

Bubble Formation and Heterogeneity of Traders: A Multi-Agent Perspective

Shu-Peng Chen · Ling-Yun He

Accepted: 12 November 2012 / Published online: 27 November 2012
© Springer Science+Business Media New York 2012

Abstract Financial bubble is an intensively discussed but quite controversial topic. In current literature, the researches usually focus on the (ir)rationality of traders and its impacts on the bubble. We thereby propose a completely different perspective, that is, of traders' heterogeneity and its impacts on the formation of bubble in financial markets. As in the real financial markets, the agents are always heterogenous. For example, some of them are fundamentalists, some are chartists, some are noise traders, etc. To model the heterogeneity of agents in the real markets, we proposed a multi-agent model to control the constitution of traders. Based on four scenarios with different constitution of traders' behaviors, we investigated three extreme situations where the market is occupied by homogeneous agents (no matter they are fundamentalists, chartists or noise traders), and one scenario where the market is made up of heterogeneous traders. By applying Log-Periodic Power-Law (LPPL) model, We studied the impacts of different investors' behaviors on the bubble formation in the market and found that: (a) the public information has an important influence on the beginning of a bubble; (b) traders' different expectations and their self-feedback is one of reasons for the existence of log-periodicity in bubble; (c) the existence of power-law growth and log-periodicity, which leads the probability of prediction for the bursting of bubble, is caused by the combined effects of public information, traders' different expectations and their self-feedback.

S.-P. Chen · L.-Y. He
Center for Futures and Financial Derivatives, College of Economics and Management,
China Agricultural University, Beijing 100083, China
e-mail: cspslchen123@163.com

L.-Y. He (✉)
Department of Economics, The University of Calgary, 2500 University Dr. N.W.,
Calgary, ABT2N 1N4, Canada
e-mail: lyhe@ucalgary.ca; lyhe@yahoo.cn

Keywords Financial bubble · Traders' heterogeneity · Multi-agent simulations

1 Introduction

Financial bubble, which seems as old as the history of financial market itself, is always one of intensively discussed but highly controversial topics. In history, a burst of major bubble was always accompanied with an economic recession and caused a huge loss. However, before our further discussion, there are some important questions waiting to be clarified: what is a bubble? what can cause a bubble? To be more specific, is traders' heterogeneity a cause of the bubble? There is still no generally accepted definition in economics. It is usually described as a large and unusual deviation from fundamental price (Stockl et al. 2010). In recent literatures, a large and growing numbers of papers focused on studying the 'rational' bubble (Głózkaynak 2008), which can be rational deviations of the price from the fundamental value (Blanchard and Watson 1983).

In the existing literature, the researchers usually focus on the traders' rational or irrational expectations on the bubble, however, few can offer the answers to the problems. The traders are usually heterogenous, e.g., there are fundamentalists, chartists, noise traders, etc, in a real market (Chiarella et al. 2009). One of the difficulties in quantifying the heterogeneity is to study the influence of different constitution of heterogeneous traders on the bubble dynamics. In this paper, Our main purposes focus on these questions: Is the constitution of heterogeneous traders one of reasons for bubble? If so, what are the different patterns of bubbles caused by different constitution of heterogeneous traders? In order to answer these questions, we investigate the dynamic cause and underlying mechanism of bubbles from a heterogeneous agent perspective.

By changing the parameters in a multi-agent model we propose, we can easily control the constitution of the traders. We generated four scenarios by simulations of multi-agent model:¹ three homogeneous scenarios, in which the traders' behaviors are mainly influenced by either Price Tendency, or Public Information, or Noise Information; meanwhile, as a comparison, we also set an integrated scenario, in which the traders are heterogeneous, making decisions based on different information sets, namely, Price Tendency, Public Information and Noise Information, respectively. Through simulations, we investigated homogeneous and heterogeneous traders' behaviors on the market.

In order to measure the bubble, many econometric tests of asset price bubbles have been proposed, such as variance bounds tests (Shiller 1981), two-step tests (West 1987), integration/cointegration based tests (Diba and Grossman 1987, 1988), Intrinsic bubbles (Froot and Obstfeld 1991), etc. However, even for the same period of bubble, different methods give completely opposite results, and all of these methods fail to detect whether there exists a bubble (Głózkaynak 2008). At the same time, in experimental context, many alternative measures have been proposed, such as price amplitude (Porter and Smith 1995), total dispersion (Haruvy and Noussair 2006), average bias (Haruvy and Noussair 2006), Haessels R^2 (Dufwenberg et al. 2005), duration

¹ This model we proposed is based on Ising-type Model (Chiarella et al. 2009; He 2010; He and Zheng 2010).

(Porter and Smith 1995), etc. However, all of these methods have a hypothesis that we know the fundamental value. But in the real market, the fundamental value is difficult to be estimated. Even for the same asset, different traders' estimated value will be different. To quantify the bubbles, we apply a method called Log-Periodic Power Law (LPPL) model, which is proposed by Sornette et al. (1995). This model was derived from the rational stochastic crashing bubble model by Blanchard (1979) and Blanchard and Watson (1983). In their definition, bubbles behave as a "super-exponential" pattern, punctuated by bursts of negative feedback spirals of crash expectations (Jiang et al. 2010).

Although this definition of bubbles remains controversial (Lux 2007; Rosser 2008), this 'bubble' model are applied and some seemingly successful predictions are reported in existing literature (Zhou and Sornette 2003, 2006; Sornette et al. 2009; Bastiaensen et al. 2009).

Our article is organized as follows: Sect. 2 describes the methodology, including Multi-Agent Model, LPPL Model, (H, q) Analysis and Lomb Spectral Analysis; Sect. 3 shows the results of simulations; Sect. 4 studies the bubbles in different scenarios; Sect. 5 is the discussion and Sect. 6 gives the main conclusions.

2 Models

2.1 Multi-agent Model

Suppose that a simplified quote-driven market is composed of a $N \times N$ two-dimensional grid network, where each grid (i, j) is occupied by a market participant. The topology of this artificial market is a torus, namely, the upper and lower boundaries are connected, so are the left and right ones. Each agent makes his own transaction (eg. buy or sell) decisions based on both his private information and influence from his close neighbors. To simplify the model, we define that each agent collects information from his nearest 4 neighbors, namely his upper, lower, left and right grid sites.

For simplicity, we assume that for every time step of simulation, agent (i, j) can be in only one of three possible states: $s_{i,j} \in \{-1, 0, +1\}$, among which -1 stands for agent selling one unit of risky assets, while $+1$ stands for agents buying one unit of risky assets, and 0 means idle state, namely neither buying or selling.

In financial market, an investor's behavior is usually based on different information sets before he makes a decision. To simplify this model, we only considered the information of price tendency T , public information G , and noise ε , so for each time step, the information of agent (i, j) can be given by

$$U_{i,j}(t) = \frac{1}{\alpha_{i,j} + \beta_{i,j} + \gamma_{i,j}} [\alpha_{i,j}T_{i,j}(t) + \beta_{i,j}G_{i,j}(t) + \gamma_{i,j}(t)\varepsilon_{i,j}(t)] \quad (1)$$

where α , β and γ represent the weights given to each information sets; for example, if $\beta = \gamma = 0$, the agent's behavior is totally influenced by the price tendency and he is more likely to be to a chartist, if $\alpha = \gamma = 0$, he prefers the public information to technical analysis; while if $\alpha = \beta = 0$, he is a pure noise trader. For each agent,

the weights are randomly chosen according to the realizations of the set of Laplace distributions (Chiarella et al. 2009), namely,

$$\begin{cases} \Pr(\alpha) = \frac{1}{\delta_1} \exp(-\alpha/\delta_1) \\ \Pr(\beta) = \frac{1}{\delta_2} \exp(-\beta/\delta_2) \\ \Pr(\gamma) = \frac{1}{\delta_3} \exp(-\gamma/\delta_3) \end{cases}$$

where the distribution is determined by the parameter δ . If the δ is larger, the weight will be a greater probability to get larger. In our simulation, the agents do not select any parameters. In different scenarios, the constitution of heterogenous traders is controlled by the parameter δ .

From the perspective of price tendency, a chartist analyzes the historical prices and then predict the future tendency before he makes a decision. To simplify our discussion, we only considered a linear tendency. For each agent (i, j) , we supposed that he has a memory of price with length of τ . The influence is measured by the slope of linear fitted curve. If the slope is steep, the positive (or negative) influence should be large, whose absolute value should be close to 1, and while if the slope is flat, the positive (or negative) influence should be small, whose absolute value should be close to 0. Then the influence of price tendency T in Eq. (1) can be given by

$$T_{i,j}(t) = \frac{1 - \exp(-\eta * Tr_{i,j}(t))}{1 + \exp(-\eta * Tr_{i,j}(t))} \tag{2}$$

where Tr is the slope which measured by the least square method, and the length of memory τ is randomly chosen according to an exponential distribution:

$$\Pr(\tau - \tau_0) = \frac{1}{\delta_4} \exp[-(\tau - \tau_0)/\delta_4]$$

Since τ is an integer, the probability then changes to be

$$\tau = [\delta_4 \ln(1 - w)] + \tau_0 \tag{3}$$

where w is randomly chosen and obeys uniform distribution $U(0, 1)$, and τ_0 is the minimum memory.

Except the chartists, there are also many other traders who make their decisions based on the public information. Since the price is determined by supply and demand, the authenticity of information is somehow less important under many circumstances; what might be more important is how many people believe and act upon a given set of information, which in this context, is a self-reinforcing process that will inevitably affect prices. In this paper, for each time step, the public information B is randomly chosen and independently distributed according to the standard normal distribution, that is, $B \sim N(0, 1)$, where the symbol of B represents the positive or negative information, while the absolute value represents the influence of the random event. Since different investors have different understandings and interpretations of the same information set (He et al. 2009), we thereby incorporate a deviation among understandings

L , where $L \sim N(0, 1)$ and then for each agent (i, j) , the information can be defined as

$$F_{i,j}(t) = B_{i,j}(t) + q * L_{i,j}(t) \tag{4}$$

where q is a control variable, and if $q = 0$, that means all of the agents can correctly understand the information; meanwhile, if the symbol of $B_{i,j}$ is different with that of $F_{i,j}$, that means agent (i, j) 's understanding is not correct. Then the influence of public information G in Eq. (1) can be given by

$$G_{i,j}(t) = \text{sign}(F_{i,j}(t))[1 - \exp(-|F_{i,j}(t)|)] \tag{5}$$

According to Black (1986), there are noise traders, who might (incorrectly) regard noise as effective information, and make their decisions accordingly. To capture this feature, we designate the ε in Eq. (1) as a Uniform distribution noise, namely, $\varepsilon \sim U(0, 1)$.

Before each agent (i, j) makes his own decisions, he is usually influenced by people around him; in this model, the nearest 4 neighbors include his upper, lower, left and right grid sites. Because usually all that each agent knows is the past transaction behaviors of the neighbors instead of their ongoing decisions, we define the final information as

$$I_{i,j}(t) = \frac{1 - m}{4} \sum_{(i,j)} J_{i,j} I_{i,j}(t - 1) + m U_{i,j}(t) \tag{6}$$

where $J(J \sim U(0, 1))$ represents the influences from a specific agent among the nearest neighbors. If $J = 0$, it means this agent has no influence on his neighbors, and nobody believe him; while if $J = 1$, it means every neighbors try to imitate his whose behaviors are among the critical information to the neighbors' decisions. Meanwhile, the parameter m is the weight of imitation, e.g., if $m = 1$, this agent only believes his own self, and completely disregard the neighbors's behaviors; on the contrary, if $m = 0$, he has no faith in himself, and follows others' behaviors blindly. Then for each agent (i, j) , the strategy can be given by

$$s_{i,j}(t) = \begin{cases} 1 & I_{i,j}(t) > v_{i,j}^p(t) \\ 0 & v_{i,j}^n(t) < I_{i,j}(t) < v_{i,j}^p(t) \\ -1 & I_{i,j}(t) < v_{i,j}^n(t) \end{cases} \tag{7}$$

where v^p and v^n are positive and negative thresholds, respectively. If an agent have collected enough information, namely $I_{i,j}$ exceeds the positive (or negative) threshold, he will buy (or sell) one unit of risky assets, otherwise he will do nothing. The thresholds are defined as

$$\begin{cases} v_{i,j}^p(1) = 1 - \exp(-|d_{i,j}^p|) \\ v_{i,j}^n(1) = -[1 - \exp(-|d_{i,j}^n|)] \end{cases} \tag{8}$$

where $d_{i,j}^p$ and $d_{i,j}^n$ are selected randomly and obey normal distribution with a unit variance and zero mean. Later on, according to the judgment of market, $v_{i,j}^p(t)$ and $v_{i,j}^n(t)$ are adjusted by the returns,

$$\begin{cases} v_{i,j}^p(t + 1) = v_{i,j}^p(t) \exp(\kappa_1 * R(t)) \\ v_{i,j}^n(t + 1) = v_{i,j}^n(t) \exp(\kappa_2 * R(t)) \end{cases} \tag{9}$$

In a market, the price can not increase or decrease unlimitedly. If the price continues to rise, the investor or agent may consider the downside risk, and the positive threshold would rise quickly; meanwhile if the price continues to decrease, some investors may worry about whether the price reaches the bottom, and the negative threshold would decrease slowly. So in Eq. (9), the parameter κ_2 should be smaller than the κ_1 .

To further simplify this problem to facilitate our discussion, let us suppose in this artificial market there is an ideal market maker who does not make profits from bid-ask price difference (He 2010), he only collects all the agents' orders and then quotes a single price at a time step to correct the imbalance between demand and supply according to the return

$$R(t) = \begin{cases} \theta(t)(D(t)/S(t) - 1) & D(t) > S(t) \\ -\theta(t)(S(t)/D(t) - 1) & D(t) \leq S(t) \end{cases} \tag{10}$$

where $D(t)$ is the demand and $S(t)$ the supply of the risky assets. Since there is a market maker, and each agent only can buy (or sell) one unit of financial assets to him at one step of simulation, the trade volume (namely, demand plus supply) can be simply given by

$$V(t) = D(t) + S(t) = \sum_{s_{i,j}(t) > 0} s_{i,j}(t) + \left| \sum_{s_{i,j}(t) < 0} s_{i,j}(t) \right| \tag{11}$$

where the parameter θ , a measure of the activity of market, is adjusted at each time by

$$\theta(t) = \lambda(V(t)/N^2) \tag{12}$$

where N^2 is the total number of agents, and the parameter λ is an adjustment rate of the market.

2.2 Log-Periodic Power Law Model

The Log-Periodic Power Law (LPPL) model (Sornette et al. 1995; Johansen et al. 2000; Jiang et al. 2010) that we use is:

$$I(t) = A + Bx^\alpha + Cx^\alpha \cos(\omega \ln x + \phi) \tag{13}$$

where $x = |t - t_c|$. The parameter t_c is a critical point of the time in which the crash is most likely to occur; $I(t)$ can be the price $P(t)$ or its logarithm $\ln P(t)$; ² A is the highest value which the $P(t)$ or $\ln P(t)$ would be most likely to reach at the critical point; α is the exponent of power-law growth; ω is an angular frequency of log-periodicity and ϕ is a phase angle.

The LPPL model is a highly non-linear function and has seven parameters that need to be estimated. Fortunately, the parameters A , B and C are in the linear part of the function. For any given values of α , t_c , ω and ϕ , the parameters A , B and C can be regarded as an expression of them by the Least Squares method. Therefore only four parameters need to be searched. In order to decrease the complexity of the fitting procedure and improve its stability, [Filimonov and Sornette \(2011\)](#) presented a simple transformation of Eq. (13) that reduces it to a function of four linear parameters and three nonlinear parameters. According to [Johansen et al. \(2000\)](#), Taboo-Search is applied to reach a global optimization.

2.3 (H , q) Analysis

As mentioned above, within the context of LPPL model, the price (or the logarithm of price) follows a super-exponential growth with log-periodic oscillations before the critical point; therefore, it is a logical necessity for us to detect whether there exist log-periodic oscillations in the generated prices.

In order to detect the log-periodic oscillations (if there exists), we applied the (H , q) analysis proposed by [Zhou and Sornette \(2002\)](#), which is based on the inverse of the q -integral. The (H , q)-derivative is given by

$$D_q^H f(x) \triangleq \frac{f(x) - f(qx)}{[(1-q)x]^H} \quad (14)$$

where H and q vary in the ranges $(-1, 1)$ and $(0, 1)$. If we chose a appropriate exponent H , it can remove the effect of power-law relationship.

2.4 Lomb Spectral Analysis

The (H , q) can obtain the part of log-periodicity in time series (if there exists). However, we can not judge whether there are log-periodic oscillations by naked eye. A method to quantify the periodicity is also needed. Usually, the Fast Fourier Transformation (FFT) is considered to be an effective method, but it is applied for the uniform distributed sample. For the log-periodicity (if any), the sample is non-uniform because the time in horizontal axis is logarithmic. The data points are intensive which are far away from critical point, while those adjacent to critical point are thin. Therefore, the Lomb periodogram may be a proper solution and is suggested to detect the log-periodic oscillations. The normalized Lomb periodogram is

² In this article, we apply the logarithm of price, namely $\ln P(t)$.

$$P_N(\omega) = \frac{1}{2\delta^2} \left[\frac{\left[\sum_{i=1}^N y(t_i) \cos \omega(t_i - \tau) \right]^2}{\sum_{i=1}^N \cos^2 \omega(t_i - \tau)} + \frac{\left[\sum_{i=1}^N y(t_i) \sin \omega(t_i - \tau) \right]^2}{\sum_{i=1}^N \sin^2 \omega(t_i - \tau)} \right] \tag{15}$$

where τ is determined by

$$\tau = \frac{1}{2\omega} \arctan \frac{\sum_{i=1}^N \sin 2\omega t_i}{\sum_{i=1}^N \cos 2\omega t_i} \tag{16}$$

If a high peak is detected in the Lomb periodogram, it means there is more likely a periodicity; in other words, the higher $P_N(\omega)$ is, the more probability of periodicity with this angular frequency.

3 Simulations

3.1 Setting of Scenarios

In this paper, we intend to investigate the relationship between the heterogeneity of traders and the bubble dynamics. To make a clear contrast, we first of all design three extreme (homogeneous) scenarios, in each of which the traders in the market are mainly influenced by one of the three information sets (price tendency, public information or noise information):

Scenario A (Price Tendency Scenario): In this scenario, the traders are influenced only by the price tendency (the parameters $\delta_1 = 5, \delta_2 = \delta_3 = 1$). Before making decisions (buy or sell a unit of asset), the investors predict the future trends of price according to the tendency of price or the history information.

Scenario B (Public Information Scenario): In this scenario, the public information plays a dominant role (the parameters $\delta_2 = 2.5^3, \delta_1 = \delta_3 = 1$). For all of agents, they can easily get free and common information, which has a large probability to have a heavy weight.

Scenario C (Noise Information Scenario): In this scenario, the agents make decisions based on the market noise (the parameters $\delta_3 = 5, \delta_1 = \delta_2 = 1$).

As a comparison, we then propose an integrated scenario, in which the traders are influenced by all of the three factors, and they are more likely to be heterogeneous, namely,

Scenario D (Integrated Scenario): In this scenario, all of the factors have an equal probability to affect each agent (the parameters $\delta_1 = \delta_2 = \delta_3 = 1$). Some of them may be mainly influenced by one of the factors and more likely to be chartist, information trader or noise trader; some of them may prefer two factors than the others; meanwhile the others may fully consider the various factors. In one word, this heterogeneous market is occupied by different kinds of traders.

³ Here we intended to set $\delta_2 = 5$, however, the price is much easily to drop to the value around 0 for a long time.

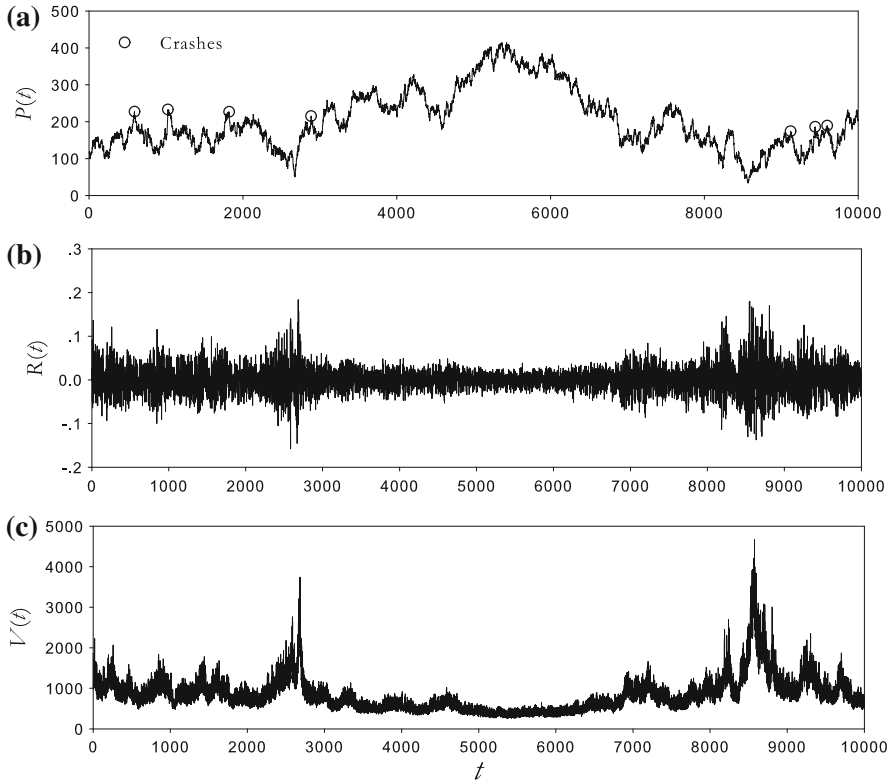


Fig. 1 Illustrates the market of Price Tendency Scenario (Scenario A). In this scenario, each agent prefers information from a price tendency (The parameters $\delta_1 = 5, \delta_2 = \delta_3 = 1$). From the *top* to the *bottom* there are prices (a), returns (b) and trading volume (c), respectively

As for the other parameters of the multi-agent model, they should be same in all of the scenarios. In our simulations, we set $N = 100$, namely, the market contains 100×100 investors; the minimum memory was set to be a week, namely, $\tau_0 = 5$, and the probability parameter $\delta_4 = 5$; meanwhile, the effect of tendency $\eta = 0.01$; misunderstanding of public information $q = 0.5$; imitation behavior $m = 0.5$; the extent of market activity $\lambda = 0.05$; adjustment speed of threshold $\kappa_1 = 1$ (positive) and $\kappa_2 = 0.9$ (negative).

3.2 Results of Simulations

Figure 1 shows the result of Scenario A. In this market the price is led by the agents who cares about the tendency. Figure 2 shows the result of Scenario B. Comparing with Scenario A, one can find an interesting phenomenon. The price changes dramatically. At some moments, it may increase rapidly and reach a high value; but at some other moments, it may drop quickly. For a long time horizon, the price crashes and fluctuates near the value of 0. Meanwhile the returns vary frequently and the volume reaches a

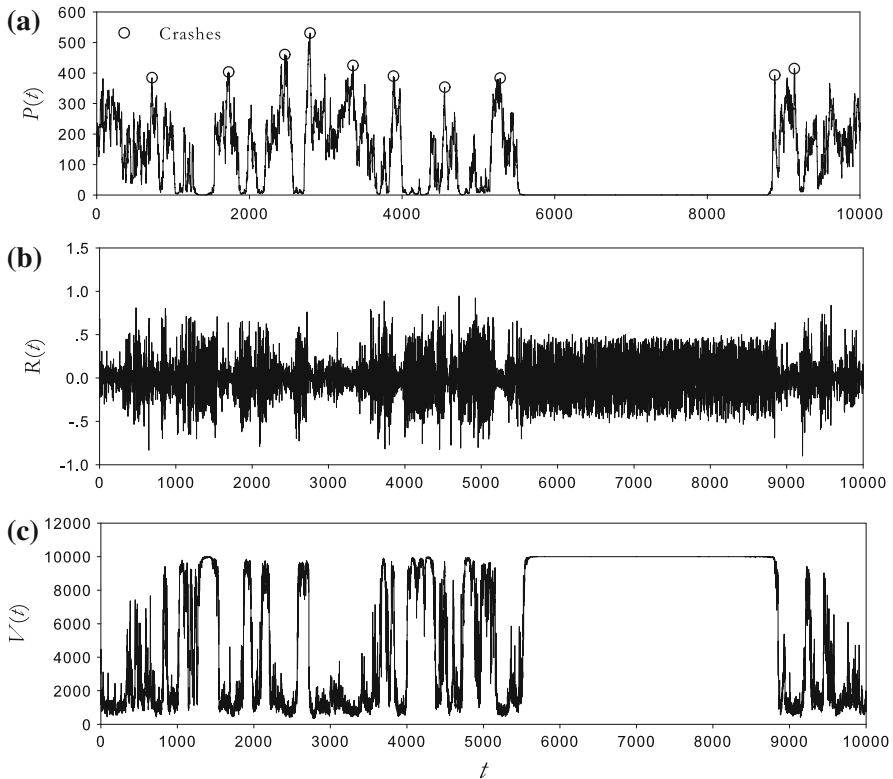


Fig. 2 Illustrates the market of Public Information Scenario (Scenario B). In this scenario, each agent prefers the public information (The parameters $\delta_2 = 2.5$, $\delta_1 = \delta_3 = 1$). From the *top* to the *bottom*, there are prices (a), returns (b) and volume (c), respectively

high value. Figure 3 shows the result of Scenario C. In this scenario the price fluctuates in a small range. Figure 4 shows the simulated results of Scenario D, in which the traders are heterogeneous. Comparing with the four scenarios, it is distinctive to find the stylized fact of volatility clustering in the returns of Scenario A, B and D; however, for Scenario C, this phenomenon is not that obvious.

4 Studying of Bubbles

4.1 Statistical Properties

To get a better understanding of the generated time series, we provided the summary statistics of the returns for different scenarios (see Table 1). Figure 5 shows the distributions of returns (rescaled by their respective standard deviation) for different scenarios. It is clear to find that, for the Price Tendency Scenario (Scenario A), Public Information Scenario (Scenario B) and Integrated Scenario (Scenario D), the distributions show a high degree of peakedness and fat tails, which have been widely detected

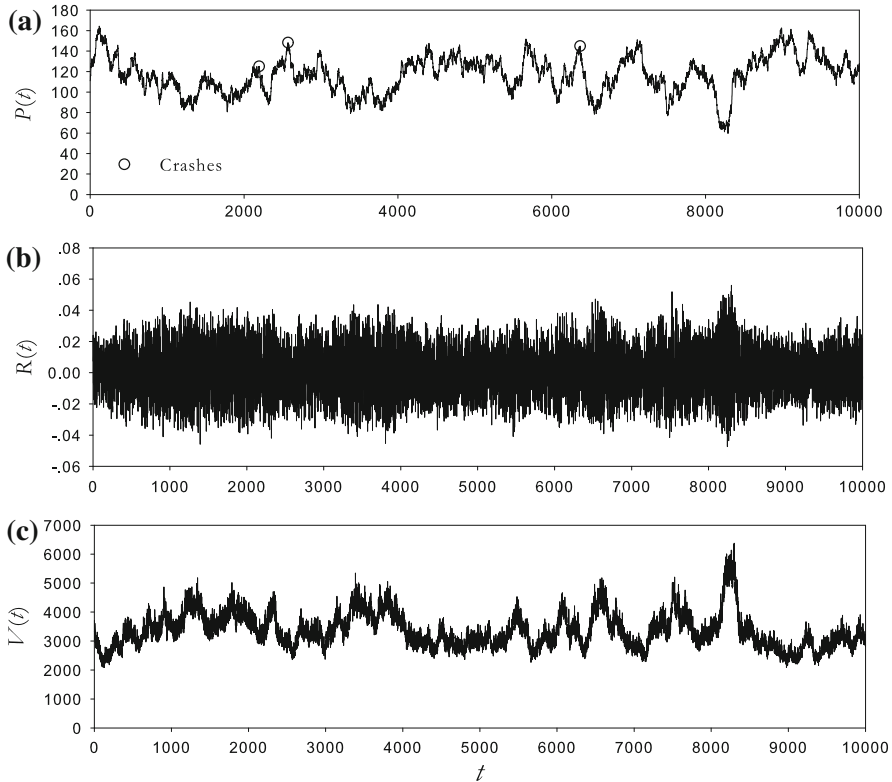


Fig. 3 Illustrates the market of Noise Information Scenario (Scenario C). In this scenario, each agent treats market noise as effective information (The parameters $\delta_3 = 5$, $\delta_1 = \delta_2 = 1$). From the *top* to the *bottom*, there are prices (a), returns (b) and volume (c) respectively

in the real financial markets, however, for the Noise Information Scenario (Scenario C), the distribution is more close to the Gaussian distribution.

Figure 6 shows the autocorrelation of returns (a) and volatilities (absolute returns) (b). The red dashed lines represent the noise levels. One can find in this figure that for all of scenarios, there is no autocorrelation in the returns. For the volatilities, there is strong long-term autocorrelation in Scenario A, B and D, while the autocorrelation seems insignificant in Scenario C.

Meanwhile, in order to detect the long memory of the time series, we estimate the Hurst exponents by means of the Detrended Fluctuation Analysis (DFA) (Kantelhardt et al. 2002). Figure 7 shows the results calculated by the returns (a) and volatilities (b) respectively. If the Hurst exponent is equal to 0.5, there is no autocorrelation and the time series are close to the random walk; if it is greater than 0.5, the time series are positive autocorrelated and show long memory; otherwise if smaller than 0.5, the time series are negative autocorrelated and show anti-persistent properties (He 2010; He and Qian 2012). From Figure 7 the results show that, for the returns, all of the Hurst exponents are close to 0.5 (namely, 0.48 ± 0.01 for Scenario A, 0.51 ± 0.01 for Scenario B, 0.50 ± 0.01 for Scenario C and 0.47 ± 0.01 for Scenario D), meanwhile,

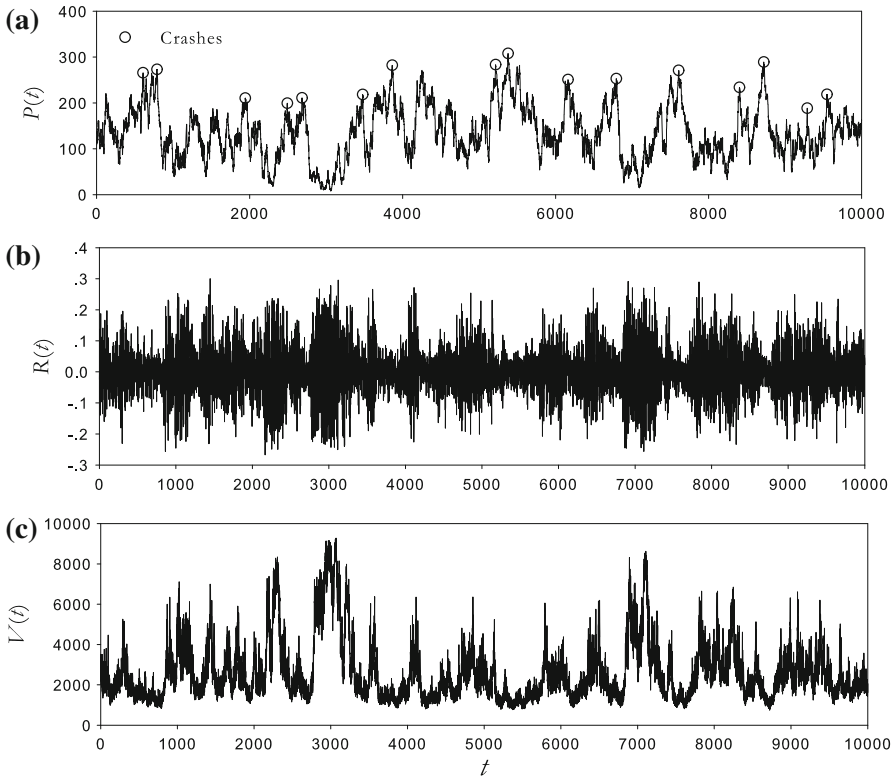


Fig. 4 Illustrates the market of Integrated (Heterogeneous) Scenario (Scenario D). In this scenario, each agent puts equal weights on the three sets of information (The parameters $\delta_1 = \delta_2 = \delta_3 = 1$). From the top to the bottom, there are prices (a), returns (b) and volume (c), respectively

for the volatilities, the exponents are greater than 0.5 and show long memory (namely, 0.69 ± 0.02 for Scenario A, 0.82 ± 0.02 for Scenario B, 0.63 ± 0.02 for Scenario C and 0.84 ± 0.03 for Scenario D), which is consistent with the stylized fact found in our previous empirical findings (Chen and He 2010; He and Chen 2011a,b,c).

From these results, one can find that, for Scenario A, B and D, our model can capture many major stylized facts observed in real world markets, such as volatility clustering, fat tails, no autocorrelation for returns and long memories for the volatilities, etc. It is very interesting that for Scenario C, it only shows weak autocorrelation for volatilities and the price is much close to the random walk for a market dominated by noise traders.

4.2 Identifying the Bubbles

Before we apply LPPL model, it is necessary to define what the bubble implies. Usually, when a bubble bursts, there is a crash, which is a very large and unusual price fall. Before a crash, the price will reach a high value (a peak), but this does not mean

Table 1 The summary statistics of returns for different scenarios

	Mean	SD	Skewness	Kurtosis
Scenario A	0.00006726	0.0254	0.3539	7.8650
Scenario B	0.00008111	0.2067	0.1078	4.3994
Scenario C	0.00001415	0.0154	0.0839	2.7509
Scenario D	0.00003845	0.0681	0.1455	5.1044

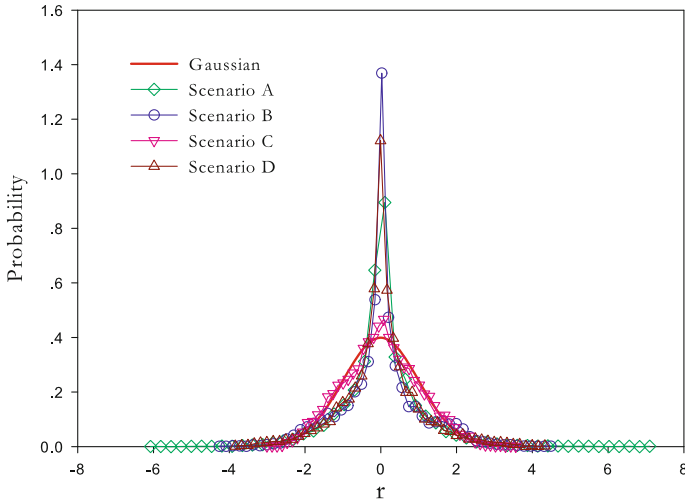


Fig. 5 The distribution of returns (rescaled by their respective standard deviation) for different scenarios. The red solid line presents the standard Gaussian distribution. The distributions of Scenario A, B and D show a high degree of peakedness and fat tails. However, for Scenario C, the distribution is more close to the Gaussian distribution

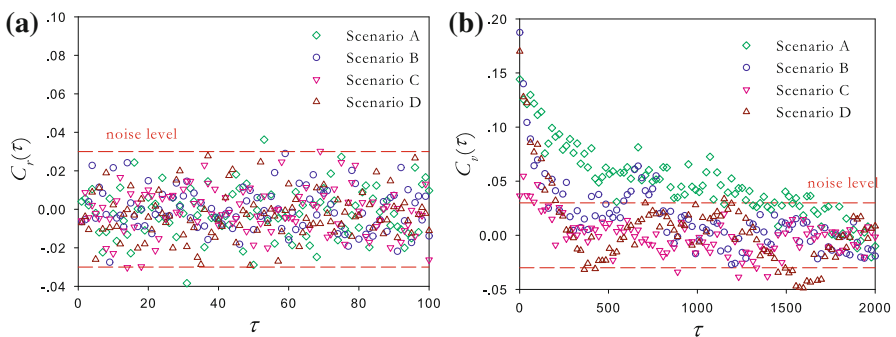


Fig. 6 The autocorrelation of returns (a) and volatilities (absolute returns) (b). The red dashed lines represent the noise levels which are computed as $\pm 3/\sqrt{L}$ (see (Bouchaud et al. 2000; Raberto et al. 2001), where L is the length of time series (namely, $L = N^2 = 10000$))

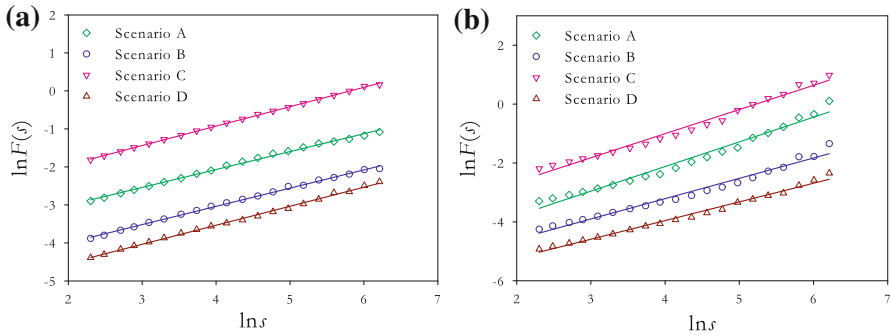


Fig. 7 The log–log plot for detrended fluctuation function $F(s)$ as a function of time lag s for the returns (a) and the volatility (b). The Hurst Exponent for the returns is close to 0.5, meanwhile, for the volatility, it ranges around 0.6 to 0.8

that every peak is a beginning of crash. For example, price may drop now and then, but may increase again rapidly when a bubble is still expanding. To avoid this situation, we apply a procedure according to [Hong and Stein \(2003\)](#), which identify a peak as one initiating a crash:

- (1) A period of 500 days⁴ prior to the peak for which there is no value higher than the peak's;
- (2) A drop in price of 25 %, namely, down to 75 % value of the peak's;
- (3) This drop must occur in 100 days after the peak.

By this procedure, 7 crashes are identified in Scenario A (see [Fig. 1](#)), 10 crashes are identified in Scenario B (see [Fig. 2](#)), 3 crashes are identified in Scenario C (see [Fig. 3](#)) and 16 crashes are identified in Scenario D (see [Fig. 4](#)).

In each of the these scenarios, we choose a typical bubble as a representative. We regard the peak as the end of a bubble, although one may not know when exactly the bubble began. In most cases, the globally lowest price is not the beginning of the bubble. We thereby use a varying window suggested by [Jiang et al. \(2010\)](#). Accordingly, the following representative bubbles are chosen for the scenarios:

- (1) Scenario A (see [Fig. 1](#)): The start time is increasing from 1,371 to 1,411 in steps of 10 (namely two weeks), while the end of the window is fixed at the time of 1,811 (the peak is at the time of 1,821);
- (2) Scenario B (see [Fig. 2](#)): The start time is increasing from 2,121 to 2,161, while the end of the window is fixed at the time of 2,444 (the peak is at the time of 2,464);
- (3) Scenario C (see [Fig. 3](#)): The start time is increasing from 6,062 to 6,102, while the end of the window is fixed at the time of 6,349 (the peak is at the time of 6,369);
- (4) Scenario D (see [Fig. 4](#)): The start time is increasing from 3,052 to 3,092, while the end of the window is fixed at the time of 3,459 (the peak is at the time of 3,479).

⁴ Here one day means a time step of our simulation.

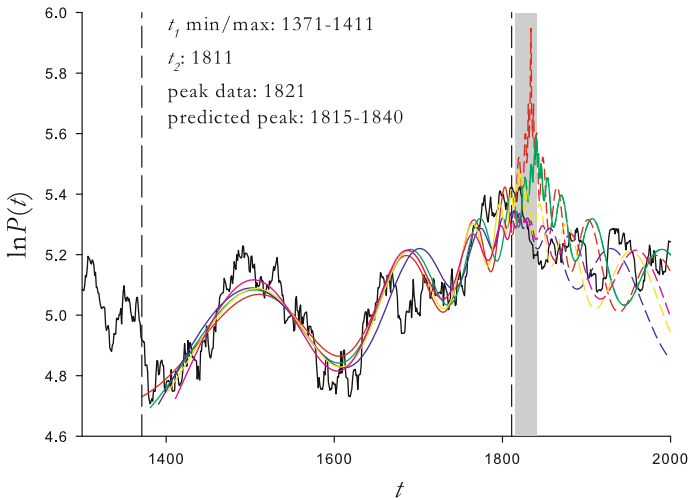


Fig. 8 The fitting of LPPL model for a representative bubble in Scenario A. The two *black dash lines* correspond to the minimum and maximum time of the sample. The *colored solid lines* are the results of fitting in sample while the *colored dash lines* are the prediction out of sample. The *shadow box* indicates the range of values of the crash dates for the fits

4.3 Fitting of LPPL Model

The selected period of time series are fitted by the LPPL model (see Figs. 8, 9, 10, and 11). Figure 8 shows the result for a representative bubble in Scenario A, from which one can find that this bubble can be described by LPPL model well with a significantly distinctive log-periodicity. The predicted crash point is between 1,815 and 1,840, while the real peak of the bubble fall into this range.

Figure 9 shows the result for a representative bubble in Scenario B. The price follows a power-law increasing process; but the log-periodicity is not significant comparing with the standard LPPL bubble. Within the sample the price can be fitted well; meanwhile out of sample, the predictions seem still good in some sense, and the real peak of the bubble is fall into the range of predicted crash data.

Figure 10 shows the result for a representative bubble in Scenario C. From this figure one can find that in sample the behavior of price can be fitted by the LPPL model very well, however, for out of sample, the predicted range is far away from the point which the real crash happened. Meanwhile, in this scenario, the log-periodicity is also not obvious.

Figure 11 illustrates the result for a bubble in Scenario D, from which one can find the power-law price increase decorated by log-periodic oscillations can be observed obviously. The bubble can be fitted well by the LPPL model while the crash can be also predicted well, and the real peak of the bubble fall into the range of predicted crash data.

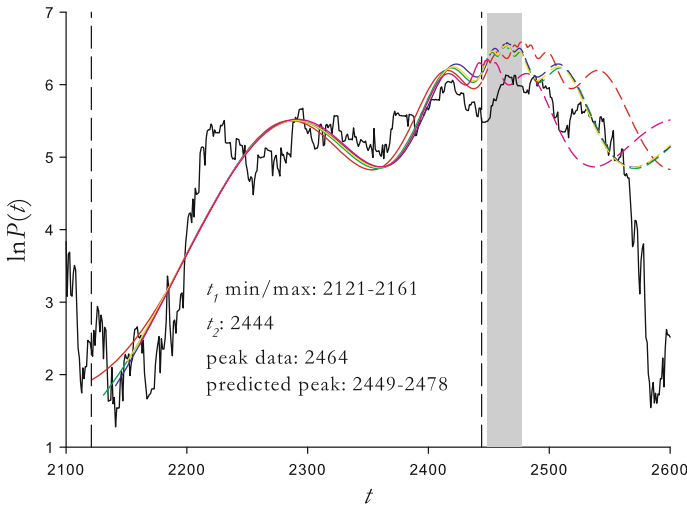


Fig. 9 The fitting of LPPL model for a representative bubble in Scenario B. The two black dash lines correspond to the minimum and maximum time of the sample. The colored solid lines are the results of fitting in sample while the colored dash lines are the prediction out of sample. The shadow box indicates the range of values of the crash dates for the fits

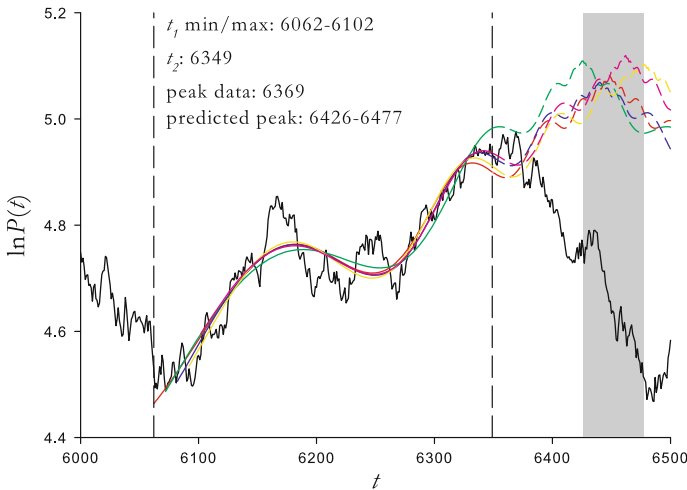


Fig. 10 The fitting of LPPL model for a representative bubble in Scenario C. The two black dash lines correspond to the minimum and maximum time of the sample. The colored solid lines are the results of fitting in sample while the colored dash lines are the prediction out of sample. The shadow box indicates the range of values of the crash dates for the fits

4.4 Analysis of Log-Periodicity

In order to check the existence of log-periodicity, we applied (H, q) analysis and Lomb spectral analysis. We set the t_c to be values which we obtained by the LPPL model,

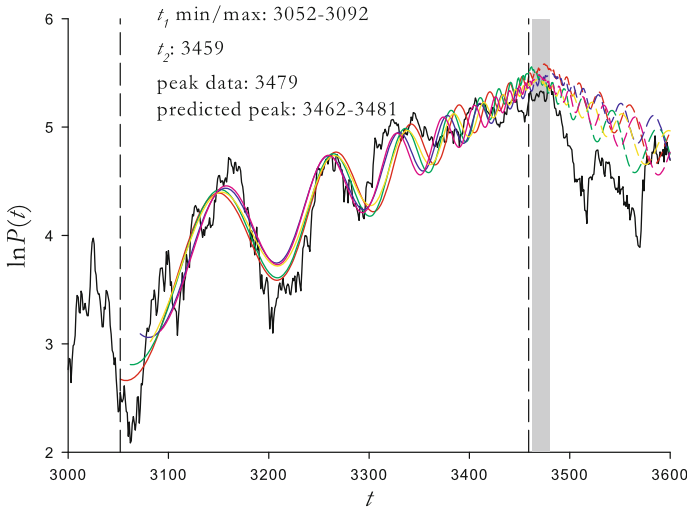


Fig. 11 The fitting of LPPL model for a representative bubble in Scenario D. The two black dash lines correspond to the minimum and maximum time of the sample. The colored solid lines are the results of fitting in sample while the colored dash lines are the prediction out of sample. The shadow box indicates the range of values of the crash dates for the fits

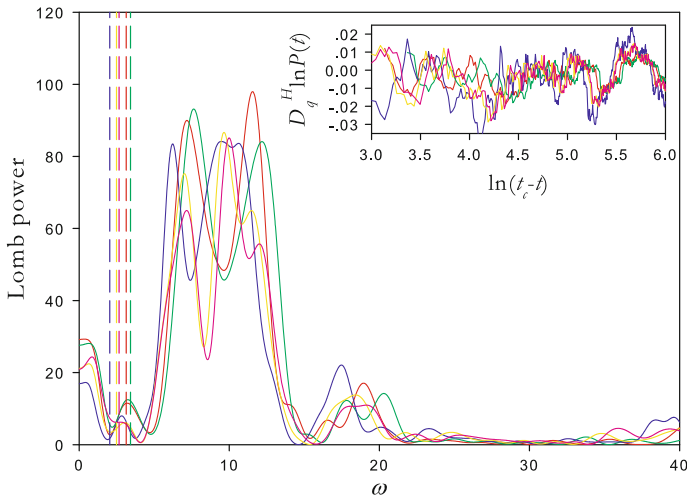


Fig. 12 Lomb periodograms for the time series obtained by (H, q) analysis with $q = 0.8$ and $H = 0.9$ for Scenario A. The inset shows the corresponding plots of $D_q^H \ln P(t)$ as a function of $\ln(t_c - t)$. The dash lines mean the most probable angular frequency. In the left of dash lines are the spurious peak associated with the most probable partial oscillations of a noisy signal. Two mainly angular frequency are found, $\omega_{Lomb}^1 = 7.04 \pm 0.59$ (the fundamental angular frequency) and $\omega_{Lomb}^2 = 10.58 \pm 1.61$ (the second harmonic)

and the results can be seen in Figs. 12, 13, 14, and 15. If there is a high Lomb peak, it means that there is a large probability to be a periodicity with the corresponding angular frequency.

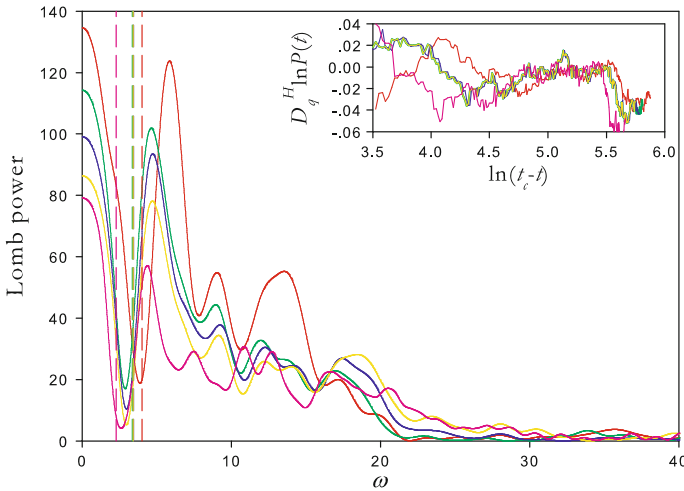


Fig. 13 Lomb periodograms for the time series obtained by (H, q) analysis with $q = 0.5$ and $H = 0.9$ for Scenario B. The *inset* shows the corresponding plots of $D_q^H \ln P(t)$ as a function of $\ln(t_c - t)$. The *dash lines* mean the most probable angular frequency. In the left of *dash lines* are the spurious peak associated with the most probable partial oscillations of a noisy signal, namely, $\omega_{Lomb}^I = 4.86 \pm 0.99$

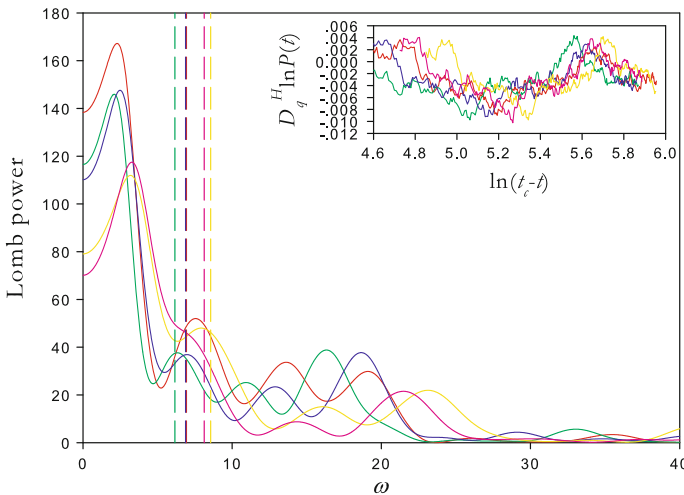


Fig. 14 Lomb periodograms for the time series obtained by (H, q) analysis with $q = 0.6$ and $H = 0.8$ for Scenario C. The *inset* shows the corresponding plots of $D_q^H \ln P(t)$ as a function of $\ln(t_c - t)$. The *dash lines* mean the most probable angular frequency. In the left of *dash lines* are the spurious peak associated with the most probable partial oscillations of a noisy signal

From each figure, one can find that there is always a high peak in a low frequency, which is dangerously close to the most probable angular frequency (Zhou and Sornette 2002). Huang et al. (2000) have found that noise decorating power laws may lead to artifactual log-periodicity with a most probable frequency corresponding roughly to 1.5 oscillations over the whole range of analysis.

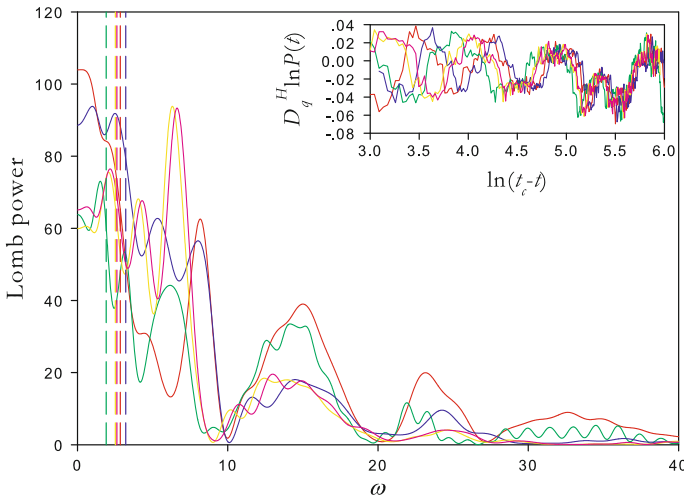


Fig. 15 Lomb periodograms for the time series obtained by (H, q) analysis with $q = 0.7$ and $H = 0.7$ for Scenario D. The *inset* shows the corresponding plots of $D_q^H \ln P(t)$ as a function of $\ln(t_c - t)$. The *dash lines* mean the most probable angular frequency. In the left of *dash lines* are the spurious peak associated with the most probable partial oscillations of a noisy signal. Two angular frequency are found, $\omega_{Lomb}^1 = 6.53 \pm 1.65$ (the fundamental angular frequency) and $\omega_{Lomb}^2 = 13.82 \pm 1.19$ (the second harmonic)

$$\omega^{mp} = \frac{2\pi * 1.5}{\ln(t_c - t_{min}) - \ln(t_c - t_{max})} \tag{17}$$

This peak is a spurious one and should be excluded from Lomb Spectral, in other words, if the peaks are near to the dash lines (see Figs. 12, 13, 14, and 15), it means there is no log-periodicity. Figure 12 shows the result of the bubble in Scenario A. The inset shows the (H, q) -derivatives of $\ln P(t)$ as a function of $\ln(t_c - t)$ with $q = 0.8$ and $H = 0.9$.⁵ The dash lines mean the most probable angular frequency. In the left of dash lines are the spurious peak associated with the most probable partial oscillations of a noisy signal. Two main angular log-frequency are found, $\omega_{Lomb}^1 = 7.04 \pm 0.59$ (the fundamental angular frequency) and $\omega_{Lomb}^2 = 10.58 \pm 1.61$ (the second harmonic).

Figure 13 shows the result of the bubble in Scenario B. The inset shows the (H, q) analysis with $q = 0.5$ and $H = 0.9$. The major angular log-frequency is found, namely, $\omega_{Lomb}^1 = 4.86 \pm 0.99$. However, from the figure one can find that those high peaks are very close to the dash lines and the log-periodicity is not obvious.

Figure 14 shows the result of the bubble in Scenario C. The inset shows the (H, q) analysis with $q = 0.6$ and $H = 0.8$. From the figure, one can find that the main angular log-frequency is smaller than that of the most probable angular frequency, which means that this periodicity (or log-periodicity) is more likely to be caused by noise. Or in other words, in this scenario, there is not a main log-periodicity.

⁵ We have tried different pairs of H and q . Here we only showed the pairs which we thought to be appropriate.

Figure 15 shows the result of the bubble in Scenario D. The inset shows the (H, q) analysis with $q = 0.7$ and $H = 0.7$. Two angular frequency are found, $\omega_{\text{Lomb}}^1 = 6.53 \pm 1.65$ (the fundamental angular frequency) and $\omega_{\text{Lomb}}^2 = 13.82 \pm 1.19$ (the second harmonic).

5 Discussions

We established a multi-agent model under three homogeneous scenarios (Scenario A, B and C), in which the traders' behaviors are mainly influenced by one single information set (Price Tendency, Public Information or Noise Information); meanwhile, as a comparison, an Integrated Scenario (Scenario D) is also proposed, in which all of the traders can be influenced by all of three information sets, and the traders are more likely to be heterogeneous. Through simulations we found some interesting results in each of the scenarios.

For Scenario A (see Fig. 1), the market is mainly lead by the chartists who infer price tendency. The traders in this market prefer the technical analysis. They are likely to collect the history information or apply their past experience to predict future trends before making a decision. To look at the whole market, there is lack of a strong common information. Even if the traders achieve synchronization or the bubble begins to form, some other interference factor (a negative public information or noise information) is easy to change the price tendency. The log-periodic power law pattern bubble or LPPL bubble is detected, however, in this market one can also find that the behavior of price seems smooth and the number of bubbles is only a few. Comparing with other scenarios (Scenario B and D), in this market the bubble is not easy to form.

For Scenario B (see Fig. 2), the market is mainly influenced by the public information. In this scenario, traders are easily to get the same information, which is a free and common knowledge for every one. The influence of public information is large enough and most of the traders are easily to obtain the same expectation or achieve synchronization, and push the price to a high (or low) level. If positive information follows by positive information (or negative information follows by negative information), A huge bubble (or crash) is more easily to be happen (see c of Fig. 11). In this market, the LPPL bubble can also be detected, however, the log-periodicity is not obvious and the formation of bubble is more likely to follow a power law increase pattern.

For Scenario C (see Fig. 3), the market is mainly influenced by the noise information. In this market, the traders are more likely to obtain the noise instead of effective information. In this market, the price fluctuates in a small range and there are only some small bubbles. There are only a few extreme events, and the behavior of price is close to the random walk, meanwhile, this market is more likely to be a weak effective market.

As for Scenario D (see Fig. 4), the market is composed of all kind of traders and it is more close to a heterogenous market. In this market, the LPPL bubble is also detected. The LPPL model is based on the hypothesis of homogeneous rational agents, which is usually criticized by some other scholars. However, from the results of our simulations, we can find that this bubble is not only detected in homogenous market, but also formed in the heterogeneous market. Moreover, Comparing with Scenario

Table 2 The summary of results in simulations

	Formation of bubbles	Stylized facts observed or not	Log-periodicity in bubble	Predicted or not
Scenario A	Hard	Yes	Yes	Yes
Scenario B	Easy	Yes	No	Yes
Scenario C	Hard	No	No	No
Scenario D	Easy	Yes	Yes	Yes

A, in which the bubble is not easy to form, and Scenario B, in which the formation of bubble is more likely to follow a power law increase pattern, the LPPL bubble is widely detected in the heterogeneous market.

Meanwhile, in order to get a better understanding, our results can be summarized in the following table (see Table 2).

From Table 2, we can find that, comparing with other two homogenous markets, the bubble in Scenario B is very easy to form. However, behind the large fluctuations of price, no log-periodicity is detected in this scenario, which means that this bubble only follows a power-law growth. For Scenario A, the LPPL bubble is detected. However, in this market the bubble is not easy to form. Comparing with Scenario B, the behavior of price seems smooth, which leads there are only a few bubbles detected. For Scenario C, the stylized facts of real market are not detected. The behavior of price in this scenario is very close to random walk. And for this reason, even if there are any bubbles, the bursting of bubble can not be predicted. Meanwhile, comparing with the three homogenous scenarios before, one can find that in a heterogenous market (Scenario D), the LPPL bubble is detected and this bubble is easy to form, which is as the same as the real market.

Combining all of the results which we studied, we can hypothesize the dynamic of the bubble: For the beginning of a bubble, the public information has an important influence. Because of good news, or positive information, public information traders believe that the asset price is underestimated. In a heterogenous market, there are many this type of traders. Their buying behavior will lead the initial increase of price. Then according to the chartists, who use the history price to predict the future price, and the imitation behaviors of traders, the bubble presents a power-law growth tendency. Meanwhile, different chartists have different expectations. Though the total tendency of price is rising, however, it is punctuated by many small and large oscillations. Because of different expectations of traders and their self-feedback, there exists a log-periodicity. Further more, because of the existence of the power-law growth and log-periodicity, the bursting of bubble has a degree of predictability, which is impossible in the random walk price. And at last, by combining the effects of all traders, or the effects of heterogeneity, the bubble represents a complex behavior in the real market.

6 Conclusions

In this article, we investigate the relationship between heterogeneity of market participants and financial bubbles. We considered three homogeneous and one heterogenous scenario, and found interesting results:

- (1) We studied the impacts of different investors' behaviors on the bubble formation in the market
 - (a) In the market which traders are mainly influenced by the noise information, the behavior of price is close to the random walk and this market is more likely to be a weak effective market;
 - (b) In the market which traders are mainly influenced by the price tendency, the behavior of price seems smooth and the bubble is not easy to form;
 - (c) In the market which traders are mainly influenced by the public information, The price changes dramatically, a huge bubble (or crash) is easily to be happen;
- (2) we studied the dynamic of the bubble and found that
 - (a) the public information has an important influence on the beginning of a bubble;
 - (b) traders' different expectations and their self-feedback is one of reasons for the existence of log-periodicity in bubble;
 - (c) the existence of power-law growth and log-periodicity, which leads the probability of prediction for the bursting of bubble, is caused by the combined effects of public information, traders' different expectations and their self-feedback.

In summary, we found that the heterogeneity of agents caused the complex pattern of bubbles in the real market.

Acknowledgments We acknowledge the research grants from National Natural Science Foundation of China (Nos. 71001101 and 71273261) and Program for New Century Excellent Talents in University (No. NCET-10-0779).

References

- Bastiaansen, K., Cauwels, P., Sornette, D., Woodard, R., & Zhou, W. X. (2009). The chinese equity bubble: Ready to burst.
- Black, F. (1996). Noise. *Journal of Finance*, *XLI*(3), 529–543.
- Blanchard, O. J. (1979). Speculative bubbles, crashes and rational expectations. *Economics Letters*, *3*(4), 387–389.
- Blanchard, O. J., & Watson, M. W. (1983). Bubbles, rational expectations and financial markets. National Bureau of economic research working paper series no. 945.
- Bouchaud, J., Potters, M., & ebrary, I. (2000). *Theory of financial risks: From statistical physics to risk management*, Vol. 217. Cambridge: Cambridge University Press.
- Chen, S. P., & He, L. Y. (2010). Multifractal spectrum analysis of nonlinear dynamical mechanisms in China's agricultural futures markets. *Physica A: Statistical Mechanics and its Applications*, *389*(7), 1434–1444.
- Chiarella, C., Iori, G., & Perellió, J. (2009). The impact of heterogeneous trading rules on the limit order book and order flows. *Journal of Economic Dynamics and Control*, *33*(3), 525–537.
- Diba, B., & Grossman, H. (1987). On the inception of rational bubbles. *The Quarterly Journal of Economics*, *102*(3), 697–700.
- Diba, B., & Grossman, H. (1988). The theory of rational bubbles in stock prices. *The Economic Journal*, *98*(392), 746–754.
- Dufwenberg, M., Lindqvist, T., & Moore, E. (2005). Bubbles and experience: An experiment. *The American Economic Review*, *95*(5), 1731–1737.
- Filimonov, V., & Sornette, D. (2011). A stable and robust calibration scheme of the log-periodic power law model. ArXiv e-prints 1108.0099.
- Froot, K., & Obstfeld, M. (1991). Intrinsic bubbles: The case of stock prices. *The American Economic Review*, *81*(5), 1189–1214.
- Glózkaynak, R. S. (2008). Econometric tests of asset price bubbles: Taking stock. *Journal of Economic Surveys*, *22*(1), 166–186.

- Haruvy, E., & Noussair, C. N. (2006). The effect of short selling on bubbles and crashes in experimental spot asset markets. *The Journal of Finance*, *61*(3), 1119–1157.
- He, L. Y. (2010). Is price behavior scaling and multiscaling in a dealer market? Perspectives from multi-agent based experiments. *Computational Economics*, *36*, 263–282.
- He, L. Y., & Chen, S. P. (2011a). Multifractal detrended cross-correlation analysis of agricultural futures markets. *Chaos, Solitons & Fractals*, *44*(6), 355–361.
- He, L. Y., & Chen, S. P. (2011b). A new approach to quantify power-law cross-correlation and its application to commodity markets. *Physica A: Statistical Mechanics and its Applications*, *390*(21–392), 3806–3814. doi:10.1016/j.physa.2011.06.013.
- He, L. Y., & Chen, S. P. (2011c). Nonlinear bivariate dependency of price-volume relationships in agricultural commodity futures markets: A perspective from multifractal detrended cross-correlation analysis. *Physica A: Statistical Mechanics and its Applications*, *390*(2), 297–308.
- He, L. Y., & Qian, W. B. (2012). A monte carlo simulation to the performance of the r/s and v/s methods statistical revisit and real world application. *Physica A: Statistical Mechanics and its Applications*, *391*(14), 3770–3782.
- He, L. Y., & Zheng, F. (2010). Detecting fractal/multifractal and asymmetric properties in an artificial quote-driven financial market. *Fractals* *18*(01):87–99. doi:10.1142/S0218348X10004762.
- He, L. Y., Fan, Y., & Wei, Y. M. (2009). Impact of speculator's expectations of returns and time scales of investment on crude oil price behaviors. *Energy Economics*, *31*(1), 77–84.
- Hong, H., & Stein, J. (2003). Differences of opinion, short-sales constraints, and market crashes. *Review of financial studies*, *16*(2), 487.
- Huang, Y., Johansen, A., Lee, M. W., Saleur, H., & Sornette, D. (2000). Artifactual log-periodicity in finite size data: Relevance for earthquake aftershocks. *Journal of Geophysical Research*, *105*(B11), 25451–25471.
- Jiang, Z. Q., Zhou, W. X., Sornette, D., Woodard, R., Bastiaensen, K., & Cauwels, P. (2010). Bubble diagnosis and prediction of the 2005–2007 and 2008–2009 Chinese stock market bubbles. *Journal of Economic Behavior & Organization*, *74*(3), 149–162.
- Johansen, A., Ledoit, O., & Sornette, D. (2000). Crashes as critical points. *International Journal of Theoretical and Applied Finance*, *3*(2), 219–255.
- Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., & Stanley, H. E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, *316*(1–4), 87–114.
- Lux, T. (2007). Applications of statistical physics in finance and economics. Tech. rep., Christian-Albrechts-University of Kiel, Department of Economics, economics working papers.
- Porter, D. P., & Smith, V. L. (1995). Futures contracting and dividend uncertainty in experimental asset markets. *The Journal of Business*, *68*(4), 509–541.
- Raberto, M., Cincotti, S., Focardi, S. M., & Marchesi, M. (2001). Agent-based simulation of a financial market. *Physica A: Statistical Mechanics and its Applications*, *299*(1–2), 319–327.
- Rosser, B. (2008). Econophysics and economic complexity. *Advances in Complex Systems (ACS)*, *11*(05), 745–760.
- Shiller, R. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *The American Economic Review*, *71*(3), 421–436.
- Sornette, D., Johansen, A., & Bouchaud, J. (1995). Stock market crashes, precursors and replicas. Arxiv, preprint cond-mat/9510036.
- Sornette, D., Woodard, R., & Zhou, W. X. (2009). The 2006–2008 oil bubble: Evidence of speculation, and prediction. *Physica A: Statistical Mechanics and its Applications*, *388*(8), 1571–1576.
- Stockl, T., Huber, J., & Kirchler, M. (2010). Bubble measures in experimental asset markets. *Experimental Economics*, *13*(3), 284–298.
- West, K. (1987). A specification test for speculative bubbles. *The Quarterly Journal of Economics*, *102*(3), 553–580.
- Zhou, W. X., & Sornette, D. (2002). Generalized q analysis of log-periodicity: Applications to critical ruptures. *Physical Review E*, *66*(4), 046–111.
- Zhou, W. X., & Sornette, D. (2003). 2000–2003 Real estate bubble in the UK but not in the USA. *Physica A: Statistical Mechanics and its Applications*, *329*(1–2), 249–263.
- Zhou, W. X., & Sornette, D. (2006). Is there a real-estate bubble in the US? *Physica A: Statistical Mechanics and its Applications*, *361*(1), 297–308.