems. The ideology of the information driving force may be applied to agent-based modeling of other open complex systems.

**Acknowledgements** This work was supported in part by the National Natural Science Foundation of China under Grant Nos. 11375149 and 11505099.

## References

1. F. Black, Studies of stock price volatility changes, Alexandria, 1976. Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section, pp 177–181
2. R. N. Mantegna and H. E. Stanley, Scaling behavior in the dynamics of an economic index, *Nature* 376(6535), 46 (1995)
3. P. Gopikrishnan, V. Plerou, L. A. N. Amaral, M. Meyer, and H. E. Stanley, Scaling of the distribution of fluctuations of financial market indices, *Phys. Rev. E* 60(5), 5305 (1999)
4. Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C. K. Peng, and H. E. Stanley, Statistical properties of the volatility of price fluctuation, *Phys. Rev. E* 60(2), 1390 (1999)
5. X. Gabaix, P. Gopikrishnan, V. Plerou, and H. E. Stanley, A theory of power-law distributions in financial market fluctuations, *Nature* 423(6937), 267 (2003)
6. T. Qiu, B. Zheng, F. Ren, and S. Trimper, Return-volatility correlation in financial dynamics, *Phys. Rev. E* 73(6), 065103 (2006)
7. X. F. Jiang, T. T. Chen, and B. Zheng, Structure of local interactions in complex financial dynamics, *Sci. Rep.* 4, 5321 (2014)
8. B. Zheng, X. F. Jiang, and P. Y. Ni, A mini-review on econophysics: Comparative study of Chinese and western financial markets, *Chin. Phys. B* 23(7), 078903 (2014)

9. L. Tan, B. Zheng, J. J. Chen, and X. F. Jiang, How volatilities nonlocal in time affect the price dynamics in complex financial systems, *PLoS One* 10(2), e118399 (2015)
10. T. Preis, J. J. Schneider, and H. E. Stanley, Switching processes in financial markets, *Proc. Natl. Acad. Sci. USA* 108(19), 7674 (2011)
11. B. Podobnik, A. Valentinčič, D. Horvatič, and H. E. Stanley, Asymmetric Lévy flight in financial ratios, *Proc. Natl. Acad. Sci. USA* 108(44), 17883 (2011)
12. W. Li, F. Z. Wang, S. Havlin, and H. E. Stanley, Financial factor influence on scaling and memory of trading volume in stock market, *Phys. Rev. E* 84(4), 046112 (2011)
13. M. Tumminello, F. Lillo, J. Piilo, and R. N. Mantegna, Identification of clusters of investors from their real trading activity in a financial market, *New J. Phys.* 14(1), 013041 (2012)
14. M. C. Münnix, T. Shimada, R. Schäfer, F. Leyvraz, T. H. Seligman, T. Guhr, and H. E. Stanley, Identifying states of a financial market, *Sci. Rep.* 2, 644 (2012)
15. X. F. Jiang and B. Zheng, Anti-correlation and subsector structure in financial systems, *Europhys. Lett.* 97(4), 48006 (2012)
16. X. F. Jiang, T. T. Chen, and B. Zheng, Time-reversal asymmetry in financial systems, *Physica A* 392(21), 5369 (2013)
17. Y. Yura, H. Takayasu, D. Sornette, and M. Takayasu, Financial Brownian particle in the layered order-book fluid and fluctuation-dissipation relations, *Phys. Rev. Lett.* 112(9), 098703 (2014)
18. A. Shleifer, Inefficient markets: An introduction to behavioral finance, *J. Inst. & Theor. Econ.* 158(2), 369 (2002)
19. H. Jo and D. M. Kim, Recent development of behavioral finance, *Int. J. Bus. Res.* 8(2), 89 (2008)
20. L. Feng, B. Li, B. Podobnik, T. Preis, and H. E. Stanley, Linking agent-based models and stochastic models of financial markets, *Proc. Natl. Acad. Sci. USA* 09(22), 8388 (2012)
21. J. J. Chen, B. Zheng, and L. Tan, Agent-based model with asymmetric trading and herding for complex financial systems, *PLoS One* 8(11), 79531 (2013)
22. V. Gontis and A. Kononovicius, Consentaneous agent based and stochastic model of the financial markets, *PLoS One* 9(7), 102201 (2014)
23. Y. Shapira, Y. Berman, and E. B. Jacob, Modelling the short term herding behaviour of stock markets, *New J. Phys.* 16(5), 53040 (2014)
24. J. J. Chen, B. Zheng, and L. Tan, Agent-based model with multi-level herding for complex financial systems, *Sci. Rep.* 5, 8399 (2015)
25. T. Kaizoji, M. Leiss, A. Saichev, and D. Sornette, Superexponential endogenous bubbles in an equilibrium model of fundamentalist and chartist traders, *J. Econ. Behav. Organ.* 112, 289 (2015)
26. R. Savona, M. Soumare, and J. V. Andersen, Financial symmetry and moods in the market, *PLoS One* 10(4), 0118224 (2015)
27. E. Samanidou, E. Zschischang, D. Stauffer, and T. Lux, Agent-based models of financial markets, *Rep. Prog. Phys.* 70(3), 409 (2007)
28. R. N. Mantegna and J. Kertész, Focus on statistical physics modeling in economics and finance, *New J. Phys.* 13(2), 25011 (2011)
29. A. Chakraborti, I. M. Toke, M. Patriarca, and F. Abergel, Econophysics review (II): Agent-based models, *Quant. Finance* 11(7), 1013 (2011)
30. D. Sornette, Physics and financial economics (1776–2014): Puzzles, Ising and agent-based models, *Rep. Prog. Phys.* 77(6), 62001 (2014)
31. T. Preis, H. S. Moat, H. E. Stanley, and S. R. Bishop, Quantifying the advantage of looking forward, *Sci. Rep.* 2, 350 (2012)
32. I. Bordino, S. Battiston, G. Caldarelli, M. Cristelli, A. Ukkonen, and I. Weber, Web search queries can predict stock market volumes, *PLoS One* 7(7), 40014 (2012)
33. T. Preis, H. S. Moat, and H. E. Stanley, Quantifying trading behavior in financial markets using google trends, *Sci. Rep.* 3, 1684 (2013)
34. H. S. Moat, C. Curme, A. Avakian, D. Y. Kenett, H. E. Stanley, and T. Preis, Quantifying wikipedia usage patterns before stock market moves, *Sci. Rep.* 3, 1801 (2013)
35. R. Hisano, D. Sornette, T. Mizuno, T. Ohnishi, and T. Watanabe, High quality topic extraction from business news explains abnormal financial market volatility, *PLoS One* 8(6), 64846 (2013)
36. L. Kristoufek, Can google trends search queries contribute to risk diversification, *Sci. Rep.* 3, 2713 (2013)
37. T. Noguchi, N. Stewart, C. Y. Olivola, H. S. Moat, and T. Preis, Characterizing the time-perspective of nations with search engine query data, *PLoS One* 9(4), e95209 (2014)
38. C. Curme, T. Preis, H. E. Stanley, and H. S. Moat, Quantifying the semantics of search behavior before stock market moves, *Proc. Natl. Acad. Sci. USA* 111(32), 11600 (2014)
39. F. Lillo, S. Micciche, M. Tumminello, J. Piilo, and R. N. Mantegna, How news affects the trading behaviour of different categories of the investors in a financial market, *Quant. Finance* 15(2), 213 (2015)
40. I. Giardina, J. P. Bouchaud, and M. Mézard, Microscopic models for long ranged volatility correlations, *Physica A* 299(1–2), 28 (2001)

41. E. Bonabeau, Agent-based modeling: methods and techniques for simulating human systems, *Proc. Natl. Acad. Sci. USA* 99(Suppl. 3), 7280 (2002)
42. T. P. Evans and H. Kelley, Multi-scale analysis of a household level agent-based model of landcover change, *J. Environ. Manage.* 72(1–2), 57 (2004)
43. F. Ren, B. Zheng, T. Qiu, and S. Trimper, Minority games with score-dependent and agent-dependent payoffs, *Phys. Rev. E* 74(4), 041111 (2006)
44. J. D. Farmer and D. Foley, The economy needs agent based modelling, *Nature* 460(7256), 685 (2009)
45. F. Schweitzer, G. Fagiolo, D. Sornette, F. V. Redondo, A. Vespignani, and D. R. White, Economic networks: The new challenges, *Science* 325, 422 (2009)
46. S. Mike and J. D. Farmer, An empirical behavioral model of liquidity and volatility, *J. Econo. Dyn. Contr.* 32(1), 200 (2008)
47. G. F. Gu and W. X. Zhou, On the probability distribution of stock returns in the mike-farmer model, *Eur. Phys. J. B* 67(4), 585 (2009)
48. G. F. Gu and W. X. Zhou, Emergence of long memory in stock volatility from a modified mike-farmer model, *Europhys. Lett.* 86, 48002 (2009)
49. H. Meng, F. Ren, G. F. Gu, X. Xiong, Y. J. Zhang, W. X. Zhou, and W. Zhang, Effects of long memory in the order submission process on the properties of recurrence intervals of large price fluctuations, *Europhys. Lett.* 98(3), 38003 (2012)
50. J. Zhou, G. F. Gu, Z. Q. Jiang, X. Xiong, W. Chen, W. Zhang, and W. X. Zhou, Computational experiments successfully predict the emergence of autocorrelations in ultrahigh-frequency stock returns, *Comput. Econ.* (2016) (in press)
51. T. T. Chen, B. Zheng, and Y. Li, Information driving forces and agent-based modelling (submitted)
52. L. Menkhoff, The use of technical analysis by fund managers: International evidence, *J. Bank. Finance* 34(11), 2573 (2010)
53. V. M. Eguíluz and M. G. Zimmermann, Transmission of information and herd behavior: An application to financial markets, *Phys. Rev. Lett.* 85(26), 5659 (2000)
54. D. Y. Kenett, Y. Shapira, A. Madi, S. Bransburg-Zabary, G. Gur-Gershgoren, and E. Ben-Jacob, Index cohesive force analysis reveals that the US market became prone to systemic collapses since 2002, *PLoS One* 6(4), e19378 (2011)
55. J. Shen and B. Zheng, On return-volatility correlation in financial dynamics, *Europhys. Lett.* 88(2), 28003 (2009)
56. B. J. Park, Asymmetric herding as a source of asymmetric return volatility, *J. Bank. Finance* 35(10), 2657 (2011)
57. K. A. Kim and J. R. Nofsinger, Institutional herding, business groups, and economic regimes: Evidence from Japan, *J. Bus.* 78(1), 213 (2005)
58. A. Walter and F. M. Weber, Herding in the German mutual fund industry, *Eur. Financ. Manag.* 12(3), 375 (2006)
59. R. Cont and J. P. Bouchaud, Herd behavior and aggregate fluctuations in financial markets, *Macroecon. Dyn.* 4(02), 170 (2000)
60. N. Blasco, P. Corredor, and S. Ferreruela, Does herding affect volatility? Implications for the Spanish stock market, *Quant. Finance* 12(2), 311 (2012)
61. J. P. Bouchaud, A. Matacz, and M. Potters, Leverage effect in financial markets: The retarded volatility model, *Phys. Rev. Lett.* 87(22), 228701 (2001)
62. Y. H. Shao, G. F. Gu, Z. Q. Jiang, W. X. Zhou, and D. Sornette, Comparing the performance of fa, dfa and dma using different synthetic long-range correlated time series, *Sci. Rep.* 2, 5225 (2012)
63. P. Gopikrishnan, M. Meyer, L. A. N. Amaral, and H. E. Stanley, Inverse cubic law for the distribution of stock price variations, *Eur. Phys. J. B* 3(2), 139 (1998)
64. G. F. Gu, W. Chen, and W. X. Zhou, Empirical distributions of Chinese stock returns at different microscopic timescales, *Physica A* 387(2–3), 495 (2008)
65. V. Plerou, P. Gopikrishnan, L. A. Nunes Amaral, M. Meyer, and H. E. Stanley, Scaling of the distribution of price fluctuations of individual companies, *Phys. Rev. E* 60(6), 6519 (1999)
66. G. H. Mu and W. X. Zhou, Tests of nonuniversality of the stock return distributions in an emerging market, *Phys. Rev. E* 82(6), 066103 (2010)
67. V. Plerou, P. Gopikrishnan, and H. E. Stanley, Two-phase behaviour of financial markets, *Nature* 421(6919), 130 (2003)
68. V. Plerou, P. Gopikrishnan, X. Gabaix, and H. E. Stanley, Quantifying stock-price response to demand fluctuations, *Phys. Rev. E* 66, 027104 (2002)
69. A. Utsugi, K. Ino, and M. Oshikawa, Random matrix theory analysis of cross correlations in financial markets, *Phys. Rev. E* 70, 026110 (2004)
70. R. K. Pan and S. Sinha, Self-organization of price fluctuation distribution in evolving markets, *Europhys. Lett.* 77(5), 58004 (2007)
71. J. Shen and B. Zheng, Cross-correlation in financial dynamics, *Europhys. Lett.* 86(4), 48005 (2009)
72. B. Podobnik, D. Wang, D. Horvatic, I. Grosse, and H. E. Stanley, Time-lag cross-correlations in collective phenomena, *Europhys. Lett.* 90(6), 68001 (2010)
73. L. Corazzini and B. Greiner, Herding, social preferences and (non-)conformity, *Econ. Lett.* 97(1), 74 (2007)