Available online at www.sciencedirect.com
SCIENCE DIRECT®

Nonlinear Analysis 63 (2005) e2009 – e2019

Nonlinear Analysis

www.elsevier.com/locate/na

Bayes estimation via filtering for a simple micro-movement model of asset price with discrete noises☆

Yong Zeng*

Department of Mathematics and Statistics, University of Missouri at Kansas City, Kansas City, MO 64110, USA

Abstract

A simple partially observed model for micro-movement of stock prices is proposed. The micro-movement refers to the transactional price behavior. The model can be framed as a filtering problem with counting process observations. Under this framework, the whole sample path is observable and is used for parameter estimation. Based on the filtering equation, we construct a recursive algorithm to compute the approximate posterior and the Bayes estimates. The consistencies of the recursive algorithm and of Bayes estimates are proven. Bayes estimates for transaction prices of Microsoft are obtained with a financial application.

Keywords: Filtering; Counting process; High frequency data; Bayes estimation

1. Introduction

Asset price models can be classified into two categories: macro- and micro-movement models. Macro-movement refers to daily, weekly, and monthly closing price behavior and micro-movement refers to transactional price behavior. There is a strong connection between the macro- and micro-movements, because the macro-movement is a daily-spaced time series sample from the micro-movement and the overall shapes of both movements

☆ The work was supported in part by a grant from University of Missouri Research Broad (2004).

* Tel.: +1 816 235 5850.
E-mail address: zeng@mendota.umkc.edu.

0362-546X/$ - see front matter
doi:10.1016/j.na.2004.02.105

are the same. However, there are also striking differences between the macro- and the micro-movements. Macro-movement can be viewed as an equally spaced time series, but micro-movement cannot. Macro-movement is low-frequency, but micro-movement is high-frequency. There are noises in stock prices. The impact of noises in macro-movement is small and noises are usually neglected. However, the impact of noises in micro-movement is large and noises must be modeled explicitly. We raise a question: how to model the micro-movement taking into account its strong connection with the macro-movement and their salient distinctions?

Stock price is distinguished from stock value and their distinction is noises. Black [2] remarks that noise is what makes our observations of a stock value imperfect, but is also a key factor that makes trading in financial markets possible so that prices can be observed. Noise moves the price of a stock just a little away from its value and usually not too far, because the further the price goes away from its value, the more aggressive the informed traders will become and the price will move back toward its value. Noise is contrasted with information, which influences stock value. Hasbrouck [7] further points out that information has a long-term, or “permanent” impact on stock price while noise has no influence on stock value and has only short-term or “transitory” impact on stock price. So, we raise another question: how to model stock price taking into account stock value, the permanent impact of information and the transitory impact of noises?

My answer to both questions is the simple model following [10]. The connection between both questions lies in the observation that a macro-movement model can serve as a model for asset value, because noises are ignored in macro-movement models. Then, the major concern left is what kinds of noise are to be modeled and how they are incorporated in the model.

The literature of market microstructure has documented many kinds of noise (see [9]). In our simple model, we focus on modeling three important types of noise in financial data: discrete, clustering and non-clustering noises. Although stocks have been traded at decimal system at NYSE since 2000, these three types of noise are still striking. Also, several theories have been proposed to explained why the price discreteness and clustering exist. For examples, see [4–6]. One natural way to model discrete noise is to round the continuous value to the nearest tick,[1] This rounding model was studied by [1]. Even the estimation of the rounding model is difficult because of the complexity of its likelihood function due to the unobservability of the underlying stock value.

The simple model has an important feature to be formulated as a filtering problem with counting process observations. Under this view, the whole sample path is observable and complete information is used for estimation. We apply Bayes estimation via filtering equation introduced in [10] to obtain the Bayes estimates.

The outline of the paper is the following. We present the simple model in two equivalent ways in Section 2. In Section 3, we review the theory of Bayes estimation via filtering and prove the consistency of the Bayes estimates. In Section 4, Bayes estimates for 2-month Microsoft transaction prices are presented with a simple application.

[1] a tick is the minimum price variation in trading.

2. The simple model

We follow this intuition to build the model: the micro-movement model is built from the macro-movement model by combining the noises in the high-frequency manner, or economically, stock price is formed from stock value by incorporating the (financial) noises when a trade occurs.

The simple model is intended to fit five sample characteristics of micro-movement that cannot be explained by the rounding model. Among these five, four are discrete-type. First, the observed frequency for price changes that are more than a tick is larger than the frequency implied by the rounding model. Second, for highly traded stocks, several trades may take place within one second. However, their prices are not observed to be the same and the difference can even be two or more ticks. Third, there are outliers. The fourth one is price clustering. If a stock value process is assumed to be diffusion or jump diffusion, and the discrete price is obtained by the rounding model, then we would expect approximately equal probability for each tick. But empirical findings are that integers are more common than halves; halves are more common than odd quarters; and odd quarters are more common than odd eighths (when the tick was one-eighth dollar) (see [6]). This means prices are clustered on even eighths.

The fifth one is the strong (the magnitude is close to but less than 0.5) negative acf(1) (autocorrelation function at lag 1) of price changes in micro-movement, which contrasts with the weak negative acf(1) in macro-movement.

Evidence for all these five sample characteristics is presented in Section 4 when we summarize a 2-month micro-movement data set of Microsoft.

Suppose that $v(t)$ is an unobserved value process for a stock, and it can be partially observed through the price process, $p(t)$. $v(t)$ lives in a continuous state space while $p(t)$ lives in a discrete state space given by the multiples of a tick, which is assumed to be $1/M$. Prices can only be observed at the irregularly spaced trading times, which are assumed to be driven by a Poisson process with a deterministic intensity, $a(t)$. The inhomogeneity assumption is more general than the naive assumption of homogeneity, it fits trade duration (waiting time) data better, and it can explain the observation that trading activity is higher near the opening and the closing than in the middle of the day.

The value process, $v(t)$, is simply assumed to be a linear Brownian motion (LBM), which in stochastic differential equation form is

$$\mathrm{d}v(t) = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W(t). \tag{1}$$

The trend is modeled by $\mu\,\mathrm{d}t$, and the information by $\sigma\,\mathrm{d}W(t)$.

There are two methods to build our model from the value process. One by constructing p from v via incorporating noises, and the other by formulating (v, p) as a filtering problem with counting process observations.

2.1. Construction of p from v

Suppose trading times $t_1, t_2, \ldots, t_i, \ldots$ are generated by a Poisson process with the intensity $a(t)$. To simplify notation, set $v = v(t_i)$, the value at time t_i, and set $p = p(t_i)$, the price at time t_i. We construct p from v in three steps.

Step 1: *Incorporate Discrete Noise* by rounding off v to its closest tick, $R[v, 1/M]$. Without other noises, trades should occur at this tick, which is the closest tick to the stock value.

Step 2: *Incorporate Non-clustering Noise* by adding: $v' = R[v, 1/M] + U$, where U is the non-clustering noise of trade i at time t_i. We assume $\{U_i\}$, are independent of the value process, and they are i.i.d. with a doubly geometric distribution:

$$P\{U = u\} = \begin{cases} (1-\rho) & \text{if } u = 0, \\ \frac{1}{2}(1-\rho)\rho^{M|u|} & \text{if } u = \pm\frac{1}{M}, \pm\frac{2}{M}, \ldots \,. \end{cases}$$

The non-clustering noise can explain three of the four discreteness-related sample characteristics. First, the non-clustering noise increases considerably the probability of the successive price changes that are more than a tick. Next, it allows the prices of trades occurring within the same second to differ and the difference can be two or more ticks. Finally, it produces outliers.

Step 3: *Incorporate Clustering Noise* by random biasing. After rounding the value process and adding the non-clustering noise, the fractional part of v' is still approximately uniformly distributed on all fractional parts. We bias v' through a random biasing function $b(\cdot)$ to produce price clustering. $\{b_i(\cdot)\}$ are assumed independent of $\{v'_i\}$ and serially independent given the sequence $\{v'_i\}$.

To be consistent with the data analyzed in Section 4, we construct a simple random biasing function only for the tick of 1/8 dollar. The data to be analyzed has this clustering phenomenon: integers and halves are most likely and have about the same frequencies; odd quarters are the second most likely and have about the same frequencies; and odd eighths are least likely and have about the same frequencies. To generate such clustering, a random biasing function is constructed based on the following rule: if the fractional part of y' is even eighths, then y stays on y' with probability one; if the fractional part of y' is odd eighth, then y stays on y' with probability $1 - \alpha - \beta$, y moves to the closest odd quarter with probability α, and moves to the closest half or integer with probability β. The detail of $b(\cdot)$ is presented in Appendix A of [11].

In summary, the construction is $p(t_i) = b_i(R[V(t_i), 1/M] + U_i)$.

Through the construction, the transition probability from v to p, denoted by $g(p|v)$, can be computed through $g(p|v) = \sum_{v'} g(p|v')g(v'|v)$ where $g(p|v')$ (or $g(v'|v)$) is transition probability from v' (or v) to p (or v'). The detail of $g(p|v)$ is in Appendix A of [11].

The simple model induces the "right" autocorrelation pattern on price change. That is, acf(1) of price changes is weak and negative in macro-movement, but strong (still less than 0.5) and negative in micro-movement; and acf(k) for $k \geqslant 2$ are zero.[2] This is confirmed by simulations in Section 3. Last but not least, the model can be framed as a filtering problem with counting process observations. This is important for statistical analysis, because under this framework, we are able to derive the filtering equation, which characterizes the posterior given the whole sample path.

[2] To see this, we look at a model close to our simple model. The value process is assumed to be LBM defined in Eq. (1) but trading times are equally spaced, that is, $t_i = t_0 + \mathrm{i}\Delta_t$ for $i = 1, 2, 3, \ldots$. Assume the price at time t_i is $p(t_i) = v(t_i) + z(t_i)$ where $z(t_i)$, the noise at trading time t_i, is assumed to be independent of the value

2.2. Counting process observations

Alternatively, we can view the asset price according to price level because of price discreteness. Namely, we view the prices as a collection of counting processes as follows:

$$\vec{p}(t) = \begin{pmatrix} N_1\left(\int_0^t \lambda_1(\vec{\theta}, v(s), s)\,\mathrm{d}s\right) \\ N_2\left(\int_0^t \lambda_2(\vec{\theta}, v(s), s)\,\mathrm{d}s\right) \\ \vdots \\ N_n\left(\int_0^t \lambda_n(\vec{\theta}, v(s), s)\,\mathrm{d}s\right) \end{pmatrix}, \tag{2}$$

where $p_k(t) = N_k(\int_0^t \lambda_k(\vec{\theta}, v(s), s)\,\mathrm{d}s)$ is the observable counting process recording the cumulative number of trades that have occurred at the kth price level (denoted by p_k) up to time t, and $\vec{\theta} = (\mu, \sigma, \rho)$.

According to the theory of multivariate point processes, we make four mild assumptions to ensure the construction of the simple model is equivalent to the specification of counting process observations. For more explanations of these assumptions, see [10].

Assumption 1. N_k*'s* are unit Poisson processes under measure P.

Assumption 2. $v, N_1, N_2, \ldots, N_n$ are independent under measure P.

Assumption 3. The structure of the intensity is $\lambda_k(\vec{\theta}, v, t) = a(t)g(p_k|v)$, where $a(t)$ is the total intensity and $g(p_k|v)$ is the transition probability from v to p_k, the kth price level at time t.

Assumption 4. The total intensity process, $a(t)$, is uniformly bounded below and above, namely, there exist positive constants, C_1 and C_2, such that $C_1 < a(t) \leqslant C_2$ for all $t > 0$.

3. Bayes estimation via filtering equation

There are five parameters, $(\mu, \sigma, \rho, \alpha, \beta)$. (μ, σ) relate to the value process, ρ relates to the non-clustering noise, and (α, β) relate to the clustering noise. (α, β) can be estimated

(Continued)
process, and the sequence of $\{z(t_i)\}$ is serially independent with zero mean and a common variance δ^2. Then, the autocorrelation function for price change at lag 1, denoted by acf(1), is,

$$\mathrm{acf}(1) = \mathrm{Corr}(p(t_{i+1}) - p(t_i), p(t_i) - p(t_{i-1})) = -\frac{\delta^2}{\sigma^2 \Delta t + 2\delta^2}.$$

So, the acf(1) is always negative and decreasing (in absolute value) with Δt. In the macro-movement setting, Δt is very large and $\sigma^2 \Delta t$ dominates. So the acf(1) produces weak negative autocorrelation. While in the micro-movement setting, Δt is hundreds of times smaller and δ^2 becomes dominant. Then acf(1) is close to but less than 0.5 (in absolute value). This explains why there is a weak negative acf(1) in daily data, but a strong negative acf(1) in tick-by-tick data. And acf after lag 1 are zero for this model, which is in line with empirical findings.

by the method of relative frequency and (μ, σ, ρ) are estimated by Bayesian approach via filtering equation.

Bayes estimate, which is the posterior mean, is the least Mean Square Errors (MSE) estimate. The core of the Bayesian estimation via filtering is to construct an algorithm to compute the conditional distribution, which becomes a posterior after a prior is assigned. The algorithm, which is based on the filtering equation, is naturally recursive with every trade. One basic requirement for the recursive algorithm is consistent, namely, the conditional distribution computed by the recursive algorithm converges to the true one determined by the filtering equation. This is guaranteed by a theorem on the convergence of conditional expectation.

We first review the filtering equation and the theorem on the convergence of conditional expectation, then construct the convergent recursive algorithm for the simple model in detail, and finally prove the consistency for the Bayes estimates.

3.1. Review of two theorems

In [10], the filtering equation for a general model including this simple model is derived and the theorem on the convergence of conditional expectation for the general model is proven. Here, we state a simpler version of the two theorems applying to the simple model. Suppose that $\vec{\theta}$ is a vector of parameters. One general assumption on $(\vec{\theta}, v)$ is made as follows.

Assumption 5. $(\vec{\theta}, v)$ is the solution of a martingale problem for a generator $\mathbf{A}_{\vec{\theta}}$ such that

$$M_f(t) = f(\vec{\theta}, v(t)) - \int_0^t \mathbf{A}_{\vec{\theta}} f(\vec{\theta}, v(s))\,\mathrm{d}s$$

is a $\mathscr{F}_t^{\vec{\theta},v}$-martingale, where $\mathscr{F}_t^{\vec{\theta},v}$ is the σ-algebra generated by $(\vec{\theta}, v(s))_{0\leqslant s\leqslant t}$.

Set $\vec{\theta} = (\mu, \sigma, \rho)$, the generator for in the simple model is

$$\mathbf{A}_{\vec{\theta}} f(\vec{\theta}, v) = \mu \frac{\partial f}{\partial v}(\vec{\theta}, v) + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial v^2}(\vec{\theta}, v). \tag{3}$$

Let $\mathscr{F}_t^{\vec{p}} = \sigma\{(\vec{p}(s))|0 \leqslant s \leqslant t\}$, let π_t be the conditional distribution of $(\vec{\theta}, v(t))$ given $\mathscr{F}_t^{\vec{p}}$ and $\pi(f, t) = E^P[f(\vec{\theta}, v(t))|\mathscr{F}_t^{\vec{p}}] = \int f(\vec{\theta}, v)\pi_t(d\vec{\theta}, dv)$. We denote $\mathbf{A}_{\vec{\theta}}$ simply as $\mathbf{A}$ in the rest.

Theorem 1. *Suppose that $\vec{p}$ is the counting process observations defined in Eq.* (2) *with Assumptions* 1 *to* 4 *and $(\vec{\theta}, v)$ satisfies Assumption* 5. *Then, π_t, is the unique solution of the filtering equation*: *for every $t > 0$ and everyf in the domain of* $\mathbf{A}$, *the generator*,

$$\pi(f, t) = \pi(f, 0) + \int_0^t \pi(\mathbf{A}f, s)\,\mathrm{d}s + \sum_{k=1}^n \int_0^t \left[\frac{\pi(f g_k, s-)}{\pi(g_k, s-)} - \pi(f, s-)\right] \mathrm{d}p_k(s), \tag{4}$$

where $p_k(s)$ is the kth price level in $\vec{p}(s)$ and $g_k = g(p_k|v)$ is the transition probability from v to p_k, the kth price level.

The filtering equation provides an effective way to characterize π_t and it is the optimum filter in the sense of least MSE. Note that $a(t)$ disappears in Eq. (4).

For the second theorem, denote $(\vec{\theta}_{\vec{\varepsilon}}, v_\varepsilon) \Rightarrow (\vec{\theta}, v)$ as the weak convergence in the Skorohod topology as $\varepsilon \to 0$. Then, $(\vec{\theta}_{\vec{\varepsilon}}, v_\varepsilon)$ is an approximate of $(\vec{\theta}, v)$. Define

$$\vec{p}_\varepsilon(t) = \begin{pmatrix} N_1\left(\int_0^t \lambda_1(\vec{\theta}_{\vec{\varepsilon}}, v_{\varepsilon_v}(s-), s-)\,\mathrm{d}s\right) \\ N_2\left(\int_0^t \lambda_2(\vec{\theta}_{\vec{\varepsilon}}, v_{\varepsilon_v}(s-), s-)\,\mathrm{d}s\right) \\ \vdots \\ N_n\left(\int_0^t \lambda_n(\vec{\theta}_{\vec{\varepsilon}}, v_{\varepsilon_v}(s0), s-)\,\mathrm{d}s\right) \end{pmatrix}, \tag{5}$$

where $\varepsilon = \max(\varepsilon_v, |\vec{\varepsilon}|)$ and $|\vec{\varepsilon}|$ is the norm of a vector. Define $\mathscr{F}_t^{\vec{p}_\varepsilon} = \sigma(\vec{p}_\varepsilon(s), 0 \leqslant s \leqslant t)$.

Theorem 2. *Suppose that $(\vec{\theta}, v, \vec{p})$ is on the probability space $(\Omega, \mathscr{F}, P)$ and Assumptions* 1–5 *hold. Suppose that $(\vec{\theta}_{\vec{\varepsilon}}, v_{\varepsilon_v}, \vec{p}_\varepsilon)$ is on $(\Omega_\varepsilon, \mathscr{F}_\varepsilon, P_\varepsilon)$, and Assumptions* 1–5 *also hold. If $(\vec{\theta}_{\vec{\varepsilon}}, v_{\varepsilon_v}) \Rightarrow (\vec{\theta}, v)$ as $\varepsilon = \max\{\varepsilon_v, |\vec{\varepsilon}|\} \to 0$, then*

(i) *$\vec{p}_\varepsilon \Rightarrow \vec{p}$, as $\varepsilon \to 0$; and*
(ii) $E^{P_\varepsilon}[F(\vec{\theta}_{\vec{\varepsilon}}, v_{\varepsilon_v}(t))|\mathscr{F}_t^{\vec{p}_\varepsilon}] \Rightarrow E^P[F(\vec{\theta}, v(t))|\mathscr{F}_t^{\vec{p}}]$ as $\varepsilon \to 0$, *for all F in the domain of* $\mathbf{A}$, *the generator.*

This theorem provides not only the theoretical foundation for consistency, but also a three-step recipe for constructing a consistent recursive algorithm based on Kushner's Markov chain approximation method to compute the continuous-time posterior. Step 1 is to construct $(\vec{\theta}_\varepsilon, v_\varepsilon)$, the Markov chain approximation to $(\vec{\theta}, v)$ with generator $\mathbf{A}_\varepsilon$. Step 2 is to obtain the filtering equation for $\pi_\varepsilon(f, t)$ corresponding to $(\vec{\theta}_\varepsilon, v_\varepsilon, Y_\varepsilon)$ by applying Theorem 1. The filtering equation (for the approximate model) can be separated into the propagation equation:

$$\pi_\varepsilon(f, t_{i+1}-) = \pi_\varepsilon(f, t_i) + \int_{t_i}^{t_{i+1}-} \pi_\varepsilon(\mathbf{A}_\epsilon f, s)\,\mathrm{d}s, \tag{6}$$

and the updating equation (assuming that a trade at *j*th price level occurs at time t_{i+1}):

$$\pi_\varepsilon(f, t_{i+1}) = \frac{\pi_\varepsilon(f g_j, t_{i+1}-)}{\pi_\varepsilon(g_j, t_{i+1}-)}. \tag{7}$$

Step 3 converts Eqs. (6) and (7) to the recursive algorithm in discrete state space and time by two substeps: (a) represents $\pi_\varepsilon(\cdot, t)$ as a finite array with the components being $\pi_\varepsilon(f, t)$ for lattice-point indicator *f* and (b) approximates the time integral in (6) with an Euler scheme. The recursive algorithm for the simple model can constructed accordingly and interested readers are referred to [10] for more details on how to do so.

3.2. Consistency of the Bayes estimates

Theorem 3. *For the simple model defined in Section* 1, *suppose that the clustering parameters* (α, β) *are known, and* (μ, σ, ρ) *has a prior. Then* $E[f(\mu,\sigma,\rho)|\mathscr{F}_t^{\vec{p}}] \to f(\mu,\sigma,\rho)$ *a.s. as* $t \to \infty$ *for any bounded continuous function f.*

Proof. Set $\vec{\theta} = (\mu, \sigma, \rho)'$. In the proof, we see $\vec{p}(t)$ as $p(t)$, the price process. We know $E[f(\vec{\theta})|\mathscr{F}_t^p]$ is a martingale converging to $E[f(\vec{\theta})|\mathscr{F}_\infty^p]$ a.s. If we can show $\vec{\theta}$ is $\mathscr{F}_\infty^p$-measurable, then $E[f(\vec{\theta})|\mathscr{F}_\infty^p] = f(\vec{\theta})$ a.s. and the result follows. □

It suffices to construct a.s. consistent estimates of μ, σ and ρ based on the whole sample path of $p(t)$ where $0 \leqslant t \leqslant \infty$.

First, we construct a consistent estimate for μ. It is easy to check $(1/t)(p(t) - p(0)) \to \mu$ a.s. as $t \to \infty$.

Next, we construct an a.s. consistent estimate for σ^2. We first define a sequence of stopping times $\{\tau_i\}$ for each integer $n > 0$. Define $\tau_1 = t_1$, which is the first trading time. Then, recursively define $\tau_{i+1} = \inf_j \{t_j, t_j > \tau_i + n\}$, which is the first trading time after $\tau_i + n$.

Define $D_{i,n} = [p(\tau_{i+1}) - p(\tau_i)]/\sqrt{n}$ for $i = 1, 2, \ldots, n$. Define $\hat{\sigma}_n^2 = (1/n)\sum_{i=1}^n (D_{i,n} - \bar{D}_n)^2$, where $\bar{D}_n = (1/n)\sum_{i=1}^n D_{i,n}$.

Lemma 1. $\hat{\sigma}_n^2 \to \sigma^2$ *a.s. as* $n \to \infty$.

The proof of Lemma 1 is available upon reader's request.

So, $\hat{\sigma}_n^2$ (or $\hat{\sigma}_n$) is an a.s. consistent of $\sigma 2$ (or σ).

Finally, to construct a consistent estimate for ρ, we choose those two consecutive trades $p(t_{i-1})$ and $p(t_i)$ such that both fractional parts of prices are odd eighths and $t_i - t_{i-1} < \delta$. For any $\delta > 0$, we can always obtain infinitely many such pairs of consecutive trades as $t \to \infty$, because the probability that two trades occur within the time span of δ is strictly positive given $C_1 < a(t)$ by Assumption 4.

Note that when $p(t_i)$ is odd eighth, $p(t_i) = R[v(t_i), 1/M] + V_i$. When $t_{i+1} - t_i < \delta$ and as δ goes to zero, $R[v(t_{i+1}), 1/M] = R[v(t_i), 1/M]$ and $p(t_{i+1}) - p(t_i) = V_{i+1} - V_i$. Then, we can have an i.i.d. sequence of $\{p(t_i) - p(t_{i-1})\}$, each has the distribution of $V_{i+1} - V_i$.

When both $p(t_{i+1})$ and $p(t_i)$ are odd eighth, $p(t_{i+1}) - p(t_i)$ must be even eighth. In such an i.i.d. sequence of $\{p(t_i) - p(t_{i-1})\}$, we can observe the empirical relative frequency that $p(t_i) - p(t_{i-1}) = 0$ for a fix time span t , which is denoted by $f_{0,t}$.

We compute that

$$P\{V_i - V_{i-1} = 0 |\, V_i - V_{i-1} = 2k, k = 0, \pm 1, \pm 2, \ldots\} = \frac{(2-\rho^2)(1-\rho^2)}{2(1+\rho^2)}.$$

Set $[(2-\rho^2)(1-\rho^2)]/[2(1+\rho^2)] = f_{0,t}$. It is easy to check that the above equation has only one root between [0, 1]. This root is an estimate of method of relative frequency. By the well-known strong consistency for the estimate of method of relative frequency, we have an a.s. consistent estimate of ρ. □

Theorem 3 together with the convergence of the recursive algorithm implies the Bayes estimates computed by the recursive algorithm should converge to their true values. This is confirmed by extensive simulation studies conducted.

4. A short simulation and real data example

First, we describe a 2-month (January and February, 1994) transaction data (also referred as *ultra-high-frequency data* in [3]) of Microsoft. Using the dataset as benchmark, we simulate a data set to demonstrate that the simple model has the fifth sample characteristics of the real data. Finally, we present the Bayes estimates for a 2-month Microsoft transaction data.

4.1. Ultra-high-frequency data of microsoft and a simulation

The data are extracted from the Trade and Quote (TAQ) database distributed by NYSE. We apply standard procedures to filter the data. The final sample has 49,937 observations. To see price clustering is captured by the modeling of noise, interested readers are referred to Section 4 of [11].

We choose the parameters close to those of the Microsoft data set. Pick $\mu = 7.0 \times 10^{-6}$, $\sigma = 0.008$ (these two are in second), $\rho = 0.25$, $\alpha = 0.2$ and $\beta = 0.4$. Assume the trading intensity is fixed and pick $a(t) = 0.054$ for all $t > 0$, which means that there is a trade about every $1/0.054 = 18.52$ s. 50,000 data are simulated. We show the acf structure of the simple model coincides with that of the Microsoft data in micro-movement and macro-movement. Table 1 presents acf of price changes in micro-movement for the simulated data and the Microsoft data assuming equally spaced in time. Both acf(1) are negative and close to but less than 0.5 in absolute value, and acf(k) are not statistically significantly different from zero for $k = 2, 3, \ldots, 7$, except acf(2) of simulated data and acf(7) of Microsoft data (comparing to $2/\sqrt{49939} = 0.0895$). For the macro-movement, we first extract the last trade of each day to make of the daily prices. Table 2 presents acf of daily price changes of the simulated data and the Microsoft data. Note that all orders of acf do not differ from zero significantly (comparing to $2/\sqrt{39} = 0.3203$).

4.2. Bayes estimates for Microsoft data

First, using method of relative frequency, we estimate $\hat{\alpha} = 2(.3707 - .25) = .2414$, $\hat{\beta} = 2(.4251 - .25) = .3502$. For more details, see [12].

Table 1
ACFs of price changes for transactions data

Data	acf(1)	acf(2)	acf(3)	acf(4)	acf(5)	acf(6)	acf(7)
Simulated	−0.4670	−0.0113	0.0001	0.0017	0.0020	−0.0089	0.0083
Microsoft	−0.4905	0.0056	0.0024	−0.0073	−0.0023	−0.0079	−0.0101

Table 2
ACFs of price changes for daily data

Data	acf(1)	acf(2)	acf(3)	acf(4)	acf(5)	acf(6)	acf(7)
Simulated	0.1759	0.0386	−0.1202	0.0061	−0.1059	0.0110	−0.2001
Microsoft	0.0782	−0.1439	0.0913	−0.0165	0.1151	−0.0304	−0.0458

Table 3
Bayes estimates of MSFT, January and February 1994

Case	Time unit	μ	σ	ρ
With clustering	Second	2.209e-6	8.823e-3	0.2264
		(8.869e-6)	(1.055e-4)	(0.0023)
	Year	\$13.44	\$21.76	0.2264
		(\$53.96)	(\$0.26)	(0.0023)
Without clustering	Second	2.272e-6	1.049e-2	0.3230
		(1.001e-5)	(1.232e-4)	(0.0019)
	Year	\$13.82	\$25.87	0.3230
		(\$60.90)	(\$0.30)	(0.0019)

Numbers in “()” are standard errors and the annualized factor is 260.

The Bayes estimates of (μ, σ, ρ) are computed by a recursive algorithm in second. Assuming 260 business days per year, the annualized estimates then are obtained. Both estimates for the Microsoft data are summarized in the upper part of Table 3. Note that the standard error of μ is much larger than those of σ and ρ (comparing to the Bayes estimates). This makes sense because μ is a trend parameter, and its accuracy of estimation depends on the length covered by the data (40 days); but the accuracy of estimating σ and ρ mainly depend on the number of observations (49,937 data).

Comparing to the usual MLE for LBM using daily closing data, our Bayes estimates are more accurate, because the model is closer to the real data and complete information is used. A more accurate estimate, especially volatility estimate, ensures more accurate option pricing and has important applications for portfolio selection.

4.3. Does price clustering matter?

Price discreteness and clustering are two striking features of stock price. Hausman et al. [8] asked the question whether price discreteness matters in parameter estimation and they provide a strong affirmative answer. It is equally important to ask whether the related price clustering matters and we supply strong evidence in the affirmative.

If clustering noise is ignored, the variation due to clustering noise moves to σ and ρ. If the estimates of σ and ρ increase, it provides supportive evidence to the question. When we ignore price clustering, we may just set $\alpha = \beta = 0$ and estimate μ, σ and ρ. The Bayes estimates are presented in the lower part of Table 3. The Bayes estimate of μ differs little, but those of σ and ρ have increased significantly. The significant increase of σ and ρ provides the strong positive evidence that price clustering does matter in estimating volatility and non-clustering noise.

References

[1] C. Ball, Estimation bias induced by discrete security price, J. Finance 43 (1988) 841–865.
[2] F. Black, Noise, J. Finance 41 (1986) 529–543.
[3] R. Engle, The econometrics of ultra-high-frequency data, Econometrica 68 (2000) 1–22.
[4] P. Godek, Why NASDAQ market makers avoid odd-eighth quotes, J. Financial Econom. 41 (1996) 465–474.
[5] S. Grossman, M. Miller, K. Cone, D. Fischel, D. Ross, Clustering and competition in asset markets, J. Law Econom. 40 (1997) 23–60.
[6] L. Harris, Stock price clustering and discreteness, Rev. Financial Studies 4 (1991) 389–415.
[7] J. Hasbrouck, Modeling market microstructure time series, in: G. Maddala, C. Rao (Eds.), Handbook of Statistics, vol. 14, North-Holland, Amsterdam, 1996, pp. 647–692.
[8] J. Hausman, A. Lo, C. Mackinlay, An ordered probit analysis of stock transaction prices, J. Financial Econom. 31 (1992) 319–379.
[9] M. O'Hara, Market Microstructure Theory, Blackwell, Oxford, 1995.
[10] Y. Zeng, A partially-observed model for micro-movement of asset prices with Bayes estimation via filtering, Math. Finance 13 (2003) 411–444.
[11] Y. Zeng, Estimating stochastic volatility via filtering for the micro-movement of asset prices, IEEE Trans. Automat. Control 49 (2004) 338–348.
[12] Y. Zeng, L.C. Scott , Bayes estimation via filtering equation for O-U process with discrete noises: application to the micro-movement of stock prices, in: Bozenna Pasik-Duncan (Ed.), Lecture Notes in Control and Information Sciences, vol. 280, Springer, Berlin, 2002, pp. 533–548.