Voter interacting systems applied to Chinese stock markets

Tiansong Wang, Jun Wang *, Junhuan Zhang, Wen Fang

Department of Mathematics, College of Science, Beijing Jiaotong University, Beijing 100044, PR China

Received 23 November 2008; received in revised form 14 March 2011; accepted 30 March 2011
Available online 17 April 2011

Abstract

Applying the theory of statistical physics systems – the voter model, a random stock price model is modeled and studied in this paper, where the voter model is a continuous time Markov process. In this price model, for the different parameters values of the intensity $\lambda$, the lattice dimension $d$, the initial density $\theta$, and the multivariate set $(\theta, \lambda)$, we discuss and analyze the statistical behaviors of the price model. Moreover, we investigate the power-law distributions, the long-term memory of returns and the volatility clustering phenomena for the Chinese stock indices. The database is from the indices of Shanghai and Shenzhen in the 6-year period from July 2002 to June 2008. Further, the comparisons of the empirical research and the simulation data are given.

Keywords: Stock price model; Voter model; Probability distribution; Return; Computer simulation

1. Introduction

As the governments deregulate the stock markets, it is becoming an important problem to model the dynamics of the forwards prices in the risk management, derivatives pricing and physical assets valuation. And the research work on the statistical properties of fluctuations of stock prices in globalized securities markets is also significant. Recently, some research work has been done, by applying the interacting particle systems, to investigate the statistical behaviors of fluctuations of stock prices, and to study the corresponding valuation and hedging of contingent claims, for instance see [11,12,17,21,22]. Stauffer and Penna [21] and Tanaka [22] have been apply the percolation model to study the market fluctuation, see [8]. In [21,22], they construct the local interaction or influence among investors in a stock market and define the cluster of investors with the same opinion about the market as a cluster of percolation. In their financial models, they suppose that the stock price fluctuation is influenced by the information in the stock market, and the investors follow the effect of sheep flock. That means that the investors decide the investment opinions by other investors’ attitudes, so the investors’ investment attitudes of the stock market lead to the stock price fluctuation. In the present paper, we apply a statistical physics model – the voter model (see [3,15]) to study the fluctuation behaviors of the return processes. Through the computer simulation on this financial model, we discuss the statistical behavior, the tail behavior, long term memory of fluctuation and the volatility clustering phenomena for the return processes.

In recent years, the empirical research in financial market fluctuations has been made. Some statistical properties for market fluctuations uncovered by the high frequency financial time series, such as fat tails distribution of price changes, the power-law of logarithmic returns and volume, volatility clustering which is described as on-off intermittency in

* Corresponding author. Fax: +86 10 51682867.
E-mail address: wangjun@bjtu.edu.cn (J. Wang).
literature of nonlinear dynamics, and multifractality of volatility, etc., see [1,5–7,9,10,12,14,16–19]. China has two stock markets, Shanghai Stock Exchange and Shenzhen Stock Exchange, and the indices studied in this paper are Shanghai Stock Exchange Composite Index (SSECI) and Shenzhen Stock Exchange Component Index (SZSECI). These two indices play an important role in Chinese stock markets and Chinese economic systems. In this paper, we analyze the daily data of SSECI and SZSECI from July 1st 2002 to June 30th 2008, and study the fluctuation properties of SSECI and SZSECI. The database of the indices is from the websites www.sse.com.cn and www.szse.cn.

First, we give the brief definitions and properties of the voter model, and also we introduce the graphical representation of the model, since the graphical representation is necessary for our computer simulations. One interpretation for the voter model is, for a collection of individuals, each of which has one of two possible positions on a political issue, at independent exponential times, an individual reassesses his view by choosing a neighbor at random with certain probabilities and then adopting his position. Specifically, the voter model is one of the statistical physics models, we think of the sites of the \( d \)-dimensional integer lattice as being occupied by persons who either in favor of or opposed to some issue. To write this as a set-valued process, we let \( \eta_s \) is the set of voters in favor, we can also think of the sites in \( \eta_s \) as being occupied by cancer cells, and the other sites as being occupied by healthy cells. We can formulate the dynamics as follows: (i) an occupied site becomes vacant at a rate equal to the number of the vacant neighbors; (ii) an vacant site becomes occupied at a rate equal to \( \lambda \) times the number of the occupied neighbors, where \( \lambda \) is a intensity which is called the “carcinogenic advantage” in voter model. When \( \lambda = 1 \), the model is called the voter model, and when \( \lambda > 1 \), the model is called the biased voter model.

For simplicity, we give the construction of graphical representation for 1-dimensional voter model (\( \lambda = 1 \)), for more general cases, see [3,15]. Thinking of 1-dimensional integer points as being laid out on a horizontal axis, with the time lines being placed vertically, above that axis. Define independent Poisson processes with rate 1 for each time lines, at each event time \((u, s)\), we choose one of its two neighbors with probability 1/2, and draw an arrow from that neighbor point to \((u, s)\), and write a \( \delta \) at \((u, s)\), see Fig. 1. To construct the process from this “graphical representation”, we imagine fluid entering the bottom at the points in \( \eta_0 \) and flowing up the structure. The \( \delta \)'s are the dams and the arrows are pipes which allow the fluid to flow in the indicated direction. For example in above Fig. 1, fluid can flow from the site \((-2, 0)\) to the site \((0, t)\). Let \( \eta^{\delta}_s \) denote the state at time \( s \) with the initial state \( \eta^A_0 = A \), then from [1,14], the voter model \( \eta^\delta_s \) approaches total consensus in \( d = 1 \) and \( d = 2 \). But in higher dimensions \( d \geq 3 \), the differences of opinion may persist. For more generally, we consider the initial distribution as \( \nu_0 \), the product measure with density \( \theta \), that is, each site is independently occupied with probability \( \theta \) and let \( \eta^\delta_s \) to denote the voter model with initial distribution \( \nu_0 \). Let \( \eta^{\delta}_0 = \{0\} \), and \( \eta^{\delta}_s(\{u\}) \) be the state of \( u \in \mathbb{Z}^d \) at time \( s \) with the initial point \( \{0\} \).

More formally, the stochastic dynamics of the voter model \( \eta_s \) is a Markov process on \( \Omega = \{0, 1\}^{\mathbb{Z}^d} \) whose generator has the form

\[
A f(\eta) = \sum_u c(u, \eta)[f(\eta^u) - f(\eta)]
\]
where the functions \( f(\Omega) \) depend on finitely many coordinates, and \( \eta^u(v) = \eta(v) \) if \( v \neq u \), \( \eta^u(v) = -\eta(v) \) if \( v = u \), for \( u, v \in \mathbb{Z}^d \). The transition rate function for the process which is given by follows (see [1,14]), for any \( \eta \in \Omega = \{0, 1\}^{\mathbb{Z}^d} \)

\[
c(\eta, \eta) = \begin{cases} 
\sum v p(u, v) \eta(v) & \text{if } \eta(u) = 0 \\
\sum v p(u, v)[1 - \eta(v)] & \text{if } \eta(u) = 1 
\end{cases}
\]

where \( p(u, v) \geq 0 \) for \( u, v \in \mathbb{Z}^d \), and \( \sum_v p(u, v) = 1 \) for all \( u \in \mathbb{Z}^d \). And we assume that \( p(u, v) \) is such that the Markov chain with those transition probabilities is irreducible.

For the biased voter model \( (\lambda > 1) \), there is a “critical value” for the process, that is, on \( \Omega = \{0, 1\}^{\mathbb{Z}^d} \) and with the corresponding probability \( P \), the critical value \( \lambda_c \) is defined as

\[
\lambda_c = \inf\{\lambda : P(\{0\}^s) > 0, \text{ for all } s \geq 0 > 0\}
\]

where \( |\eta_s^0| \) is the cardinality of \( \eta_s^0 \). Suppose \( \lambda > \lambda_c \), then there is a convex set \( C \) so that on \( \Omega_\infty = \{\eta_s^0 \neq \emptyset \}, \text{ for all } s \}, \) we have for any \( \epsilon > 0 \) and for all \( s \) sufficiently large (see [3,15])

\[
(1 - \epsilon)sC \cap \mathbb{Z}^d \subset \eta_s^0 \subset (1 + \epsilon)sC \cap \mathbb{Z}^d
\]

If \( \lambda < \lambda_c \), for some positive \( \rho(\lambda) \), we have

\[
P(\{0\}^0 \neq \emptyset) \leq e^{-\rho s}
\]

The above results imply that, on \( d \)-dimensional lattice, if \( \lambda < \lambda_c \), the process dies out (becomes vacant) exponentially fast, if \( \lambda > \lambda_c \), the process survives with the positive probability.

### 2. Modeling a stock price model by the voter dynamic systems

In this part, the return process of a stock price on \( d \)-dimensional integer lattice is constructed, based on the voter model. We assume that the information in the stock market lead to the fluctuation of a stock price, and there are three kinds of information: buying information, selling information and neutral information. And the fluctuation of a stock price relies on the investor’s investment attitudes, which accordingly classify buying stock, selling stock and holding stock. Consider a model of auctions for the same stock defined above, we can derive the stock price process from the auctions. Assume that each trader can trade the stock several times at each day \( t \in \{1, 2, \ldots, n\} \), but at most one unit number of the stock at each time. Let \( t \) be the time length of trading time in each trading day, we denote the stock price at time \( s \) in the \( t \) th trading day by \( S_t(s) \), where \( s \in [0, l] \). Suppose that this stock consists of \( m_2 + 1 \) (\( m_2 \) is large enough) investors, who are located in a line \( \{-m_2/2, \ldots, -1, 0, 1, \ldots, m_2/2\} \subset \mathbb{Z} \) (similarly for \( d \)-dimensional lattice \( \mathbb{Z}^d \)). At the beginning of trading in each day, suppose that only the investor at the site \( \{0\} \) receives some news. We define a random variable \( \xi_t(\{0\}) \) for this investor, suppose that this investor taking buying position (\( \xi_t(\{0\}) = 1 \)), selling position (\( \xi_t(\{0\}) = -1 \)) or neutral position (\( \xi_t(\{0\}) = 0 \)) with probability \( p_1, p_{-1} \) or \( 1 - (p_1 + p_{-1}) \) respectively. Then this investor sends bullish, bearish or neutral signal to his nearest neighbors. According to \( d \)-dimensional voter process system, investors can affect each other or the news can be spread, which is assumed as the main factor of price fluctuations for the investors. Moreover, here the investors can change their buying positions or selling positions to neutral positions independently at a constant rate. More specifically, (I) when \( \xi_t(\{0\}) = 1 \) and if \( \eta_t^0(u) = 1 \), we say that the investor at \( u \) takes buying position at time \( s \), and this investor recovers to neutral position 0 at a rate equal to the number of the vacant neighbors; if \( \eta_t^0(u) = 0 \), we think the investor at \( u \) takes neutral position at time \( s \), and this investor is changed to take buying position by his nearest neighbors at rate \( \lambda \sum_{v:|u-v|=1} \eta_t^0(v) \). In this case, the more investors with taking buying positions, the more possible the stock price goes up. (II) When \( \xi_t(\{0\}) = -1 \) and if \( \eta_t^0(u) = 1 \), we say that the investor at \( u \) takes selling position at time \( s \), also this investor recovers to neutral position 0 at a rate equal to the number of the vacant neighbors; if \( \eta_t^0(u) = 0 \), the investor is changed to take selling position by
his nearest neighbors at rate $\lambda \sum_{v:|v-u|=1} \eta_3^{[0]}(v)$. (III) When the initial random variable $\xi_t(\{0\}) = 0$, the process $\eta_3^{[0]}(u)$ is ignored, this means that the investors do not affect the fluctuation of the stock price.

For a fixed $s \in [0, l]$ ($l$ large enough), let

$$B_t(s) = \xi_t([0]) \frac{|\eta_s^{[0]}|}{m_2(n)} , \quad s \in [0, l]$$

where $|\eta_s^{[0]}| = \sum_{u=-m_2/2}^{m_2/2} \eta_s^{[0]}(u)$, and $m_2$ depends on the trading days $n$. From above definitions and [1,12,19,22], we define the stock price at $t$ th trading day $t(t=1, 2, \ldots, n)$ as

$$S_t(s) = e^{\alpha B_t(s)} S_{t-1}(s) , \quad s \in [0, l]$$

where $\alpha > 0$, represents the depth parameter of the market. Then we have

$$S_t(s) = S_0 \exp(\alpha \sum_{k=1}^{t} B_k(s)) , \quad s \in [0, l]$$

where $S_0$ is the stock price at time 0. According to the theory of the voter process (see [15]) and above definitions, if $\lambda > \lambda_c$, the news will spread widely, so this will affect the investors’ positions, and finally will affect the fluctuation of the stock price. If $\lambda \leq \lambda_c$, the influence on the stock price by the investors is limited.

Next we define the returns of stock prices, see [13,20,23], we have the formula of the stock logarithmic return as follows

$$r(t) = \ln \frac{S_{t+1}(s)}{S_t(s)} , \quad t = 1, 2, \ldots, n$$

In this paper, we analyze the logarithmic returns for the daily price changes. In the followings, we let $s = 2000$, then we plot the fluctuations of the stock prices and the corresponding returns by simulating the one-dimension voter model, see Figs. 2 and 3.

3. Data analysis by the computer simulation

In recent years, the probability distribution in the financial market fluctuation has been studied, the empirical research results show that the distribution of the large returns follow a power-law distribution with the exponent 3, that is $P(r(t) > x) \sim x^{-\mu_r}$, where $r(t)$ is the returns of the stock prices in a given time interval $\Delta t$, $\mu_r \approx 3$, for example see [5,9,10,14,19]. In this section, Section 3.1 presents the statistical analysis of the simulation data for univariate process.
Section 3.2 investigates the returns for multivariate process. Finally, in Section 3.3, the conclusions of the analysis between univariate process and multivariate process are summarized.

3.1. Analysis for univariate process

In this section, we discuss the three parameters of the financial model, the intensity $\lambda$, the lattice dimension $d$, the initial density $\theta$ (or the initial distribution $\nu_{\theta}$ of the model). And we show some figures and tables to describe the power-law distribution of the returns for the simulation data. Further, we compare the power-law distribution of the returns with the corresponding Gaussian distribution, and study the statistical properties of the price model by the different values of one parameter. First, we discuss the behaviors of the returns for the different values of $\lambda$.

According to the computer computations in Table 1, the kurtosis is increasing steadily with the parameter $\lambda$ increasing, while the exponent $\mu$ is decreasing with the value $\lambda$ increasing, for the fixed values $d = 1$, $\theta = 0.1$. When $\lambda = 5$, the kurtosis value is the smallest while the value $\mu$ is the biggest in the five groups simulation data, $3.248$ and $5.444$ respectively; when $\lambda = 25$, the situation is reverse. Moreover, we can see that the fat tail behavior of the returns is more significantly as $\lambda$ increases in Fig. 4. These results imply that the parameter $\lambda$, which represents the rate of information spread in the price model, can affect the power-law distributions of the price returns. Fig. 4 also shows that the probability distributions of the returns deviate from the corresponding Gaussian distribution.

In Fig. 5 and Table 2, we continue to study the statistical properties of the price returns with the different lattice dimensions $d$ for the fixed values $\lambda = 5$, $\theta = 0.1$. In Fig. 5, the tails of the distributions are deviating from that of the corresponding Gaussian distribution as the dimension $d$ increases, this implies the fat-tails phenomena is much obvious in this case. Similarly to the results of Tables 1 and 2 also show that the kurtosis of the model is increasing and the exponent $\mu$ is decreasing for the parameter $d$. When $d = 3$, the value of kurtosis is $8.033$, which is the biggest, and $\mu$ is $3.658$, the smallest value in three groups data. The reason of these increasing (or decreasing) properties is that the

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Variance</th>
<th>Kurtosis</th>
<th>Skewness</th>
<th>$\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>$-0.0317$</td>
<td>0.0227</td>
<td>$-6.135 \times 10^{-5}$</td>
<td>$5.711 \times 10^{-5}$</td>
<td>3.248</td>
<td>$-0.0156$</td>
<td>5.444</td>
</tr>
<tr>
<td>10</td>
<td>$-0.0368$</td>
<td>0.0396</td>
<td>$2.989 \times 10^{-4}$</td>
<td>$1.641 \times 10^{-4}$</td>
<td>3.701</td>
<td>$-0.0231$</td>
<td>5.181</td>
</tr>
<tr>
<td>15</td>
<td>$-0.0345$</td>
<td>0.0296</td>
<td>$1.707 \times 10^{-4}$</td>
<td>$4.194 \times 10^{-5}$</td>
<td>4.163</td>
<td>0.0129</td>
<td>4.213</td>
</tr>
<tr>
<td>20</td>
<td>$-0.0413$</td>
<td>0.0404</td>
<td>$-4.562 \times 10^{-4}$</td>
<td>$4.558 \times 10^{-4}$</td>
<td>4.304</td>
<td>$-0.0159$</td>
<td>3.661</td>
</tr>
<tr>
<td>25</td>
<td>$-0.0491$</td>
<td>0.0518</td>
<td>$-3.996 \times 10^{-4}$</td>
<td>$5.886 \times 10^{-4}$</td>
<td>5.076</td>
<td>0.0182</td>
<td>3.477</td>
</tr>
</tbody>
</table>

Table 1
The statistical analysis of the returns for the intensity $\lambda$. 

Fig. 3. The fluctuations of market returns.
interaction among the investors becomes more active in this financial model (which is defined by the voter dynamic system) as the dimension $d$ increases.

Finally, we discuss the effects of the parameter $\theta$ on the returns. $\theta$ is the initial density of the price model, and it represents the proportion of the investors obtaining the information at the beginning of each trading day. When the

![Figure 4](image_url_4)

**Fig. 4.** The cumulative distributions of the normalized price returns with the different values of $\lambda$ when $d=1, \theta=0.1$.

![Figure 5](image_url_5)

**Fig. 5.** The cumulative distributions of the normalized price returns with the different values of $d$ when $\lambda=5, \theta=0.1$.

<table>
<thead>
<tr>
<th>$d$</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Variance</th>
<th>Kurtosis</th>
<th>Skewness</th>
<th>$\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$-0.0317$</td>
<td>0.0227</td>
<td>$-6.135 \times 10^{-5}$</td>
<td>$5.711 \times 10^{-5}$</td>
<td>3.248</td>
<td>$-0.0156$</td>
<td>5.444</td>
</tr>
<tr>
<td>2</td>
<td>$-0.0332$</td>
<td>0.0314</td>
<td>$4.870 \times 10^{-5}$</td>
<td>$9.220 \times 10^{-5}$</td>
<td>5.963</td>
<td>$-0.0123$</td>
<td>4.794</td>
</tr>
<tr>
<td>3</td>
<td>$-0.0558$</td>
<td>0.0517</td>
<td>$3.074 \times 10^{-4}$</td>
<td>$5.282 \times 10^{-4}$</td>
<td>8.033</td>
<td>$-0.0222$</td>
<td>3.658</td>
</tr>
</tbody>
</table>

**Table 2**

The statistical analysis of the returns for the lattice dimension $d$. 
initial density value $\theta$ increases, the more investors will share the same opinion about the market, then the stock price should fluctuate more greatly. This property can been seen in Fig. 6 and Table 3.

In fact, the simulation results of the returns for the initial density $\theta$ can be extended in this financial model. For example, we can suppose that the initial state of the model is some random variable $\xi_t$ (or $\xi_t^\theta$) following a certain distribution $\nu$. In a stock market, the investors usually decide their investment positions by analyzing the historical data on the stock, the nearer the time of the historical data is to the present, the stronger impact the data have on the initial distribution $\nu$. In this case, we can introduce the Brownian motion in order to make the initial distribution have the effect of random movement while maintaining the original trend, for the related work see [13]. It is also interesting to do the computer simulations of the returns for the different random initial distributions $\nu$, it can improve our model because the fluctuations in the stock markets are random.

3.2. Analysis for multivariate process

In the above section, we have discussed the returns for univariate process. In this part, we study the returns for multivariate process. We analyze the statistical behaviors of the returns for the multivariate set $(\theta, \lambda)$. For the simplicity, let $d = 1$, and suppose that $\theta$ and $\lambda$ satisfy a linear function $a\theta + b\lambda + c = 0$, where $a, b, c$ are three constants. In Table 4, we consider the multivariate set

$$\{(\theta, \lambda) : a\theta + b\lambda + c = 0, \quad a = 100, \ b = -1, \ c = 5\}.$$

By the computer computation, we have Table 4.

From Table 4, we have the similar statistical properties as that in Section 3.1, when $\theta$ and $\lambda$ increase simultaneously, the kurtosis value increases while the value of $\mu$ decreases accordingly. However, the rise and the fall of the values are more dramatic. The kurtosis in the multivariate processing on $(\theta, \lambda)$ increases more quickly than that in the single

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Variance</th>
<th>Kurtosis</th>
<th>Skewness</th>
<th>$\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>-0.0320</td>
<td>0.0294</td>
<td>$7.008 \times 10^{-5}$</td>
<td>$6.206 \times 10^{-5}$</td>
<td>3.312</td>
<td>-0.0224</td>
<td>7.032</td>
</tr>
<tr>
<td>0.05</td>
<td>-0.0345</td>
<td>0.0296</td>
<td>$9.506 \times 10^{-5}$</td>
<td>$4.259 \times 10^{-5}$</td>
<td>3.493</td>
<td>-0.0181</td>
<td>5.079</td>
</tr>
<tr>
<td>0.10</td>
<td>-0.0413</td>
<td>0.0404</td>
<td>$-4.562 \times 10^{-4}$</td>
<td>$4.558 \times 10^{-4}$</td>
<td>4.304</td>
<td>-0.0159</td>
<td>3.661</td>
</tr>
<tr>
<td>0.15</td>
<td>-0.0398</td>
<td>0.0466</td>
<td>$5.101 \times 10^{-4}$</td>
<td>$8.574 \times 10^{-5}$</td>
<td>4.836</td>
<td>-0.0274</td>
<td>3.497</td>
</tr>
<tr>
<td>0.20</td>
<td>-0.0511</td>
<td>0.0502</td>
<td>$4.085 \times 10^{-5}$</td>
<td>$9.612 \times 10^{-5}$</td>
<td>7.974</td>
<td>-0.0209</td>
<td>3.125</td>
</tr>
</tbody>
</table>

Table 3 The statistical analysis of the returns for the initial density $\theta$.
Table 4
The statistical analysis of the returns for the multivariate set ($\theta, \lambda$).

<table>
<thead>
<tr>
<th>($\theta, \lambda$)</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Variance</th>
<th>Kurtosis</th>
<th>Skewness</th>
<th>$\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.01, 5)</td>
<td>-0.0392</td>
<td>0.0382</td>
<td>3.118 $\times 10^{-4}$</td>
<td>5.273 $\times 10^{-5}$</td>
<td>3.127</td>
<td>-0.0119</td>
<td>7.579</td>
</tr>
<tr>
<td>(0.01, 10)</td>
<td>-0.0224</td>
<td>0.0191</td>
<td>4.5287 $\times 10^{-5}$</td>
<td>3.2289 $\times 10^{-5}$</td>
<td>3.1589</td>
<td>-0.0666</td>
<td>7.307</td>
</tr>
<tr>
<td>(0.01, 15)</td>
<td>-0.0214</td>
<td>0.0255</td>
<td>2.3494 $\times 10^{-4}$</td>
<td>3.7238 $\times 10^{-5}$</td>
<td>3.2851</td>
<td>-0.0856</td>
<td>7.101</td>
</tr>
<tr>
<td>(0.01, 25)</td>
<td>-0.0373</td>
<td>0.0373</td>
<td>-1.7212 $\times 10^{-4}$</td>
<td>7.7171 $\times 10^{-5}$</td>
<td>3.3711</td>
<td>-0.0746</td>
<td>6.890</td>
</tr>
<tr>
<td>(0.05, 5)</td>
<td>-0.0181</td>
<td>0.0130</td>
<td>-2.2706 $\times 10^{-4}$</td>
<td>1.8717 $\times 10^{-5}$</td>
<td>3.3953</td>
<td>-0.0737</td>
<td>6.526</td>
</tr>
<tr>
<td>(0.05, 15)</td>
<td>-0.0248</td>
<td>0.0261</td>
<td>-7.0518 $\times 10^{-5}$</td>
<td>5.1835 $\times 10^{-5}$</td>
<td>3.4371</td>
<td>-0.0682</td>
<td>5.903</td>
</tr>
<tr>
<td>(0.05, 25)</td>
<td>-0.0287</td>
<td>0.0213</td>
<td>1.5817 $\times 10^{-4}$</td>
<td>3.8764 $\times 10^{-5}$</td>
<td>3.8845</td>
<td>-0.0553</td>
<td>4.702</td>
</tr>
<tr>
<td>(0.1, 5)</td>
<td>-0.0317</td>
<td>0.0227</td>
<td>-6.135 $\times 10^{-5}$</td>
<td>5.711 $\times 10^{-5}$</td>
<td>3.248</td>
<td>-0.0156</td>
<td>5.444</td>
</tr>
<tr>
<td>(0.1, 10)</td>
<td>-0.0368</td>
<td>0.0396</td>
<td>2.989 $\times 10^{-4}$</td>
<td>1.641 $\times 10^{-4}$</td>
<td>3.701</td>
<td>-0.0231</td>
<td>5.181</td>
</tr>
<tr>
<td>(0.1, 15)</td>
<td>-0.0345</td>
<td>0.0296</td>
<td>1.707 $\times 10^{-4}$</td>
<td>4.194 $\times 10^{-5}$</td>
<td>4.163</td>
<td>-0.0129</td>
<td>4.213</td>
</tr>
<tr>
<td>(0.1, 20)</td>
<td>-0.0413</td>
<td>0.0404</td>
<td>-4.562 $\times 10^{-4}$</td>
<td>4.558 $\times 10^{-4}$</td>
<td>4.304</td>
<td>-0.0159</td>
<td>3.661</td>
</tr>
<tr>
<td>(0.1, 25)</td>
<td>-0.0491</td>
<td>0.0518</td>
<td>-3.996 $\times 10^{-4}$</td>
<td>5.886 $\times 10^{-4}$</td>
<td>5.076</td>
<td>0.0182</td>
<td>3.477</td>
</tr>
<tr>
<td>(0.15, 5)</td>
<td>-0.0104</td>
<td>0.0087</td>
<td>-7.9715 $\times 10^{-5}$</td>
<td>4.6671 $\times 10^{-6}$</td>
<td>5.1226</td>
<td>-0.1011</td>
<td>4.308</td>
</tr>
<tr>
<td>(0.15, 25)</td>
<td>-0.0176</td>
<td>0.0182</td>
<td>-3.4408 $\times 10^{-5}$</td>
<td>1.4799 $\times 10^{-5}$</td>
<td>5.9906</td>
<td>-0.0065</td>
<td>3.004</td>
</tr>
<tr>
<td>(0.20, 5)</td>
<td>-0.0064</td>
<td>0.0046</td>
<td>-1.3812 $\times 10^{-5}$</td>
<td>9.3924 $\times 10^{-7}$</td>
<td>6.2413</td>
<td>-0.2876</td>
<td>3.709</td>
</tr>
<tr>
<td>(0.20, 25)</td>
<td>-0.0476</td>
<td>0.0412</td>
<td>6.101 $\times 10^{-4}$</td>
<td>3.277 $\times 10^{-4}$</td>
<td>7.632</td>
<td>-0.0217</td>
<td>3.134</td>
</tr>
</tbody>
</table>

processing on $\theta$ or on $\lambda$, similarly to the case for $\mu$, this implies that the fat-tails phenomena is more significant. This shows that, when the investment information is more widely spread and the more investors share the same investment attitudes, the effect of sheep flock is more obvious. In this section we only consider the case that $\theta$ and $\lambda$ have the linear relation, for the nonlinear case, we can analyze the simulative data of the price model by the same statistical methods.

3.3. The conclusions of the analysis between univariate process and multivariate process

Through two parts above, we know that the information decides the fluctuation of return. More quickly and more widely the information are extended, more probable it is that returns have fierce fluctuation. Whichever univariate process or multivariate process, if the change increases the spread velocity of information, the fluctuation of returns will be more fiercer; if the change decreases the speed of information, the fluctuation of returns will become weaker.

4. The comparison of SSECI, SZSECI and the financial model

In this section, we compare the returns of SSECI and SZSECI (from July 1st 2002 to June 30th 2008) with the empirical data from the financial model constructed by the voter model, where $\lambda = 20$, $\theta = 0.1$ and $d = 1$. According to the statistical methods and the data analyzing methods (see [6,9,20]), we will study the cumulative probability distributions of the daily returns and the power-law character of the daily returns for Shanghai stock market and Shenzhen stock market, and we also simulate the corresponding cumulative probability distributions of the returns by the financial model, which is constructed based on the voter model.

In recent years, the fluctuation of the returns is believed to follow a Gaussian distribution for long time intervals but to deviate from it for short steps, called the fat-tail phenomena. We try to study the probability distribution of the returns for SSECI, SZSECI and the simulation data. In the following, the statistical properties of skewness and kurtosis of the distributions for the returns are given. The kurtosis represents the data’s centrality, and the skewness shows the symmetry of the data. And we know that the skewness and the kurtosis of a standard Gaussian distribution is 0 and 3 respectively. The following is the statistics of SSECI, SZSECI and the simulation data.

From Table 5, the kurtosis values of SSECI, SZSECI and the simulation data are 6.819, 6.257 and 4.304 respectively, and they are all larger than 3; the corresponding skewness values are $-0.230$, $-0.154$ and $-0.016$ respectively, and they are all on the left of 0. These indicate that the returns of three financial series fluctuate more violently than that of the corresponding Gaussian distribution. Next we take the single-sample Kolmogorov–Smirnov test for three financial series in Table 6. Because the values of two-tail test $P$ is 0.000, the hypothesis is refused that the return distributions of SSECI, SZSECI and the simulation data follow the corresponding Gaussian distribution. Therefore, the return distributions of all three series obviously have fat-tails character and are biased.
Table 5
The statistical analysis of the returns for SSECI, SZSECI and the simulation data.

<table>
<thead>
<tr>
<th></th>
<th>Sample size</th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>( \mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSECI</td>
<td>1455</td>
<td>4.599 \times 10^{-4}</td>
<td>2.756 \times 10^{-4}</td>
<td>-0.230</td>
<td>6.819</td>
<td>2.865</td>
</tr>
<tr>
<td>SZSECI</td>
<td>1454</td>
<td>8.399 \times 10^{-4}</td>
<td>3.309 \times 10^{-4}</td>
<td>-0.154</td>
<td>6.257</td>
<td>2.734</td>
</tr>
<tr>
<td>Simulation</td>
<td>5000</td>
<td>-4.562 \times 10^{-4}</td>
<td>4.558 \times 10^{-4}</td>
<td>-0.016</td>
<td>4.304</td>
<td>3.661</td>
</tr>
</tbody>
</table>

4.1. The estimates of the long term memory for the returns

The evaluation of the long memory of one time series can be made through various methodologies. In this section, the long memory is measured by the Hurst exponent \( H \), calculated through the classic Rescaled Range \((R/S)\) Analysis and modified Rescaled Range \((R/S)\) statistic.

4.2. Rescaled Range Analysis

Rescaled Range \((R/S)\) Analysis is a very effective nonparametric test to analyze the durative of time sequences and do not need special distributions of time sequences studied. \( R/S \) is a new statistics which can be used to distinguish whether time sequence relies on the time, that is, the exponent of Hurst. Green and Fielitz firstly applied \( R/S \) to study the mathematical finance and analysis the common stock returns. The research shows that, when \( H = 0.5 \), the time sequence belongs to Brownian motion and the variables are independent and the correlation coefficients are 0; when \( 0 < H < 0.5 \), the variables are negative correlation and are anti-persistence; when \( 0.5 < H < 1 \), that means the time sequence is persistent and has the quantification of long-term memory. In order to calculate the Hurst exponent \( H \), the time series \( x_t \) with \( N \) numbers is divided by \( A \) parts. And each part marks as \( x_{a,k} \), \( a = 1, 2, \ldots, A \). For \( t = 1, 2, \ldots, N \), let

\[
\bar{x}_a = \frac{1}{n} \sum_{k=1}^{n} x_{a,k}, \quad X_{a,t} = \sum_{k=1}^{t} (x_{a,k} - \bar{x}_a)
\]

\[
R_a = \max_{1 \leq t \leq N} X_{a,t} - \min_{1 \leq t \leq N} X_{a,t}, \quad S_a = \left[ \frac{1}{N} \sum_{k=1}^{N} (x_{a,k} - \bar{x}_a)^2 \right]^{1/2}
\]

By computing \( R_a/S_a \) by each part data, then we have the mean as follows

\[
(R/S)_N = \frac{1}{A} \sum_{a=1}^{A} (R_a/S_a)
\]

The Hurst exponent \( H \) is obtained by the expression \( \ln (R/S)_N = \ln C + H \ln N \), which is given by Mandelbrot and Wallis [18], and the values of \( H \) can be calculated by the lease ordinary squares. By applying the method of the \( R/S \) test, the exponent \( H \) of the returns for SSECI, SZSECI are 0.6363, 0.6367 respectively, and the exponent \( H \) of the returns for the price model is 0.5979, see Fig. 7. The \( H \) of three groups’ data are all larger than 0.5, which indicate that these three financial time sequences are all long-term memory sequences.

Table 6
The Kolmogorov–Smirnov test.

<table>
<thead>
<tr>
<th></th>
<th>Sample size</th>
<th>( H )</th>
<th>( \text{CV} )</th>
<th>Two-tail test ( P )</th>
<th>Statistics of K–S</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSECI</td>
<td>1455</td>
<td>1</td>
<td>0.029</td>
<td>0.000</td>
<td>0.475</td>
</tr>
<tr>
<td>SZSECI</td>
<td>1454</td>
<td>1</td>
<td>0.028</td>
<td>0.000</td>
<td>0.472</td>
</tr>
<tr>
<td>Simulation</td>
<td>5000</td>
<td>1</td>
<td>0.025</td>
<td>0.000</td>
<td>0.406</td>
</tr>
</tbody>
</table>
Fig. 7. The Hurst exponent of the returns for the simulation data.
4.3. Modified R/S analysis

Lo [16] pointed out that the classical R/S statistic is sensitive to the presence of the short-term dependence and, unless the short-term memory in the data is accounted for, the R/S analysis may not be reliable. Then a modified R/S statistic is given as follows:

\[ Q_N = \frac{R_N}{\sigma_N(q)} \]

where

\[ \sigma_N^2(q) = \gamma_0 + 2 \sum_{j=1}^{q} \left( \frac{1 - j}{1 + q} \gamma_j \right) \]

and \( \gamma_j \) is the \( j \)-th order autocovariance of \( x_n \). Then, the Hurst exponent \( H \) is given by the limit of the ratio \( \log Q_n / \log n \).

A practical difficulty concerns the choice of the truncation lag \( q \). One suggestion for the truncation lag is Andrew’ data-dependent formula:

\[ q = \left\lceil \left( \frac{3N}{2} \right)^{1/3} \left( \frac{2\rho}{1 - \rho^2} \right)^{2/3} \right\rceil \]

(\( \lceil \cdot \rceil \) denotes the integer part and \( \rho \) is the correlation coefficient), which is the first order autocorrelation of the data series. If \( q \) is set to 0 for the modified R/S statistic, then we have the classical R/S statistic of Hurst. Then we derive the limiting distribution of modified R/S statistic (see [16]), in the proper situations, for \( N \to \infty \)

\[ \max_{1 \leq k \leq N} \frac{1}{\sigma_N(q) \sqrt{N}} \sum_{i=1}^{k} (X_i - \overline{X}_N) \Rightarrow \max_{1 \leq w \leq 1} W^0(w) = M^0 \]

\[ \min_{1 \leq k \leq N} \frac{1}{\sigma_N(q) \sqrt{N}} \sum_{i=1}^{k} (X_i - \overline{X}_N) \Rightarrow \min_{1 \leq w \leq 1} W^0(w) = m^0 \]

\[ \frac{1}{\sqrt{N}} Q_N(q) \Rightarrow M^0 - m^0 = V \]

where “\( \Rightarrow \)” denotes the weak convergence, and \( W^0 = B(w) - wB(1) \) is Brownian Bridge in [0, 1], where \( B(w) \) is the standard Brownian motion. The corresponding distribution function is

\[ F_V(x) = 1 + 2 \sum_{k=1}^{\infty} (1 - 4k^2x^2)e^{-2(kx)^2} \]

Then the mean and the variance of \( V \) are follows: \( E(V) = \sqrt{\pi/2} \approx 1.25 \), \( \Var(V) = \pi^2/6 - \pi/2 \approx 0.074 \). When the statistical significance is at 5% level, the confidence interval of statistic \( V \) is [0.809, 1.862]. If the value of \( V \) is within the interval, null hypothesis that time series have the long range dependence is accepted. Conversely, alternative hypothesis is accepted.

Applying the modified R/S statistic, we test whether the returns of SSECI, SZSECI and the simulation data have the behavior of the long-term memory. By the calculation, the values of the modified R/S statistic of the absolute returns for SSECI, SZSECI and the financial model are 2.9853, 2.7655 and 3.8484 respectively, which are not in the intervals [0.809, 1.862]. Therefore, three financial series have the long-term memory.

5. The volatility clustering phenomena of the financial model

Many financial market time series show the volatility clustering phenomena. The empirical research in financial market fluctuations indicates that the daily returns of stock markets and the derivation markets show this phenomenon frequently. The volatility clustering phenomena’s essence is that the fluctuation in the current will impact on the certain day of fluctuation expected value in the future. In this section, We study the simulation data of the price model to
check whether it has this property, in the meantime, comparing with the actual Chinese Stock market indices, SSECI, SZSECI.

In order to test the volatility clustering phenomena of the daily returns sequence, we test the significance of the autocorrelation coefficients of the square returns by the Ljung-Box Q-statistics. Firstly, we compare the autocorrelation coefficients of the daily returns sequence with those of the square returns sequence. We have calculated the different order autocorrelation coefficients of the daily returns sequence and the square returns sequence of the simulation data, SSECI and SZSECI. The plots of the above two kinds of the returns sequences autocorrelation coefficient succession are given in Fig. 8. In Fig. 8, the ordinate axis expresses autocorrelation coefficients, and the abscissa axis expresses the number of autocorrelation coefficient lag. The two horizontal dashed lines in plots express the marginal value level of 95% confidence interval.

The comparisons of the autocorrelation coefficient for SSECI, SZSECI and the financial model are given in Fig. 8. The fluctuating patterns of the autocorrelation coefficient of two kinds of the returns sequences are entirely different when the lag exponent number increases. For the square returns sequence, the initial several order autocorrelation coefficients are quite large, and significantly different from zero. It is weakening slowly with the lag exponent number increasing. For the daily returns sequence’s autocorrelation coefficient, it has not displayed any fixed pattern. The autocorrelation coefficient of the daily returns sequence is a bit like a steady sequence. At the same time, we can see that, the movement of the actual markets graphs in Fig. 8(a) and (b) is very similar to that of the simulation data for the model in the figure (c). Moreover, the obvious results in Fig. 8 are that the autocorrelation coefficients of the daily returns sequence falls generally in the region composed by two parallel lines of 95% confidence interval. This indicates that the overwhelming majority autocorrelation coefficient is not significant. In addition, on the signs of the autocorrelation coefficients, the different order autocorrelation coefficients of the square returns sequence are larger than zero, but the signs of different order autocorrelation coefficients of the daily returns sequence actually are half-and-half. This indicates that the volatility of the returns have the significant autocorrelation. The above results show that the researched data show the volatility clustering phenomena. Similar to the realistic stock market, the simulation data also show the volatility clustering phenomena, which shows that the model is reasonable.

6. TARCH model

TARCH model was first proposed by Zakoian and Glosten et al. For the TARCH model we use the following specification of the conditional variance (see [2,4]):

$$\delta_t^2 = \omega + \alpha u_{t-1}^2 + \gamma \delta_{t-1}^2 I_{t-1} - 1 + \beta \delta_{t-1}$$

where the variance $\delta_t^2$ is a function of the past squared residuals $u_{t-1}^2$, and of its own lagged values $\delta_{t-1}^2$. The variable $I_{t-1}$ is a dummy variable equal to one if $u_{t-1} < 0$, and equal to zero otherwise. $\omega$, $\alpha$, $\gamma$, and $\beta$ are the parameters of the conditional variance equation that will be estimated.

In this model, we obtain good news $(u_t > 0)$ and bad news $(u_t < 0)$. This means that the model has differential effects on the conditional variance – good news has an impact on $\alpha$, while bad news has an impact on $\alpha + \gamma$. Asymmetry is a feature that is intended to capture the empirical regularity that positive and negative shocks of equal magnitude have different impacts on volatility, so we say that the asymmetric effect exists if $\gamma \neq 0$, the asymmetric effect makes the fluctuation of stock price larger when $\gamma > 0$, and smaller when $\gamma < 0$.

Next, we study the asymmetric effect of the data of SSECI, SZSECI and simulation by the TARCH model. The results of TARCH model are reported in Table 7.

From Table 7, the coefficient of asymmetric effect $\gamma$ are equal to 0.136, 0.090, and 0.015, not equal to 0, so asymmetric effect exist in these three markets. The empirical evidence suggests that the conditional variance of the three markets respond asymmetrically to past information. Since the leverage is intended to capture the possibility that negative shocks increase volatility while positive shocks decrease volatility, and the $\alpha$ (or ARCH) coefficient associated with positive shocks is not negative, there will not be a decrease in volatility arising from a positive shock, so there is no leverage effect exist in TARCH model (see [2]).
Fig. 8. The autocorrelation coefficient succession of the returns for SSECI (a), SZSECI (b) and the simulation data (c).
Table 7
TARCH model test.

<table>
<thead>
<tr>
<th></th>
<th>Sample size</th>
<th>$u$</th>
<th>$a$</th>
<th>$\gamma$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSECI</td>
<td>1455</td>
<td>$7.97 \times 10^{-6}$</td>
<td>0.106</td>
<td>0.136</td>
<td>0.814</td>
</tr>
<tr>
<td>S.E.</td>
<td></td>
<td>($1.23 \times 10^{-6}$)</td>
<td>(0.015)</td>
<td>(0.024)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>SZSECI</td>
<td>1454</td>
<td>$7.09 \times 10^{-6}$</td>
<td>0.103</td>
<td>0.090</td>
<td>0.838</td>
</tr>
<tr>
<td>S.E.</td>
<td></td>
<td>($7.28 \times 10^{-6}$)</td>
<td>(0.020)</td>
<td>(0.019)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Simulation</td>
<td>5000</td>
<td>$5.37 \times 10^{-7}$</td>
<td>0.133</td>
<td>0.015</td>
<td>0.743</td>
</tr>
<tr>
<td>S.E.</td>
<td></td>
<td>($1.39 \times 10^{-7}$)</td>
<td>(0.019)</td>
<td>(0.007)</td>
<td>(0.023)</td>
</tr>
</tbody>
</table>

Note: S.E. means standard error of regression.

7. Conclusion

In the paper, we introduce a stock price model by applying the voter system to study the behaviors of the fluctuations for the stock markets. We discuss the statistical properties of the model for the intensity $\lambda$, the lattice dimension $d$, the initial density $\theta$, and the multivariate set $(\theta, \lambda)$. And we compare the properties of the simulation data with those of the two groups of data from Shanghai stock market and Shenzhen stock market, including the power-law behavior, the long term memory and the volatility clustering phenomena. Through the comparisons between the real markets and the simulation market, we hope to show that the financial model of the present paper is reasonable to some extent.

Acknowledgments

The authors would like to thank the reviewers very much for helpful comments and suggestions. The reviewers kindly suggest us to discuss the long memory of market returns, the leading conditional and stochastic volatility models, and the univariate versus multivariate processes in this article. The manuscript is modified accordingly. Further we revise the results of the asymmetric effect and the leverage effect for TARCH model by following their suggestions. The reviewers’ comments and suggestions are very significant for this work.

The authors were supported in part by National Natural Science Foundation of China Grant No. 70771006 and Grant No. 10971010.

References