Available online at www.sciencedirect.com
SCIENCE DIRECT®

Nonlinear Analysis 64 (2006) 295–309

Nonlinear Analysis

www.elsevier.com/locate/na

A micro-movement model with Bayes estimation via filtering: Application to measuring trading noises and costs

Robert Spalding[a], Kam-Wah Tsui[b], Yong Zeng[a,*,1]

[a] *Department of Mathematics and Statistics, The University of Missouri at Kansas City, 5100 Rockhill Road, Kansas City, MO 64110, USA*

[b] *University of Wisconsin—Madison, WI, USA*

Received 16 June 2005; accepted 23 June 2005

Abstract

We highlight a general hybrid system as the micromovement model for asset price using counting processes recently introduced with its Bayes estimation via filtering. We construct a new simple micromovement model and apply it to analyze trade-by-trade stock price data in the light of the series of works initiated by Christie and Schultz [Why do NASDAQ market makers avoid odd-eighth quotes?, Finance 49 (1994) 1813–1840]. Through the new model, we propose more reasonable, but computationally intensive measures for trading noise including clustering noise and non-clustering noise, and for trading cost. We employ Bayes estimation via filtering to obtain parameter estimates of the new model and to provide numerical measures of trading noise and trading cost for three stocks from four chosen periods. Our empirical results support the important findings in [Christie, Harris, Schultz, Why did NASDAQ market makers stop avoiding odd-eighth quotes?, Finance 49 (1994) 1841–1860; Barclay, Christie, Harris, Kandel, Schultz, The effects of market reform on the trading costs and depths of NASDAQ stocks, J. Finance 54(1) (1999) 1–34].
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Filtering; Counting process; Bayesian inference; Price clustering; High-frequency data; Trading cost

* Corresponding author. Tel.: +1 816 235 5850.
E-mail address: zeng@mendota.umkc.edu (Y. Zeng).

[1] Yong acknowledged the partial financial support from the University of Missouri Research Broad (2004).

0362-546X/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.na.2005.06.051

1. Introduction

Since transactions data (or trade-by-trade data) became available in late 1980s, many micromovement models, describing the transactional price behaviors, have been proposed and developed to answer questions concerning the assessment of market quality, market regulations, and the understanding of market behavior and market microstructure. These data are referred to as *ultra-high-frequency* (UHF) data by Engle [4]. Other recent developments include Goodhart and O'Hara [9], Engle and Russell [6], Zhang et al. [16], Engle and Lunde [5], Rydberg and Shephard [12], among others.

Zeng [14] developed a general class of partially observed models using counting processes for the micromovement of asset price. This class of models is a hybrid system of continuous state space and discrete state space, which is able to capture the difference between the micromovement (in a discrete state space) and the macromovement (referring to the daily, weekly or monthly price behaviors, and in a continuous state space) caused by trading noise, and sustain the strong relationship between the micro- and macromovements. In the models, the price and the intrinsic value of an asset are separated. The value process is supposed not to be observed directly, but can be observed through the prices, whose trading times follow a conditional Poisson process. Due to price discreteness, the prices are formulated as a collection of counting processes, each of which counts the number of trade at a price level. The model can then be framed as a filtering problem with counting process observations. The advantage of this framework is that it is not only able to fit the real world situation, but also able to contain the complete information of prices and trading times and to use them in Bayesian parameter estimation via filtering also developed in Zeng [14]. Transactions data are unevenly spaced in time, discrete in value and huge in size. Bayes estimation via filtering illustrates a substantial improvement in real-time estimation of parameters and value process.

Good measures of trading noise and cost are important, because they quantify the degree of competition and the market efficiency, and hence can be used to assess market quality. Many researches have been done on measuring trading noise. The models commonly used are time series models (see [8], and the references therein). However, due to the limitation of time series models, they ignore important features of transaction data such as the irregularity in time, the price discreteness, price clustering and other micromovement characteristics. Christie and Schultz [3] propose a simple measure for trading cost and their works have important impact (see below). This paper demonstrates one important application of the models and the Bayes estimation via filtering in [14] by providing more sophisticated measures and actual estimates for both trading noise and trading cost.

Christie and Schultz [3] questioned the competitiveness of the NASDAQ market, based on their observation that NASDAQ dealers avoided odd-eighth quotes in 70 of the 100 largest NASDAQ stocks and further comparison of the trading noise and trading costs of NASDAQs dealer markets and those of NYSE market. They suggested that NASDAQ dealers had tacitly colluded to sustain wide ask-bid spread, leading to unnecessarily large trading noise and trading cost for investors. Subsequentially, Christie et al. [2] described an interesting event: after several national newspapers reported the discovery of Christie and Schultz [3] on May 26 and 27, 1994, four dealers out of 10 most active NASDAQ stocks, dramatically increased their use of odd-eighth quotes on May 27 and the next trading day, May 31, 1994, resulting in nearly 50% drop in trading noise and cost.

Furthermore, the results of Christie and Schultz [3] led to regulatory investigations, legal activities, and numerous academic studies. This culminated with the Securities and Exchange Commission (SEC) imposed a series of market reforms in NASDAQ. First, market-makers consented to stop the agreement of escaping odd-eighth quotes. Second, the SEC established new rules governing trading on NASDAQ. These reforms provide investors more competitive quotes via the compulsory exposition of customer limit orders and the dissemination of superior prices through a proprietary trading system. The new trading environment was established gradually starting on January 20, 1997 for the first group of 50 stocks, then for a second group of 50 stocks on February 10, 1997. It was completed for all the NASDAQ stocks by October 13, 1997. Barclay et al. [1] surveyed the outcomes of these market reforms by calculating the changes in several elementary quantitative measures on trading noise, trading cost and others. Their results validated the SECs many goals had been achieved. The trading noise had become overall smaller, resulting in the trading cost measured by the effective spreads further dropped by approximate 30 per cent.

The data analyzed in these papers have a large amount of trading noise caused by trading mechanism including discrete noise (trade at multiples of $\$\frac{1}{16}$), clustering noise (for example, more prices at quarters than at odd-eighths), and non-clustering noise which includes all other trading noise. In this paper, in the spirit of Zeng [14], we construct a new simple micromovement model (we call this model NSMM) to analyze the UHF data in the works of Christie and Schultz. The NSMM not only explicitly models the discrete noise, clustering noise and none clustering noise, but also provides natural, but more sophisticated measures for different types of trading noise and for trading costs. We illustrate the usefulness of the NSMM and its Bayes estimation via filtering by examining the UHF data for three stocks over four periods. Two stocks, AMGEN and APPLE (Computer) are from NASDAQ and they were among the four stocks identified by Christie et al. [2] that suddenly adopted the use of odd-eighth quotes on May 27 or May 31, 1994. The first data set chosen is one month UHF data (usually, 22 business days) up to the last day avoiding odd-eighth quotes. The second data set is one month UHF data starting from the first day adopting odd-eighth quotes (May 27, 1994 for AMGEN, and May 31, 1994 for APPLE). The third period chosen is of the month of June, 1995, which is one year after the publicity of the finding of Christie and Schultz [3], but is before the market reforms in 1997. The fourth period is of the month of June, 1999, which is about two years after the market reforms were completed. Our empirical studies confirm the dramatic drop in trading noise and in trading cost for AMGEN and APPLE from the first period to the second period. The trading noise and trading cost in the third period were about the same as those in the second period, but there was further significant drop in the four period after the market reforms. The third stock is IBM from New York Security Exchange (NYSE), a continuous double-auction market. The results of IBM are consistent over the four periods showing the stability and maturity of NYSE.

The rest of the paper proceeds as follows. Section 2 highlights the general class of micromovement models, presents the NSMM and defines the new measures for trading noises and cost. Section 3 briefly describes the related Bayes estimation via filtering procedure and reports the empirical results for the three stocks over the four periods of time. Section 4 concludes.

2. The micromovement model

2.1. The general class of micromovement models

The underlying economic intuition is that the price is constructed from an intrinsic value process by incorporating trading noises. The intrinsic value process, X, has a continuous state space. X can only be partially observed through the price process, Y, at unevenly spaced trading times modeled by a conditional Poisson process. Y has a discrete state space given by the multiples of the minimum price variation (set by trading regulation), *a tick*, which is assumed to be $1/M$. The couple, (X, Y), furnishes a reasonable hybrid system for the trade-by-trade price process.

Suppose that θ is a vector of parameters in the model and (θ, X, Y) is defined in a complete filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{\geqslant 0}, P)$ (see for example, [11]), where P is the physical probability measure. We first make a mild assumption on (θ, X). The martingale problem and the generator (or semigroup operator) approach [7] offer a general and useful methodology for the characterization of Markov processes. Examples include diffusion processes, jump-diffusion processes, and regime-switching jump-diffusion processes.

Assumption 2.1. (θ, X) is the unique solution of a martingale problem for a generator $\mathbf{A}$ such that for every f in the domain of $\mathbf{A}$

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

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

We have two equivalent ways to build the model. The first produces Y from X by incorporating noises. The second put (X, Y) into a filtering problem with counting process observations. The first approach is economically sound and the second is useful for statistical inference.

2.1.1. Construction of price from value

After specifying the value process $X(t)$ by its generator, there are two more steps in constructing Y from X. First, the distribution of trading times $t_1, t_2, \ldots, t_i, \ldots$ is assumed to follow a conditional Poisson process with an intensity $a(X(t), \theta(t), t)$. Namely, the current intensity is allowed to depend on the current asset value and the current parameter values. Then, $Y(t_i)$, the price at time t_i, is specified by

$$Y(t_i) = F(X(t_i)),$$

where $y = F(x)$ is a random transformation with the transition probability $p(y|x)$. In Section 2.2, $F(x)$ is constructed to fit the three types of trading noise found in the data of Christie and Schultz [3].

2.1.2. Counting process observations

Due to price discreteness, the prices can be viewed according to price levels, that is, prices are viewed as a collection of counting processes in the following form:

$$\vec{Y}(t) = \begin{pmatrix} N_1(\int_0^t \lambda_1(\theta(s), X(s), s)\,\mathrm{d}s) \\ N_2(\int_0^t \lambda_2(\theta(s), X(s), s)\,\mathrm{d}s) \\ \vdots \\ N_n(\int_0^t \lambda_n(\theta(s), X(s), s)\,\mathrm{d}s) \end{pmatrix}, \tag{1}$$

where $Y_k(t) = N_k(\int_0^t \lambda_k(\theta(s), X(s), s)\,\mathrm{d}s)$ is the counting process indicating the cumulative number of trades that have happened at the kth price level (denoted by y_k) up to time t.

Four mild assumptions are made to ensure that the previous construction is equivalent in probability measure to the counting process observations in Eq. (1). The equivalence guarantees that the statistical inference based on the counting process observations can be exploited to the previous construction.

Assumption 2.2. N_j's are unit Poisson processes under the probability measure P.

Assumption 2.3. $(\theta, X), N_1, N_2, \ldots, N_n$ are independent under P.

Assumption 2.4. For the total trading intensity, $a(\theta, x, t)$, there exist a constant, C, such that $0 < a(\theta, x, t) \leqslant C$ for all θ, x and $t > 0$.

Assumption 2.5. The intensities are of the form: $\lambda_k(\theta, x, t) = a(\theta, x, t)p(y_k|x)$, where $p(y_k|x)$ is the transition probability from x to y_k, the kth price level, as in $F(x)$.

In this framework, $(\theta(t), X(t))$ is the unobserved signal process, and $\vec{Y}(t)$ is the observation process polluted by trading noise, which is modeled by $p(y|x)$. Hence, $(\theta, X, \vec{Y})$ becomes a *filtering problem with counting process observations*.

2.2. The new simple micromovement model

Now, we build a model for analyzing the trading noise and trading cost in the works of Christie and Schultz by following three steps in construction. First, as the simplified model in [14], we let the intrinsic value process be a geometric Brownian motion (GBM), which is a standard model in finance (for the case of linear Brownian motion, see [15]). In stochastic differential equation (SDE) form, $X(t)$ follows:

$$\frac{\mathrm{d}X(t)}{X(t)} = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W(t),$$

where $W(t)$ is a standard Brownian motion, μ is the instantaneous expected return and σ is the instantaneous volatility. Its generator for Assumption 2.1 is

$$\mathbf{A}f(\mu, \sigma, \rho, x) = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2}(\mu, \sigma, \rho, x) + \mu x \frac{\partial f}{\partial x}(\mu, \sigma, \rho, x), \tag{2}$$

where ρ is the parameter for non-clustering noise to be defined and the parameter vector in the model is $\theta = (\mu, \sigma, \rho)$.

We assume a deterministic trading intensity $a(t)$. The trading duration (or waiting time) data support this assumption instead of the elementary one that the trading intensity is constant. Finally, we incorporate the three types of trading noise: discrete, non-clustering and clustering, to obtain the price process. Here we construct new trading noises to fit the data whose tick, the minimum price variation, is $\frac{1}{16}$ instead of $\frac{1}{8}$ as in Zeng [14]. Furthermore, the clustering noise is more flexible and reasonable. For simple notation, at a trading time t_i, set $x = X(t_i)$, $y = Y(t_i)$. We specify the random transformation, $y = F(x)$, in three steps:

Step 1: Add non-clustering noise V; $x' = x + V$, where V is the non-clustering noise of trade i at time t_i. $\{V_i\}$, are assumed to be independent of the value process, and they are i.i.d. with a doubly geometric distribution:

$$P\{V = v\} = \begin{cases} (1-\rho) & \text{if } v = 0, \\ \frac{1}{2}(1-\rho)\rho^{16|v|} & \text{if } v = \pm\frac{1}{16}, \pm\frac{2}{16}, \ldots \end{cases}.$$

Step 2: Integrate discrete noise by rounding off x' to its closest sixteenth, $y' = R[x', \frac{1}{16}] = R[x + V, \frac{1}{16}]$.

Step 3: Integrate clustering noise by biasing y' through a random biasing function $b(\cdot)$. $\{b_i(\cdot)\}$ are assumed independent of $\{y'_i\}$ and serially independent. The data to be analyzed have this clustering phenomenon: integers of prices are most likely, then are halves, odd quarters, odd eighths, and odd sixteenths. Odd quarters have roughly the same probabilities, and similarly for odd eighths and odd sixteenths.

To generate such clustering, a biasing function is constructed based on the following biasing rules in words: if the fractional part of y' is zero, then y stays on y' with probability one. If the fractional part of y' is an half, then y stays on y' with probability $1 - \alpha$, and moves to the closest integers each with probability $\frac{1}{2}\alpha$. If the fractional part of y' is an odd quarter, then y stays on y' with probability $1 - \alpha - \beta$, and moves to the closest integer with probability α or to the closest half with probability β. If the fractional part of y' is an odd eighth, then y stays on y' with probability $1 - \alpha - \beta - \gamma$, and moves to the closest integer with probability α, to the closest half with probability β and the closest odd quarter with probability γ. If the fractional part of y' is an odd sixteenth, then y stays on y' with probability $1 - \alpha - \beta - \gamma - \delta$, and moves to the closest integer with probability α, to the closest half with probability β, to the closest odd quarter with probability γ and to the closest odd eighth with probability δ.

This is a more flexible and reasonable biasing function than the one in [14], because it further allows the halves to move to the closest integers, the odd quarters to the closest halves or integers, and the odd eighths to the closest odd quarters, halves or integers.

To sum up, the price at a trading time t_i is

$$Y(t_i) = b_i\left(R\left[X(t_i) + V_i, \frac{1}{M}\right]\right) = F(X(t_i)).$$

Through these steps, the random transformation, $F(x)$, modeling the impact of trading noise, is constructed. Appendix A provides the details on $b(\cdot)$, and the explicit transition probability $p(y|x)$ of F.

2.2.1. New measures for trading noises and trading cost

Christies and Schultz focused on the clustering of ask and bid quotes and ignored the trading noise in price. In our NSMM, the non-clustering noise is evaluated by the parameter, $\rho \in [0, 1)$. When there is no non-clustering noise, $\rho = 0$. The larger ρ, the more non-clustering noise exists. The clustering noise is measured by the vector: $(\alpha, \beta, \gamma, \delta)$, which are the probabilities moving to the closest integers, halves, odd quarters, and odd eighths, respectively. When there is no clustering noise, all are zeros. However, because the impact of each probability is different, we cannot just choose the usual vector norm as a summary measure for the clustering noise. We need to weight each probability according to its impact for the trading noise. Clearly, α, the probability moving to the closest integers, has the heaviest impact, because it moves the price furthest from the value, causing the largest trading noise. Moreover, we observe that the impact of α is about twice of that of β, because α moves the prices about twice the distance from the value as β does. Therefore, we propose the following summary measure, ψ, for the clustering noise:

$$\psi = \tfrac{9}{16}\alpha + \tfrac{1}{4}\beta + \tfrac{1}{8}\gamma + \tfrac{1}{16}\delta.$$

Christie and Schultz [3] define the *effective spread* (*ES*) as a measure of trading cost relative to the "best price" as the following: $ES = 2|\textit{Trade Price} - \text{“}\textit{Best Price}\text{”}|$.

They then proposed the mid-point of the bid-ask quote as a simple estimate for the "Best Price", that is, $\textit{Best Price} = (\textit{Ask} + \textit{Bid})/2$. This estimate, though simple, however, is very rough and not informative, because the ask and bid offer sizes may not be the same. Moreover, in the data they analyzed, there were more than a thousand trade prices in a day, but there were usually only about a hundred quotes within that day (sometimes even only about an half hundred quotes). This implies much useful information was discarded.

In our model, at a trading time t_i, clearly, the intrinsic value, $X(t_i)$, is the best price. Following their ideas, we define our *effective spread* (*ES*) as

$$ES = 2|Y(t_i) - X(t_i)|.$$

The intrinsic value, $X(t_i)$, can be estimated from the transactions data for each trade using Bayes estimation via filtering. This "best price" is more accurate because the model is more sound and much more information is used in obtaining the estimate. Moreover, the average ES for trades with size in some range can be defined as $\sum(ES \times \textit{size})/(\sum \textit{size})$ where the summation goes over all the related trades.

3. The empirical studies

We first highlight the Bayes estimation via filtering and then present the empirical studies for the three stocks over the four periods of time.

3.1. Bayes estimation via filtering

In the modified simple micromovement model, the parameters are (μ, σ) for the value process, ρ for the non-clustering noise and $(\alpha, \beta, \gamma, \delta)$ for the clustering noise. All of them can be estimated by Bayes estimation via filtering. However, in order to reduce the number of parameters in Bayes estimation, the clustering noise parameters is estimated by the method of relative frequency, a variant of method of moments.

3.1.1. Estimating parameters for clustering noise

Let f_0 (f_1, f_2, f_3, and f_4, respectively) be the empirical relative frequency of integer (half, odd quarter, odd eighth, and odd sixteenth, respectively) price. We suppose that the fractional parts of X are uniformly distributed. Then, the method of relative frequency implies $f_0=\frac{1}{16}+\frac{15}{16}\alpha$, $f_1=\frac{1}{16}(1-\alpha)+\frac{7}{8}\beta$, $f_2=\frac{1}{8}(1-\alpha-\beta)+\frac{3}{4}\gamma$, $f_3=\frac{1}{4}(1-\alpha-\beta-\gamma)+\frac{1}{2}\delta$, $f_4=\frac{1}{2}(1-\alpha-\beta-\gamma-\delta)$, with the unique solutions: $\hat{\alpha}=(16f_0-1)/15$, $\hat{\beta}=(16f_1-1+\hat{\alpha})/14$, $\hat{\gamma}=(8f_2-1+\hat{\alpha}+\hat{\beta})/6$, $\hat{\delta}=(4f_3-1+\hat{\alpha}+\hat{\beta}+\hat{\gamma})/2$.

Standard results on multinomial distribution (see two theorems in [13, pp. 109,122]) ensure these estimates are consistent and asymptotically normal. The standard errors can be calculated also. Moreover, they are reliable because of the large amount of data used. Other parameters (μ, σ, ρ) and the value at each trading time, $X(t_i)$, are estimated by the Bayes estimation via filtering. The Bayes estimates for (μ, σ, ρ) are consistent by Theorem 5.1 of Zeng [14].

Bayes estimates, the marginal posterior mean, are determined by the filtering equation characterizing the evolution of the joint posterior. Central to Bayes estimation via filtering is to construct a consistent (or robust) recursive algorithm to compute the posterior.

3.1.2. Filtering equation

Let $\mathscr{F}_t^{\vec{Y}}=\sigma\{(\vec{Y}(s))|0\leqslant s\leqslant t\}$ be the σ-algebra generated by the observed sample path of $\vec{Y}$. $\mathscr{F}_t^{\vec{Y}}$ is all the available information up to time t. Let π_t be the conditional distribution of $(\theta(t), X(t))$ given $\mathscr{F}_t^{\vec{Y}}$. Assuming priors on $(\theta(0), X(0))$, π_t becomes the joint posterior of $(\theta(t), X(t))$. Let $\pi(f,t)=E^P[f(\theta(t), X(t))|\mathscr{F}_t^{\vec{Y}}]=\int f(\theta,x)\pi_t(\mathrm{d}\theta, \mathrm{d}x)$. A simple version of the filtering equation from Zeng [14] is reviewed below.

Theorem 3.1. *Suppose that* (θ, X) *satisfies Assumption* 2.1, *and* $\vec{Y}$ *is defined in Eq.* (1) *with Assumptions* 2.2–2.4 *with a deterministic trading intensity*, *that is*, $a(\theta(t), X(t), t)=a(t)$. *Then, for every* $t>0$ *and every* f *in the domain of generator* $\mathbf{A}$, $\pi(f,t)$ *is the unique solution of the SDE*, *the normalized filtering equation*,

$$\pi(f,t)=\pi(f,0)+\int_0^t \pi(\mathbf{A}f,s)\,\mathrm{d}s+\sum_{k=1}^{n}\int_0^t\left[\frac{\pi(fp_k,s-)}{\pi(p_k,s-)}-\pi(f,s-)\right]\mathrm{d}Y_k(s). \tag{3}$$

3.1.3. A convergence theorem and recursive algorithm

We review a convergence theorem in Zeng [14], which provides a blueprint for constructing a recursive algorithm and guarantees the consistency of the algorithm.

Let $(\theta_\varepsilon, X_\varepsilon)$ be an approximation of (θ, X). Then, we define

$$\vec{Y}_\varepsilon(t) = \begin{pmatrix} N_1(\int_0^t \lambda_1(\theta_\varepsilon(s), X_\varepsilon(s), s)\,\mathrm{d}s) \\ N_2(\int_0^t \lambda_2(\theta_\varepsilon(s), X_\varepsilon(s), s)\,\mathrm{d}s) \\ \vdots \\ N_n(\int_0^t \lambda_n(\theta_\varepsilon(s), X_\varepsilon(s), s)\,\mathrm{d}s) \end{pmatrix} \tag{4}$$

and $\mathscr{F}_t^{\vec{Y}_\varepsilon} = \sigma(\vec{Y}_\varepsilon(s), 0 \leqslant s \leqslant t)$. We use the notation, $X_\varepsilon \Rightarrow X$, to mean X_ε converges weakly to X in the Skorohod topology as $\varepsilon \to 0$.

Theorem 3.2. *Suppose that* $(\theta, X, \vec{Y})$ *is on* $(\Omega, \mathscr{F}, P)$ *with Assumptions* 2.1–2.5, *and* $(\theta_\varepsilon, X_{\varepsilon_x}, \vec{Y}_\varepsilon)$ *is on* $(\Omega_\varepsilon, \mathscr{F}_\varepsilon, P_\varepsilon)$ *with also Assumptions* 2.1–2.5. *If* $(\theta_\varepsilon, X_{\varepsilon_x}) \Rightarrow (\theta, X)$ *as* $\varepsilon \to 0$, *then as* $\varepsilon \to 0$ (i) $\vec{Y}_\varepsilon \Rightarrow \vec{Y}$; *and* (ii) $E^{P_\varepsilon}[F(\theta_\varepsilon(t), X_{\varepsilon_x}(t))|\mathscr{F}_t^{\vec{Y}_\varepsilon}] \Rightarrow E^P[F(\theta(t), X(t))|\mathscr{F}_t^{\vec{Y}}]$ *for all bounded continuous function* F.

Another proof can be found in Kouritzin and Zeng [10]. This theorem furnishes a recipe for producing a recursive algorithm based on the Markov chain approximation method to compute the continuous-time posteriors and Bayes estimates. There are three steps. Step 1 is to build $(\theta_\varepsilon, X_\varepsilon)$, the Markov chain approximation to (θ, X). Step 2 is to obtain the filtering equation for $\pi_\varepsilon(f, t)$ corresponding to $(\theta_\varepsilon, X_\varepsilon, Y_\varepsilon)$ by applying Theorem 3.1. The filtering equation can be divided into two parts: the propagation equation;

$$\pi_\varepsilon(f, t_{i+1}-) = \pi_\varepsilon(f, t_i) + \int_{t_i}^{t_{i+1}-} \pi_\varepsilon(\mathbf{A}_\varepsilon f, s)\,\mathrm{d}s \tag{5}$$

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

$$\pi_\varepsilon(f, t_{i+1}) = \frac{\pi_\varepsilon(f p_{\varepsilon,j}, t_{i+1}-)}{\pi_\varepsilon(p_{\varepsilon,j}, t_{i+1}-)}. \tag{6}$$

Note that the observation gets into the updating equation through the *j*th observed price level. Step 3 transforms Eqs. (5) and (6) to the recursive algorithm in discrete state space and in discrete time by two substeps: (a) expresses $\pi_\varepsilon(\cdot, t)$ as a finite array with elements being $\pi_\varepsilon(f, t)$ for lattice-point indicator f and (b) approaches the time integral in (5) with an Euler scheme.

For detailed construction of the recursive algorithm for the NSMM, the readers are referred to [14, Section 5.2].

3.2. The empirical study

3.2.1. Data

The data for the three stocks: AMGEN, APPLE Computer and IBM for the four periods of time are extracted from the Trade and Quote (TAQ) database distributed by NYSE. We apply standard procedures to filter the data. Each period consists of 22 business days (usually, a month) with two exceptions in IBM. Although June 1995 had 22 business days, the first two days' data of IBM are excluded to ensure no large jump in price, because a dividend

Table 1
Relative frequency table for fractional parts of price for AMGEN

Fractional parts	0	1/16	1/8	3/16	1/4	5/16	3/8	7/16
4/27–5/26, 94	**0.2232**	0.0072	0.0509	0.0110	**0.1878**	0.0053	0.0413	0.0051
5/27–6/29, 94	**0.1471**	0.0170	**0.0962**	0.0100	**0.1103**	0.0135	**0.1195**	0.0144
June, 95	**0.1549**	0.0125	**0.1177**	0.0089	**0.1123**	0.0090	**0.0982**	0.0064
June, 99	**0.1109**	0.0522	0.0718	0.0464	0.0773	0.0422	0.0660	0.0466
Fractional parts	1/2	9/16	5/8	11/16	3/4	13/16	7/8	15/16
4/27–5/26, 94	**0.1654**	0.0059	0.0343	0.0064	**0.1843**	0.0075	0.0570	0.0074
5/27–6/29, 94	**0.1036**	0.0119	**0.0933**	0.0105	**0.1182**	0.0129	**0.1108**	0.0108
June, 95	**0.1045**	0.0113	**0.1117**	0.0135	**0.1069**	0.0081	**0.1135**	0.0105
June, 99	0.0800	0.0441	0.0645	0.0479	0.0772	0.0460	0.0774	0.0494

The relative frequencies higher than 9% are in bold.

Table 2
Relative frequency table for fractional parts of price for APPLE

Fractional parts	0	1/16	1/8	3/16	1/4	5/16	3/8	7/16
4/27–5/26, 94	**0.2424**	0.0083	0.0465	0.0034	**0.1828**	0.0040	0.0416	0.0057
5/27–6/29, 94	**0.1743**	0.0138	**0.1191**	0.0138	**0.1210**	0.0090	**0.1012**	0.0132
June, 95	**0.1747**	0.0137	**0.1093**	0.0096	**0.1046**	0.0118	**0.1078**	0.0117
June, 99	**0.1263**	0.0505	0.0637	0.0365	0.0627	0.0478	0.0579	0.0423
Fractional parts	1/2	9/16	5/8	11/16	3/4	13/16	7/8	15/16
4/27–5/26, 94	**0.1789**	0.0073	0.0430	0.0054	**0.1762**	0.0057	0.0419	0.0068
5/27–6/29, 94	**0.1133**	0.0080	0.0794	0.0089	**0.1014**	0.0126	**0.1022**	0.0089
June, 95	**0.1139**	0.0102	**0.1012**	0.0074	**0.0990**	0.0099	**0.1028**	0.0123
June, 99	0.0819	0.0394	0.0629	0.0474	0.0865	0.0591	0.0768	0.0583

The relative frequencies higher than 9% are in bold.

was paid out at the beginning of the third date for IBM. The other exception is June 1997 of IBM, because that month has only 21 business days. For IBM, June, 1997 instead of June, 1998 or 1999 is chosen for comparison because IBM's tick, the minimum price variation, had changed to ticks smaller than $\frac{1}{16}$ after June 1997.

According to Christie et al. [2], AMGEN started odd eighths quotes on May 27, 1994. APPLE started on the next trading day, May 31, 1994. The four periods for each stock with the number of trading days and the number of trades occurring in each period are presented in the first parts of Tables 4–6 for AMGEN APPLE and IBM, respectively. Note that the number of trades is the number of data used for estimation. After extraction, the data consist of trading time, price, and size for each trade. Tables 1–3 reports the relative frequencies for all fractional parts of price in the four periods for AMGEN, APPLE and IBM. For AMGEN and APPLE, clearly, there are price clustering. The first period has a highest degree of clustering while the last period has the least. The middle two are about the same. IBM, however, has about the same degree of clustering in all the four periods.

Table 3
Relative frequency table for fractional parts of price for IBM

Fractional parts	0	1/16	1/8	3/16	1/4	5/16	3/8	7/16
4/27–5/26, 94	**0.1837**	0.0000	**0.1287**	0.0001	**0.1126**	0.0000	0.0771	0.0001
5/27–6/29, 94	**0.1967**	0.0001	**0.1116**	0.0001	0.0817	0.0000	**0.0905**	0.0001
June, 95	**0.1600**	0.0000	**0.1145**	0.0001	**0.1401**	0.0001	**0.1274**	0.0001
June, 99	**0.1518**	0.0081	**0.1049**	0.0161	**0.1220**	0.0101	**0.1045**	0.0107
Fractional parts	1/2	9/16	5/8	11/16	3/4	13/16	7/8	15/16
4/27–5/26, 94	**0.1046**	0.0002	**0.1170**	0.0001	**0.1323**	0.0000	**0.1431**	0.0001
5/27–6/29, 94	**0.1193**	0.0000	**0.1119**	0.0001	**0.1295**	0.0002	**0.1579**	0.0003
June, 95	**0.1412**	0.0002	**0.1127**	0.0000	**0.1008**	0.0001	**0.1026**	0.0001
June, 99	**0.1278**	0.0100	**0.1023**	0.0115	**0.0981**	0.0121	**0.0955**	0.0147

The relative frequencies higher than 9% are in bold.

Table 4
Parameter estimates for AMGEN in 4 periods

Period	4/27–5/26, 94	5/27–6/29, 94	June, 95	June, 99
No. of trades	16916	16242	23136	120829
μ	105.25%	−29.37%	120.58%	−43.47%
	(101.11%)	(98.44%)	(79.94%)	(145.41%)
σ	31.92%	29.60%	23.31%	41.32%
	(0.64%)	(0.50%)	(0.33%)	(0.26%)
ρ	0.1873	0.1690	0.1862	0.1527
	(0.0058)	(0.0045)	(0.0039)	(0.0012)
α: 1	0.1714(0.0030)	0.0902(0.0026)	0.0986(0.0022)	0.0516(0.0008)
β: 1/2	0.1298(0.0025)	0.0534(0.0021)	0.0550(0.0018)	0.0237(0.0007)
γ: 1/4	0.3797(0.0028)	0.1619(0.0025)	0.1513(0.0020)	0.0520(0.0008)
δ: 1/8	0.2077(0.0060)	0.4922(0.0077)	0.5348(0.0065)	0.1231(0.0025)
Overall: ψ	0.1893	0.1151	0.1216	0.0491
Size (shares)		Average	Effective	Spread
100	34.62	9.37	10.08	5.01
200	33.70	8.57	10.15	4.76
300	33.36	8.63	8.96	4.56
400	33.13	8.43	8.32	4.44
500	34.23	8.32	7.65	4.56
501–1000	33.25	6.94	8.20	4.26
1001–5000	28.01	7.58	8.51	5.64
5001–10 000	26.48	10.41	12.29	9.91
> 10 000	28.96	12.67	14.20	13.45
Total average	29.45	9.20	9.90	7.61

The estimates for μ and σ are annualized. Standard errors are in parentheses. The overall clustering measure: $\psi = \frac{9}{16}\alpha + \frac{1}{4}\beta + \frac{1}{8}\gamma + \frac{1}{16}\delta$. Average effective spreads are in cents.

Table 5
Parameter estimates for APPLE in 4 periods

Period	4/28–5/27, 94	5/31–6/29, 94	June, 95	June, 99
No. of trades	20074	27967	40529	57263
μ	−29.18%	−130.93%	121.56%	42.45%
	(160.62%)	(157.33%)	(119.67%)	(140.54%)
σ	46.77%	45.48%	35.37%	35.54%
	(0.86%)	(0.72%)	(0.49%)	(0.36%)
ρ	0.1605	0.1553	0.2317	0.1087
	(0.0053)	(0.0036)	(0.0030)	(0.0017)
α: 1	0.1919(0.0028)	0.1193(0.0021)	0.1197(0.0018)	0.0680(0.0013)
β: 1/2	0.1467(0.0024)	0.0666(0.0017)	0.0673(0.0014)	0.0271(0.0010)
γ: 1/4	0.3684(0.0025)	0.1607(0.0019)	0.1360(0.0015)	0.0481(0.0011)
δ: 1/8	0.1997(0.0053)	0.4771(0.0059)	0.5035(0.0049)	0.0942(0.0037)
Overall: ψ	0.2032	0.1337	0.1326	0.0569
Size (shares)		Average	Effective	Spread
100	20.34	9.58	10.49	4.18
200	18.91	9.07	10.61	4.30
300	18.42	9.07	9.93	3.95
400	17.42	9.28	8.86	3.85
500	12.73	7.87	9.12	4.09
501–1000	11.34	6.90	8.75	4.10
1001–5000	11.07	7.49	8.98	4.55
5001–10 000	14.65	8.79	11.88	6.86
> 10 000	17.67	11.07	11.86	8.62
Total average	13.79	8.63	10.05	5.72

The estimates for μ and σ are annualized. Standard errors are in parentheses. The overall clustering measure: $\psi = \frac{9}{16}\alpha + \frac{1}{4}\beta + \frac{1}{8}\gamma + \frac{1}{16}\delta$. Average effective spreads are in cents.

3.2.2. *Empirical results*

A Fortran program for the recursive algorithm is constructed to calculate, at each trading time t_i, the joint posterior of (μ, σ, ρ, X), their marginal posteriors, their Bayes estimates and their standard errors, and the trading effective spread, respectively. Then, the effective spreads are cumulated according to trading size (volume). Finally, we obtain the average effective spread over trading size and the overall average effective spread. The Fortran program runs much faster than the real-time computation required and they are tested extensively by simulated data as in Zeng [14] before applying to actual UHF data.

Tables 4–6 present the Bayes estimates for (μ, σ, ρ), the estimates for the clustering parameters $(\alpha, \beta, \gamma, \delta)$ with summary measure ψ, and the average effective spread across size for the four periods for AMGEN, APPLE and IBM, respectively.

For the first two periods of AMGEN and APPLE, the estimates for σ and ρ are close. The estimates for μ are quite different, but because their standard deviations (or errors) (SEs) are large, their difference is not significant. The estimates for $(\alpha, \beta, \gamma, \delta)$, however, are significantly different. For AMGEN, α, β and γ, the probabilities to move to integers, halves, and odd quarters, respectively, dropped significantly, ranged from 47.4–58.9%. δ, the

Table 6
Parameter estimates for IBM in 4 periods

Period	4/29–5/31, 94	6/3–6/30, 94	June, 95	June, 97
No. of trades	27745	23940	35471	53021
μ	118.27%	−83.23%	60.90%	43.72%
	(79.40%)	(70.39%)	(70.09%)	(83.78%)
σ	21.22%	20.53%	19.18%	23.52%
	(0.32%)	(0.31%)	(0.22%)	(0.21%)
ρ	0.0911	0.0843	0.1275	0.1233
	(0.0034)	(0.0036)	(0.0030)	(0.0024)
α: 1	0.1293(0.0022)	0.1431(0.0024)	0.1040(0.0018)	0.0952(0.0015)
β: 1/2	0.0574(0.0016)	0.0751(0.0018)	0.0974(0.0016)	0.0814(0.0013)
γ: 1/4	0.1911(0.0019)	0.1514(0.0020)	0.1885(0.0017)	0.1563(0.0013)
δ: 1/8	0.6207(0.0059)	0.6286(0.0065)	0.6090(0.0053)	0.4808(0.0043)
Overall: ψ	0.1498	0.1575	0.1445	0.1235
Size (shares)	Average	Effective	Spread	
100	7.54	7.49	9.02	7.22
200	7.51	7.22	8.90	7.25
300	7.40	7.40	8.43	6.93
400	6.81	6.80	8.53	7.13
500	7.14	7.00	8.63	7.01
501–1000	7.08	6.82	7.94	6.85
1001–5000	7.56	7.57	8.04	7.60
5001–10 000	8.76	8.70	9.65	8.87
> 10 000	13.12	13.26	14.78	14.21
Total average	9.99	9.84	10.64	9.76

The estimates for μ and σ are annualized. Standard errors are in parentheses. The overall clustering measure: $\psi = \frac{9}{16}\alpha + \frac{1}{4}\beta + \frac{1}{8}\gamma + \frac{1}{16}\delta$. Average effective spreads are in cents.

probability to move to odd eighth, however, rose to more than double. Similar phenomenon happened for APPLE. Overall, the summary measure for clustering noise, ψ dropped 39.20% for AMGEN and 34.20% for APPLE in the second period. Next, we compare the *ES* for these two periods. For AMGEN, we find that the average *ES* decreased 56.25 to 79.12% across trading size and the overall average decreased 68.76%; and for APPLE, the average *ES* decreased 37.35 to 52.90% across trading size and the overall average decreased 37.42%. All of these support the findings of Christie et al. [2] that after odd eighth quotes were adopted in the second period, although the basic characteristics of stocks such as μ and σ were not affected, more prices, however, occurred at odd eighths. Consequently, the trading noise and trading cost measured by ES reduced substantially.

The clustering estimates for the third period are close to those of the second period for AMGEN and APPLE. Comparing the third and fourth periods, we observe substantial drops in every parameter estimate of clustering noise and the summary measure, ψ, dropped more than 50% for both stocks. For the trades with size not more than 5000 shares, the effective spread decreased 33.73 to 53.10% for AMGEN and decreased 49.35 to 60.15% for APPLE. All of these confirm the findings of Barclay et al. [1] that after the market reforms on

NASDAQ, there were further significant reductions in trading noise and trading cost and NASDAQ became a more efficient market.

As for IBM, although the estimates and the effective spreads vary a bit in different periods, they are close and the effective spreads are low. This shows the NYSE auction market is more mature and stable over this four periods.

4. Conclusions

In this paper, a hybrid system—the general partially observed micromovement model and its Bayes estimation via filtering are reviewed. A new model is constructed for transaction data with tick $\$\frac{1}{16}$ and with more flexible clustering noise. Through the new model, new measures for trading noise and trading cost are proposed. Bayes estimation via filtering is then applied to three stocks from four chosen periods to obtain empirical measures of trading noise and cost. Our results confirm the major findings in [2,1].

Appendix A. More on modeling clustering noise

In order to formulate the biasing rule, we first define a classifying function $r(\cdot)$,

$$r(y) = \begin{cases} 4 & \text{if the fractional part of } y \text{ is an odd sixteenth,} \\ 3 & \text{if the fractional part of } y \text{ is an odd eighth,} \\ 2 & \text{if the fractional part of } y \text{ is an odd quarter,} \\ 1 & \text{if the fractional part of } y \text{ is a half,} \\ 0 & \text{if the fractional part of } y \text{ is a zero.} \end{cases} \tag{7}$$

The biasing rules determine the transition probabilities from y' to y, $p(y|y')$. Then, $p(y|x)$, the transition probability for the random transformation, F, can be computed through $p(y|x) = \sum_{x'} p(y|y')p(y'|x)$ where $p(y'|x) = P\{V = 16(y' - R[x, \frac{1}{16}])\}$. Suppose $D = 16|y - R[x, \frac{1}{16}]|$. Then, $p(y|x)$ can be calculated as the following when $r(y) = 4$ or 3. Similarly, the more complicated cases when $r(y) = 2$, 1 or 0 can be computed

$$p(y|x) = \begin{cases} (1-\alpha-\beta-\gamma-\delta)(1-\rho) & \text{if } r(y)=4 \text{ and } D=0, \\ \frac{1}{2}(1-\alpha-\beta-\gamma-\delta)(1-\rho)\rho^D & \text{if } r(y)=4 \text{ and } D\geqslant 1, \\ (1-\rho)(1-\alpha-\beta-\gamma+\delta\rho) & \text{if } r(y)=3 \text{ and } D=0, \\ \frac{1}{2}(1-\rho)[\rho(1-\alpha-\beta-\gamma) & \\ \quad +\delta(2+\rho^2)] & \text{if } r(y)=3 \text{ and } D=1, \\ \frac{1}{2}(1-\rho)\rho^{D-1}[\rho(1-\alpha-\beta-\gamma) & \\ \quad +\delta(1+\rho^2)] & \text{if } r(y)=3 \text{ and } D\geqslant 2. \end{cases} \tag{8}$$

References

[1] M. Barclay, W. Christie, J. Harris, E. Kandel, P.H. Schultz, The effects of market reform on the trading costs and depths of NASDAQ stocks, J. Finance 54 (1) (1999) 1–34.

[2] W. Christie, J.H. Harris, P.H. Schultz, Why did NASDAQ market makers stop avoiding odd-eighth quotes?, J. Finance 49 (1994) 1841–1860.

[3] W. Christie, P.H. Schultz, Why do NASDAQ market makers avoid odd-eighth quotes?, J. Finance 49 (1994) 1813–1840.
[4] R. Engle, The econometrics of ultra-high-frequency data, Econometrica 68 (2000) 1–22.
[5] R. Engle, A. Lunde, Trades and quotes: a bivariate point process, J. Financial Econometrics 1 (2003) 159–188.
[6] R. Engle, J. Russell, Autoregressive conditional duration: a new model for irregularly spaced transaction data, Econometrica 66 (1998) 1127–1162.
[7] S. Ethier, T. Kurtz, Markov Processes: Characterization and Convergence, Wiley, New York, 1986.
[8] T.J. George, C.Y. Hwang, Information flow and pricing errors: a unified approach to estimation and testing, Rev. Financial Stud. 14 (2001) 979–1020.
[9] C. Goodhart, M. O'Hara, High frequency data in financial markets: issues and applications, J. Empirical Finance 4 (1997) 73–114.
[10] M. Kouritzin, Y. Zeng, Weak convergence for a type of conditional expectation: application to the inference for a class of asset price models, Nonlinear Anal. Ser. A: Theory Methods 60 (2005) 231–239.
[11] P. Protter, Stochastic Integration and Differential Equations, Springer, New York, 1990.
[12] T.H. Rydberg, N. Shephard, Dynamics of trade-by-trade price movements: decomposition and models, J. Financial Econometrics 1 (2003) 2–25.
[13] R.J. Serfling, Approximation Theorems of Mathematical Statistics, Wiley, New York, 1980.
[14] Y. Zeng, A partially-observed model for micro-movement of asset prices with Bayes estimation via filtering, Math. Finance 13 (2003) 411–444.
[15] Y. Zeng, Bayes estimation via filtering for a simple micro-movement model of asset price with discrete noises, Nonlinear Analysis Series A: Theory Methods, 2005, in press.
[16] M.Y. Zhang, J.R. Russell, R.S. Tsay, A nonlinear autoregressive conditional duration model with applications to financial transaction data, J. Econometrics 104 (2001) 179–207.