Price Dynamics in an Order-Driven Market with Bayesian Learning

Jiahua Wang,1 Hongliang Zhu,1 and Dongxin Li2

1School of Management and Engineering, Nanjing University, Nanjing, Jiangsu 210093, China
2School of Finance, Nanjing University of Finance and Economics, Nanjing, Jiangsu 210046, China

Correspondence should be addressed to Hongliang Zhu; hlzhu@nju.edu.cn and Dongxin Li; leedongxin@nju.edu.cn

Received 4 April 2018; Revised 12 August 2018; Accepted 30 September 2018; Published 7 November 2018

Academic Editor: Dan Selişteanu

Copyright © 2018 Jiahua Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we have developed a model of limit order book with learning mechanism and investigated its price dynamics. In this model, continuous Bayesian learning is introduced to describe the dynamics of self-adjusting learning mechanism of agents, which can result in some important stylized facts of limit order markets. This study also provides some behavioral explanations for these well-known stylized facts that are commonly observed in the financial markets.

1. Introduction

In order-driven markets investors can submit either market or limit orders. Market orders will be immediately executed against the quoted bid or ask, whereas limit orders are stored in the limit order book, which is the list of all buy and sell limit orders with their corresponding price and size. Buy and sell orders are matched continuously subject to price and time priorities. Nowadays, most of the security exchanges around the world are considered to be purely electronic limit order markets, and hence it is obviously important to understand the dynamics of the limit order book [1–3].

However, the decision-making rules of the investors are so complicated that it is very difficult to derive the asset prices analytically as well as to evaluate the mechanism of market empirically in behavioral finance. Complex phenomena on asset prices emerge from the interactions between the micro-rules of the behavior of investors [4, 5]. In the last two decades, a number of interesting agent-based models have been proposed to explain the origin of some phenomena in order-driven markets [6–12]. Agent-based modeling has provided a powerful tool to investigate the effect of the mechanism of market. These models successfully describe the statistical properties of asset prices and some important stylized facts of financial markets, such as fat tails for the distribution of returns and volatility clustering [10, 13–35].

In stock markets, heterogeneous traders exhibit different investment behaviors that lead to the time evolution of market prices. Research on heterogeneous agents is an important area in behavioral finance and in the agent-based modeling [36–44]. Many agent-based models focus on analyzing the impact of heterogeneous traders on the fluctuations in price. Chiarella and Iori [43] classified different trading strategies into three parts: noise trading, fundamentalism, and Chartism, which defined all the three trading strategies, and employed heterogeneous agent models to investigate how different traders may affect the dynamics of price, bid-ask spread, trading volume, and volatility. Furthermore, Chiarella et al. [37] continued to modify their models by introducing the utility function. Takahashi et al. [44] employed an agent-based approach to analyze how the asset prices are affected by the investors and investment systems that are based on behavioral finance. Maskawa [45] considered the mimetic behavior of traders in a continuous double auction market and studied the fluctuations in the stock price and the power-law tail of the distribution of returns. LeBaron et al. [46] introduced an order-driven market with heterogeneous investors, analyzed the markets with and without learning and adaptation, and showed that the interaction of purely fundamentalist agents with heterogeneous estimates is sufficient to generate fat tails and volatility clustering.

In order to track the dynamic demands and supplies in order-driven markets and to clarify the rational and irrational
behaviors of heterogeneous investors, this study introduced few new assumptions from the psychological views into the microstructure modeling and tested them in an artificially built market. A model of limit order book has been proposed with learning mechanism which was developed using the concept of Chiarella et al. [37]. First, the agents are supposed to be able to update their orders, including the prices, volumes, and trading sides, when their arrival in the market is modeled as a Poisson’s process; second, in our model, the Bayesian learning process is introduced to distribute the strategic elements rationally for each investor along with the running time of market; third, the trading characteristics of each agent are not static after applying the Bayesian learning mechanism. Thus, this paper provides all the agents equal access to enter into the market and to update their orders, including the prices, volumes, and trading sides, when their arrival in the market occurs. Additionally, we assumed that each agent makes their investment decisions according to the combination of three pieces of information: the fundamental information, the technical information or price charts, and the market noise. Therefore, we can describe each agent by the combination of three investing styles: fundamentalist, chartist, and noise investors. If an agent is a fundamentalist, then there is a greater possibility for him to enter the market if the fundamental price is higher than the market price, irrespective of the intensity of market volatility. Based on the assumption, the expected rate of return of the risky asset that the agent i considers is set as follows:

\[
\bar{r}_{t+1,t} = \frac{1}{\omega_f} \sum_{j=1}^{\tau_f(t)} \omega_f \ln \left( \frac{p_f}{p_t} \right) + \omega_n \epsilon_t
\]

where \(\omega_f\), \(\omega_m\), \(\omega_n\) represent the weights of fundamental, chartist, and noise components, respectively, and all the three weights were set to be greater than or equal to zero. \(p_t\) and \(p_f\) are the market and fundamental prices of the asset, respectively, at time “t” and for every agent, \(\tau_f\) could be estimated by the time interval between time “t” and the next time at which the agent i arrives at the market. (This approximation is easy to understand as follows: when an agent arrives in the market, his expectation about the future price always focuses on time interval between time “t” and last time. Thus, the time horizon is set to be the interval between the two times at which the agent arrived.) \(\tau_f\), as discussed by Chiarella et al. [37], is the time scale over which the fundamentalist component for the mean reversion of price to the fundamental is calculated; \(\tau_i\) is defined as follows:

\[
\tau_{i} = \frac{1}{T} \sum_{j=1}^{T} \epsilon_{t-j} = \frac{1}{T} \sum_{j=1}^{T} \ln \left( \frac{p_{t-j}}{p_{t-j-1}} \right)
\]

where T is the total number of time steps before the time point t.

The paper is organized as follows. Section 2 describes the main model considered in this study which gives the detailed decision-making rules of the agents; Section 3 presents the simulated results of the model which include the parameters, market and mid-point prices/returns series, and the market depth, and a discussion related to the stylized facts is also considered in each subsection; Section 4 gives a conclusion of this study and provides some prospects for future research work.

2. The Model

As discussed by Chiarella et al. [37], we assume that every agent in the market knows the public information, the fundamental value of the risky asset which is assumed to follow the geometric Brownian motion (GBM), and submits orders based on their own past trading history. Each order contains volume and price that the agent plans to buy or to sell. There are only three options for the agent: sell order, buy order, or do not do anything. The number of agents is finite, and the agent arrival in the market is modeled as a Poisson’s process. The market price is set as follows: it is the past trading or quoted price when a real transaction occurs; the average of best bid or ask is considered if no real deal occurs.

Additionally, we assumed that each agent makes their investment decisions according to the combination of three pieces of information: the fundamental information, the technical information or price charts, and the market noise. Therefore, we can describe each agent by the combination of three investing styles: fundamentalist, chartist, and noise investors. If an agent is a fundamentalist, then there is a greater possibility for him to enter the market if the fundamental price is higher than the market price, irrespective of the intensity of market volatility. Based on the assumption, the expected rate of return of the risky asset that the agent i considers is set as follows:

\[
\bar{r}_{t+1,t} = \omega_f \ln \left( \frac{p_f}{p_t} \right) + \omega_n \epsilon_t
\]

where \(\omega_f\), \(\omega_m\), \(\omega_n\) represent the weights of fundamental, chartist, and noise components, respectively, and all the three weights were set to be greater than or equal to zero. \(p_t\) and \(p_f\) are the market and fundamental prices of the asset, respectively, at time “t” and for every agent, \(\tau_f\) could be estimated by the time interval between time “t” and the next time at which the agent i arrives at the market. (This approximation is easy to understand as follows: when an agent arrives in the market, his expectation about the future price always focuses on time interval between time “t” and last time. Thus, the time horizon is set to be the interval between the two times at which the agent arrived.) \(\tau_f\), as discussed by Chiarella et al. [37], is the time scale over which the fundamentalist component for the mean reversion of price to the fundamental is calculated; \(\tau_i\) is defined as follows:

\[
\tau_{i} = \frac{1}{T} \sum_{j=1}^{T} \epsilon_{t-j} = \frac{1}{T} \sum_{j=1}^{T} \ln \left( \frac{p_{t-j}}{p_{t-j-1}} \right)
\]

where T is the total number of time steps before the time point t.
After the initiation, the agents who arrive at the market will adjust their expectation based on their trading history, which could be described as an information set, $\delta_i^t$, that includes the history of fundamental rates, $r_i^{t\rightarrow}$, chartist rates, $\pi_s$ (s = 1, 2, ..., $t - r_i^t$), and the noise rates, $\epsilon_t$, whose continuous distribution could be assumed as normal. (This assumption is consistent with the assumption of geometric Brownian motion for the fundamental value of the asset.) Thus, from the historical information of the three types about agent i, the continuous distribution could be easily estimated based on the experience of the agent for $r_i^{t\rightarrow}$ and $\pi_s$ as shown in the following.

$$\delta_i^t = \{ r_i^{t\rightarrow}, \pi_s, \epsilon_t, s = 1, 2 \ldots t - r_i^t \}$$

$$= \{ f_i \left( r_i^{t\rightarrow} | w_{f, t-\tau_i} \right), f_i \left( \pi_s | w_{mn, t-\tau_i} \right), f_i \left( \epsilon_t | w_{n, t-\tau_i} \right) \}$$

(5)

From (5), it could be seen that the information set $\delta$ should present the past trading experience of agent “i” by time “$t$” and the normal distributed form of each random rate’s pdf could also be easily inferred from the previous assumption. (For convenience, the experimental distribution of $r_i^{t\rightarrow}$ and $\pi_s$ could also be assumed as normal, and from the assumption of geometric Brownian motion, it could be easily inferred that the difference in the price is also log normally distributed.)

Following this, (6), (7), and (8) could be set as below.

$$f_i \left( r_i^{t\rightarrow} | w_{f, t-\tau_i} \right) = \frac{1}{\sqrt{2\pi}\sigma_f} \exp \left[ -\frac{1}{2\sigma_f^2} \left( r_i^{t\rightarrow} - \bar{w}_{f, t-\tau_i} \right)^2 \right]$$

(6)

$$f_i \left( r_i^{mn} | w_{mn, t-\tau_i} \right) = \frac{1}{\sqrt{2\pi}\sigma_m} \exp \left[ -\frac{1}{2\sigma_m^2} \left( r_i^{mn} - \bar{w}_{mn, t-\tau_i} \right)^2 \right]$$

(7)

$$f_i \left( r_i^{n} | w_{n, t-\tau_i} \right) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp \left[ -\frac{1}{2\sigma_n^2} \left( r_i^{n} - \bar{w}_{n, t-\tau_i} \right)^2 \right]$$

(8)

In (6)-(8), $\sigma_k$ (k=f, m, n) denotes the standard deviation of the time series of $r_i^{t\rightarrow}, \pi_s, \epsilon_t$; $\bar{w}_t$ represents the mean of all the history values of $r_i^{t\rightarrow}$ by time “$t$” which is defined by (1); and $E_i[w_f + w_m + w_n]$ denotes the expectation of the sum of the three random weights by time “$t$”.

Based on $\delta_i^t$ at time “$t$”, it could be assumed that agent $i$ will adjust the investing strategy through the following Bayesian equations:

$$w_{k, t} | \delta_i^t \sim g \left( w_{k, t} | \delta_i^t \right) \quad (k = f, m, n)$$

(9)

where

$$g \left( w_f | \delta_i^t \right) = \frac{f_i \left( r_i^{t\rightarrow} | w_{f, t-\tau_i} \right) g \left( w_{f, t-\tau_i} \right)}{\int_0^{w_f} f_i \left( r_i^{t\rightarrow} | w_{f, t-\tau_i} \right) g \left( w_{f, t-\tau_i} \right) dw_f}$$

(10)

$$g \left( w_m | \delta_i^t \right) = \frac{f_i \left( \pi_s | w_{mn, t-\tau_i} \right) g \left( w_{mn, t-\tau_i} \right)}{\int_0^{w_m} f_i \left( \pi_s | w_{mn, t-\tau_i} \right) g \left( w_{mn, t-\tau_i} \right) dw_m}$$

(11)

$$g \left( w_n | \delta_i^t \right) = \frac{f_i \left( \epsilon_t | w_{n, t-\tau_i} \right) g \left( w_{n, t-\tau_i} \right)}{\int_0^{w_n} f_i \left( \epsilon_t | w_{n, t-\tau_i} \right) g \left( w_{n, t-\tau_i} \right) dw_n}$$

(12)

The initial distribution of the three components has been assumed as normal by (4).

Theorem 1. Based on the assumption of (4), the three random weights $(w_f, w_m, w_n)$ after Bayesian adjustment, which are defined by (10)-(12), will still follow the normal distribution at any point of time line. The mean and volatility after the adjustment are $(k=f, m, n)$

$$E \left[ w_k | \delta_i^t \right] = \frac{\mu_k \sigma_f^2 E_i \left[ \left( w_f + w_m + w_n \right)^2 \right] + r_k \sigma_w \sigma_f \sigma_m \left( w_f + w_m + w_n \right)}{\sigma_f^2 E_i \left( w_f + w_m + w_n \right) + \sigma_w^2 \sigma_f^2}$$

(13)

$$\text{var} \left[ w_k | \delta_i^t \right] = \frac{\sigma_w^2 \sigma_f^2 E_i \left[ \left( w_f + w_m + w_n \right)^2 \right] + \sigma_w^2 \sigma_f^2 \sigma_m^2}{\sigma_f^2 E_i \left( w_f + w_m + w_n \right) + \sigma_w^2 \sigma_f^2}$$

where $\sigma_w$ denotes the volatility of weight before Bayesian adjustment.

(In the appendix a proof has been given.)

Theorem 1 provides the specific formulation of the distribution of three random weightings after each time of Bayesian leaning, which assists in easily simulating the weightings during the process.

Based on the assumption of compound interest, given the expected rate of return, the following equation for the expected price could be obtained.

$$\bar{p}_{t+\tau_i} = p_t \exp \left( \overline{r}_{t+\tau_i} \right)$$

(14)

Following the assumption used by Chiarella et al. [37] and the constant absolute risk aversion (CARA) utility function, the number of shares that the agent “i” wishes to hold at time “$t$” could be obtained as follows:

$$V_i^* = \pi \left( \frac{\bar{p}_{t+\tau_i}}{p_t} \right) = \frac{\ln \left( \frac{\bar{p}_{t+\tau_i}}{p_t} \right)}{\alpha_i \text{var} \overline{r}_t p_t}$$

(15)

where $p_t$ is the market price at time “$t$”, var$_t$ is the variance at time “$t$”, and $\alpha_i$ is the risk aversion coefficient which is adjusted from the initial setting of $\alpha$ as given below

$$\alpha_i = \frac{1 + \overline{w}_{t, f}}{1 + \overline{w}_{t, m}}$$

(16)
so, the volume that each agent wants to trade is given as

\[ V_{ib}^i = V_i^{i*} - V_{i-t_i}^{i} \quad \text{if } V_i^{i*} > V_{i-t_i}^{i} \]
\[ V_{is}^i = V_i^{i} - V_i^{i*} \quad \text{if } V_i^{i*} < V_{i-t_i}^{i} \]  (17)

where \( V_{i}^{i} \) is the risky asset share that the agent “i” holds at time “t”, \( V_{ib}^i \) and \( V_{is}^i \) are the trading volumes of a buy order and a sell order, respectively, and \( V_{i-t_i}^{i} \) is the risky asset share that an agent holds before trading at time “t”. Besides, the price at which the agent submits orders to buy or to sell the risky asset is a uniformly distributed random variable as follows.

\[ p_i^t \sim U \left( p_i^{i_\text{low}}, p_i^{i_\text{high}} \right) \]  (18)

The two bounds of \( p_i^t \) are set as follows.

\[ p_i^{i_\text{low}} = \pi \left( p_i^{i_\text{low}}, V_{i-t_i}^{i} \right) = C_i^t \]
\[ p_i^{i_\text{high}} = \hat{p}_{i-t_i}^t \]  (19)

We can get another bound, \( p_i^{i*} \), as follows:

\[ p_i^{i*} = \pi^{-1} \left( V_{i-t_i}^{i} \right), \]  (20)

where the \( \pi \) function is determined by (15).

Table 1 presents the summary of the trading mechanism of agent “i” at time “t”.

### Table 1: Summary of the trading mechanism of agent “i” at time “t”.

<table>
<thead>
<tr>
<th>Order price</th>
<th>Type of order</th>
<th>Order volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_i^{i_\text{low}} &lt; p_i^t &lt; a_i )</td>
<td>Buy/Limit</td>
<td>( V_{ib}^i )</td>
</tr>
<tr>
<td>( a_i \leq p_i^t &lt; p_i^{i*} )</td>
<td>Buy/Market</td>
<td>( V_{i}^{i} )</td>
</tr>
<tr>
<td>( p_i^t = p_i^{i*} )</td>
<td>Do nothing</td>
<td></td>
</tr>
<tr>
<td>( p_i^{i*} &lt; p_i^t )</td>
<td>Sell/Market</td>
<td>( V_{is}^i )</td>
</tr>
<tr>
<td>( b_i &lt; p_i^t )</td>
<td>Sell/Limit</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2: Market parameters.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Economic meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Total number of agents who might arrive into the market</td>
</tr>
<tr>
<td>Pf</td>
<td>Initial price of the fundamental value of asset</td>
</tr>
<tr>
<td>Fsigma</td>
<td>Volatility for the fundamental value of the asset</td>
</tr>
<tr>
<td>Lamdap</td>
<td>Poisson’s probability to enter the market</td>
</tr>
<tr>
<td>P</td>
<td>Initial market price</td>
</tr>
<tr>
<td>Mstock</td>
<td>The maximum number of stock shares that investors could hold initially</td>
</tr>
<tr>
<td>Mcash</td>
<td>Maximum cash that investors could hold initially</td>
</tr>
<tr>
<td>Tick</td>
<td>Tick size for price in the submitted order</td>
</tr>
<tr>
<td>Time</td>
<td>Total time to run the market (seconds)</td>
</tr>
</tbody>
</table>

### Table 3: Bayesian parameters.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Economic meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_{rf}, \sigma_{rm}, \sigma_{rn} )</td>
<td>Standard deviation of the three rates: fundamental, chartism and noise rate</td>
</tr>
<tr>
<td>( \mu_{rf}, \mu_{rm}, \mu_{rn} )</td>
<td>Average of the three random weights after last adjustment</td>
</tr>
<tr>
<td>( \sigma_{rf}, \sigma_{rm}, \sigma_{rn} )</td>
<td>Volatility of the three random weights (expressed by standard deviation)</td>
</tr>
</tbody>
</table>

### Table 4: Values of the parameters.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>200</td>
</tr>
<tr>
<td>Pf(1)</td>
<td>125</td>
</tr>
<tr>
<td>Fsigma</td>
<td>0.01</td>
</tr>
<tr>
<td>Lamdap</td>
<td>0.3</td>
</tr>
<tr>
<td>P(1)</td>
<td>100</td>
</tr>
<tr>
<td>Mstock</td>
<td>50</td>
</tr>
<tr>
<td>Mcash</td>
<td>12500</td>
</tr>
<tr>
<td>Tick</td>
<td>0.0005</td>
</tr>
<tr>
<td>Time</td>
<td>3000</td>
</tr>
<tr>
<td>( \sigma_{rf}, \sigma_{rm}, \sigma_{rn} )</td>
<td>0.001/0.01</td>
</tr>
<tr>
<td>( \mu_{rf}, \mu_{rm}, \mu_{rn} )</td>
<td>100, 10, 500</td>
</tr>
<tr>
<td>( \sigma_{rf}, \sigma_{rm}, \sigma_{rn} )</td>
<td>0.01, 0.1, 0.01</td>
</tr>
<tr>
<td>A</td>
<td>0.1</td>
</tr>
</tbody>
</table>

### 3. Simulation Results

#### 3.1. Parameters.

Tables 2 and 3 present all the used exogenous parameters during the simulation process as well as their economic meaning.

By running the sensitivity analysis program, an initial value fitted for every parameter could be found out (Table 4).
3.2. Stylized Facts. With the settling of all the exogenous parameters, the model of limit order book has been presented and evaluated as "Bayes-EUT model". As clarified in the introduction, some statistical properties of order books, in particular, the following four well-known stylized facts in the order-driven markets, were investigated: (1) heavy-tailed return distribution; (2) the absence of autocorrelation of returns; (3) volatility clustering; (4) nontrivial Hurst exponent.

According to the observations of Cont and Bouchaud, “the unconditional distribution of mid-price returns displays tails that are heavier than a normal distribution (i.e. they have positive excess kurtosis), but thinner than a stable Pareto-Levy distribution” [47]. Chakraborti et al. [11] and Cont [48] found that “Except on very short timescales, when it exhibits weak negative autocorrelation, the time series of mid-price returns does not display any significant autocorrelation”. Cont et al. [19, 33] found that “time series of absolute or square mid-price returns displays positive autocorrelation that persists up to timescales of weeks or even months. The observed autocorrelation is slowly decaying (i.e. decays like a power-law function of time lag)”, which is defined as the consequence of periods with large price movements clustering together in time and periods with small price movements clustering together in time. Bouchaud and Potters [49] and Gençay et al. [50] found that “In a variety of different markets, including the S&P 500, foreign exchange markets, and interest rate markets, the mid-price return series was found to be sub-diffusive on timescales of up to months, and diffusive on longer timescales”, where a time series with a Hurst exponent, $H = 0.5$, is said to be diffusive, and $H < 0.5$ is said to be subdiffusive, whereas $H > 0.5$ is said to be superdiffusive.

It has been found that the focus of all the four stylized facts is on the analysis of the time series of mid-price, which is the average of the largest bid and the smallest ask and could be a little different from the market price. The price may not be the mid-point between the best ask and bid, though the smallest ask and the largest bid could change after each transaction. Therefore, both the characteristics of the return series of mid-price and the market price were explored to evaluate whether the outcome of the Bayes-EUT Model could depict the four important features of the order-driven markets.

### Table 5: Descriptive statistics of the return series.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.00025940</td>
</tr>
<tr>
<td>Variance</td>
<td>0.005735</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.07573</td>
</tr>
<tr>
<td>Maximum value</td>
<td>0.4313</td>
</tr>
<tr>
<td>Median</td>
<td>-0.00020391</td>
</tr>
<tr>
<td>Minimum value</td>
<td>-0.4331</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>8.1401</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.007674</td>
</tr>
</tbody>
</table>

3.3. Market Price and Return. Figure 1 presents the varying curve of the market price considering the assumptions of Bayes-EUT model, and Figure 2 presents the logarithmic return of market. Furthermore, Figures 3(a) and 3(b) demonstrate the distribution of return in the Bayes-EUT market. From these figures, it could be observed that the return has a heavier tail and larger kurtosis than the normal distribution.

Table 5 gives the descriptive statistics of the return series.

Since the kurtosis of the return series is 8.1401, which is larger than the kurtosis of standard normal distribution, it could be concluded that the Bayes-EUT market's return has a heavy tail. Then, the autocorrelation of the return series was calculated to check whether there is any significant existence of autocorrelation.

Figure 4 presents the varying process of the autocorrelation and partial correlation related to the market return series. From this figure, it becomes obvious that the absence of autocorrelation is perfectly satisfied for the market generated by the Bayes-EUT model. Figures 5 and 6 show the autocorrelation of the square return and absolute value of the return series, respectively. From these figures it could be
concluded that the volatility clustering of the return series is also satisfied for the Bayes-EUT market.

To test whether the return series is subdiffusive on short timescales and diffusive on longer timescales, the Hurst exponent of the market return along with different timescales was computed.

From Figure 7, it could be noted that for return on small timescales, the Hurst exponent of the market return series is obviously smaller than 0.5, which indicates that the return series is subdiffusive on short timescales, whereas for long timescales, the Hurst exponent is varying around 0.5. In general, the return could be diffusive on long timescales.

FIGURE 2: Logarithmic return of market.

FIGURE 3: Distribution of return in the Bayes-EUT market (normal distribution has zero mean and the same std as that of the return series).

FIGURE 4: Varying process of the autocorrelation and partial correlation related to the market return series.
Thus, the market return series based on the assumptions of Bayes-EUT model satisfies the fourth stylized facts.

3.4. Mid-Point Price and Return. Furthermore, in the Bayes-EUT market, the smallest ask and largest bid price are described as shown in Figure 8.

From Figure 8, it could be found that the bid and ask price become gradually closer to each other along with the time. After this, the mid-price, the mid-point of the smallest ask and largest bid, was considered. Figure 9 presents the mid-price and market price series in one coordinate plane. It could be noted that the trend of mid-price is similar to the market price, though the market price could vary much more significantly than the mid-price.

Figure 10 is the logarithmic return series of the mid-price. It could be observed that the standard deviation of normal distribution is the same as that of the mid-price return series as noted from Figures 11(a) and 11(b). Table 6 shows the descriptive statistical analysis for the mid-price return series.

<table>
<thead>
<tr>
<th></th>
<th>Table 6: Descriptive statistical analysis for the mid-price return series.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.00028624</td>
</tr>
<tr>
<td>Variance</td>
<td>0.000048348</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.006953</td>
</tr>
<tr>
<td>Maximum value</td>
<td>0.07114</td>
</tr>
<tr>
<td>Median</td>
<td>-0.00017934</td>
</tr>
<tr>
<td>Minimum value</td>
<td>-0.1801</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>206.2640</td>
</tr>
<tr>
<td>Skewness</td>
<td>-8.2835</td>
</tr>
</tbody>
</table>
Since the kurtosis of the mid-price return series is dramatically larger than that of the standard normal distribution, it could be concluded that the mid-price return series has a stylized “heavy tail”.

Considering the autocorrelation of the mid-price return, as observed in Figure 12, it could be seen that there are significant autocorrelation which exists on long timescales and on short timescales, and the negative autocorrelation is
not significant. This result reveals that the mid-price return series generated by the Bayes-EUT model does not satisfy the stylized fact as reported by Chakraborti et al. [11] and Cont [48].

Figure 13 is the autocorrelation of the absolute value of the mid-price return and the square of the mid-price return values. It could be noticed that the absolute return has the positive return persisting to long timescales, with a trend of gradual decay, which is consistent with the findings of Cont [33] and Liu et al. [19]. However, the square return does not display the persisting long time autocorrelation, where the positive autocorrelation of that series only appears in a very short time interval at the beginning of market's initialization. Therefore, the mid-price return does not satisfy volatility clustering.

From Figure 14, it could be observed that no matter it is a short or long timescales, the Hurst exponent is always larger than 0.5, which indicates that the mid-price return series is superdiffusive under the assumption of this model. Thus, the mid-price return generated by the Bayes-EUT model does not satisfy the fourth fact too.

Overall, for the market generated by the Bayes-EUT model, the return series of market price satisfies all the stylized facts found out by the empirical studies, but for the return of mid-price series, except for the first fact having a heavy tail, it does not satisfy all of the other three facts.

3.5. Market Depth. Figure 15 (log-log scale) and Figure 16 (log-linear scale) give the general shape of market depth distributed along with the distance from the mid-point price in the Bayes-EUT limit order book. Figure 15 displays an obvious decreasing trend along with the increasing the price distance, whereas Figure 16 displays an obvious hump shape and the depth is relatively higher in the middle of the graph. Both phenomena indicate that the simulated limit order book has some important features in the order-driven markets.

4. Conclusions

In this study, a model of limit order book has been proposed following the idea of Chiarella et al. [37]. By incorporating a Bayesian learning mechanism into the decision-making
process of agents and replacing the original time horizon with a random interval between the two time points when each agent arrives in the market, the proposed model from this study could give rise to some important stylized facts of limit order markets. It has been found that the market price and return series totally satisfy the four stylized facts, but the series of mid-point price and return have only a stylized heavy tail. Besides, the simulated market prices are far smaller than the fundamental values but have a much higher volatility, which reflects the evaluation activity of each agent regarding their own ideal value of the risk asset through their learning process. Moreover, the vast difference between
the behaviors of market price and mid-point price indicates that many transactions occur at a price far from the best ask and bid.

Generally speaking, our experimental results have two significant meanings: (1) the model successfully outputs the market price and return series similar to that of the real word, which means our assumption of the Bayesian learning mechanism could be somewhat reasonable; (2) according to our model, the market price of risky asset could dramatically deviate from its underlying fundamental value, and this case also happens in the real world. Therefore we could say that we find some reason regarding the deviation of market price in the real equity market.

Further research could focus on introducing other self-adjusting mechanisms to describe the statistical properties of order books, and the application of some behavioral theories could provide more valuable insights into the dynamics of the order-driven markets.

Appendix

Proof of Theorem 1. From (6)-(8), the information of the experience set for every agent is assumed as a set of some conditional normal probability distribution functions as follows.

\[
f(r_k | w_k) = \frac{1}{\sqrt{2\pi}\sigma_r} \exp\left( -\frac{1}{2\sigma^2_r} \left( r_k - \frac{E\bar{r}}{EW_i}w_k \right)^2 \right) \tag{A.1}
\]

Here, for convenience, \( r_k \) is used to represent three different information rates, \( \sigma_r \) and \( E\bar{r}_k \) denote the volatility and mean of \( r_k \), and \( EW_i \) denotes the expectation of the sum of the three random weights. On the other hand, \( w_k \) is also a normal random variable.

\[
g(w_k) = \frac{1}{\sqrt{2\pi}\sigma_w} \exp\left( -\frac{1}{2\sigma_w^2} (w_k - \mu_k)^2 \right) \tag{A.2}
\]

From (10)-(12), we have the abstract Bayesian posterior probability distribution function for each random weight, \( w_k \).

\[
g(w_k \mid r_k) = \frac{g(w_k) f(r_k \mid w_k)}{\int_{-\infty}^{\infty} f(r_k \mid w_k) g(w_k) dw_k} \tag{A.3}
\]

By adding a coefficient in front of \( w_k \), (A.2) is rewritten as follows (If \( x \sim N(0, 1) \), \( Ax + B \sim N(B, A^2)(A > 0) \), and then for any \( P > 0 \), we have \( P(Ax + B) \sim N(PB, (PA)^2) \)),

\[
g\left( \frac{E\bar{r}}{EW_i}w_k \right) = \frac{1}{\sqrt{2\pi} \left( \sigma_w \left( \frac{E\bar{r}}{EW_i} \right) \right)} \cdot \exp\left[ -\frac{1}{2 \left( \sigma_w \left( \frac{E\bar{r}}{EW_i} \right) \right)^2} \left( \frac{E\bar{r}}{EW_i}w_k - \frac{E\bar{r}}{EW_i}\mu_k \right)^2 \right] \tag{A.4}
\]

Substituting (A.1) and (A.4) into (A.3), then we have the following.
\[
g\left(\frac{E_r}{EW_t} \mu_k \mid r_k\right) = \frac{\exp\left[-\left(1/2 \left(\sigma_w (E_r/EW_t)\right)^2\right) \left(\left(\sigma_w (E_r/EW_t)\right) \mu_k - (E_r/EW_t) \mu_k\right)^2 - (1/2 \sigma_r^2) \left(r_k - (E_r/EW_t) \mu_k\right)^2\right]}{\int_{-\infty}^{+\infty} \exp\left[-\left(1/2 \left(\sigma_w (E_r/EW_t)\right)^2\right) \left(\left(\sigma_w (E_r/EW_t)\right) \mu_k - (E_r/EW_t) \mu_k\right)^2 - (1/2 \sigma_r^2) \left(r_k - (E_r/EW_t) \mu_k\right)^2\right] d \left((E_r/EW_t) \mu_k\right)}
\]

(A.5)

In (A.5), we set the following.

\[
\begin{align*}
\sigma_x &= \sigma_r, \\
\sigma_\mu &= \sigma_w \frac{E_r}{EW_t}, \\
x &= r_k, \\
m &= \frac{E_r}{EW_t} \mu_k, \\
\mu &= \frac{E_r}{EW_t} \mu_k
\end{align*}
\]

(A.6)

Also, (A.5) could be written as follows.

\[
g(\mu \mid x)
= \frac{\exp\left[-\left(1/2 \sigma_\mu^2\right) \left(\mu - m\right)^2 - (1/2 \sigma_x^2) \left(x - \mu\right)^2\right]}{\int_{-\infty}^{+\infty} \exp\left[-\left(1/2 \sigma_\mu^2\right) \left(\mu - m\right)^2 - (1/2 \sigma_x^2) \left(x - \mu\right)^2\right] d\mu}
\]

(A.7)

Then, we are going to calculate (A.7). First, since

\[
\begin{align*}
\exp\left[-\frac{1}{2 \sigma_\mu^2} \left(\mu - m\right)^2 - \frac{1}{2 \sigma_x^2} \left(x - \mu\right)^2\right] & = \exp\left[-\frac{\sigma_\mu^2}{2 \sigma_x^2 \sigma_\mu^2} \left(\mu - m\right)^2 + \frac{\sigma_\mu^2}{2 \sigma_x^2 \sigma_\mu^2} \left(x - \mu\right)^2\right] \\
& = \exp\left[-\frac{m^2 \sigma_\mu^2 + x^2 \sigma_\mu^2}{2 \sigma_x^2 \sigma_\mu^2}\right] \cdot \exp\left[\frac{m \sigma_\mu^2 + x \sigma_\mu^2}{\sigma_x^2 \sigma_\mu^2} \cdot \mu\right] \\
& \cdot \exp\left[\frac{\sigma_\mu^2 + \sigma_\mu^2}{2 \sigma_x^2 \sigma_\mu^2} \mu^2\right]
\end{align*}
\]

(A.8)

then,

\[
a = \frac{\sigma_x^2 + \sigma_\mu^2}{2 \sigma_x^2 \sigma_\mu^2}, \\
b = \frac{m \sigma_x^2 + x \sigma_\mu^2}{\sigma_x^2 \sigma_\mu^2}, \\
c = \exp\left[-\frac{m^2 \sigma_\mu^2 + x^2 \sigma_\mu^2}{2 \sigma_x^2 \sigma_\mu^2}\right]
\]

(A.9)

and, therefore, the integral in the denominator part of (A.7) could be expressed as follows.

\[
c \int_{-\infty}^{+\infty} \exp(-a \mu^2 + b \mu) d\mu = A \int_{-\infty}^{+\infty} e^{-t} dt
\]

(A.10)

Let \(A = (c/\sqrt{a})e^{b^2/4a}\), \(t = \sqrt{a} \mu - b/2\sqrt{a}\), then we have the following.

\[
c \int_{-\infty}^{+\infty} \exp(-a \mu^2 + b \mu) d\mu = A \int_{-\infty}^{+\infty} e^{-t} dt
\]

(A.11)

Since

\[
\left(\int_{-\infty}^{+\infty} e^{-t^2} dt\right)^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dx dy
\]

\[
= \int_{0}^{+\infty} \int_{-\pi}^{\pi} e^{-r^2} r dr d\theta
\]

\[
= \left(-\frac{1}{2}\right) \int_{0}^{+\infty} de^{-r^2} \left(\int_{-\pi}^{\pi} d\theta\right)
\]

\[
= \frac{1}{2} \cdot (0 - 1) \cdot 2\pi = \pi
\]

(A.12)

we have the following.

\[
\int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2 \sigma_\mu^2} \left(\mu - m\right)^2 - \frac{1}{2 \sigma_x^2} \left(x - \mu\right)^2\right] d\mu = A \sqrt{\pi} = \frac{c}{\sqrt{a}} \exp\left[\frac{b^2}{4a}\right] \sqrt{\pi}
\]

(A.13)
Applying (A.13) into (A.7), the following equation could be obtained.

\[
g(\mu | x) = \sqrt{\frac{\sigma^2 + \sigma^2}{2\pi\sigma^2\sigma^2}} \exp \left[-\left(\frac{\sigma^2 + \sigma^2}{2\sigma^2\sigma^2}\right) \mu^2 \right] - 2 \frac{m \sigma^2 + x \sigma}{\sigma^2 + \sigma^2}\mu + m \sigma^2 + x \sigma^2_{\mu} \left(\frac{\sigma^2 + \sigma^2}{\sigma^2 + \sigma^2}\right) \right]\]

\[
+ \left(\frac{\sigma^2 + \sigma^2}{\sigma^2 + \sigma^2}\right)^2 m \sigma^2 + x \sigma^2_{\mu} \left(\frac{\sigma^2 + \sigma^2}{\sigma^2 + \sigma^2}\right) \right] \right] (A.14)
\]

Then, by retaking the variables in (A.6), the posterior density function could be obtained.

\[
g\left(\frac{E \bar{w}_k}{EW_i} | r_k \right) = \left(\frac{\sigma^2_{E \bar{w}_k} + \sigma^2_{\sigma^2_{\bar{w}_k}E \bar{w}_k}^2}{2\pi\sigma^2_{\sigma^2_{\bar{w}_k}}E \bar{w}_k} \right)^{1/2} \cdot \exp \left[-\left(\frac{\sigma^2_{E \bar{w}_k} + \sigma^2_{\sigma^2_{\bar{w}_k}E \bar{w}_k}^2}{2\sigma^2_{\sigma^2_{\bar{w}_k}}E \bar{w}_k} \right) \mu \right] \]

\[
\cdot \left(\frac{E \sigma^2_{\bar{w}_k} - \sigma^2_{E \bar{w}_k} + \sigma^2_{\sigma^2_{\bar{w}_k}E \bar{w}_k}^2}{2\sigma^2_{\sigma^2_{\bar{w}_k}}E \bar{w}_k} \right)^{1/2} \left(\frac{\sigma^2_{E \bar{w}_k} + \sigma^2_{\sigma^2_{\bar{w}_k}E \bar{w}_k}^2}{2\sigma^2_{\sigma^2_{\bar{w}_k}}E \bar{w}_k} \right) \right] (A.15)
\]

From (A.7), the following equation could be obtained.

\[
g(\omega_k | r_k) = \left(\frac{\sigma^2_{E \bar{w}_k} + \sigma^2_{\sigma^2_{\bar{w}_k}E \bar{w}_k}^2}{2\pi\sigma^2_{\sigma^2_{\bar{w}_k}}E \bar{w}_k} \right)^{1/2} \cdot \exp \left[-\left(\frac{\sigma^2_{E \bar{w}_k} + \sigma^2_{\sigma^2_{\bar{w}_k}E \bar{w}_k}^2}{2\sigma^2_{\sigma^2_{\bar{w}_k}}E \bar{w}_k} \right) \mu \right] \]

\[
\cdot \left(\frac{\mu_0 + \sigma^2_{E \bar{w}_k} + \sigma^2_{\sigma^2_{\bar{w}_k}E \bar{w}_k}}{\sigma^2_{E \bar{w}_k} + \sigma^2_{\sigma^2_{\bar{w}_k}E \bar{w}_k}} \right)^{1/2} \left(\frac{\sigma^2_{E \bar{w}_k} + \sigma^2_{\sigma^2_{\bar{w}_k}E \bar{w}_k}^2}{2\sigma^2_{\sigma^2_{\bar{w}_k}}E \bar{w}_k} \right) \right] (A.16)
\]

Thus, the following equation could be obtained finally.

\[
\omega_k | r_k \sim N\left(\frac{\mu_0 + \sigma^2_{E \bar{w}_k} + \sigma^2_{\sigma^2_{\bar{w}_k}E \bar{w}_k}}{\sigma^2_{E \bar{w}_k} + \sigma^2_{\sigma^2_{\bar{w}_k}E \bar{w}_k}}, \frac{\sigma^2_{\sigma^2_{\bar{w}_k}E \bar{w}_k}^2}{\sigma^2_{E \bar{w}_k} + \sigma^2_{\sigma^2_{\bar{w}_k}E \bar{w}_k}} \right) (A.17)
\]

\[\square\]

**Data Availability**

All the data used to support the findings of this study are included within the article.

**Additional Points**

*Highlights.* (i) We developed a model of limit order book with learning mechanism. (ii) A mechanism for continuous Bayesian learning is introduced to describe the decision-making rules. (iii) Some important stylized facts and price dynamics are investigated. (iv) Interesting features of the mid-price return series in the order-driven market are described.

**Conflicts of Interest**

The authors declare that they have no conflicts of interest.

**Acknowledgments**

Hongliang Zhu acknowledges the financial support from the National Natural Science Foundation (NNSF) of China under grant No. 71717109, and Humanity and Social Science Foundation of Ministry of Education of China under grant No. 17YJA790101. Dongxin Li acknowledges the financial support from the NNSF of China under the grant No. 71203144. We also would like to thank the kind help from Professor Carole Bernard from the University of Waterloo and Jian Feng from the Department of Mathematics, Nanjing University. We thank editor and two anonymous reviewers for their helpful comments on an earlier draft of this paper.

**References**


