Understanding the complex dynamics of stock markets through cellular automata

G. Qiu, D. Kandhai, and P. M. A. Sloot
Section Computational Science, Faculty of Science, University of Amsterdam, Kruislaan 403, 1098 SJ Amsterdam, The Netherlands
(Received 13 February 2006; revised manuscript received 20 November 2006; published 26 April 2007)

We present a cellular automaton (CA) model for simulating the complex dynamics of stock markets. Within this model, a stock market is represented by a two-dimensional lattice, of which each vertex stands for a trader. According to typical trading behavior in real stock markets, agents of only two types are adopted: fundamentalists and imitators. Our CA model is based on local interactions, adopting simple rules for representing the behavior of traders and a simple rule for price updating. This model can reproduce, in a simple and robust manner, the main characteristics observed in empirical financial time series. Heavy-tailed return distributions due to large price variations can be generated through the imitating behavior of agents. In contrast to other microscopic simulation (MS) models, our results suggest that it is not necessary to assume a certain network topology in which agents group together, e.g., a random graph or a percolation network. That is, long-range interactions can emerge from local interactions. Volatility clustering, which also leads to heavy tails, seems to be related to the combined effect of a fast and a slow process: the evolution of the influence of news and the evolution of agents' activity, respectively. In a general sense, these causes of heavy tails and volatility clustering appear to be common among some notable MS models that can confirm the main characteristics of financial markets.

DOI: 10.1103/PhysRevE.75.046116 PACS number(s): 89.65.Gh

I. INTRODUCTION

The complex dynamics of financial markets can be characterized by some "stylized facts," which are common across many financial instruments, markets, and time horizons. Most of them are counterintuitive and contrary to the expectations of traditional financial theories. These stylized facts have been observed or discussed in many independent studies [1–7]. On long time scales (typically a week or longer), empirical distributions of financial return[1] generally fit the Gaussian distribution. However, most financial returns over short time scales are described well by a non-Gaussian (heavy-tailed or fat-tailed) distribution. A commonly used, although not rigorous, criterion for the normality of a distribution is its kurtosis (k): $k=3$ corresponds to a Gaussian distribution, whereas $k>3$ indicates a so-called leptokurtic distribution with a sharp peak and heavy tails. The kurtosis of financial returns is far from that of a Gaussian distribution. For instance, our estimate for the kurtosis of Standard and Poor's 500 (S&P 500)[2] daily returns over the period June 1950 to June 2005 is around 38. Figure 1(a) shows the distribution of these returns, together with a Gaussian probability density function (PDF) and a Lorentzian PDF for comparison. Clearly, daily returns of S&P 500 follow a non-Gaussian (fat-tailed) distribution, implying a greater frequency of extreme events than would be expected if they followed a normal distribution. However, the variance of the distribution is finite, whereas that of a Lorentz distribution (or a stable Lévy distribution in general) is infinite. Furthermore, the autocorrelation function (ACF) of the daily returns quickly converges to the noise range, whereas the corresponding ACF of volatility[3] decays slowly [see Fig. 1(b)]. The long-term autocorrelation of volatility is the reflection of the phenomenon termed "volatility clustering"—high (positive or negative) returns tend to group together. Figure 1(c) shows the time series of return over the period. In this figure, the effect of volatility clustering is clearly illustrated.

Traditional (analytical) approaches in finance and economics to study aggregative phenomena either are purely macroscopic, or rely on top-down construction based on a number of unrealistic assumptions mainly for the sake of analytical tractability. Interactions between traders play no role in the explanation of the phenomena [8]. In fact, markets consist of a large number of agents. The interactions between

[1]Generally, return is defined as $R_1^{t+1}=\ln P^{t+1}-\ln P^t$, where R_1^{t+1} is the return at time $t+1$, P^t is the price at time t, and so on. The basic relation is $R_2^{t+1}=(P^{t+1}-P^t)/P^t$. Sometimes, return is defined as price change, $R_3^{t+1}=P^{t+1}-P^t$. For high-frequency data, $|R_3^{t+1}|\ll P^t$. Hence, $R_1^{t+1}=\ln(1+R_3^{t+1}/P^t)\simeq R_3^{t+1}/P^t=R_2^{t+1}$. Since R_3^t is a fast variable and P^t is a slow variable, $R_2^t\simeq CR_3^t$, where the time dependence of C is negligible. (See pp. 35–39 of Ref. [6].) In our simulations, $|P^{t+1}-P^t|\ll P^t$, so $R_1^t\simeq R_2^t\simeq CR_3^t$. Within our cellular automation model, these three indicators are alternatives to each other in analyzing the regularity in return distributions.

[2]An index is a sample list of stocks that is representative of a whole stock market. It is used by investors to track the performance of the stock market. Different methods are used for calculating the price of an index. For example, the Dow Jones Industrial Average, which contains 30 of the most influential companies in the United States, is the price-based weighted average of the prices of the included stocks. The Standard and Poor's 500 Index (S&P 500), which includes 500 large publicly held companies that trade on major U.S. stock exchanges, weights companies by market capitalization (the overall value of a company's stock on the market).

[3]In the finance literature, volatility refers to the spread of asset returns measured as the standard deviation of a sample of returns over a period of time, i.e., $\sigma=\sqrt{(1/T)\Sigma_{t=1}^{T}(R^t-\bar{R})^2}$, where T is the length of the period, R^t is the return at time t, and $\bar{R}$ is the average return over the period. Substituting $T=1$ and $\bar{R}=0$ into this equation, we obtain the absolute value of the return over a period of one time unit, $|r|$, which is the most commonly used proxy for volatility in practice. The other commonly used proxy is r^2. We adopt $|r|$ as the volatility.

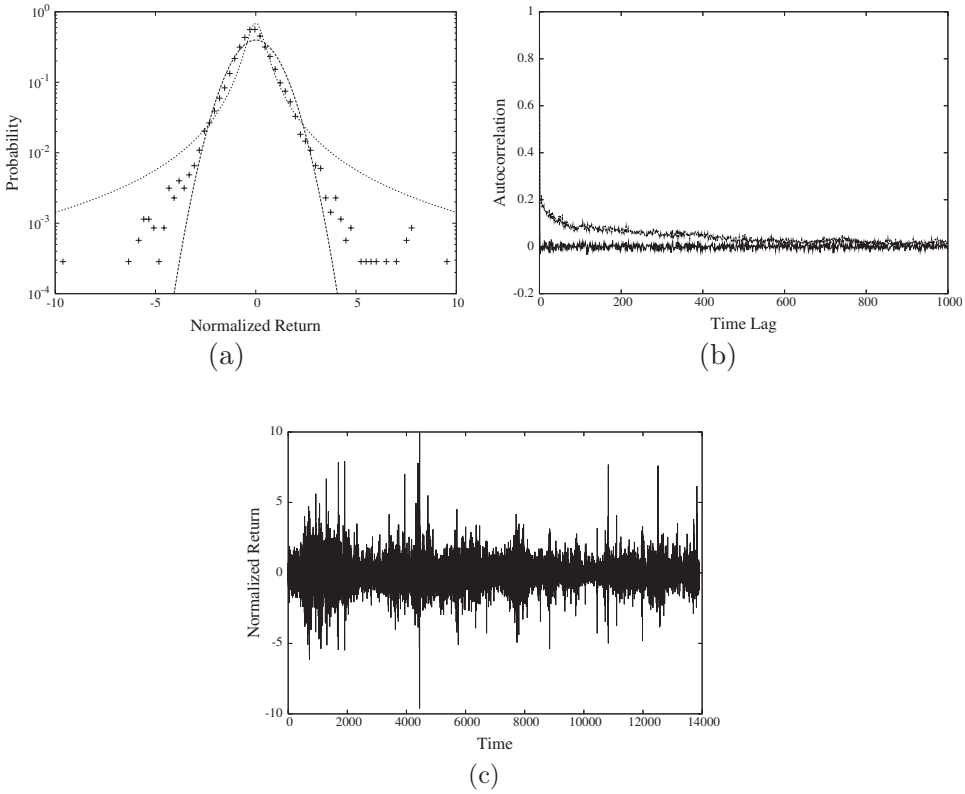


FIG. 1. (a) Distribution of the daily returns of S&P 500 over the period from June 1950 to June 2005 (the points), compared with a Gaussian PDF (the curve that decays faster) and a Lorentzian PDF. A logarithmic scale is used for the vertical axis. (b) Autocorrelation function of the daily returns (the lower line) and the corresponding ACF of volatility. (c) Time series of the daily returns.

them give rise to some macroeconomic regularity, which in turn influences the microscopic interactions. This dynamics is highly nonlinear and difficult to describe analytically.

In recent years, researchers have used microscopic simulation (MS) to explore complex economic dynamics from the bottom up. With MS, we study a complex system by directly modeling its individual elements and their interactions. The macroscopic behavior of the system will eventually emerge from the microdynamics. MS has shown great potential for more realistically modeling complex dynamical systems in economics and finance [9]. In addition, it facilitates the testing of existing economic or financial models and theories, and the development of new theories and models [8]. At present, most research on MS in finance focuses on understanding the characteristics of financial markets. To achieve this objective, many MS models of financial markets have been developed during the last decade.

However, as will be shown in Sec. II, researchers in the field have not yet reached an agreement on explaining the complex dynamics of financial markets. In addition, as recently pointed out by Cont [10], due to the complexity of the existing (agent-based) models, it is often not clear which aspects of the models are responsible for generating the stylized facts, and whether all their ingredients are indeed required for explaining empirical observations.

In view of the facts, our general motivation is to develop a MS model with a simple structure that can reproduce the main stylized facts. More importantly, the causalities of the dynamics generated by the model can be clearly identified. In particular, we wish to confirm the main stylized facts within a parsimonious cellular automaton (CA) framework by dealing with local agent interactions and adopting simple rules for representing agents' behavior and a simple rule for price updating. Building on the previous work presented in Ref. [11], we have constructed a particular CA model in order to achieve these objectives. In our simulations, we make sure that the model parameters have clear economic relevance or interpretations.

In Sec. II, we first give an overview of the most notable MS models of financial markets that have been reported in the literature. We then give a detailed description of our CA model in Sec. III, presenting it in order of increasing sophistication so that the cause(s) of certain stylized fact(s) can be identified at each level of sophistication. The simulation results of this model are shown in Sec. IV. Section V provides a thorough investigation of the simulated dynamics through computational experiments and mathematical analysis. In Sec. VI, we present our conclusions.

II. OVERVIEW OF SOME MS MODELS

Bak *et al.* have developed a MS model [12] in which the stock market contains fundamental value traders and noise traders. The former set their prices based on a utility function, whereas the behavior of the latter is characterized by drifting their prices to spot prices and copying other traders' prices (imitation). The strength of this drifting is proportional to recent price variation. This mechanism is called "volatility feedback." This model can generate non-Gaussian distributions of return and volatility clustering. Its volatility feedback is similar to the mechanism of activity adjustment within our proposed CA model, whereas other features are different.

Within the model of Lux and Marchesi [13], the market consists of two groups of traders: fundamentalists and noise traders. Fundamentalists buy (sell) the asset when its market

price is below (above) its fundamental value. Noise traders are further differentiated into optimists and pessimists. The former believe in a rising market and buy the asset, whereas the latter believe in a declining market and sell. Agents move to the other group or subgroup when they believe that the traders in that (sub)group are more successful (imitation or herding). Volatility clustering and fat-tailed distributions of return can be produced by this model. While they are different in other aspects, this model and our model are virtually identical with regard to the behavior of fundamentalists and price updating.

The model of Cont and Bouchaud [14] deals with homogeneous traders who group together in clusters through binary links between them. Such a structure is known as a "random graph." When certain conditions are satisfied, the sizes of the clusters follow a power law distribution. The members in each cluster coordinate their individual demands to decide whether to buy, sell, or not trade (herding). This model gives rise to heavy-tailed probability distributions of price change. However, it cannot generate volatility clustering. In principle, the price updating rule of this model is identical to the corresponding rule within our model, although these two models are very different in other aspects.

Cellular automata have been widely applied to study complex phenomena in different fields such as physics, chemistry, biology, and social and economic sciences, etc. [15]. Recently, some researchers have used cellular automata to model financial markets for studying their complex dynamics.

Iori has developed a CA-type model [16] that represents a market as a two-dimensional lattice, of which each node is an agent connected with four neighbors. The decision making of each agent is driven by his own signal and the signals of his neighbors (imitation). When the aggregate signal exceeds his activation threshold, he will transact. The thresholds are adjusted over time in response to price changes. This model generates fat-tailed distributions of return and volatility clustering. Differing from each other in other aspects (such as the nonlinear price updating rule employed here compared to the simple rule used within our model), both this model and our model adopt only local interactions.

In [17], Bartolozzi and Thomas present a stochastic CA model of stock markets. In simulations using this model, clusters of active traders form and evolve over time through percolation dynamics (herding). This process produces a power law distribution of cluster size. Similar to the process of a random Ising model, traders within each cluster exchange information and update their states. The model can produce heavy tails and volatility clustering. The price updating schemes adopted within this model and our model are similar, although they are distinct in other aspects. One important difference is that within this model the traders act in groups, whereas within our model long-range interactions emerge from local interactions.

The model of Bandini *et al.* [11] adopts agents of two types: fundamentalists and imitators. The former trade in quantities proportional to the differences between their perceived fundamental values of the asset and spot prices. The latter follow the actions of their neighbors (imitation). Reference [11] does not report any stylized facts. This model and our CA model are similar with regard to the basic behavior of the agents and the price updating, whereas other features are completely different.

We can see that, although most of these MS models can confirm the stylized facts, they are different in agent types, description of agents' behavior, and ultimately market dynamics. However, it should be noticed that all these models share a typical kind of behavior, namely, imitation. Some of these models have been studied in Ref. [18].

III. DESCRIPTION OF OUR CA MODEL

A. Modeling stock markets with cellular automata

We represent a stock market as a two-dimensional $L \times L$ lattice. Each vertex of the lattice denotes an agent (trader) who has a Moore neighborhood.[4] Within our model, speculative traders of only two types are adopted: fundamentalists and imitators. All the agents trade in a single stock.

Fundamentalists are those traders who are informed of the nature of the stock being traded and act according to its fundamental value. They believe that the price of the stock may temporarily deviate from, but will eventually return to, the fundamental value. They therefore buy (sell) the asset whenever its price is lower (higher) than the fundamental value as perceived by them. In stock markets, there are also some traders who do not know or do not care about fundamental values. Instead, they follow their acquaintances and adopt the trading opinions of the majority. An agent of this type is referred to as an imitator.

News influences both fundamentalists and imitators. However, the ways news affects them are distinct in many aspects. For example, fundamentalists pay relatively more attention to news about the specific company that has issued the stock, while imitators respond comparatively more frequently to news related to the stock market as a whole.

We can adopt other types of agents to model stock markets more realistically. However, we think that the two kinds of behavior discussed here are the most typical. The behavior of other speculative traders has no obvious characteristics. For example, we cannot find a general trait for chartists, because, even using the same data, they may come to different conclusions due to differences in the techniques used. We can treat these agents as noise traders who randomly influence the price to different extents. However, because their adoption within our model does not fundamentally influence the dynamics characterized by the stylized facts, we choose to ignore them.

The real fundamental value of a stock is related to the current and prospective states of the company that has issued the stock, among many other factors. The modeling of its variations is beyond the scope of this work. Instead, we are more interested in the reason(s) for excess volatility, i.e., the extra factor(s) causing the price of a stock to be more volatile than its real fundamental value. For this reason, we assume

[4] A Moore neighborhood $\mathrm{NB}_{i,j}$ in our two-dimensional lattice is defined as a set whose members are the eight cells surrounding a given cell located at (i,j), i.e., $\mathrm{NB}_{i,j}=\{(i-1,j-1),(i,j-1),(i+1,j-1),(i-1,j),(i+1,j),(i-1,j+1),(i,j+1),(i+1,j+1)\}$.

that the real fundamental value of the asset F is a constant. (Tests showed that adding a drift to F to model the time value of money does not influence the characteristics of the returns. Drifts are therefore excluded from our model.)

B. Level I model

1. Fundamentalists

Empirically, the larger the difference between the price of a stock and its fundamental value as perceived by a fundamentalist, the more likely he will trade it. We assume for the moment that the fundamentalists perceive the real fundamental value accurately. We can then adopt Eq. (1) to express the transaction quantity based on the current price level at time $t+1$ of a fundamentalist when he is the ith agent, $V_{i,fu}^{t+1}$, and his actual transaction quantity at the same time, $q_{i,fu}^{t+1}$:

$$q_{i,fu}^{t+1} = V_{i,fu}^{t+1} = F - P^t, \tag{1}$$

where P^t is the price at time t. Notice that we have assumed for the moment that the two transaction quantities are equivalent. [The other factor determining (actual) transaction quantities will be introduced in Sec. III D.]

2. Imitators

We take the average transaction quantity based on the current price level of an imitator's neighbors at the previous time step as his corresponding quantity at present, i.e., $V_{i,im}^{t+1} = \langle V_{i,nb}^t \rangle$. We can then use Eq. (2) to express his (actual) transaction quantity at time $t+1$, $q_{i,im}^{t+1}$:

$$q_{i,im}^{t+1} = V_{i,im}^{t+1} = \langle V_{i,nb}^t \rangle. \tag{2}$$

C. Level II model

1. Fundamentalists

News influences fundamentalists' perceptions of fundamental values. Positive (negative) news can cause them to overestimate (undervalue) assets. Within our model, we assume that at each time step all the fundamentalists perceive the fundamental value identically. (We can alternatively assume that their perceived values at each time step are normally distributed, without fundamentally influencing the dynamics.)

We express the perceived fundamental value at time t as $F\eta_{fu}^t$, in which η_{fu}^t denotes the influence of the news at that time. We assume that $\eta_{fu}^t = 1 + c_{fu}\phi_{fu}^t$, where ϕ_{fu}^t is an independent Gaussian random variable with mean 0 and standard deviation 1 and c_{fu} is a positive parameter indicating the fundamentalists' sensitivity to news. At this point, we have a modified expression for the transaction quantity of a fundamentalist,

$$q_{i,fu}^{t+1} = V_{i,fu}^{t+1} = F\eta_{fu}^{t+1} - P^t. \tag{3}$$

2. Imitators

We assume that news influences all the imitators identically. (We can alternatively assume that the effects of news at each time step are normally distributed, without fundamentally influencing the dynamics.) Significant (unimportant) news can make an imitator trade more (less) than his neighbors, and vice versa. We reformulate the transaction quantity of an imitator as

$$q_{i,im}^{t+1} = V_{i,im}^{t+1} = \langle V_{i,nb}^t \rangle \eta_{im}^{t+1}, \tag{4}$$

in which η_{im}^{t+1} indicates the influence of the news at time $t+1$ and is equal to $1 + c_{im}\phi_{im}^{t+1}$, where ϕ_{im}^t is an independent Gaussian random variable with mean 0 and standard deviation 1 and c_{im} is a positive parameter indicating the imitators' sensitivity to news.

D. Level III model

A common strategy used by traders is to buy low and sell high (BLASH). It aims for capital gains by taking advantage of changes in prices. Price fluctuations are therefore indispensable for this strategy.

Based on the BLASH strategy, capitals of traders move among different assets pursuing larger profits at lower risks. When the price fluctuation level of a stock is at the two extremes, i.e., very low and very high, the asset is the least desirable: If it is very low, traders who hold the asset will not be able to find an opportunity to sell it profitably and will not even be able to cover their opportunity costs.[5] If it is very high, traders will consider the investment in the asset too risky. Within the range between the two extremities, as the price fluctuation level rises, the asset will be first more favorable and then, after a certain level, less attractive.

When a stock is more favorable compared to other alternatives, traders will trade it more frequently. We therefore assume that the trading activity of the agents is equivalent to the desirability of the stock. However, BLASH is a risky approach itself, because there is no way to predict price changes accurately. Frequently, traders just end up selling at a loss. In order to reduce this risk, traders typically consider previous price changes of a stock for a longer period.

We represent the price fluctuation level of a stock at time t as

$$L^t = \frac{1}{k}\sum_{i=t-k}^{t-1} |P^i - \bar{P}|/\bar{P}, \tag{5}$$

where k is the length of a period before t, P^i is the price of the asset at time i in the period, and $\bar{P}$ is the average price over the period. (We can alternatively assume that agents take different values of k that are normally distributed, without fundamentally influencing the dynamics.)

For the sake of simplicity, we adopt a straightforward linear function for the trading activity of the agents,

[5] Opportunity cost, or cost of capital, is the rate of return that a business could earn if it chose another investment with equivalent risk [26].

$$M^t(L^t) = \begin{cases} c_l L^t, & L^t \leqslant L_m, \\ c_l(-L^t + 2L_m), & L^t > L_m, \end{cases} \tag{6}$$

where L_m is the fluctuation level where the stock becomes less favorable and c_l is a positive parameter. (Simulations show that adopting other concave functions leads to similar results.)

Within the level III model we consider that the (actual) transaction quantity of an agent is the product of his transaction quantity based on the current price level and his current trading activity. The transaction quantity of a fundamentalist is therefore

$$\begin{aligned} q_{i,fu}^{t+1} &= V_{i,fu}^{t+1} M^{t+1} \\ &= (F\eta_{fu}^{t+1} - P^t) M^{t+1}, \end{aligned} \tag{7}$$

whereas the transaction quantity of an imitator is

$$\begin{aligned} q_{i,im}^{t+1} &= V_{i,im}^{t+1} M^{t+1} \\ &= \langle V_{i,nb}^t \rangle \eta_{im}^{t+1} M^{t+1}. \end{aligned} \tag{8}$$

Considering the fact that agents always have a number of exceptional reasons to transact, we adopt a lower bound for M^t.

E. Rule of price updating

The price is updated according to the following rule:

$$P^{t+1} = P^t + \frac{c_p Q^t}{N}, \tag{9}$$

where Q^t is the total transaction quantity or the excess demand for the asset at time t and N is the number of traders. Since Q^t is proportional to N, we rescale it with N. We adopt a positive parameter c_p to indicate the sensitivity of the price to the excess demand. Due to the fact that stock prices cannot be negative, the lower bound of P^t is 0.

Equation (9) can be explained as the action of a market maker[6] to balance the supply and the demand of the stock. In principle, however, it is merely the translation of the classic theory of supply and demand stating that price will move toward the point that equalizes supplied and demanded quantities.[7]

[6]To ensure liquidity, many organized exchanges use market makers, individuals who maintain inventories of their chosen securities and stand ready to buy or sell whenever the public wishes to sell or buy.

[7]In 1890, Marshall published *Principles of Economics* [27], in which he discussed how both supply and demand interact to determine price. His supply-demand model has become one of the fundamental concepts of economics. According to the model, if all other factors remain equal, the higher the price, the lower the quantity demanded and the higher the quantity supplied, vice versa. In a price (ordinate)–quantity (abscissa) chart the curve of demand is a downward slope, and the supply relationship shows an upward slope. Equilibrium occurs at the intersection point of the two curves. In the chart, if straight lines are drawn instead of the more general curves (the shapes of the curves do not change the general relationships), we immediately obtain Eq. (9).

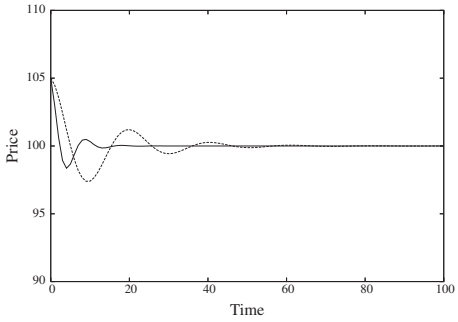


FIG. 2. Price trajectories obtained through the simulation using the level I model when $\alpha_{im}=20\%$ and 80%, respectively. The curve that decays faster is of the first instance. The parameter settings used are $N=1\times10^4$, $c_p=0.5$, $F=100$, and $P^0=105$.

IV. SIMULATION RESULTS

A. Simulation results of the level I model

In simulations using the level I model, we set an initial price ($P^0=105$) that deviates from the fundamental value ($F=100$). The number of agents is 1×10^4. Figure 2 displays the price trajectories corresponding to two fractions of imitators: $\alpha_{im}=20\%$ and 80%, respectively.

The parameter c_p has an important impact on the price: When its value is increased up to 1 while other parameters are kept constant, the price process may start to switch from a convergent process to a divergent one, depending on the value of c_p itself and the value of α_{im}. We provide a theoretical analysis of this issue in Sec. V. To model a stable market, we adopt only those values of c_p smaller than 1.

As shown in Fig. 2, although the level I model is not completely identical to the model of Bandini *et al.* [11], it does generate similar price trajectories. Starting from an initial deviation from the fundamental value, the price either directly converges to it, or fluctuates around it for some time and eventually overlaps. Obviously, both models cannot produce sustained price movement. Since the price quickly dies out, we cannot obtain any stylized facts by using the level I model.

B. Simulation results of the level II model

Within the level II model, we have added random factors η_{fu}^t and η_{im}^t, so that it can produce sustained price fluctuations. When the fraction of imitators is set to 70%, we obtain simulation results shown in Fig. 3. In our simulations, return is represented by the difference between two successive natural logarithms of price, i.e., the log-return is used.

We see that the level II model can generate a non-Gaussian (fat-tailed) distribution of return, but is not able to confirm another important stylized fact, namely, volatility clustering. It therefore has the same problem presented by the Cont-Bouchaud model [14]. Nevertheless, through simulations using this model, we can further study how the fraction of imitators influences the distribution of return. Figure 4 shows the results for different instances: $\alpha_{im}=20\%$, 50%, and 80%, respectively. If the fraction is small, returns will follow a Gaussian distribution; increasing it enlarges the tails of the return distribution.

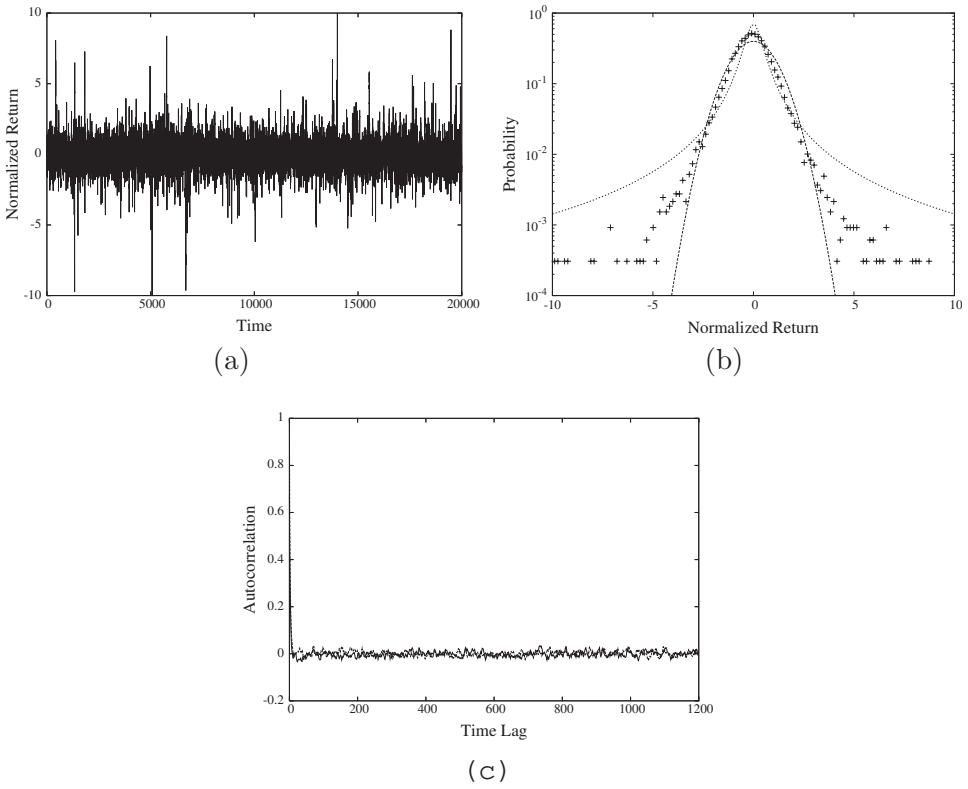


FIG. 3. Simulation results of the level II model when $\alpha_{im}=70\%$. (a) Normalized return. (b) Distribution of return (the scale of the vertical axis is logarithmic). (c) Autocorrelation function of return and that of volatility. The parameter settings used are $N=1\times10^4$, $c_p=0.005$, $F=100$, $c_{fu}=0.2$, $c_{im}=0.7$, and $P^0=100$.

C. Simulation results of the level III model

Within the level III model we have further added a mechanism through which agents' activity is adjusted over time. Fixing α_{im} to 70%, we obtain simulation results shown in Fig. 5. (1) Figure 5(a) records the price process. Some large "flights" can be observed, which correspond to large (positive or negative) returns. (2) Figure 5(b) illustrates the time series of return. The effect of volatility clustering is clear. (3) Figure 5(c) shows the probability distribution of return, together with a Gaussian PDF and a Lorentzian PDF for comparison. The tails of the distribution are clearly heavier than those of a Gaussian PDF. (4) Figure 5(d) displays the ACF of return (the lower curve) and that of volatility. The former converges quickly to the noise range, whereas the latter decays much more slowly. (5) Figure 5(e) shows the time evolution of trading volume.[8] (6) Figure 5(f) illustrates the time evolution of trading activity. It is a slow process in comparison with the fast evolution of the influence of news.[9]

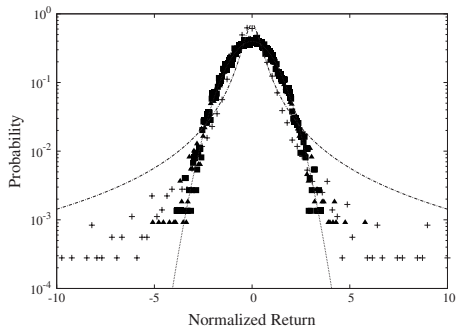


FIG. 4. Return distributions of the level II model for different fractions of imitators. The ■ points, the ▲ points, and the + points refer to $\alpha_{im}=20\%$, 50%, and 80%, respectively. The scale of the vertical axis is logarithmic. The parameter settings used are $N=1\times10^4$, $c_p=0.005$, $F=100$, $c_{fu}=0.2$, $c_{im}=0.7$, and $P^0=100$.

These simulation results indicate that our CA model (level III) is able to reproduce the main stylized facts. In addition, as shown below, this model is robust with regard to the stylized facts for wide ranges of the parameters.

Table I compiles the kurtosis values of the return distributions for different values of α_{im}, c_{im}, and c_{fu}, respectively. We see that imitators have a strong influence on kurtosis, while the relation between fundamentalists and kurtosis is not explicit. Specifically, the fraction of imitators α_{im} and the sensitivity of imitators to news c_{im} are positively correlated with kurtosis. When either of them increases to a certain level, kurtosis suddenly becomes very large, implying that the system becomes unstable. For example, when $c_{im}=0.9$, we obtain a price pattern with frequent dramatic "flights" and a time series of return with many striking strokes. These are shown in Fig. 6.

Keeping other parameters constant and adopting different values of k, we obtain the autocorrelation functions of vola-

[8]Volume is defined as the sum of absolute aggregate demand and absolute aggregate supply.

[9]Here, we define the rate of time evolution of a variable X as $(|\Delta X|/|X|)/\Delta t$, where ΔX is the change of X within time increment Δt.

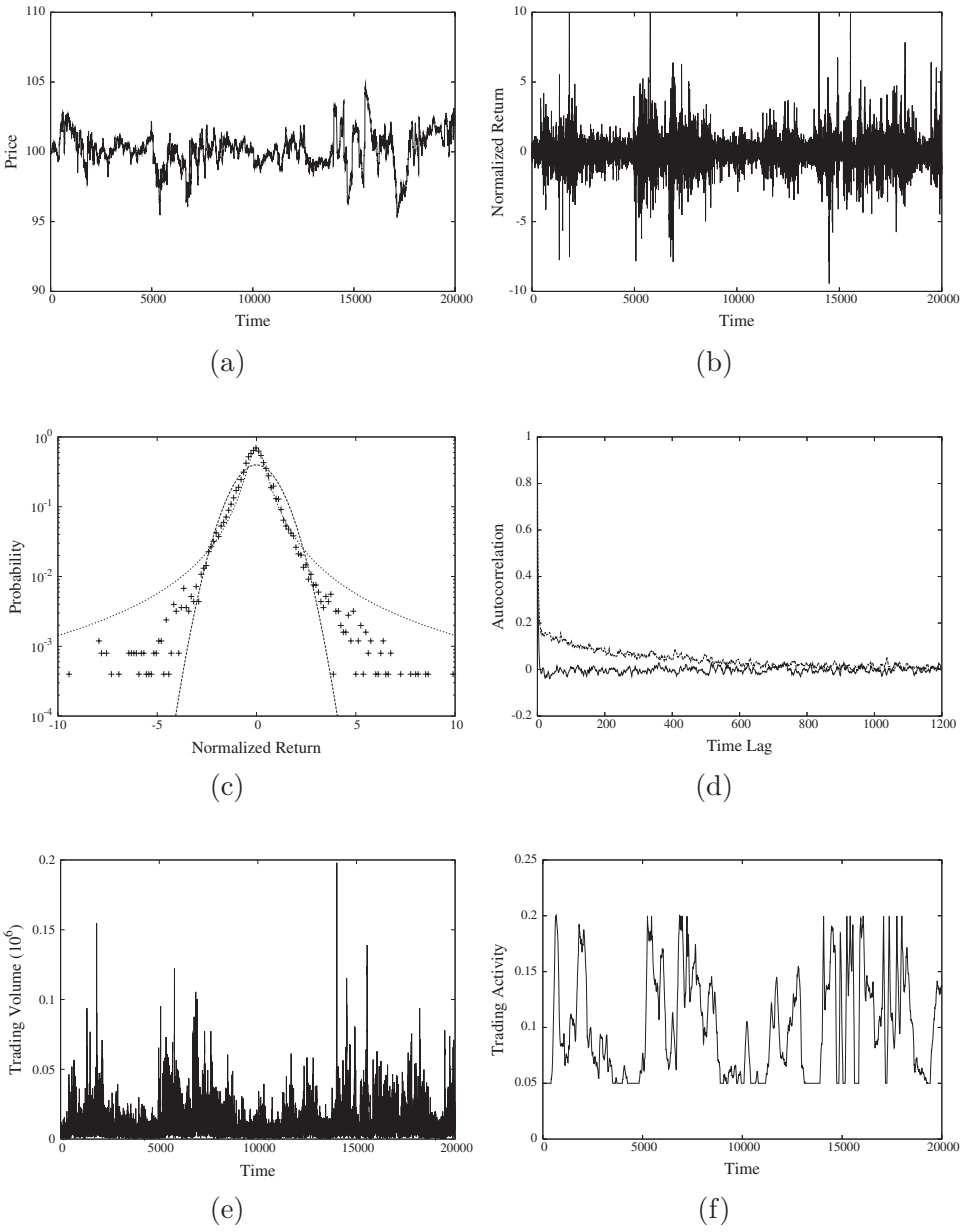


FIG. 5. Simulation results of the level III model when 70% of the agents are imitators. (a) Price. (b) Normalized return. (c) Distribution of return (the scale of the vertical axis is logarithmic). (d) Autocorrelation function of return (the lower curve) and that of volatility. (e) Trading volume. (f) Trading activity. The parameter settings used are $N=1\times10^4$, $c_p=0.05$, $F=100$, $c_{fu}=0.2$, $c_{im}=0.7$, $P^0=100$, $k=400$, $c_l=20$, and $L_m=0.01$. The lower bound of M^t is 0.05.

tility shown in Fig. 7(a). When k is smaller than 50, ACFs of volatility quickly drop to the noise range and the effects of volatility clustering are correspondingly negligible. Volatility clustering becomes significant when k is increased to around 100. The importance of k to volatility clustering will further manifest itself in Sec. V D.

Choosing three values for the number of agents (lattice sizes) and keeping other parameters constant, simulations give the ACFs of volatility shown in Fig. 7(b). All these ACFs are qualitatively similar to that of S&P 500 shown in Fig. 1(b), indicating that the model can reproduce the stylized facts not only for markets with small numbers of agents, but also for markets with many agents. At this point, the model differs from some MS models that behave realistically only for limited numbers but not large numbers of traders [22].

V. DISCUSSION: THE MARKET DYNAMICS REVEALED BY OUR CA MODEL

A. Long-range interactions can form from local interactions

In this section, for the sake of simplicity, we take a one-dimensional version of our CA model to derive analytical expressions. An agent located at i then has two neighbors at $i-1$ and $i+1$, respectively. Statistically, the total quantity at time $t+1$ can be expressed as

$$Q^{t+1}=\sum_{i=1}^{N} q_{i,(\cdot)}^{t+1}=\sum_{i=1}^{N}[u_i q_{i,fu}^{t+1}+(1-u_i)q_{i,im}^{t+1}], \qquad (10)$$

where u_i is determined in the following way: We sample a variable γ $(0\leqslant\gamma\leqslant1)$ that is uniformly distributed. If $0\leqslant\gamma\leqslant\alpha_{fu}$, $u_i=1$, else $u_i=0$. Here, α_{fu} is the fraction of fundamentalists.

TABLE I. Kurtosis values of the return distributions produced by the level III model for increasing values of α_{im}, c_{im}, and c_{fu}, respectively. The parameter settings used are $N=1\times10^4$, $c_p=0.05$, $F=100$, $c_{fu}=0.2$, $c_{im}=0.7$, $P^0=100$, $k=400$, $c_l=20$, and $L_m=0.01$. The lower bound of M^t is 0.05.

α_{im}	Kurtosis	c_{im}	Kurtosis	c_{fu}	Kurtosis
0%	4.44	0.1	4.36	0.1	15.53
10%	4.36	0.2	5.37	0.2	17.22
20%	4.57	0.3	6.37	0.3	42.94
30%	5.33	0.4	7.44	0.4	37.28
40%	6.07	0.5	8.60	0.5	35.39
50%	6.64	0.6	10.55	0.6	46.71
60%	8.51	0.7	17.22	0.7	43.98
70%	17.22	0.8	32.57	0.8	42.11
80%	69.33	0.9	119.14	0.9	35.79
90%	190.46	1.0	422.77	1.0	30.72

The terms $q_{i,fu}^{t+1}$ and $q_{i,im}^{t+1}$ in Eq. (10) are determined by Eqs. (7) and (8), respectively. However, because M^t changes much more slowly than Q^t, we can consider the former as a constant to study the basic dynamics of the latter. We set $M^{(\cdot)}=1$; then $q_{i,fu}^{t+1}$ and $q_{i,im}^{t+1}$ are respectively determined by Eq. (3) and Eq. (4). Therefore,

$$q_{i,im}^{t+1} = \eta_{im}^{t+1}\left(\frac{1}{2}\right)[V_{i-1,(\cdot)}^{t} + V_{i+1,(\cdot)}^{t}]$$
$$= \eta_{im}^{t+1}\left(\frac{1}{2}\right)\{[u_{i-1}V_{i-1,fu}^{t} + (1-u_{i-1})V_{i-1,im}^{t}]$$
$$+ [u_{i+1}V_{i+1,fu}^{t} + (1-u_{i+1})V_{i+1,im}^{t}]\}. \quad (11)$$

Similarly, the terms $V_{i-1,im}^{t}$ and $V_{i+1,im}^{t}$ in Eq. (11) can be respectively expressed as

$$V_{i-1,im}^{t} = \eta_{im}^{t}\left(\frac{1}{2}\right)\{[u_{i-2}V_{i-2,fu}^{t-1} + (1-u_{i-2})V_{i-2,im}^{t-1}]$$
$$+ [u_{i}V_{i,fu}^{t-1} + (1-u_{i})V_{i,im}^{t-1}]\} \quad (12)$$

and

$$V_{i+1,im}^{t} = \eta_{im}^{t}\left(\frac{1}{2}\right)\{[u_{i}V_{i,fu}^{t-1} + (1-u_{i})V_{i,im}^{t-1}]$$
$$+ [u_{i+2}V_{i+2,fu}^{t-1} + (1-u_{i+2})V_{i+2,im}^{t-1}]\}. \quad (13)$$

Following the same scheme, we can further express the terms $V_{i-2,im}^{t-1}$, $V_{i,im}^{t-1}$, and $V_{i+2,im}^{t-1}$ in Eqs. (12) and (13) in terms of the corresponding quantities at time step $t-2$ of the neighbors of the agents located at $i-2$, i, and $i+2$, respectively, and so on. Basically, in this way, we can replace each imitator's transaction quantity based on the current price level at each time step with the fundamentalists' corresponding quantities at the preceding time steps, noting that $V_{(\cdot),fu}^{t}=V_{i,fu}^{t}$. After substitutions, we have

$$Q^{t+1} = \sum_{i=1}^{N}[A_i^{t+1}V_{i,fu}^{t+1} + A_i^{t}(\eta_{im}^{t+1})V_{i,fu}^{t} + A_i^{t-1}(\eta_{im}^{t+1}\eta_{im}^{t})V_{i,fu}^{t-1}$$
$$+ A_i^{t-2}(\eta_{im}^{t+1}\eta_{im}^{t}\eta_{im}^{t-1})V_{i,fu}^{t-2} + \cdots$$
$$+ A_i^{t-\tau}(\eta_{im}^{t+1}\eta_{im}^{t}\eta_{im}^{t-1}\cdots\eta_{im}^{t-\tau+1})V_{i,fu}^{t-\tau} + \cdots], \quad (14)$$

where $\tau=-1,0,1,2,\ldots$. The first few instances of $A_i^{t-\tau}$ are

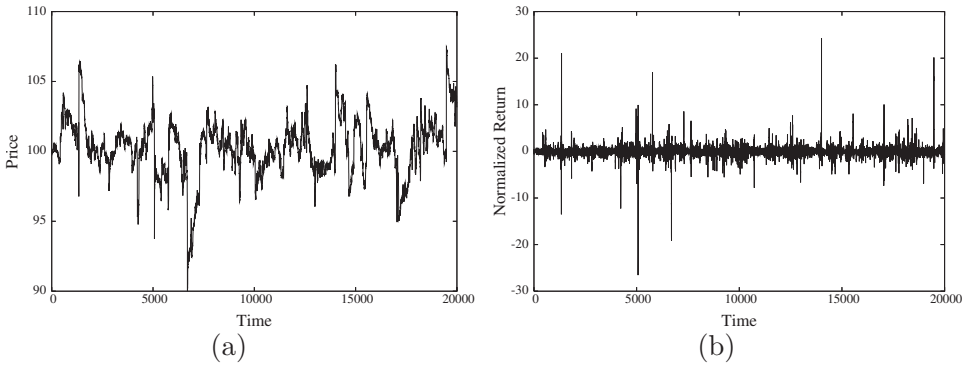
$$A_i^{t+1} = u_i,$$

$$A_i^{t} = \frac{1}{2}[(1-u_i)u_{i-1} + (1-u_i)u_{i+1}],$$

$$A_i^{t-1} = \frac{1}{2^2}[(1-u_i)(1-u_{i-1})u_{i-2} + (1-u_i)(1-u_{i-1})u_i$$
$$+ (1-u_i)(1-u_{i+1})u_i + (1-u_i)(1-u_{i+1})u_{i+2}],$$

$$A_i^{t-2} = \frac{1}{2^3}[(1-u_i)(1-u_{i-1})(1-u_{i-2})u_{i-3} + (1-u_i)(1-u_{i-1})$$
$$\times(1-u_{i-2})u_{i-1} + (1-u_i)(1-u_{i-1})(1-u_i)u_{i-1}$$
$$+ (1-u_i)(1-u_{i-1})(1-u_i)u_{i+1} + (1-u_i)(1-u_{i+1})$$
$$\times(1-u_i)u_{i-1} + (1-u_i)(1-u_{i+1})(1-u_i)u_{i+1} + (1-u_i)$$
$$\times(1-u_{i+1})(1-u_{i+2})u_{i+1} + (1-u_i)(1-u_{i+1})$$
$$\times(1-u_{i+2})u_{i+3}].$$

In each product within $A_i^{t-\tau}$, the sequence of $1-u_i$ terms indicates the propagation of imitation over time (backward) and space (agents). However, if at least one of the terms is



(a) (b)

FIG. 6. Simulation results of the level III model when $\alpha_{im}=70\%$ and $c_{im}=0.9$. (a) Price. (b) Normalized return. The parameter settings used are $N=1\times10^4$, $c_p=0.05$, $F=100$, $c_{fu}=0.2$, $P^0=100$, $k=400$, $c_l=20$, and $L_m=0.01$. The lower bound of M^t is 0.05.

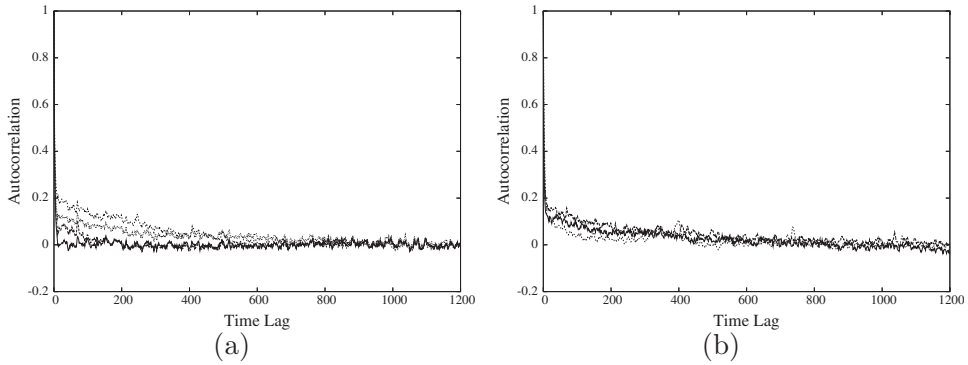


FIG. 7. Autocorrelation functions of volatility produced by the level III model when $\alpha_{im}=70\%$. (a) ACFs when $k=10$ (the lowest curve), 100 (the second lowest), 300 (the highest), and 500 (the second highest), respectively. (b) ACFs when $N=10\times10$ (the middle curve), 100×100 (the highest one), and 1000×1000 (the lowest one), respectively. The parameter settings used are $N=1\times10^4$, $c_p=0.05$, $F=100$, $c_{fu}=0.2$, $c_{im}=0.7$, $P^0=100$, $k=400$, $c_l=20$, and $L_m=0.01$. The lower bound of M^t is 0.05.

equal to zero, which corresponds to a fundamentalist, the whole product will be zero. As the fraction of imitators (fundamentalists) increases (decreases), some $A_i^{t-\tau}$ terms with larger τ values are greater than zero.

The imitation chains show that long-range interactions can form from local imitations. In the resultant networks, each agent is influenced, directly or indirectly, by some other near or remote agents. Here, the strengths and time lags of influence differ. In this respect, our CA model is different from the Cont-Bouchaud model, where any two agents can be directly linked, and agents in a group behave identically. It is also distinct from the model of Bartolozzi *et al.*, within which agents in a cluster influence each other with an equivalent strength.

B. Price and volatility are mean reverting

Fundamentalists behave according to price while imitators follow other agents but do not directly respond to price. We therefore argue that it is the fundamentalists' behavior that determines the price trend. This argument can be confirmed by our simulations: If $\alpha_{fu}=0$ (all the agents are imitators), price fluctuations die out; in other cases we obtain price trajectories similar in shape but distinct only in amplitude.

Therefore, for the sake of simplicity, we can take the special instance that all the agents are fundamentalists to study the basic dynamics of the price. In such an instance, $\alpha_{fu}=1$; hence $u_{(\cdot)}=1$. Then, Eq. (14) gives

$$Q^t = NV_{i,fu}^t = N(F\eta_{fu}^t - P^{t-1}). \tag{15}$$

The noise term η_{fu}^t is indispensable for a sustained price process, but is not responsible for any regularity in price trends. We therefore set $\eta_{fu}^{(\cdot)}=1$ for the sake of simplicity. Then, Eq. (15) becomes

$$Q^t = N(F - P^{t-1}). \tag{16}$$

Equation (9) gives

$$Q^t = \left(\frac{N}{c_p}\right)(P^{t+1} - P^t). \tag{17}$$

Substituting Eq. (17) into Eq. (16), we obtain

$$P^{t+1} - P^t + c_p P^{t-1} = c_p F. \tag{18}$$

Equation (18) is a second-order difference equation. Depending on the value of c_p, the price can follow a monotonically decaying process ($c_p<0.25$), a damped fluctuating process ($0.25<c_p<1$), or an explosive fluctuating process ($c_p>1$). Figure 8 shows the three typical price trajectories when the initial price is 105. In all these instances, the price is mean reverting. Some researchers have studied the mean-reverting nature of price processes for different behavioral types, as well as different stabilizing-destabilizing endogenous mechanisms of financial markets [19–21].

To analyze the process of volatility generated by our model, we need to consider the mechanism as well as the noise. First of all, we have assumed that the noise, which causes the volatility, follows an independent Gaussian random process. Within this process, those values more close to the mean have higher probabilities. Second, by examining Eqs. (5) through (9), we can recognize that when L^t is smaller (greater) than L_m, a positive (negative) feedback loop will form between L^t and M^t; namely, small (large) values of L^t tend to be enlarged (lessened). Therefore, the nature of the noise, the trading behavior of the agents, and the rule of price updating ensure that the volatility also follows a mean-reverting process.

C. Heavy tails due to large price variations are caused by imitations

In this section, for the sake of simplicity, we adopt price change as return, i.e., $R^{t+1}=P^{t+1}-P^t$. According to Eq. (9), we can then examine Eq. (14) in order to investigate the cause of the resultant non-Gaussian return distributions.

In the simulations demonstrated in Sec. IV B, if $\alpha_{fu}=1$, we cannot generate fat tails. In this case, the total quantity is described by Eq. (15), a special instance of Eq. (14) when all the terms with a product of η_{im}^t terms are equal to zero.

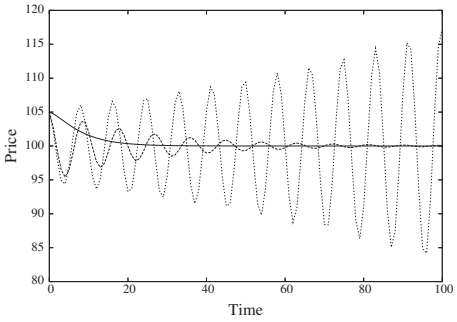


FIG. 8. Price trajectories given by Eq. (18) for different values of c_p. The monotonically decaying curve, the convergent fluctuating curve, and the divergent fluctuating curve correspond to $c_p=0.1$, 0.9, and 1.02, respectively. The parameter settings used are $F=100$ and $P^0=105$.

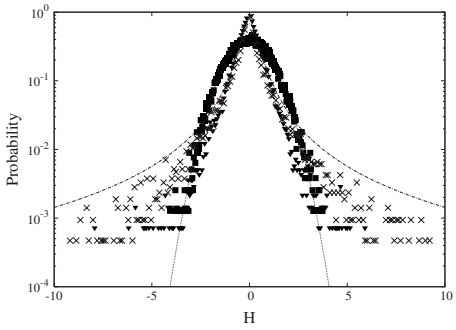


FIG. 9. Probability distributions of H^t defined by Eq. (19). Different values of ϵ are taken for comparison. The ■ points, the ▼ points, and the × points refer to the instances of $\epsilon=1$, 0.5, and 0, respectively. The scale of the vertical axis is logarithmic.

Heavy tails are generated when $\alpha_{fu}<1$ and some of these terms are present in Eq. (14). We therefore suppose that it is the multiplication of the various ϕ^t_{im} terms in different η^t_{im} terms that is responsible for the non-Gaussian distributions, although all these terms themselves follow a Gaussian distribution.

To confirm this supposition, we define a simple reference model:

$$H^t = \epsilon\phi^t + (1-\epsilon)\phi^t\phi^{t-1}\phi^{t-2}, \tag{19}$$

where ϕ^t is an independent Gaussian random variable with mean 0 and standard deviation 1 and ϵ is a parameter. Recall that, in Eq. (14), $\eta^t_{im}=1+c_{im}\phi^t_{im}$. Since Eq. (19), as with Eq. (14), deals with the sum of products of Gaussian terms, it represents the basic structure of the latter.

Figure 9 presents the experimental probability distributions of H^t obtained when choosing different values of ϵ for comparison: 1, 0.5, and 0. In this figure we see that, when ϵ decreases, the distribution of H^t gradually changes from being pure Gaussian to being very fat-tailed non-Gaussian. Thus, the more the product of ϕ^t terms is weighted, the heavier the tails of the consequent distribution. From the discussion in Sec. V A we know that, if α_{fu} is small, the products of more η^t_{im} factors in Eq. (14) will have more weight. This experiment therefore explains the regularity discussed in Secs. IV B and IV C: Larger fractions of imitators correspond to return distributions with heavier tails. In addition, products of η^t_{im} terms give rise to continued products of c_{im}. The multiplication of c_{im} explains the exponential growth of kurtosis following the increase of c_{im}, as shown in Sec. IV C.

D. Volatility clustering is related to the evolution of trading activity

According to Eq. (9) and the definition of return adopted in this section,

$$R^{t+1} \propto Q^t. \tag{20}$$

Since M^t changes much more slowly than Q^t, we have $M^t \simeq M^{t-1} \simeq \cdots \simeq M^{t-\tau}$ for small values of τ. (Note that the analysis here is by no means rigorous.) Then, for a small value of τ, according to Eqs. (7), (8), and (10), as well as the scheme conveyed by Eqs. (11)–(13), Eq. (20) gives

$$R^{t+1} \propto M^t U^t, \tag{21}$$

where

$$U^t = \sum_{i=1}^{N} [A_i^t(F\eta^t_{fu} - P^{t-1}) + A_i^{t-1}(\eta^t_{im})(F\eta^{t-1}_{fu} - P^{t-2}) + \cdots + A_i^{t-\tau-1}(\eta^t_{im}\eta^{t-1}_{im}\cdots\eta^{t-\tau}_{im})(F\eta^{t-\tau-1}_{fu} - P^{t-\tau-2})].$$

In Eq. (21), M^t is a factor that emerges from the agents' trading and in turn reinforces it. Because it changes more slowly than U^t, successive values of $|R^{t+1}|$ are positively correlated with each other. However, consecutive values of R^{t+1} are only weakly correlated due to the fast variation in its sign, which is caused by the fast variation in the sign of U^t due to news and the mean-reverting nature of the price. These explain the stylized facts of long-term autocorrelation of volatility and short-term autocorrelation of return.

E. The regularity can be identified within some other MS models

In a general sense, the MS models discussed in Sec. II that can confirm the stylized facts agree with our CA model on the origins of large price variations and volatility clustering.

Although explaining imitation from different angles, all these MS models and our CA model show that the fraction of abnormally large price variations is much larger when agents imitate each other than when they are mutually independent. In the latter instance and in the limit of a large number of agents, returns follow a Gaussian distribution.

Within these MS models and our CA model, we can ultimately attribute volatility clustering to the evolution of agents' activity, although the corresponding processes of the models that indicate activity are quite distinct. (Notice that all these processes are positively correlated with the evolution of trading volume.) These processes are, respectively, the evolution of volatility (Bak *et al.* [12]), the development of the fraction of noise traders (Lux *et al.* [13]), the evolution

of agents' activation thresholds (Iori [16]), the percolation process (Bartolozzi *et al.* [17]), and the progression of the desirability of an asset (our CA model). In addition, these processes are slower than their corresponding "source" processes. Therefore, volatility clustering generated by these MS models and our CA model is the combined effect of two processes on different time scales.

In the literature, on the one hand, there is not yet a common agreement on the origins of the stylized facts [10]. On the other hand, various analytical models for describing the phenomena, e.g., generalized autoregressive conditional heteroskedasticity (GARCH) models,[10] stochastic volatility models,[11] and a recently published Itô-Langevin model described in [25], do not provide explicit economic explanations for the underlying dynamics. The regularity discussed here can help us achieve a more accurate understanding of the complex dynamics of stock markets.

[10]Here "heteroskedasticity" means a situation in which the variance of a variable varies. It is a technique for modeling economic time series with time-varying volatility. A GARCH(p,q) process is defined as $\epsilon(t)=\sigma(t)e(t)$. Here $e(t)$ is an independent and identically distributed variable with zero mean and unit variance, $\sigma^2(t)=\alpha_0+\Sigma_{i=1}^{q}\alpha_i\epsilon^2(t-i)+\Sigma_{j=1}^{p}\beta_j\sigma^2(t-j)$, where $\alpha_{(\cdot)}$ and $\beta_{(\cdot)}$ are parameters [7,23,24].

[11]A class of stochastic volatility models considers volatility to be independent of return. Price is then assumed to follow a geometric Brownian motion with a time-dependent volatility: $dS(t)=\mu S(t)dt+\sigma(t)S(t)dz_1$. Here dz_1 describes a Wiener process. With $v(t)=\sigma^2(t)$, the time-dependent variance follows a different stochastic process $dv(t)=m(v(t))dt+s(v(t))dz_2$, where dz_2 is another Wiener process. Different forms of $m(v(t))$ and $s(v(t))$ correspond to some popular models of this type [7].

VI. CONCLUSION

In this paper, a cellular automaton model for simulating the complex dynamics of stock markets has been described. The model can confirm the main stylized facts observed in empirical financial time series. Our simulations and analysis suggest that price and volatility are mean reverting. Long-range agent interactions, which are responsible for large price variations, can form from local interactions. Volatility clustering is associated with the variation in agents' trading activity, a slower process compared with the variation in the influence of news. Heavy-tailed distributions of return are related to both large price variations and volatility clustering. Finally, these non-Gaussian distributions are produced by agents' behavior in response to the arrival of news, even though the influence of news is assumed to follow a Gaussian random process.

ACKNOWLEDGMENTS

The authors wish to thank Breanndán Ó Nualláin, Zhengdao Wang, Joost Geerdink, Ketan Maheshwari, Dhiraj Hegde, Thomas Miedema, and Stanley Jaddoe for their valuable suggestions and discussions about microscopic simulations in finance. This work was carried out in the context of VLe (Virtual Laboratory for eScience), a project supported by a BSIK grant from the Dutch Ministry of Education, Culture and Science.

[1] B. Mandelbrot, J. Business **36**, 394 (1963).
[2] A. Pagan, J. Empirical Finance **3**, 15 (1996).
[3] Z. Ding, C. W. J. Granger, and R. F. Engle, J. Empirical Finance **1**, 83 (1993).
[4] R. Cont, Quant. Finance **1**, 223 (2001).
[5] D. M. Guillaume, M. M. Dacorogna, R. R. Davé, U. A. Müller, R. B. Olsen, and O. V. Pictet, Finance Stoch. **1**, 95 (1997).
[6] R. N. Mantegna and H. E. Stanley, *An Introduction to Econophysics* (Cambridge University Press, Cambridge, U.K., 2000).
[7] J. Voit, *The Statistical Mechanics of Financial Markets* (Springer-Verlag, Berlin, 2003).
[8] L. Tesfatsion, Artif. Life **8**, 55 (2002).
[9] M. Levy, H. Levy, and S. Solomon, *Microscopic Simulation of Financial Markets* (Academic Press, San Diego, 2000).
[10] R. Cont, in *Long Memory in Economics*, edited by A. Kirman and G. Teyssiere (Springer, Berlin, 2005), p. 289.
[11] S. Bandini, S. Manzoni, A. Naimzada, and G. Pavesi, in *Proceedings of the 6th International Conference on Cellular Automata for Research and Industry, ACRI 2004, Amsterdam*, edited by P. M. A. Sloot, B. Chopard, and A. G. Hoekstra (Springer-Verlag, Berlin, 2004), p. 591.
[12] P. Bak, M. Paczuski, and M. Shubik, Physica A **246**, 430 (1997).
[13] T. Lux and M. Marchesi, Nature (London) **397**, 498 (1999).
[14] R. Cont and J.-P. Bouchaud, Macroecon. Dyn. **4**, 170 (2000).
[15] D. Talia and P. M. A. Sloot, FGCS, Future Gener. Comput. Syst. **16**, v (1999).
[16] G. Iori, J. Econ. Behav. Organ. **49**, 269 (2002).
[17] M. Bartolozzi and A. W. Thomas, Phys. Rev. E **69**, 046112 (2004).
[18] Z. Wang, MSc thesis, University of Amsterdam, 2005.
[19] W. J. Baumol, Rev. Econ. Stat. **39**, 263 (1957).
[20] A. Beja and M. B. Goldman, J. Financ. **35**, 235 (1980).
[21] R. H. Day and W. Huang, J. Econ. Behav. Organ. **14**, 299 (1990).
[22] E. Egenter, T. Lux, and D. Stauffer, Physica A **268**, 250 (1999).
[23] T. Bollerslev, J. Econometr. **31**, 307 (1986).
[24] Z. Ding and C. W. J. Granger, J. Econometr. **73**, 185 (1996).
[25] C. Anteneodo and R. Riera, Phys. Rev. E **72**, 026106 (2005).
[26] J. Downes and J. E. Goodman, *Dictionary of Finance and Investment Terms* (Barron's Educational Series, New York, 1998).
[27] A. Marshall, *Principles of Economics* (Macmillan and Co., London, 1890).