Physica A 496 (2018) 593–601

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Information driving force and its application in agent-based modeling

Ting-Ting Chen [a,b], Bo Zheng [a,b,*], Yan Li [a,b], Xiong-Fei Jiang [c,a,**]

[a] Department of Physics, Zhejiang University, Hangzhou 310027, PR China
[b] Collaborative Innovation Center of Advanced Microstructures, Nanjing 210093, PR China
[c] School of Information Engineering, Ningbo Dahongying University, Ningbo 315175, PR China

HIGHLIGHTS

- We define an information driving force based on the 'big data' in the public media.
- An agent-based model driven by the information driving force is proposed.
- The stationary and non-stationary properties of the financial markets are simulated.
- We propose a few-body model to simulate the financial market in the laboratory.

ARTICLE INFO

Article history:
Received 16 June 2017
Received in revised form 1 November 2017
Available online 30 December 2017

Keywords:
Complex systems
Econophysics
Information driving force
Agent-based modeling

ABSTRACT

Exploring the scientific impact of online big-data has attracted much attention of researchers from different fields in recent years. Complex financial systems are typical open systems profoundly influenced by the external information. Based on the large-scale data in the public media and stock markets, we first define an information driving force, and analyze how it affects the complex financial system. The information driving force is observed to be asymmetric in the bull and bear market states. As an application, we then propose an agent-based model driven by the information driving force. Especially, all the key parameters are determined from the empirical analysis rather than from statistical fitting of the simulation results. With our model, both the stationary properties and non-stationary dynamic behaviors are simulated. Considering the mean-field effect of the external information, we also propose a few-body model to simulate the financial market in the laboratory.

1. Introduction

Open complex systems are many-body systems whose interplay with the external environment should not be ignored. It is well known in physics that the external force plays a crucial role in dealing with the open systems. As an important example of open complex systems, the financial market is substantially influenced by the external information. However, our understanding of the external information and its controlling effects in the agent-based modeling is still limited [1–4]. Furthermore, the interactions among the agents may arise from or be influenced by their response to the external information. On the other hand, it is rather challenging in the laboratory to capture the statistical features of the financial

* Corresponding author at: Department of Physics, Zhejiang University, Hangzhou 310027, PR China.
** Corresponding author at: School of Information Engineering, Ningbo Dahongying University, Ningbo 315175, PR China.
E-mail addresses: zhengbo@zju.edu.cn (B. Zheng), xfjiang@nbdhyu.edu.cn (X.-F. Jiang).

https://doi.org/10.1016/j.physa.2017.12.128

system only with the internal interactions. This may be due to the small number of experimental subjects. Thus the external information shows its very importance in the human experiments [5–8].

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 the external information on the complex financial system [9–17]. For example, the trading behavior of the traders can be quantified with the Google search volumes and Wikipedia topic view times [11,12]. The web search volumes and news are positively correlated with the trading volumes [10,13]. The time perspective of the traders can be investigated with the Google query data [9,15]. Although the agent-based modeling has become one of the key tools to study the financial system [18,19], the external information is usually either ignored or simply treated as the exogenous noises in previous agent-based models [20–27].

To understand the financial market comprehensively, it is also important to investigate the non-stationary dynamic properties [28–30]. Phenomenologically, the dynamic relaxation processes both before and after the large fluctuations are characterized by a power law, and are asymmetric at the daily time scale [31,32]. It has been suggested that the external forces could drive the financial system to a non-stationary state and induce the time-reversal asymmetry of the dynamic relaxation processes. With our model, we reproduce not only the stationary statistical properties, but also the non-stationary dynamic behavior of the financial system. In Section 2, we analyze the information driving force based on the web query data. In Section 3, we construct an agent-based model driven by the information driving force. In Section 4, taking the mean-field effect of the external information into account, we propose a few-body model to simulate the financial market in the laboratory.

2. Empirical information driving forces

2.1. Data

From Google Trends (http://trends.google.com), we retrieve the weekly Google search volumes for the S&P 500 components from September 7, 2008 to July 19, 2014. The searching keywords are in the form of "ticker symbol + stock". The ticker symbol is used to ensure that the searched term is related to the stock market [14]. The search volumes provided by Google Trends are given as the adjusted ratios to the total number of searches on Google in the corresponding time interval. The zero search volume actually is not zero, but below a threshold.

Further, we collect the historical market data, such as the daily and weekly closing prices and the weekly trading volumes of the S&P 500 components in the same period from Yahoo! Finance (http://finance.yahoo.com). We exclude the stocks which report either zero search volumes frequently, or no available market data in the corresponding time period. Finally, there are 108 stocks.

At time t, the Google search volume, trading volume and closing price for the ith stock are denoted by $G_i(t)$, $V_i(t)$ and $Y_i(t)$ respectively. The logarithmic price return is defined as

$$R_i(t) = \ln[Y_i(t)/Y_i(t - \Delta t)], \tag{1}$$

where Δt is the time interval. For the daily closing prices, $\Delta t = 1$ day, and for the weekly closing prices, $\Delta t = 1$ week. For simplicity, the volatility is defined as the absolute value of the return in this paper. The calculations in Section 2 are based on the weekly Google search volumes, trading volumes and closing prices for the S&P 500 components. In Section 3, the daily closing prices are treated as comparisons.

2.2. Definitions and results

The complex financial system is an open system with many-body interactions [33–36], and its interplay with the external environment should not be ignored. We define an information driving force based on the web query data combining with the stock market data, and analyze how it drives the financial system. The web query data can not only reflect the arrival of news, but also provide a proxy measurement of the information gathering process of the investors before their trading decisions. In this section, the calculations are based on the weekly Google search volumes, and the corresponding historical market data for the S&P 500 components.

The states of the external information, i.e., the Google search volume $G_i(t)$, may be complicated [13,17]. Here, the information states are classified into two states, i.e.,

$$S_i(t) = \begin{cases} 1 & G_i(t) > \bar{G}_i \\ 0 & G_i(t) \le \bar{G}_i, \end{cases} \tag{2}$$

where $\bar{G}_i$ is the mean value of $G_i(t)$ for the ith stock. The investors are more influenced by the external information at $S_i(t) = 1$, while less at $S_i(t) = 0$.

For the comparison of different time series, we normalize a time series $Q(t')$ by

$$q(t') = [Q(t') - \langle Q(t')\rangle]/\sigma, \tag{3}$$

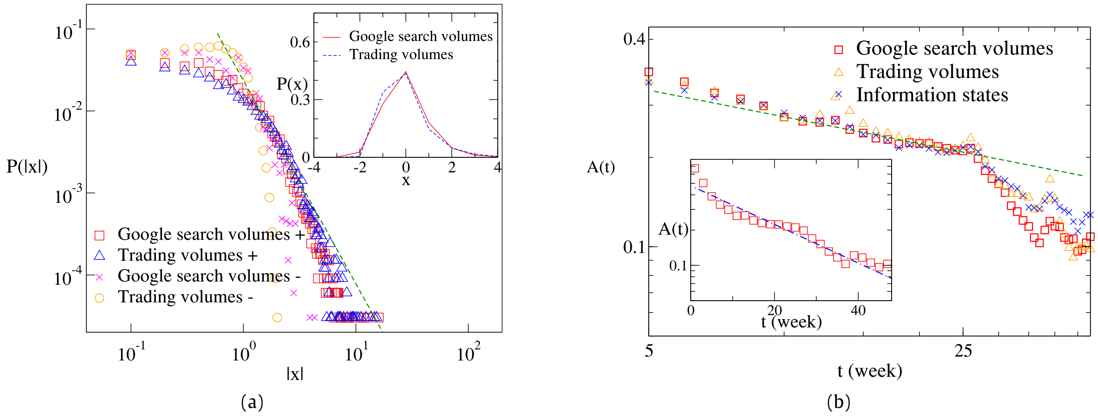

Fig. 1. (a) The probability distributions of the Google search volumes and trading volumes of the S&P 500 components. The dashed line corresponds to a power law fit with the slope $= 3.5$. (b) The average auto-correlation functions of the Google search volumes, trading volumes and the information states of the S&P 500 components. The curve for the trading volumes is shifted down slightly. A power-law fit with the slope $= 0.30$ is given by the dashed line. As shown in the dot-dashed line of the inset, we also fit the curve with an exponential law, i.e., $A(t) = c\exp(-t/\tau)$ with $\tau = 26$.

where $\langle \cdots \rangle$ denotes the average over time t', and $\sigma = \sqrt{\langle Q(t')^2\rangle - \langle Q(t')\rangle^2}$ is the standard deviation of $Q(t')$. The normalized variables for $G_i(t')$, $V_i(t')$, and $R_i(t')$ are denoted by $g_i(t')$, $v_i(t')$ and $r_i(t')$ respectively. As displayed in Fig. 1(a), the probability distributions of both the Google search volumes and the trading volumes are roughly symmetric with power-law-like tails. The similar scaling phenomena of $g_i(t')$ and $v_i(t')$ cast a new insight on the relationship between the external information and the trading behavior of the investors. To investigate the temporal dynamics, we define the auto-correlation function of the variable $q(t')$ as

$$A(t) = [\langle q(t') \cdot q(t'+t)\rangle]/A_0, \tag{4}$$

where $A_0 = \langle q(t')^2\rangle$. For each stock, the auto-correlation functions of $g_i(t')$, $s_i(t')$ and $v_i(t')$ are computed. Then we average the auto-correlation functions over i. As displayed in Fig. 1(b), the average auto-correlation functions of $g_i(t')$ and $s_i(t')$ exhibit a power-law-like behavior in a certain period of time. This is similar to that of $v_i(t')$. On the other hand, all the three curves start deviating from the power law at about $t = 26$ weeks. A more careful analysis such as fitting the curve to an exponential law shows that the correlating time of the Google search volumes is just about $\tau = 26$ weeks, which represents the approximate persisting time of the information states.

To investigate the influence of external information on the trading behavior of the investors, we compute the moving time averages of the trading volumes in different information states. Here we adopt the correlating time τ of the Google search volumes as the length of the moving time window. Since τ represents the approximate persisting time of the information states, the average values in the moving time window $t' \in (t, t+\tau)$ should capture the characteristic dynamic behavior of the stock market. In fact, the results are robust for τ in the range from 20 to 35 weeks. Denoting the moving time averages of the trading volumes at $S_i(t') = 1$ and $S_i(t') = 0$ by $V_i^1(t)$ and $V_i^0(t)$ respectively, one can simply compute

$$\begin{aligned} V_i^1(t) &= \langle V_i(t')\rangle_\tau |_{S_i(t)=1}, \\ V_i^0(t) &= \langle V_i(t')\rangle_\tau |_{S_i(t)=0}, \end{aligned} \tag{5}$$

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 ith stock on time t

$$\tilde{F}_i(t) = V_i^1(t)/V_i^0(t) - 1. \tag{6}$$

If $\tilde{F}_i(t) > 0$, i.e., $V_i^1(t) > V_i^0(t)$, the investors trade more frequently at the state $S_i(t) = 1$, and the external information does drive the market to be more active. The positive information driving forces reflect the information gathering process of the investors before their trading decisions. If $\tilde{F}_i(t) < 0$, i.e., $V_i^1(t) < V_i^0(t)$, the investors trade less frequently at the state $S_i(t) = 1$, and the market is not driven to be more active. The negative information driving forces may be related to the ambiguous or uncertain information which does not play a key role in the trading behavior. As displayed in Fig. 2, both the positive and negative $\tilde{F}_i(t)$ are observed. The probability distribution of $\tilde{F}_i(t)$ is obviously asymmetric with a heavier positive tail. Particularly, the maximum positive $\tilde{F}_i(t)$ is up to 4.8, while no negative $\tilde{F}_i(t)$ is smaller than -0.6. This result indicates that the external information usually drives the market to be more active, which is consistent with the previous empirical findings for the web query data or news [13,10].

$R_i'(t)$ is defined as the moving time average of $R_i(t)$ over the time window $t' \in (t, t+\tau)$. Similarly, the moving time average of $V_i(t)$ is denoted by $V_i'(t)$. The stock market is defined to be in a bullish state if $R_i'(t) > 0$, and in a bearish one if

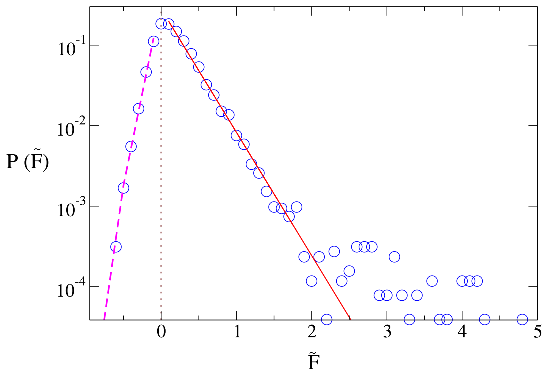

Fig. 2. The probability distribution of the information driving forces for the S&P 500 components. An exponential fit for $\tilde{F}_i(t) > 0$ is displayed in the solid line, i.e., $P(\tilde{F}_i) = a_1 \exp(-b_1\tilde{F}_i)$ with $b_1 = 3.5$. The dashed line corresponds to an exponential fit for $\tilde{F}_i(t) < 0$, i.e., $P(\tilde{F}_i) = a_2 \exp(b_2\tilde{F}_i)$ with $b_2 = 10.5$.

$R_i'(t) < 0$. To investigate the information driving forces in different market states, we define the average information driving forces $\tilde{F}^{bear}$ and $\tilde{F}^{bull}$ for the bull and bear markets respectively,

$$\begin{aligned}\tilde{F}^{bear} &= \frac{1}{N_{R_i'(t)<0}} \sum_{t,i} \tilde{F}_i(t)|_{R_i'(t)<0},\\ \tilde{F}^{bull} &= \frac{1}{N_{R_i'(t)>0}} \sum_{t,i} \tilde{F}_i(t)|_{R_i'(t)>0}.\end{aligned} \tag{7}$$

Here $N_{R_i'(t)>0}$ and $N_{R_i'(t)<0}$ are the numbers of positive and negative $R_i'(t)$ respectively. Thus the difference of the average information driving forces $\tilde{F}^{bear}$ and $\tilde{F}^{bull}$ for the bull and bear market states is computed,

$$\Delta\tilde{F} = (\tilde{F}^{bear} - \tilde{F}^{bull})/\langle\tilde{F}\rangle, \tag{8}$$

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

3. Agent-based herding model driven by information driving forces

3.1. Model framework

As an application, we propose an agent-based model driven by the information driving force in this section. In our model, there are N agents, and each agent operates one share at time t. Because of the highly incomplete information for the investors in the real market, the decisions of agents to buy, sell or hold are assumed to be random. At time t, the m-agent makes a trading decision

$$\psi_m(t) = \begin{cases} 1 & \text{buy} \\ -1 & \text{sell} \\ 0 & \text{hold.} \end{cases} \tag{9}$$

The probabilities for the mth agent to buy, sell and hold at time t are denoted by $p_m^{buy}(t)$, $p_m^{sell}(t)$ and $p_m^{hold}(t)$ respectively, and $p_m^{hold}(t)+p_m^{buy}(t)+p_m^{sell}(t) = 1$. In Ref. [20], it is assumed that $p_m^{buy}(t) = p_m^{sell}(t) = p$, and p is a constant estimated to be 0.0154. In our model, we still assume $p_m^{buy}(t) = p_m^{sell}(t)$, but the trading probability of the mth agent $p_m(t) = p_m^{buy}(t) + p_m^{sell}(t)$ evolves with time.

The price return $R(t)$ in our model is defined as the difference between the demand and supply of the stock, i.e., the difference between the numbers of agents buying and selling the stock,

$$R(t) = \sum_{m=1}^{N} \psi_m(t). \tag{10}$$

Based on the results of the empirically-defined driving forces in the previous section, we introduce the information driving force $F_m(t)$ of the mth agent in our model. We assume that $F_m(t)$ induces a dynamic fluctuation of the trading probability $p_m(t)$,

$$p_m(t) = [1 + F_m(t)]p(0), \tag{11}$$

where $p(0)$ is the initial value of $p(t)$. We set $p(0) = 2p/(1+\bar{F})$ to ensure the time average of the trading probabilities for the mth agent $\langle p_m(t)\rangle = 2p$. Here $\bar{F}$ is the mean value for the distribution of the information driving forces $F_m(t)$.

The simplest form of $F_m(t)$ should be

$$F_m(t) = s_m(t)\eta(t), \tag{12}$$

where $s_m(t)$ is the state of the mth agent, and $\eta(t)$ is a stochastic variable. According to Section 3, we suppose $\eta(t)$ obeys the probability distribution $P(\eta) \sim \exp(-\alpha\eta)$ with $\alpha = 3.5$. We only consider the positive information driving forces, since the negative ones are not dominating.

As mentioned in Section 3, there are two information states for the market, i.e., $S(t) = 1$ and $S(t) = 0$. The information driving force plays an important role only at the state $S(t) = 1$. The information state will flip between $S(t) = 1$ and $S(t) = 0$ with an average transition probability $p_t = 1/\tau$, where τ is the correlating time of the Google search volume. Each time t, we set the states for a dominating percentage of the agents to be $s_m(t) = S(t)$, and the states for the others to be $s_m(t) = 1 - S(t)$. This percentage will impact the trading probability, and thus influence the simulation results. The simulation results, however, remain robust when the percentage is "dominating", e.g., above 80%. In our simulations, the percentage is set to be 95%.

The investment horizon is defined for a better description of agents' behavior based on the previous stock performance of different time scales [21]. The integrated investment basis $R'(t)$ is defined as

$$R'(t) = \sum_{l=1}^{M}\left[\gamma_l \sum_{j=0}^{l-1} R(t-j)\right]. \tag{13}$$

Here $\gamma_l \propto l^{-1.12}$ represents the relative portion of the investors with a l-days investment horizon, and $\Sigma_{j=0}^{l-1}R(t-j)$ represents their simplified basis for estimating the previous price movements [20]. The maximum investment horizon is denoted by M. In our model we estimate M to be τ, which is the correlating time of Google search volumes. For M between 50 and 500, the simulation results remain qualitatively robust.

The stock market is assumed to be in a bullish state if $R'(t) > 0$, and in a bearish one if $R'(t) < 0$. To describe the asymmetric trading behavior of agents in the bull and bear markets, we complete the form of $F_m(t)$,

$$F_m(t) = s_m(t)\eta(t)[1 + c \cdot sign(R'(t-1))], \tag{14}$$

where c is the asymmetric coefficient. We assume that the asymmetric coefficient $c = \Delta\tilde{F}/2$. Here $\Delta\tilde{F}$ is the difference of the empirically-defined information driving forces in the bull and bear markets, which is computed from Eq. (8).

The herding behavior can be accounted for by the information dispersion [37,38]. The agents behave similarly since they are exposed to the same information. We assume that the agents whose information driving force $F_m(t) > 0$ are divided into groups. The average number of agents in each group $n(t)$ should be related to the information driving force, and we set $n(t) = p_t^{-1}\Sigma_{m=1}^{N}F_m(t)/N$. Here N is the number of agents.

3.2. Simulation

With the number of agents set to be $N = 10{,}000$ in our model, we perform the numerical simulation in the following procedures.

(1) Initially, $R(t)$ of τ days are set to be zero.
(2) On day t, we first set the information state $S(t)$ of the market and the state $s_m(t)$ of each agent.
(3) Then we compute $R'(t-1)$ with Eq. (13) and $F_m(t)$ with Eq. (14). Thus $p_m(t)$ can be calculated with Eq. (11).
(4) The agents whose $F_m(t) > 0$ are divided into $N/n(t)$ groups. The agents in the same group will adopt the same trading choice, i.e., buy, sell, or hold.
(5) After all agents have made the trading decisions with the probability $p_m(t)$ according to Eq. (9), the return $R(t)$ is computed with Eq. (10).

Repeating procedures (1)–(4), we obtain the return time series $R(t)$. The first 10,000 data points are abandoned for equilibration, and then 10,000 data points of $R(t)$ are collected.

To reduce the fluctuations, the calculations for the empirical data are averaged over different stocks. The price return $R(t)$ is normalized to $r(t)$ according to Eq. (3). The model reproduces the statistical features of the real stock markets. The cumulative probability distributions $P(|r| > x)$ of the absolute values of price returns are displayed in Fig. 3(a), and the fat tails are observed. The volatility clustering is characterized by the auto-correlation function $A(t)$ of volatilities, seen in Eq. (4). As shown in Fig. 3(b), $A(t)$ from the simulation is in agreement with that from the empirical data. The hurst exponent of $A(t)$ is calculated to be 0.90, which also indicates the long-range correlation of volatilities [39].

To describe how past returns affect future volatilities, the return-volatility correlation function $L(t)$ is computed [40],

$$L(t) = [\langle r(t')|r(t'+t)|^2\rangle]/Z, \tag{15}$$

where $Z = \langle |r(t')|^2\rangle^2$. The average $L(t)$ over the S&P 500 components shows a negative correlation, which is the well-known leverage effect [41]. As displayed in Fig. 4, $L(t)$ from our simulation is in agreement with that from empirical data.

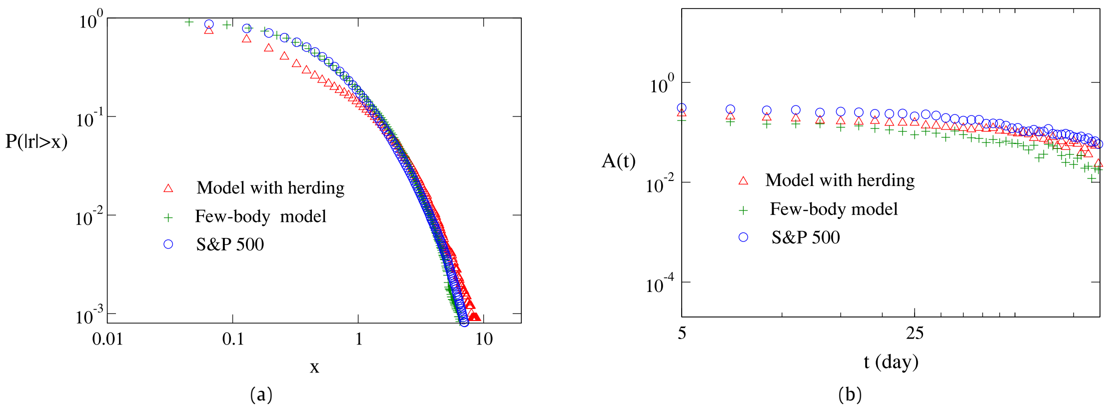

Fig. 3. (a) The cumulative probability distributions of the absolute values of returns for the S&P 500 components, and for the simulation. (b) The auto-correlation functions of volatilities for the S&P 500 components, and for the simulation.

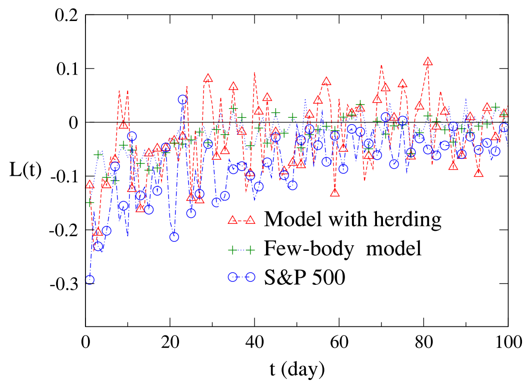

Fig. 4. The return-volatility correlation functions for the S&P 500 components and for the simulation.

3.3. *Non-stationary relaxation processes*

It is important to investigate the non-stationary dynamic properties to understand the financial market comprehensively. A typical example is the so-called financial crash. It has been suggested that the external forces could drive the financial system to a non-stationary state and induce the time-reversal asymmetry of the dynamic relaxation processes [31,42].

To study the dynamic relaxation processes after and before the large fluctuations, we introduce the remanent and anti-remanent volatilities,

$$v_c^+(t) = [\langle r(t'+t)|\rangle_c - \langle |r(t')|\rangle]/Z, \tag{16}$$

$$v_c^-(t) = [\langle r(t'-t)|\rangle_c - \langle |r(t')|\rangle]/Z, \tag{17}$$

where $Z = \langle |r(t')|\rangle_c - \langle |r(t')|\rangle$. Here $\langle\cdots\rangle_c$ represents the average over those t' with specified large volatilities, and $\langle\cdots\rangle$ represents the average over time t'. The large volatilities are selected by the condition $|r(t')| > \zeta$, and the threshold ζ is well above σ, i.e., the standard deviation of $r(t')$. When ζ is sufficiently large, the selected events correspond to the financial crashes or rallies. To reduce the fluctuations, the cumulative function of $v_c^{\pm}(t)$ is computed as

$$V_c^{\pm}(t) = \sum_{t''=1}^{t} v_c^{\pm}(t''). \tag{18}$$

We collect the daily closing prices of the S&P 500 index from November 27, 1974 to July 18, 2014. The previous study shows that the dynamic relaxation processes both before and after the large fluctuations are characterized by the power law and are time-reversal asymmetric at the daily time scale [31]. Especially, the time-reversal asymmetry and ζ-dependent $V_c^{\pm}(t)$ are mainly induced by exogenous events [31,32]. With our model, we generate 100,000 data points in the simulation. As displayed in Fig. 5, the simulation results are consistent with those of the empirical data. Particularly, we reproduce the

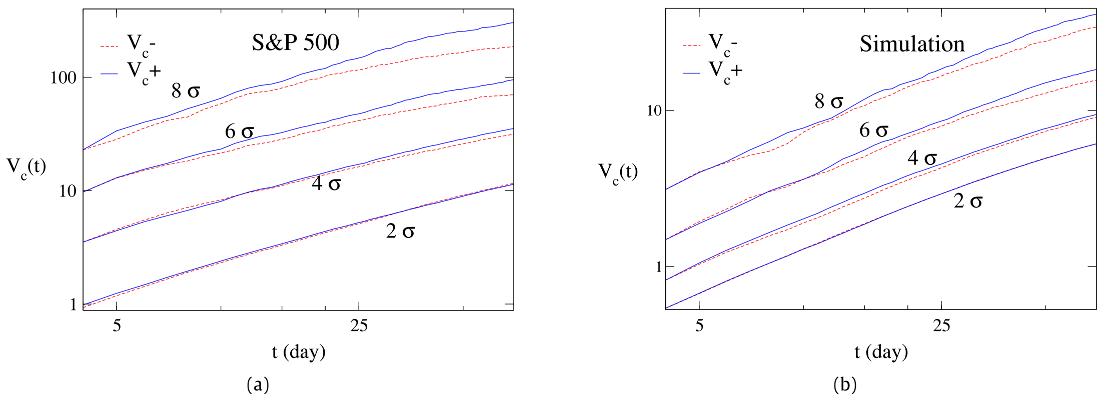

Fig. 5. The cumulative remanent and anti-remanent volatilities for (a) the S&P 500 index and for (b) the simulation. The threshold is $\zeta = 2\sigma, 4\sigma, 6\sigma$, and 8σ respectively. For clarity, the curves are shifted up or down.

time-reversal asymmetry and ζ-dependent $V_c^{\pm}(t)$, which are the characteristics of the non-stationary dynamics. Here we emphasize that, driven by the information driving force, our model naturally reproduces the exogenous events. This implies that the web query data reflect the arrival of news for the exogenous events.

To summarize, our model reproduces not only the well-known stationary statistical properties of the real markets, but also the non-stationary relaxation processes after and before the large fluctuations. These results provide a better understanding of the controlling effects of the information driving force on the financial system, and illustrate the potential of combining the interactions among the agents with their response to the external information. The ideology of the information driving force may be applied to the agent-based modeling of other open complex systems.

4. Mean-field effect of information driving forces

How to reproduce the statistical properties of financial markets in the laboratory is an appealing topic. However, due to the limitations of laboratory conditions, it is difficult to recruit very many subjects in the experiments. In the laboratory, therefore, herding is usually controlled by the external information that represents a kind of mean-field effects [5–8]. Now we would like to take the mean-field effect of information driving forces into account, and propose a few-body model to simulate the financial market in the laboratory.

Let us denote agents' trading behavior on day t by a vector

$$p(t) = \begin{bmatrix} p_{buy}(t) \\ p_{sell}(t) \end{bmatrix}. \tag{19}$$

The probabilities for agents to buy and sell are assumed to be constants and equal to each other when there is no external information driving force. The initial trading behavior is denoted by

$$p(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} p_0. \tag{20}$$

Involving the herding effect to the mean-field effect of information driving force, we define the information driving force $F^{\pm}(t)$ for the whole market. There are mainly two kinds of information driving forces, i.e., $F^{+}(t)$ and $F^{-}(t)$. $F^{+}(t)$ and $F^{-}(t)$ drive the dynamic fluctuation of $p_{buy}(t)$ and $p_{sell}(t)$ respectively,

$$p^{buy/sell}(t) = (1 + F^{\pm}(t))p(0). \tag{21}$$

Here $F^{\pm}(t)$ takes the form

$$F^{\pm}(t) = S(t)\eta(t)[1 + c \cdot sign(F'(t-1))]. \tag{22}$$

c is the asymmetric coefficient, which describes the asymmetric trading behavior of agents who have received the positive or negative external information. We also assume $c = \Delta\tilde{F}/2$, where $\Delta\tilde{F}$ is the difference of the empirically-defined information driving forces in the bull and bear markets.

In a real market, the trading behavior of agents will be influenced by not only the instant information but also the previous one [10,20]. To quantify the effect of the historical external information, we introduce a historical force $F'(t)$. At time t, for an agent with an investment horizon of l time steps, $\Sigma_{j=0}^{l-1} g(t)F^{\pm}(t-j)$ represents the basis for estimating the previous information driving force. Thus the historical force $F'(t)$ is defined as

$$F'(t) = \sum_{l=1}^{M} \left[\gamma_l \sum_{j=0}^{l-1} g(t) F^{\pm}(t-j) \right]. \tag{23}$$

Here $g(t) = 1$ for $F^{+}(t)$, and $g(t) = -1$ for $F^{-}(t)$. Since the auto-correlation of price returns is extremely weak [43], $g(t)$ is randomly set to be 1 or -1 at each time t. γ_l and M is defined in Section 3.1.

With $N = 100$, $p(0) = 0.1$ and $\mu = 1$, we perform the numerical simulation in the following procedures. Initially, $F(t)$ of the first τ times are set to be zero. On time t, we compute $F'(t-1)$ with Eq. (23), then $F^{\pm}(t)$ with Eq. (22). Next, we set the sign of $g(t)$ randomly, and calculate $p(t)$ with Eq. (21). After all agents have made the trading decisions with the probability $p(t)$ according to Eq. (9), the return $R(t)$ is computed with Eq. (10). Repeating the procedures, we obtain the return time series $R(t)$. The first 1000 data points are abandoned for equilibration, and then 1000 data points of $R(t)$ are collected. We obtain similar results with the simplified mean-field model, as shown in Figs. 3 and 4.

5. Summary

Open complex systems are many-body systems whose interplay with the external environment should not be ignored. As an important example of the open complex systems, the financial system is substantially influenced by the external information. In this paper, we investigate the information driving force based on the Google query and stock market data, and analyze how it affects the financial system. The information driving forces are stronger in the bear market state than in the bull one, i.e., the investors are more sensitive to the external information in the bear market.

As an application, we propose an agent-based model driven by the information driving force, and determine the key parameters from the empirical analysis rather than from statistical fitting of the simulation results. On one hand, the stationary properties of the real markets, such as the fat-tailed distribution of price returns, volatility clustering, and leverage effect are reproduced with the model. On the other hand, considering the mean-field effect of the external information, we also propose a few-body model to simulate the financial market in the laboratory. The asymmetric dynamic relaxation processes after and before the large fluctuations at the daily time scale are also simulated. Driven by the information driving force, the model naturally produces the exogenous events, which induce the time-reversal asymmetry of the dynamic relaxation processes.

Acknowledgment

This work was supported in part by NNSF of China under Grant Nos. 11775186, 11375149 and 11505099.

References

[1] E. Samanidou, E. Zschischang, D. Stauffer, T. Lux, Agent-based models of financial markets, Rep. Progr. Phys. 70 (2007) 409.
[2] R.N. Mantegna, J. Kertsz, Focus on statistical physics modeling in economics and finance, New J. Phys. 13 (2011) 25011.
[3] A. Chakraborti, I.M. Toke, M. Patriarca, F. Abergel, Econophysics review: II. Agent-based models, Quant. Finance 11 (2011) 1013.
[4] D. Sornette, Physics and financial economics (1776–2014): puzzles, ising and agent-based models, Rep. Progr. Phys. 77 (2014) 62001.
[5] M. Kirchlera, J. Huber, Fat tails and volatility clustering in experimental asset markets, J. Econom. Dynam. Control 31 (2007) 1844.
[6] A. Morone, Financial markets in the laboratory: an experimental analysis of some stylized facts, Quant. Finance 8 (2008) 513.
[7] K.N. An, X.H. Li, G. Yang, J.P. Huang, A controllable laboratory stock market for modeling real stock markets, Eur. Phys. J. B 86 (2013) 436.
[8] S. Palan, A review of bubbles and crashes in experimental asset markets, J. Econ. Surv. 27 (2013) 570.
[9] T. Preis, H.S. Moat, H.E. Stanley, S.R. Bishop, Quantifying the advantage of looking forward, Sci. Rep. 2 (2012) 350.
[10] I. Bordino, S. Battiston, G. Caldarelli, M. Cristelli, A. Ukkonen, I. Weber, Web search queries can predict stock market volumes, PLoS One 7 (2012) 40014.
[11] T. Preis, H.S. Moat, H.E. Stanley, Quantifying trading behavior in financial markets using google trends, Sci. Rep. 3 (2013) 1684.
[12] H.S. Moat, C. Curme, A. Avakian, D.Y. Kenett, H.E. Stanley, T. Preis, Quantifying Wikipedia usage patterns before stock market moves, Sci. Rep. 3 (2013) 1801.
[13] R. Hisano, D. Sornette, T. Mizuno, T. Ohnishi, T. Watanabe, High quality topic extraction from business news explains abnormal financial market volatility, PLoS One 8 (2013) 64846.
[14] L. Kristoufek, Can Google Trends search queries contribute to risk diversification, Sci. Rep. 3 (2013) 2713.
[15] T. Noguchi, N. Stewart, C.Y. Olivola, H.S. Moat, T. Preis, Characterizing the time-perspective of nations with search engine query data, PLoS One 9 (2014) 95209.
[16] C. Curme, T. Preis, H.E. Stanley, H.S. Moat, Quantifying the semantics of search behavior before stock market moves, Proc. Natl. Acad. Sci. 111 (2014) 11600.
[17] F. Lillo, S. Micciche, M. Tumminello, J. Piilo, R.N. Mantegna, How news affects the trading behaviour of different categories of the investors in a financial market, Quant. Finance 15 (2015) 213.
[18] E. Bonabeau, Agent-based modeling: methods and techniques for simulating human systems, Proc. Natl. Acad. Sci. 99 (2002) 7280.
[19] J.D. Farmer, D. Foley, The economy needs agent-based modelling, Nature 460 (2009) 685.
[20] L. Feng, B. Li, B. Podobnik, T. Preis, H.E. Stanley, Linking agent-based models and stochastic models of financial markets, Proc. Natl. Acad. Sci. 109 (2012) 8388.
[21] J.J. Chen, B. Zheng, L. Tan, Agent-based model with asymmetric trading and herding for complex financial systems, PLoS One 8 (2013) 79531.
[22] V. Gontis, A. Kononovicius, Consentaneous agent-based and stochastic model of the financial markets, PLoS One 9 (2014) 102201.
[23] Y. Shapira, Y. Berman, E. Ben-Jacob, Modelling the short term herding behaviour of stock markets, New J. Phys. 16 (2014) 53040.
[24] J.J. Chen, B. Zheng, L. Tan, Agent-based model with multi-level herding for complex financial systems, Sci. Rep. 5 (2015) 8399.

[25] T. Kaizoji, M. Leiss, A. Saichev, D. Sornette, Super-exponential endogenous bubbles in an equilibrium model of fundamentalist and chartist traders, J. Econ. Behav. Organ. 112 (2015) 289.

[26] R. Savona, M. Soumare, J.V. Andersen, Financial symmetry and moods in the market, PLoS One 10 (2015) 118224.

[27] T.T. Chen, B. Zheng, Y. Li, X.F. Jiang, New approaches in agent-based modeling of complex financial systems, Front. Phys. 12 (6) (2017) 128905.

[28] D. Sornette, R. Woodard, W.X. Zhou, The 2006–2008 oil bubble: Evidence of speculation, and prediction, Physica A 388 (2009) 1571–1576.

[29] T. Qiu, B. Zheng, G. Chen, Financial networks with static and dynamic thresholds, New J. Phys. 12 (2010) 43057.

[30] L. Tan, B. Zheng, J.J. Chen, X.F. Jiang, How volatilities nonlocal in time affect the price dynamics in complex financial systems, PLoS One 10 (2015) 118399.

[31] X.F. Jiang, T.T. Chen, B. Zheng, Time-reversal asymmetry in financial systems, Physica A 392 (2013) 5369.

[32] X.F. Jiang, B. Zheng, T. Qiu, F. Ren, Extreme-volatility dynamics in crude oil markets, Eur. Phys. J. B 30 (2017) 90.

[33] A.Z. Górski, S. Drożdż, J. Speth, Financial multifractality and its subtleties: an example of dax, Physica A 316 (2002) 496.

[34] J. Shen, B. Zheng, Cross-correlation in financial dynamics, Europhys. Lett. 86 (2009) 48005.

[35] X.F. Jiang, B. Zheng, Anti-correlation and subsector structure in financial systems, Europhys. Lett. 97 (2012) 48006.

[36] X.F. Jiang, T.T. Chen, B. Zheng, Structure of local interactions in complex financial dynamics, Sci. Rep. 4 (2014) 5321.

[37] V.M. Eguiluz, M.G. Zimmermann, Transmission of information and herd behavior: An application to financial markets, Phys. Rev. Lett. 85 (2000) 5659.

[38] L. Corazzini, B. Greiner, Herding, social preferences and (non-)conformity, Econom. Lett. 97 (2007) 74.

[39] Y.H. Shao, G.F. Gu, Z.Q. Jiang, W.X. Zhou, D. Sornette, Comparing the performance of fa, dfa and dma using different synthetic long-range correlated time series, Sci. Rep. 2 (2012) 5225.

[40] B. Zheng, X.F. Jiang, P.Y. Ni, A mini-review on econophysics: Comparative study of Chinese and western financial markets, Chin. Phys. B 23 (2014) 78903.

[41] F. Black, Studies of stock price volatility changes, in: Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section, Alexandria, 1976. pp. 177–181.

[42] X.F. Jiang, B. Zheng, F. Ren, T. Qiu, Localized motion in random matrix decomposition of complex financial systems, Physica A 471 (2017) 154.

[43] Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.K. Peng, H.E. Stanley, Statistical properties of the volatility of price fluctuation, Phys. Rev. E 60 (1999) 1390.