NEW YORK UNIVERSITY
STERN SCHOOL OF BUSINESS
FINANCE DEPARTMENT

Working Paper Series, 1996

The Dynamics of Discrete Bid and Ask Quotes

Hasbrouck, Joel

FIN-96-26

The Dynamics of Discrete Bid and Ask Quotes

Joel Hasbrouck

First Draft: February 3, 1996
Current draft: December 2, 1996

Preliminary Draft
Not for Quotation
Comments Welcome

Professor of Finance
Stern School of Business
New York University
Suite 9-190
44 West Fourth St.
New York, NY 10012-1126

Tel: (212) 998-0310
Fax: (212) 995–4901
E-mail: jhasbrou@stern.nyu.edu
Web: http://www.stern.nyu.edu/~jhasbrou

For comments on an earlier draft, I am indebted to James Angel, Darrell Duffie, Mark Rubinstein, Matt Spiegel, and seminar participants at Cornell University, Georgetown University, New York University, Stanford University and the University of California at Berkeley. All errors are my own responsibility.

Copyright 1996, Joel Hasbrouck. All rights reserved.

The Dynamics of Discrete Bid and Ask Quotes

Abstract

This paper describes a general approach to the estimation of security price dynamics when the phenomena of interest are of the same scale or smaller than the tick size. The model views discrete bid and ask quotes as arising from three continuous random variables: the efficient price of the security, a cost of quote exposure (information and processing costs) on the bid side and a similar cost of quote exposure on the ask side. The bid quote is the efficient price less the bid cost rounded *down* to the next tick; the ask quote is the efficient price plus the ask cost rounded *up* to the next tick. To deal with situations in which the cost of quote exposure possesses both stochastic and deterministic components and the increments of the efficient price are nonstationary, the paper employs a nonlinear state-space estimation method. The method is applied to intraday quotes at fifteen-minute intervals for Alcoa (a randomly chosen Dow stock). The results confirm the existence of persistent intraday volatility. More importantly they establish the existence of a persistent stochastic component of quote exposure costs that is large relative to the deterministic intraday "U" component.

1. Introduction

Although most determinants of a security price are likely to be continuous variables, institutional arrangements generally constrain the security to a discrete grid. This grid may be coarse relative to the price variation over brief intervals, and also relative to the economic costs of order submission and execution. This paper suggests that the bid and ask quotes arise from an implicit efficient price and quote-exposure costs, all of which are continuous random variables. The discrete bid quote is the implicit efficient price less the continuous bid exposure cost rounded down to the next tick; the discrete ask is the efficient price plus the continuous ask exposure cost rounded up. The paper proposes a nonlinear state-space procedure for estimation and (in real time applications) online filtering, and applies this model to fifteen-minute bid and ask quotes for a New York Stock Exchange stock.

This paper is most closely related to earlier empirical studies of discreteness in stock prices. The first analyses in this area focused on estimation of long-term stock return variances from transaction prices, a concern motivated by option pricing applications (see Ball (1988), Cho and Frees (1988), Gottlieb and Kalay (1985), and Marsh and Rosenfeld (1986)). The emphasis in later studies of transaction prices shifted to microstructure phenomena (see Angel (1994), Dravid (1991), Glosten and Harris (1988), Harris (1990, 1991, 1994), Hausman, Lo and MacKinlay (1992), and Madhavan, Richardson and Roomans (1994)). Discreteness is often encountered as a "nuisance" effect, a data characteristic that must be addressed en route to confronting more interesting economic hypotheses. But since the minimum tick size may affect trading activity, discreteness is also of theoretical and policy interest (Ahn, Cao and Choe (1996), Anshuman and Kalay (1994), Bernhardt and Hughson (1996), Brown, Laux and Schacter (1991), Chordia and Subrahmanyam (1995), Cordella and Foucault (1996), Glosten (1994), Harris (1990, 1991)).

The present study seeks to model bid and ask quotes as opposed to transaction prices. Quotes are of particular interest in microstructure studies because they can be updated in the absence of trades to reflect changing information and also because they reflect perceived asymmetric information costs. Discreteness has different effects on quotes and transaction prices. A risk-neutral trader would presumably be indifferent to the fair-game perturbation associated with symmetrically rounding the unobserved continuous price to the nearest tick. A market maker posting bid and offer quotes, on the other hand, must round his bid price down and his offer price up in order avoid the expectation of losing money on the next trade. For this asymmetric rounding Dravid (1991) computes the first two moments of the transaction price changes under the assumptions of stationarity and a constant quote exposure cost. The present approach allows for nonstationarity and stochastic quote exposure cost.

The paper is also closely related to studies of discreteness in the bid-ask spread. Harris (1994) and Bollerslev and Melvin (1994) model the discrete spread using ordered qualitative-data approaches. In these studies, the spread is a continuous function of observable variables and a random disturbance that is transformed onto a discrete grid. The spread in the present model, in contrast, is driven by a underlying variables that are continuous, unobservable, stochastic and autocorrelated. Furthermore, in modeling the bid and ask separately, the present analysis incorporates a rich specification of the efficient price dynamics.

The model possesses a state-space representation in which a continuous unobserved efficient price and quote exposure costs are the state variables and the discrete bid and ask prices are the observations. This framework is appealing because the recursive procedure used to compute the likelihood function is a Bayesian updating that mimics agents' inferences. Furthermore, state-space models are natural and convenient tools with which to investigate deterministic and stochastic time variation in parameters. The paper implements several such generalizations.

In comparison with reduced-from vector autoregressive (VAR) microstructure models (e.g., Hasbrouck (1991a, 1991b, 1993), the present design assumes more structure in the form of the probability densities and the discrete-valued functions that map the continuous state variables onto prices. These assumptions suffice to identify a (nonlinear) state-space model. Although tightly structured in its discreteness aspects, this model remains general in other regards, notably those related to time variation in the parameters. The conventional approach to characterizing intraday parameter variation involves estimating fixed-parameter models over intraday subsamples (e.g., as in Hasbrouck (1991), the first hour of trading). In contrast, the present approach admits stochastic and deterministic parameter variation in a comprehensive statistical model.

Thus, although the paper deals primarily with discreteness, the ultimate aim in this line of inquiry is a modeling framework flexible enough to accommodate parameter shifts resulting from the start and finish of trading and random variation in the underlying informational and liquidity determinants of trading activity.

The analysis does not extend to clustering (the affinity of transaction prices and quotes for integers, halves, quarters, etc., in decreasing frequency). Clustering in transaction prices is examined by Niederhoffer (1965, 1966) and Harris (1991), and in quotes by Christie and Shultz (1995a, 1995b). In a dynamic setting, clustering requires specification of a stochastic mapping from continuous state variables to discrete observations that is more complicated than the simple rounding functions employed here.

The paper is organized as follows. The next section describes the underlying economic model that generates the bid and ask quotes. The paper then turns to the problem of inference: how to estimate the underlying model from the observed discrete bid and ask prices. Section 3 discusses the restrictions imposed on the underlying variables by the discrete observations. Section 4 introduces the nonlinear filtering algorithm, the associated maximum likelihood procedure and computational techniques. The full dynamic model, which incorporates stochastic and deterministic time variation in the cost and efficient price volatility, is presented in Section 5. The model is estimated for a representative NYSE stock in Section 6. Section 7 discusses the role of discreteness in microstructure analyses and the costs and benefits of the proposed technique. A brief summary concludes the paper in section 8.

2. The economic model

Denote by m the implicit efficient price of the security (the expectation of the security's terminal value, conditional on all public information). The agent establishing the bid quote is assumed to be subject to a nonnegative cost of quote exposure $\beta \geq 0$ for small trades, such that in the absence of discreteness restrictions she would quote a bid price of $m-\beta$. This cost is assumed to impound fixed transaction costs and asymmetric information costs. With a one-unit tick size in the market, she is assumed to quote a bid price of $b = \text{Floor}(m-\beta)$, where $\text{Floor}(\cdot)$ rounds its argument down to the next whole integer. Similarly, the agent establishing the ask quote is assumed to be driven by a quote exposure cost $\alpha \geq 0$ (also for small trades), such that in the absence of discreteness restrictions he would quote an ask price of $m+\alpha$. Constrained by discreteness, he offers at an ask quote of $a = \text{Ceiling}(m+\alpha)$, where $\text{Ceiling}(\cdot)$ rounds its argument up to the next whole integer. In summary, the bid and ask prices are given by

$$\begin{aligned} b &= \text{Floor}(m-\beta) \\ a &= \text{Ceiling}(m+\alpha) \end{aligned} \tag{1}$$

If the tick size is not unity, all variables may simply be rescaled.

This construct can be motivated by most simple models of dealer behavior. In the framework of Glosten and Milgrom (1985) quote setters face a population of informed and uninformed traders. m is the expectation of the final value of the security conditional on all public information (including the transaction price history). The quote exposure costs are defined implicitly by the conditions that $m-\beta$ and $m+\alpha$ ensure the quote-setter(s) zero expected profits and no ex post regret, an outcome supported by Bertrand competition.

By asymmetrically rounding up on the ask and down on the bid, the market maker avoids the possibility of loss on the incoming trade. If the rounding were symmetric (all prices rounded up, all prices rounded down or all prices rounded to the nearest integer), then one or both sides of the quotes might be associated with an expected loss. For example, if the efficient price is 5 and the cost is 1.1, nearest-integer rounding yields a bid of 4 and an ask of 6, both of which yield expected losses. Furthermore, symmetric rounding may imply degenerate quotes (identical bid and ask prices) if α and β are small.

Due to the asymmetric rounding, a Glosten-Milgrom dealer will achieve a profit (both ex ante and ex post) on each trade. These profits need not lead to competitive price cutting because the discreteness restriction ensures that any such action, if feasible, will result in a loss. Nor need these profits lead to a surge of new entrants. Even markets (such as the NYSE) that allow nondealers to enter limit orders usually enforce local time priority. The probability of execution and therefore the incentives for limit order placement diminish with the length of the queue.

More generally, β is the quote setter's marginal cost on the bid side of the market at a particular time. From an economic perspective, it is useful to recognize that some of the components of this cost may be negative, *as long as the total* β *is nonnegative*. An example of this arises in the context of inventory control. Suppose that the cost of clearing a trade is 0.5 (ticks). A dealer who is short (relative to her desired holdings) might nevertheless bid as if β=0.2, reflecting a greater propensity to accumulate a position. There is an implicit benefit of accumulation that may be viewed as a negative cost of -0.3. If the same dealer were also offering the security, we might also expect her α to be high relative to the clearing cost, reflecting her reluctance to accommodate further sales. (Similar remarks apply, of course, to the ask exposure cost.)

Negative cost components may also arise in the case of quotes established by public limit order traders. Their principal alternative to a limit order is a market order. They are not seeking to realize a dealer's profit on average, but merely to reduce their costs of trading (Harris and Hasbrouck (1996) and Harris (1994)). The quote exposure cost may also impound the quote setter's private information.

From a modeling viewpoint, nonnegative α and β serve to prevent the bid and ask implied by (1) from coinciding or crossing. In general practice, the bid and ask quotes prevailing at a point in time reflect entered orders that have been subjected to a matching procedure according to the rules of the market. The assumptions of a common *m* and nonnegativity of α and β are expedients that avoid the necessity of explicitly modeling this matching process.

Most interesting applications will involve situations where the quote exposure costs are random. This randomness can be viewed as arising from several sources. Along the lines of the Glosten-Milgrom model, there may be random time variation in the determinants of this cost, such as perceived exposure to adverse information or holding costs. In this view all dealers and potential dealers are subject to the same cost. Alternatively, we may view the quote setter as an agent drawn from a population of traders with random cost functions. If more than one such agent is active at an instant, then the relevant costs are the minimum α's and β's in the set.

Although this model allows for randomness in α, β and m, the discreteness aspect of the model arises from a nonstochastic transformation. There is no discreteness "error" or disturbance that is required to impound the effect of discreteness.

As noted in the introduction, most studies of discreteness in security markets have focused on transaction prices. Quotes and transaction prices are obviously related, however, and the transaction price models therefore offer useful points of comparison. In this connection, there is at the outset one obvious incompatibility. If transactions arise as uncorrelated equiprobable realizations of the bid and ask quotes determined by equation (1), these prices cannot be described as a symmetrically-rounded random-walk. In the present model discreteness is imposed at the point at which the quotes are set, not the subsequent point at which trade occurs.

There is no assumption that all trades take place at the posted quotes. Following Rock (1996), the posted quotes modeled here are viewed as the best available prices absent knowledge of the full size of the incoming order. A trader (such as a specialist or floor trader) who can bid or offer conditional on the incoming order size may better the posted quotes. In practice, such agents bid or offer after the order has been received, and these implicit quotes do not prevail after the transaction has occurred.

The models actually estimated in the paper allow the quote exposure costs to exhibit both deterministic and stochastic dynamic behavior. It is useful to point out, however, that even when these costs are equal and constant, random variation in m suffices to induce randomness in the spread. For example, if $\alpha = \beta = 1/4$, then the spread is one tick as long as the fractional part of m is between 1/4 and 3/4; and the spread is two ticks otherwise. Therefore, variability in the discrete spread may be an erroneous proxy for variability in the spread's continuous determinants. In addition, price transitions will sometimes be marked by quotes that appear to move "one leg at a time". (Consider the quotes associated with assuming $\alpha = \beta = 1/4$ and the m_t sequence {0.4, 0.9, 1.3}.) U.S. stock quotes often exhibit this behavior.

In the present model the quote setter's solution to an implicit continuous optimization problem (α or β) is subjected to a transformation to yield discrete quote placement strategies. This must be viewed as an approximation to a decision process in which discreteness is more fundamentally incorporated into the calculation, i.e., an integer programming problem. Models along these lines include Anshuman and Kalay (1994), Glosten (1994), Chordia and Subrahmanyam (1995), Bernhardt and Hughson (1996) and Cordella and Foucault (1996). These models are stylized in numerous respects (typically allowing a restricted set of traders and permissible interactions) and focus almost exclusively on information costs. In these characterizations, a continuous "pre-rounding" cost constructs (such as the present α and β) do not explicitly arise. It could nevertheless be argued that such quantities exist implicitly, and that they impound the costs of quote-setting mentioned above (although they would also incorporate discreteness effects).

3. Inference from observed bid and ask quotes.

Viewed as a transformation of continuous random inputs (m, α and β) into discrete bid and ask prices, the model described by (1) is a very simple one. From the perspective of the econometrician (and that of many market participants), however, the observed bid and ask prices are given, and inference focuses on the unobserved inputs. Viewed in this direction, the model is more complex.

As a function of the observed bid and ask quotes (b, a), the feasible region for (m, α, β) consistent with model (1) is:

$$Q(b,a) = \{(m,\alpha,\beta): \alpha > 0, \beta > 0, b \le m - \beta < b+1 \text{ and } a-1 < m+\alpha \le a\} \tag{2}$$

The inequalities define a convex polytope (geometric solid) of up to six faces. Figure 1 depicts the region $Q(b=0, a=1)$ (a one-tick spread), along with several rotated perspectives.

Although the estimations in this paper are based on (2), it is useful to consider the special case in which the quote exposure costs are the same on both bid and offer sides. Letting $c = \alpha = \beta$, the feasible region is

$$Q(b,a) = \{(m,c): c > 0,\ b \le m - c < b+1 \text{ and } a-1 < m+c \le a\} \tag{3}$$

These inequalities define a two-dimensional region. Figure 2 depicts $Q(b=0, a=1)$ (a one-tick spread), $Q(b=0, a=2)$ (a two-tick spread), and $Q(b=0, a=3)$ (a three-tick spread). The diamond shape of the region $Q(b=0, a=2)$, for example, can be viewed as arising in the following way. When c is just slightly greater than zero or slightly less than one, the range of m consistent with $b=0$ and $a=2$ is a small neighborhood about one. When c is 1/2, m can range from 1/2 to 3/2.

(Figures 1 and 2 are related as follows. The condition that $\alpha=\beta$ defines a vertical plane in Figure 1 lying at a forty-five degree angle with respect to the α and β axes. This plane now defines two-dimensional region containing m and c ($=\alpha=\beta$). The intersection of the plane and the tetrahedron defines the half-diamond shape in Figure 2 associated with $Q(b=0, a=1)$.)

Given a prior probability density function $f(m,a,b)$, the posterior density conditional on observing bid and ask quotes b and a is:

$$f(m,\alpha,\beta|b,a) = \begin{cases} \dfrac{f(m,\alpha,\beta)}{\Pr(b,a)} & \text{if } (m,\alpha,\beta) \in Q(b,a) \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where $\Pr(b, a)$ is the probability of observing discrete bid and ask quotes b and a:

$$\Pr(b,a) = \int_{(m,\alpha,\beta)\in Q(b,a)} f(m,\alpha,\beta)\, dm\, d\alpha\, d\beta \tag{5}$$

Since this conditioning imposes a truncation on the ranges of the variables, it might seem that the conditional densities would be simple truncated versions of the priors. The truncations defined by $Q(b,a)$, however, apply to linear combinations of the variables, not the variables themselves. The shape of $Q(b,a)$ effectively forces a nonlinear transformation on the priors.

As an example, consider the case of equal exposure costs where $c=\alpha=\beta$ is lognormally distributed: $\ln(c)$ is assumed to be distributed normally with mean $\mu=-1$ and standard deviation $\sigma=0.6$. From equation (5) this implies $\Pr[a-b=1] = 0.29$, $\Pr[a-b=2] = 0.58$, $\Pr[a-b=3] = 0.11$, and $\Pr[a-b>3] = 0.03$, i.e., frequencies of one-, two- and three- tick spreads that might be observed for a typical NYSE stock.

For simplicity, assume a uniform diffuse prior on m: i.e., a probability density that is constant over some suitably large region. Formally it suffices to take

$$f(m) = \begin{cases} \kappa^{-1} \text{ for } m \in (0,\kappa) \\ 0, \text{ otherwise} \end{cases} \tag{6}$$

where κ is "large" (but not infinite). The choice of κ is arbitrary; it integrates out of all calculations. The cost parameter is assumed to be independent of the price level, which implies that the prior density of the latent variables may be written as $f(m, c)=f(m)\,f(c)$.

Figure 3 depicts the prior and conditional density functions conditional on observing bid and ask quotes $b=0$ and $a=2$. The prior for m is drawn as a flat line of height κ^{-1}. The conditional density for m is not uniform over the allowable range of m ($1/2<m<3/2$). If m is near an endpoint, the range of feasible c values is a small one, with correspondingly low probability. If m is near the center of the range, the feasible set for c is larger. Similarly the conditional density for c is not simply a truncated log normal, but slopes down gradually to the boundary defined by $c=1$. When c lies on this boundary, the set of m values consistent with the observed quotes is a single point (of probability zero, given the continuous prior assumed for m). As we move inward from this constraint, the set of feasible m becomes larger. (The peak in the conditional cost density arises from the "corners" in the diamond $Q(a=0, b=2)$ in Figure 2.)

4. Maximum Likelihood Estimation.

Suppose now that the quote generation process occurs over time periods t=1, . . ., T with state variable realizations $z_t=\{m_t, \alpha_t, \beta_t\}$ and corresponding observed bid and ask quotes $q_t=\{b_t, a_t\}$. In most applications the state variables will not be i.i.d.. Typically m_t might follow a random walk with non i.i.d. increments, and the latent cost variables might also exhibit serial correlation. The model in such cases is neither linear nor Gaussian. The general estimation approach follows Hamilton (1994a, 1994b) and Harvey (1991). The numerical technique is due to Kitagawa (1987), which is summarized in Hamilton (1994b). (Glosten and Harris (1988) employ another variant of this method.)

The essence of the procedure is a recursive likelihood calculation. Suppose that the probability density function of the current state variables conditional on current and past observations, $f_t(z_t|q_t, q_{t-1},\ldots;\theta)$ is known for some time t. Looking ahead to t+1,

$$f(z_{t+1}|q_t, q_{t-1},\ldots) = \int f(z_{t+1}|z_t) f(z_t|q_t, q_{t-1},\ldots) dz_t \tag{7}$$

where $f(z_{t+1}|z_t)$ is the state transition density function.

The conditional probability of observing q_{t+1} is

$$\Pr(q_{t+1}|q_t, q_{t-1},\ldots) = \int_{z_{t+1}\in Q_{t+1}} f(z_{t+1}|q_t, q_{t-1},\ldots) dz_{t+1} \tag{8}$$

where $Q_{t+1} = Q(b_{t+1}, a_{t+1})$ as defined in equation (2). The sequence of these probabilities may be used to construct the likelihood function.

The range of the integration in (8) is a distinctive feature of the present problem. In typical filtering applications the integration region is $\mathbf{R}^d$ where d is the dimension of the state vector. In the present application, however, the quotes serve to bound the possible values of the state variables: Q_t defines a small region of z_t space. In an online forecasting application we would be interested in computing the probabilities given by (8) for a number of possible realizations of q_{t+1}. In an estimation situation, however, we need only compute the probability for the value of q_{t+1} that actually occurs in the sample.

Next, note that the joint density of next period's state variables and quotes is:

$$f\left(z_{t+1}, q_{t+1} \middle| q_t, q_{t-1}, \ldots\right) = \begin{cases} f\left(z_{t+1} \middle| q_t, q_{t-1}, \ldots\right), \text{ if } z_{t+1} \in Q_{t+1} \\ 0, \text{ otherwise} \end{cases} \tag{9}$$

This too reflects a simplification peculiar to the present problem: computation of the right-hand-side density usually involves integration over a density function of the observational errors. Here, the observations (the quotes) are a deterministic function of the state variables. Therefore

$$f\left(z_{t+1} \middle| q_{t+1}, q_t, q_{t-1}, \ldots\right) = \begin{cases} \dfrac{f\left(z_{t+1}, q_{t+1} \middle| q_t, q_{t-1}, \ldots\right)}{\Pr\left(q_{t+1} \middle| q_t, q_{t-1}, \ldots\right)}, \text{ if } z_{t+1} \in Q_{t+1} \\ 0, \text{ otherwise} \end{cases} \tag{10}$$

This completes the update. Maximum likelihood estimation proceeds by maximizing the sum of the log of the conditional probabilities $\Pr\left(q_{t+1} \middle| q_t, q_{t-1}, \ldots\right)$. It should be noted, however, that the procedure does not provide estimated residuals that might be used to test the specification of the transition densities.

Although straightforward in principle, this update requires the evaluation of two integrals for which closed-form solutions are not readily available. In the standard Kalman filter, all joint, marginal and conditional densities are normal, and the results of the integrations are summarized by update formulae for the conditional means and variances. In the present case, successive updates involves computation of nested, truncated densities of increasing dimension.

Computational Considerations

The present analysis follows Kitagawa (1987) in approximating the conditional density $f\left(z_t \middle| q_t, q_{t-1}, \ldots\right)$ by a numerical grid or lattice. For state variables $z_t = \left\{m_t, \alpha_t, \beta_t\right\}$, let lattice cell C_t^i define a rectangular solid in z_t space. The conditional density $f\left(z_t \middle| q_t, q_{t-1}, \ldots\right)$ is now represented by the set of numerical values defined over the lattice cells $\Pr\left(C_t^i \middle| q_t, q_{t-1}, \ldots\right)$ for $i = 1, \ldots$. The state transition densities $f\left(z_{t+1} \middle| z_t\right)$ become the discrete transition probabilities $\Pr\left(C_{t+1}^j \middle| C_t^i\right)$ for $i = 1, \ldots$ and $j = 1, \ldots$. The integration in (7) is now replaced by the summation:

$$\Pr\left(C_{t+1}^j \middle| q_t, q_{t-1}, \ldots\right) = \sum_i \Pr\left(C_{t+1}^j \middle| C_t^i\right) \Pr\left(C_t^i \middle| q_t, q_{t-1}, \ldots\right) \tag{11}$$

The integration in (8) becomes:

$$\Pr\left(q_{t+1}\middle|q_t,q_{t-1},\ldots\right)=\sum_j \Pr\left(C_{t+1}^j\middle|q_t,q_{t-1},\ldots\right)\frac{Vol\left(C_{t+1}^j\cap Q_{t+1}^j\right)}{Vol\left(C_{t+1}^j\right)} \tag{12}$$

When the lattice cell lies entirely within the feasible region Q_{t+1}, the summand in (12) is simply $\Pr\left(C_{t+1}^j\middle|q_t,q_{t-1},\ldots\right)$. But when only a portion of the cell lies within the feasible region, the summand is weighted to reflect this. The calculation of the intersection volume was performed using the computational geometry algorithms and software routines discussed in O'Rourke (1994).

As with the integration in (8), computation of the summation in (12) is facilitated by the restrictions implied by Q_{t+1}. The intersection $C_{t+1}^j\cap Q_{t+1}$ is empty for virtually all of cells in the z_t space, and it is easy to specify the small set of nonempty cells. Furthermore, if the purpose of the calculation is "off-line" estimation (rather than real-time forecasting), we can economize on the calculation of the $\Pr\left(C_{t+1}^j\middle|q_t,q_{t-1},\ldots\right)$ in equation (12) by computing only the values that will be used in the subsequent probability calculation. These simplifications greatly reduce the computational burden.

There are several approaches to approximating the discrete transition probabilities $\Pr\left(C_{t+1}^j\middle|C_t^i\right)$. If the density function is relatively constant over C_{t+1}^j, a useful approximation is

$$\Pr\left(C_{t+1}^j\middle|C_t^i\right)\approx \mathrm{Vol}\left(C_{t+1}^j\right)f\left(z_{t+1}^j\middle|z_t^i\right) \tag{13}$$

where z_{t+1}^j and z_t^i are the cell midpoints. If the density function is not relatively constant over C_{t+1}^j, one may employ

$$\Pr\left(C_{t+1}^j\middle|C_t^i\right)\approx \int_{z_{t+1}\in C_{t+1}^j} f\left(z_{t+1}\middle|z_t^i\right)dz_{t+1} \tag{14}$$

The method makes no assumptions that the state variables evolve independently. The framework accommodates, for example, models in which quote exposure costs covary or depend on the change in the efficient price. As will be discussed in the next section, however, the present implementation is more restrictive.

The model is "discretized" in two respects. The first stems from the discreteness fundamental to the problem at hand, that of the bid and ask observations. The second aspect, the discrete lattice used to update the probabilities, is only a computational device. The two roles for discreteness are unrelated. If there existed efficient algorithms for computing the required definite integrals of large dimension, the lattice approach would be unnecessary.

The use of a lattice for numerical evaluation obviously introduces misspecification into the model. The adequacy of the approximation depends on the cell size relative to the scale of the phenomena involved.

5. The full dynamic model

Overview

The present implementation assumes that the efficient price evolves independently of the quote exposure costs. Under this assumption, the dynamics of the two components may be discussed separately. The efficient price evolves as:

$$m_t = m_{t-1} + u_t \tag{15}$$

As discussed below, the increments to the efficient price u_t are assumed to follow an ARCH-type process.

The bid and ask quote exposure costs are assumed to evolve as:

$$\begin{aligned} \ln(\alpha_t) &= \mu_t + \phi\left(\ln(\alpha_{t-1}) - \mu_{t-1}\right) + v_t^\alpha \\ \ln(\beta_t) &= \mu_t + \phi\left(\ln(\beta_{t-1}) - \mu_{t-1}\right) + v_t^\beta \end{aligned} \tag{16}$$

where v_t^α and v_t^β are independently distributed as $N\left(0, \sigma_v^2\right)$. This specification allows for deterministic time variation in the mean, and also a persistent stochastic component.

Although both the bid and ask exposure costs are driving by a common deterministic component, the stochastic components are assumed independent. This assumption is most appropriate to a market setting in which the bid and ask quotes are set by limit orders of different traders. In the case where the quotes reflect the interests of a single dealer, it would probably be more appropriate to allow for positive correlation between the two costs. From a computational viewpoint, it is simple to restrict the model to perfect correlation. In this case there is a single cost of quote exposure $c=\alpha=\beta$. In this case (discussed in Section 3), the state variables are m_t and c_t. While one would certainly seek to allow for general correlation, attempts in this direction have not yet proved promising.

I now turn to a detailed discussion of the two components.

The Efficient Price Evolution

The modeling of the increments to the efficient price follows the exponential generalized autoregressive conditional heteroskedasticity (EGARCH) approach suggested by Nelson (1991). To allow for leptokurtosis and time-varying volatility in the efficient price increments, the standardized increment is assumed to be distributed as generalized error distribution (GED) variate with parameter ν: $\zeta_t \equiv u_t/\sigma_t$ is distributed as *GED(ν)*. The GED distribution is given by:

$$f_{GED}(\zeta;\nu)=\frac{\nu\exp\left[-\left(\frac{1}{2}\right)|\zeta/\lambda|^{\nu}\right]}{\lambda 2^{(1+1/\nu)}\Gamma(1/\nu)},\text{ where }\lambda\equiv\sqrt{\frac{2^{(-2/\nu)}\Gamma(1/\nu)}{\Gamma(3/\nu)}}\tag{17}$$

In the case where $\nu=2$, this reduces to the standard normal density.

A standard EGARCH specification models time-varying variances as:

$$\ln\left(\sigma_t^2\right)=\eta+\varphi\left(\ln\left(\sigma_{t-1}^2\right)-\eta\right)+\gamma\left(|\zeta_{t-1}|-\mathrm{E}|\zeta_{t-1}|\right)\tag{18}$$

where the terms on the right hand side reflect a mean, an autoregressive adjustment toward the mean, and a disturbance component driven by the prior period's shock. The expected absolute value is unconditional and time-invariant, depending only on the tail-thickness parameter. It is given by $\mathrm{E}|\zeta|=\lambda 2^{(1/\nu)}\Gamma(2/\nu)/\Gamma(1/\nu)$. (The asymmetry term suggested by Nelson is omitted.)

In the present application, a problem arises from the fact that the m_t (and therefore u_t and $\zeta_t=u_t/\sigma_t$) are not observable. Since knowledge of the bid and offer quote history is insufficient to compute equation (18), σ_t must be carried as an unobservable state variable. This is not computationally feasible. As a more tractable alternative, I assume that the variance process is driven by the conditional expectation of the absolute efficient price increment. That is, $|\zeta_{t-1}|$ in (18) is replaced by its conditional expectation $E_{t-1}[|\zeta_{t-1}|] = E_{t-1}[|u_{t-1}|]/\sigma_{t-1}$. In addition, the mean is allowed to be time-varying:

$$\ln\left(\sigma_t^2\right) = \eta_t + \varphi\left(\ln\left(\sigma_{t-1}^2\right) - \eta_{t-1}\right) + \gamma\left(E_{t-1}|\zeta_{t-1}| - E|\zeta_{t-1}|\right) \tag{19}$$

where $E_{t-1}|\zeta_{t-1}| = E\left[|u_{t-1}||q_{t-1}, q_{t-2}, \ldots\right]/\sigma_t$. This quantity is easily computed in the course of the iterative update.

Deterministic time variation

Both the quote exposure cost function in (16) and the variance specification in (19) allow for deterministic effects. At a bare minimum it appears necessary to allow for the intraday "U" shapes frequently exhibited by market data. A parsimonious function that permits end-point elevation can be built from exponential decay functions. The deterministic component of the cost process is:

$$\mu_t = k_1 + k_2^{open} \exp\left(-k_3^{open} \tau_t^{open}\right) + k_2^{close} \exp\left(-k_3^{close} \tau_t^{close}\right) \tag{20}$$

where τ_t^{open} is the elapsed time since the opening quote of the day (in hours) and τ_t^{close} is the time remaining before the scheduled market close (in hours). The deterministic component of the variance is similarly modeled as:

$$\eta_t = \begin{cases} l_1 + l_2^{open} \exp\left(-l_3^{open} \tau_t^{open}\right) + l_2^{close}, & \text{if } t \text{ is an intraday interval} \\ \eta^{overnight}, & \text{if } t \text{ is an overnight interval} \end{cases} \tag{21}$$

(Alternative specifications employed an end-of-day exponential function similar to that used in (20). The results were statistically indistinguishable from those based on (21).)

Alternative specifications

The full model described above is a joint description of the bid and ask quote exposure costs and the efficient price. Given the complexity of the model and its computational burdens, however, it is useful to investigate the performance of simpler models with more modest aims. For example, if only the quote exposure costs are of interest, the model might be estimated assuming at each point in time a diffuse prior for the efficient price (cf. the development in section 3). This variant, consisting solely of the cost equations (16) is termed the "cost model". When *m* is eliminated as a state variable, the numerical grid approach is still necessary due to the stochastic variation in the cost, but the reduction in dimension speeds computation. Alternatively, if the efficient price dynamics are the sole concern, one might estimate (15) under the assumption that the quote exposure costs are diffusely distributed over the positive real line. This variant is termed the "discrete EGARCH model".

6. Estimation

Data

I estimate the specifications described in the last section to NYSE bid and ask quotes for Alcoa (ticker symbol AA) for all trading days in 1994. Alcoa is the first Dow Stock (in alphabetical ordering) and is viewed as a representative high-activity security. Bid and ask quotes are those prevailing at the close of 15-minute intervals. The first observation of a day generally corresponds to 9:45, the last to 16:00 (26 points). There are 6,780 observations.

Table 1 reports descriptive statistics for the absolute value of the bid first-differences. (Results for ask first-differences were virtually identical.) The proportion of intervals for which the bid change is zero is 39% (intraday) and 18% (overnight). Not reported in the table is the additional finding that in 24% of the intraday intervals and in 8% of the overnight intervals, there was no change in either the bid or the ask quote. If the underlying changes in the efficient price are viewed as arising from a continuous distribution of modest leptokurtosis, these figures suggest that the efficient price changes are not large relative to the tick size. The extreme values in the sample lie roughly seven standard deviations from the mean for the intraday intervals and five standard deviations from the mean for the overnight intervals.

(For a normally-distributed variate, the probability of observing an extreme value seven standard deviations from the mean in a sample of 6,528 observations is approximately 1×10^{-10}; that of an extreme value five standard deviations from the mean in a sample of 251 observations is approximately 1×10^{-7}.)

Table 2 reports descriptive statistics for the bid-ask spread. There is clear variation in the spread. In a sense, one purpose of the present model is the allocation of this variation to deterministic and stochastic effects.

Computational details

Computation of the likelihood function for the model of Section 5 followed the method described in Section 4. The three-dimensional integration lattice was constructed from one-dimensional lattices, one for each of the state variables α, β, and m. The break-points for the α lattice were (in ticks): 0., 0.01, 0.2, 0.04, 0.07, 0.13, 0.24, 0.46, 0.88, 1.67, 3.16 and 6 (the maximum spread in the analysis). These breakpoints approximate fixed intervals in $\ln(\alpha)$. The lattice for m ranged from the lowest bid in the sample to the highest ask, in 0.2-tick increments.

The number of cells necessary to cover a given quote region $Q(a,b)$ depends only on the spread a-b. For spread sizes of one through six ticks, the corresponding cell counts are: 282, 352, 218, 218, 198, 162 (for the full model); 5, 10, 15, 25, 30 (for the restricted EGARCH-only model); 74, 95, 42, 40, 40, 25 (for the restricted cost-only model). From equations (11) and (12), it is apparent that the number of computations involved in the recursive update is the product of the number of cells for the time t observation and that of the time t+1 observation.

Given the structure of the model, the transition probability density function factors as $f(z_{t+1}|z_t) = f(\alpha_{t+1}|\alpha_t)f(\beta_{t+1}|\beta_t)f(m_{t+1}|m_t)$, where the components may readily be derived from equations (15) and (16). The transition probabilities between lattice cells were generally computed using the midpoint approximation described in (13). However, if the m transition contained zero, approximation (14) was used. (The GED distribution is peaked at zero for low values of ν.)

Estimates of the full model

Table 3 reports parameter estimates. For purposes of exposition, these may be grouped as cost- and EGARCH (variance)-related. The EGARCH-related parameter estimates suggest a strong persistent stochastic component of the return variance. The autoregressive variance parameter estimate of φ=0.88, however, implies a half-life of about six (15-minute) periods. The intraday persistence reflects, therefore, phenomena different from those underlying daily and longer-term volatility persistence. The GED tail-thickness parameters of $\nu^{day} = 0.86$ and $\nu^{overnight} = 1.02$. For comparison purposes, Figure 4 graphs the GED density with ν=0.86 against the standard normal.

The ν estimates are lower than Nelson's estimate for daily CRSP returns (about 1.6). The present estimates imply a more pronounced leptokurtosis, consistent with a "lumpy" intraday information arrival process for individual stocks. This is less pronounced in the daily index returns due to aggregation over firms and time.

Turning now to the quote-exposure cost estimates, the deterministic parameters depict the usual U-shaped intraday pattern, although the standard errors of the decay rates are large. Of more interest is the characterization of the stochastic component. Both the disturbance variance σ_v and the autoregressive parameter ϕ are strongly positive. The autoregressive parameter suggests that 37% of the excess log cost persists at the subsequent time point (fifteen minutes later).

The relative importance of the deterministic and stochastic sources of variation in the quote exposure cost can be ascertained from simulations of the model using the parameter estimates. For a simulation of 2,500 days, Figure 5 depicts the time path of the 10^{th}, 50^{th} and 90^{th} percentiles of the cost expressed in dollars per share. The 50^{th} percentile (the median) displays the "U" shape characteristically found in spreads. The median is roughly four cents per share at the open and two cents thereafter, rising slightly at the close. Most importantly, the elevation associated with the beginning and end of trading is modest compared with the stochastic variation implied by the 10^{th} and 90^{th} percentile bands. This suggests that the stochastic component is relatively large.

The finding of a mean-reverting persistent component in the cost of quote exposure is consistent with several models of economic behavior. It may reflect mean-reversion in the underlying cost determinants (such as asymmetric information exposure costs or inventory holding costs related to risk) that are common across all actual and potential quote-setters. It may also reflect, however, the arrivals and departures of individual traders with differing costs. A buyer anxious to trade, for example, might enter a limit order that betters the prevailing bid. At some point this limit order is likely to be removed, either because it has been hit or else because the trader has withdrawn it and replaced it with a market buy order.

Estimates of the cost and EGARCH models.

Both the cost and EGARCH models are computationally simpler subcases of the full model. The first follows from an ongoing assumption of a diffuse prior for the efficient price; the second assumes a diffuse prior on the quote exposure costs. The resulting estimates are given in the last two columns of Table 3. Not only are the estimates virtually identical to those obtained for the full model, but so are the estimated standard errors. Although one might have hoped that the full specification would result in more precise estimates, this does not appear to be the case.

Several considerations could account for this failure. One possibility is simply general model misspecification. But it is also possible that even in a correctly specified model, the information about α and β contributed by m is small. There is no strong economic presumption supporting correlation between m and the cost variables. (For example, when we are in a full diamond region of Figure 2, knowledge of m is more informative about the dispersion of c than expected value of c.)

7. Discussion: does discreteness matter?

The question was posed by Hausman, Lo and MacKinlay (1992) in discussing their ordered probit model of transaction price changes. It is relevant here for similar reasons. At the cost of complexity and computational expenditure, both papers employ refined statistical methods to characterize the discreteness aspect of security prices. It is fair to inquire whether and in what circumstances, these costs might be justified by the benefits. The following discussion suggests some general principles.

If the microstructure time series can be assumed covariance stationary for the purposes at hand, and if the features of interest are functions of the first and second moments, then one can apply the methods of linear multivariate time series analysis without making an explicit provision for discreteness. Broadly speaking, vector moving-average and autoregressive models are based on the Wold representation theorem, which doesn't rely on assumptions contrary to discreteness. Hasbrouck (1996) discusses the use of these techniques in microstructure analyses. Reduced-form vector autoregressions are capable of modeling key aspects of the interactions between order flow, transactions and quotes without specifying structural features of the market. As an example, Hasbrouck (1991a) discusses a cross-firm analysis of estimated market depth (price responsiveness to order flow).

However, if one wishes to move beyond the characterizations offered by the reduced form models, then one is forced to confront the fact that most of the underlying determinants of security prices and trading costs are continuous variables and most of the sample data are discrete. If the features of interest are large relative to the tick size (e.g., annual movements in a security's value), then there is little lost by taking the discrete data as realizations from a continuous distribution. As the scale of these features becomes comparable to or smaller than the tick size (e.g., intraday value changes), then the model becomes progressively more misspecified. The scale of the quote exposure costs investigated in this paper, for example, is so much smaller than the tick size that ignoring discreteness would lead to meaningless results.

It may be, furthermore, computationally treacherous to fit discrete data to a continuous likelihood function. When the object of analysis is the modeling of security price variances, for example, the econometrician will often attempt to increase the power of the analysis by using data of higher frequency. The generalized error distribution used here provides a case in point. Attempts to model actual and simulated fifteen-minute discrete bid changes in this fashion were numerically unstable. (The distribution of discrete bid changes has a large mode at zero. The likelihood maximization procedure kept attempting values of the kurtosis parameter ν tending toward zero, in a presumed effort to capture this peak.)

The present procedure possesses the additional merit that it accommodates both deterministic nonstationarity and stochastic parameter variation. Both are frequently found in microstructure models that attempt to capture behavior of diverse agents in non-time-homogeneous settings (e.g., around market openings and closures). Although the usual Gaussian Kalman filter estimates are generally consistent and "best-linear" when the disturbances are non-Gaussian, neither of these properties obtains when the transition probabilities are time-varying and stochastic (Hamilton (1994)).

Finally, the present procedure shares with most state-space estimators the feature that it is essentially a Bayesian forecasting algorithm. As such, the econometrician's computation parallels the inference that might plausibly be taken in real time by a market participant.

8. Conclusion

This paper has presented a dynamic model of discrete bid and ask quotes. The discrete quotes are rounded transformations of a continuous efficient price and continuous quote exposure costs. The latter are presumed to capture most of the costs usually associated with market-making or limit order placement, such as fixed transaction costs and asymmetric information costs. The full statistical model is a rich one, allowing for stochastic and deterministic time variation in the efficient price volatility and the quote exposure cost. The model is estimated by maximum likelihood using a nonlinear state-space filtering approach due to Kitagawa (1987).

This specification is estimated for NYSE bid and ask quotes collected at the end of 15-minute intervals for Alcoa over 1994. The estimates confirm the existence of deterministic "U" shapes in the quote cost and efficient price volatility. More importantly, however, the estimates confirm the existence of a persistent stochastic component of the quote exposure cost. The magnitude of this component is roughly comparable to the variation associated with the "U" shapes.

In extending the model to incorporate other aspects of the market process, there are several guidelines. It is relatively easy to incorporate deterministic effects and observed exogenous variables into either the cost or efficient price specifications. Such developments usually require additional parameters in the likelihood function, which does not significantly affect the time required for the numerical calculation of this function (although it will probably increase the number of iterations required for convergence).

It is more difficult to add endogenous variables, such as quote sizes (number of shares at the bid and ask) or trades that are determined in part by prevailing quotes. These developments require an expansion of the set of state variables and a large accompanying increase in the computational burden. One might also want to specify a model for the quote exposure cost that is more complicated than the first-order autoregressive process employed here, by including additional autoregressive or moving average terms. These modifications also require additional state variables. Expansion of the state variable set runs into the "curse of dimensionality" because of the requirement that the integration of the conditional probabilities be computed numerically over all variables.

References

Ahn, H.-J, C. Q. Cao, and H. Choe, 1996, Tick size, spread and volume, Journal of Financial Intermediation, 5.

Angel, J., 1994, Tick size, share prices and stock splits, working paper, Georgetown University.

Anshuman, V. R., and Avner Kalay, 1994, Market making rents under discrete prices: Theory and evidence, Working paper, Boston College.

Ball, C. A., 1988, Estimation bias introduced by discrete security prices, Journal of Finance, 43, 841-865.

Bernhardt, D. and E. Hughson, 1996, Discrete pricing and the design of dealership markets, Journal of Economic Theory, 71, 148-182.

Bollerslev, T., and M. Melvin, 1994, Bid-ask spreads in the foreign exchange market: an empirical analysis, Journal of International Economics.

Bollerslev, T., R. F. Engle and D. B. Nelson, 1994, ARCH models, in R. F. Engle and D. L. McFadden, eds., Handbook of Econometrics, Volume IV, (Elsevier).

Brown, S., P. Laux and B. Schacter, 1991, On the existence of an optimal tick size, Review of Futures Markets, 10, 50-72.

Cho, D. C. and E. W. Frees, 1988, Estimating the volatility of discrete stock prices, Journal of Finance, 43, 451-466.

Chordia, T. and A. Subrahmanyam, 1995, Market making, the tick size, and payment-for-order flow: theory and evidence, Journal of Business, 68, 543-575.

Christie, W. G. and P. H. Schultz, 1994a, Why do NASDAQ market makers avoid odd-eighth quotes?, Journal of Finance, 49, 1813-40.

Christie, W. G. and P. H. Schultz, 1994b, Why did NASDAQ market makers stop avoiding odd-eighth quotes?, Journal of Finance, 49, 1841-60.

Cordella, Tito and Thierry Foucault, 1996, Minimum price variations, time priority and quote dynamics, Working paper, Universitat Pompeu Fabra.

Dravid, A. R., 1991, Effects of bid-ask spreads and price discreteness on stock returns, Working Paper, Rodney L. White Center, Wharton School.

Easley, D., N. Kiefer, and M. O'Hara, 1994, A day in the life of a very common stock, working paper, Cornell University Graduate School of Management.

Glosten, L. R. and L. E. Harris, 1988, Estimating the components of the bid-ask spread, Journal of Financial Economics, 21, 123-42.

Glosten, L. R. and P. R. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, Journal of Financial Economics, 14, 71-100.

Glosten, L. R., 1994, Is the electronic open limit order book inevitable?, Journal of

Finance, 49, 1127-1161.

Gottlieb, G. and A. Kalay, 1985, Implications of the discreteness of observed stock prices, Journal of Finance, 40, 135-153.

Hamilton, J. D., 1994, "State-Space Models," in R. F. Engle and D. L. McFadden, eds., Handbook of Econometrics, Volume IV, (Elsevier).

Hamilton, J. D., 1994, Time Series Analysis, (Princeton University Press, Princeton).

Harris, L. E. and J. Hasbrouck, 1996, Market vs. Limit orders: the SuperDOT evidence on order submission strategy, Journal of Financial and Quantitative Analysis, 31, 213-232.

Harris, L. E., 1990, Estimation of stock variances and serial covariances from discrete observations, Journal of Financial and Quantitative Analysis, 25, 291-306.

Harris, L. E., 1990, Liquidity, trading rules and electronic trading systems, Salomon Center Monograph Series in Finance and Economics 1990-4 (Stern School of Business, New York University).

Harris, L. E., 1991, Stock price clustering and discreteness, Review of Financial Studies, 4, 389-415.

Harris, L. E., 1994, Limit Orders, working paper, University of Southern California.

Harris, L. E., 1994, Minimum price variations, discrete bid-ask spreads and quotation sizes, Review of Financial Studies, 7, 149-178.

Harvey, A. C., 1991, Forecasting, Structural Time Series Models and the Kalman Filter, (Cambridge).

Hasbrouck, J., 1991a, Measuring the information content of stock trades, Journal of Finance, 46, 179-207.

Hasbrouck, J., 1991b, The summary informativeness of stock trades: an econometric investigation, Review of Financial Studies, 4, 571-95.

Hasbrouck, J., 1993, Assessing the quality of a security market: a new approach to measuring transaction costs, Review of Financial Studies, 6, 191-212.

Hausman, J. A., A. W. Lo and A. C. MacKinlay, 1992, "An Ordered Probit Analysis of Transaction Stock Prices," Journal of Financial Economics, 31, 319-330.

Kitagawa, G., 1987, "Non-Gaussian State-Space Modeling of Nonstationary Time Series," Journal of the American Statistical Association, 82, 1032-1041.

Madhavan, A., M. and S. Smidt, 1991, A Bayesian Model of Intraday Specialist Pricing, Journal of Financial Economics, 30, 99-134.

Madhavan, A., M. and S. Smidt, 1993, An Analysis of Daily Changes in Specialist's Inventories and Quotations, Journal of Finance, 48, 1595-1628.

Madhavan, A., M. Richardson and M. Roomans, 1994, Why do security prices change? A transaction level analysis of NYSE stocks, Working Paper, Wharton School.

Marsh, T. A. and E. R. Rosenfeld, 1986, Non-trading, market making and estimates of stock price volatility, Journal of Financial Economics, 15, 359-372.

Nelson, D. B., 1991, Conditional Heteroskedasticity in asset returns: a new approach, Econometrica, 59, 347-370.

Niederhoffer, V., 1965, Clustering of stock prices, Operations Research, 13, 258-265.

Niederhoffer, V., 1966, A new look at clustering of stock prices, Journal of Business, 39, 309-313.

O'Rourke, Joseph., 1994, Computational Geometry in C. Cambridge ; New York : Cambridge University Press.

Rock, K., 1996, The specialist's order book and price anomalies, forthcoming, Review of Financial Studies.

Table 1.

Descriptive statistics for 15-minute bid changes, Alcoa, 1994.

Bid changes (in 1/8 dollar ticks) were computed for Alcoa for all trading days in 1994 (plus the overnight change).

	Intraday	Overnight
N	6,528	251
Min (ticks)	-10	-15
Max (ticks)	11	19
Mean (ticks)	0.03	-0.21
Std. Dev. (ticks)	1.58	3.70
Distribution		
% with no change	39%	18%
% with 1-tick change	37%	27%
% with >1-tick change	25%	47%

Table 2.

Descriptive statistics for bid-ask spread at 15-minute intervals, Alcoa, 1994.

Spreads (in 1/8 dollar ticks) were computed for Alcoa at fifteen minute intervals during the trading day, for all trading days in 1994

N	6,780
Min	1
Max	5
Mean	1.65
Std	0.67
Distribution	
1-Tick	45.9%
2-Tick	43.5%
3-Tick	10.3%
4 or more ticks	0.3%

Table 3.

The state variables in the model are $z_t = \{m_t, \alpha_t, \beta_t\}$ where m_t is the implicit efficient price, α_t is the quote exposure cost on the ask or offer side of the market, and β_t is the quote exposure cost on the bid side of the market. t indexes 15-minute intraday intervals (plus the overnight period). The dynamics of the state variables are:

$$m_t = m_{t-1} + u_t \tag{15}$$

and

$$\begin{aligned} \ln(\alpha_t) &= \mu_t + \phi(\ln(\alpha_{t-1}) - \mu_{t-1}) + v_t^\alpha \\ \ln(\beta_t) &= \mu_t + \phi(\ln(\beta_{t-1}) - \mu_{t-1}) + v_t^\beta \end{aligned} \tag{16}$$

where v_t^α and v_t^β are independently distributed as $N(0, \sigma_v^2)$. (Equation numbers are those given in the text.) The efficient price disturbance, u_t, has standard deviation σ_t and after standardization is distributed as a generalized error distribution variate with tail-thickness parameter ν:

$$\zeta_t \equiv u_t / \sigma_t \quad f_{GED}(\zeta; \nu) = \frac{\nu \exp\left[-\left(\frac{1}{2}\right)|\zeta/\lambda|^\nu\right]}{\lambda 2^{(1+1/\nu)}\Gamma(1/\nu)}, \text{ where } \lambda \equiv \sqrt{\frac{2^{(-2/\nu)}\Gamma(1/\nu)}{\Gamma(3/\nu)}} \tag{17}$$

The efficient price variance follows a modified EGARCH process:

$$\ln(\sigma_t^2) = \eta_t + \varphi(\ln(\sigma_{t-1}^2) - \eta_{t-1}) + \gamma(\mathrm{E}_{t-1}|\zeta_{t-1}| - \mathrm{E}|\zeta_{t-1}|) \tag{19}$$

where $E_{t-1}|\zeta_{t-1}| = E[|u_{t-1}||q_{t-1}, q_{t-2}, \ldots]/\sigma_t$ is the filtered estimate conditional on the bid and ask prices through t-1.

The deterministic component of the cost process is:

$$\mu_t = k_1 + k_2^{open}\exp(-k_3^{open}\tau_t^{open}) + k_2^{close}\exp(-k_3^{close}\tau_t^{close}) \tag{20}$$

where τ_t^{open} is the elapsed time since the opening quote of the day (in hours) and τ_t^{close} is the time remaining before the scheduled market close (in hours). The deterministic component of the variance is:

$$\eta_t = \begin{cases} l_1 + l_2^{open}\exp(-l_3^{open}\tau_t^{open}) + l_2^{close}, & \text{if } t \text{ is an intraday interval} \\ \eta^{overnight}, & \text{if } t \text{ is an overnight interval} \end{cases} \tag{21}$$

Table 3 (Continued).

The observations are the quotes, which comprise a bid and ask price, $q_t = \{b_t, a_t\}$. These are functions of the state variables:

$$\begin{aligned} b_t &= \text{Floor}(m_t - \beta_t) \\ a_t &= \text{Ceiling}(m_t + \alpha_t) \end{aligned} \tag{1}$$

The column corresponding to the "full" model gives parameter estimates based the Kitagawa nonlinear filtering procedure. The "cost" estimates reflect an estimate of the cost-related parameters assuming a diffuse prior for the efficient price (also using the Kitagawa procedure). The "EGARCH" estimates refer to maximum likelihood estimation of a discretized EGARCH specification.

The models are estimated for Alcoa over all trading days in 1994, with t indexing 15-minute intervals within the day (and the overnight interval). Standard errors are reported in parentheses.

Table 3 (Continued).

		Model		
		Full	Cost	EGARCH
Quote exposure cost parameters:	k_1	-1.67 (0.03)	-1.68 (0.03)	
	k_2^{open}	0.45 (0.06)	0.46 (0.06)	
	k_3^{open}	2.42 (0.72)	2.40 (0.69)	
	k_2^{close}	0.21 (0.07)	0.19 (0.07)	
	k_3^{close}	3.48 (2.50)	3.50 (2.68)	
	ϕ	0.37 (0.03)	0.39 (0.03)	
	σ_v	0.86 (0.03)	0.86 (0.02)	
EGARCH parameters:	l_1	0.39 (0.06)		0.35 (0.07)
	l_2^{open}	1.73 (0.23)		1.76 (0.24)
	l_3^{open}	1.11 (0.23)		1.11 (0.23)
	l_2^{close}	0.59 (0.14)		0.61 (0.15)
	$\eta^{overnight}$	2.73 (0.13)		2.73 (0.14)
	φ	0.88 (0.02)		0.90 (0.02)
	γ	0.29 (0.03)		0.28 (0.03)
	ν^{day}	0.86 (0.02)		0.81 (0.02)
	$\nu^{overnight}$	1.02 (0.12)		1.01 (0.11)

Figure 1

As a function of the efficient price m, bid quote exposure cost β, and ask exposure cost α, the discrete bid and ask quotes are given by b=Floor($m-\beta$) and a=Ceiling($m+\alpha$). Given the observed discrete quotes, the feasible region for m, α and β is $Q(b,a) = \{(m,\alpha,\beta): \alpha, \beta > 0, b < m - \beta < b+1 \text{ and } a-1 < m+\alpha < a\}$. The figure depicts the region $Q(b$=0, a=1). The figure shows a detailed view and rotated perspectives.

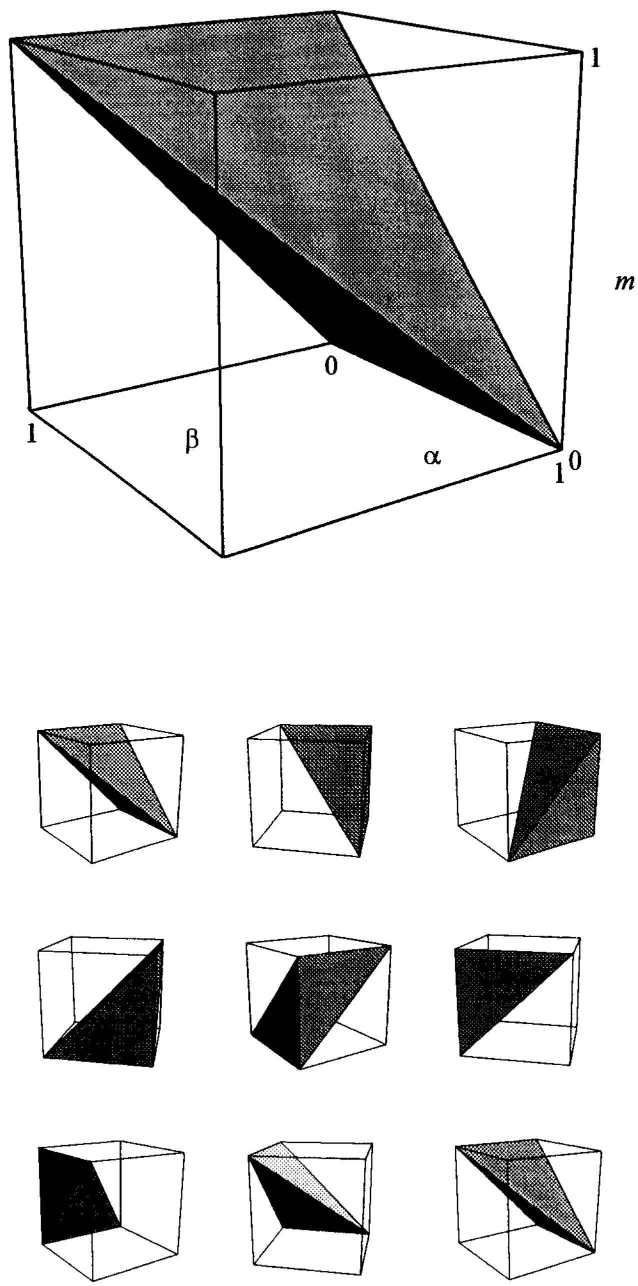


Figure 2

As a function of the efficient price m and quote exposure cost c, the discrete bid and ask quotes are given by b=Floor(m–c) and a=Ceiling(m+c). Given bid and ask quotes a and b, the region of feasible m and c is:

$$Q(b,a)=\{(m,c): c>0, b\leq m-c<b+1 \text{ and } a-1<m+c\leq a\}$$

The figure depicts the regions $Q(0, 1)$, $Q(0, 2)$ and $Q(0, 3)$.

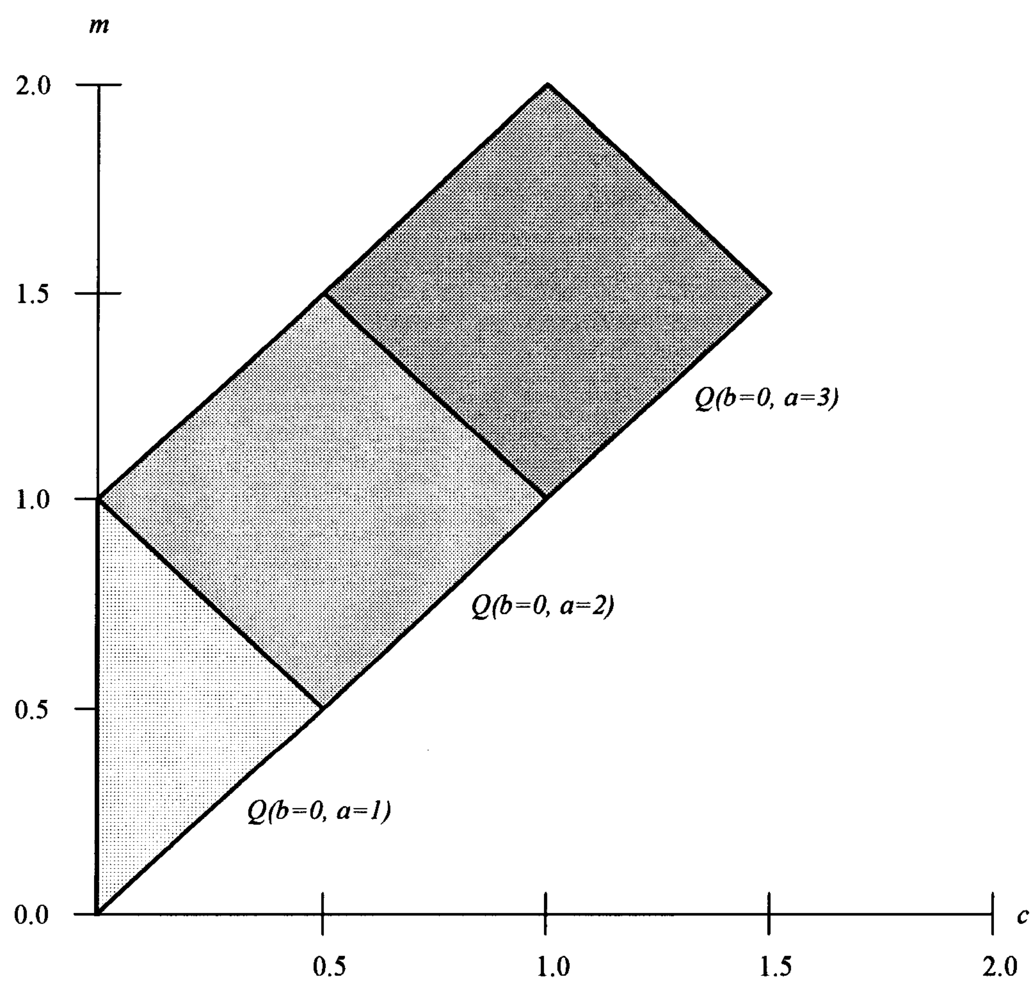
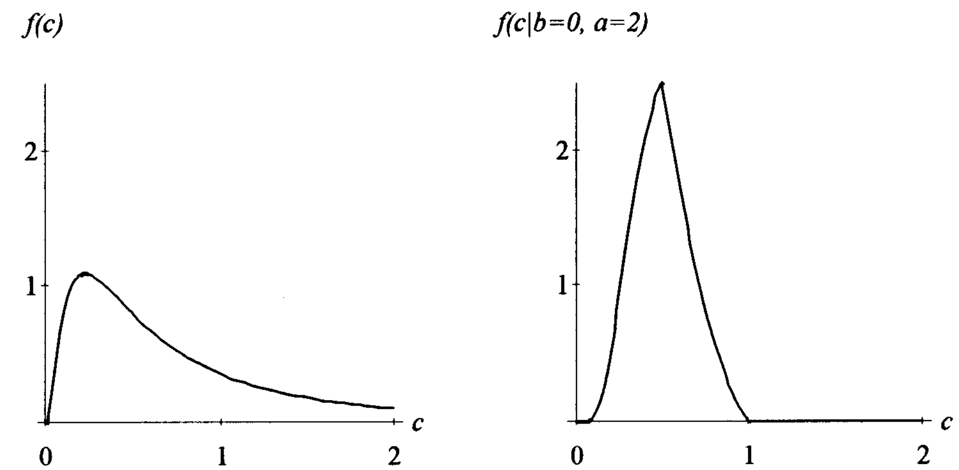


Figure 3

Figure depicts unconditional and condition probability densities for the efficient price m and quote exposure cost c. The unconditional density of c is lognormal: Log[c] is normally distributed with mean –1.0 and standard deviation 0.6. The unconditional density for m is a uniform diffuse prior on the interval (0, κ), where κ is an arbitrary positive constant (and does not appear in the conditional densities). The conditional densities are conditional on observing bid and ask quotes of b=0 and a=2.

Panel A. Unconditional and conditional densities of the quote exposure cost c.



Panel B. Unconditional and conditional densities for the efficient price m.

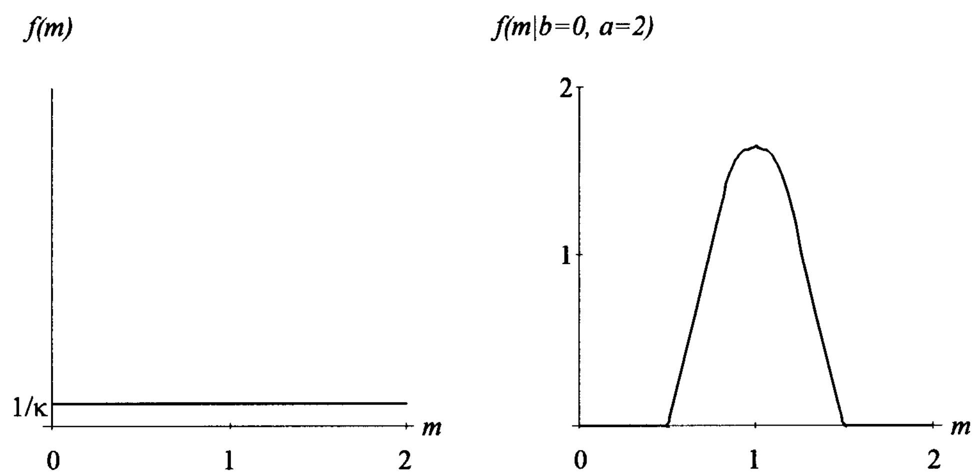


Figure 4

Figure depicts the probability density functions for the standard normal and standard GED with tail-thickness parameter ν=0.86.

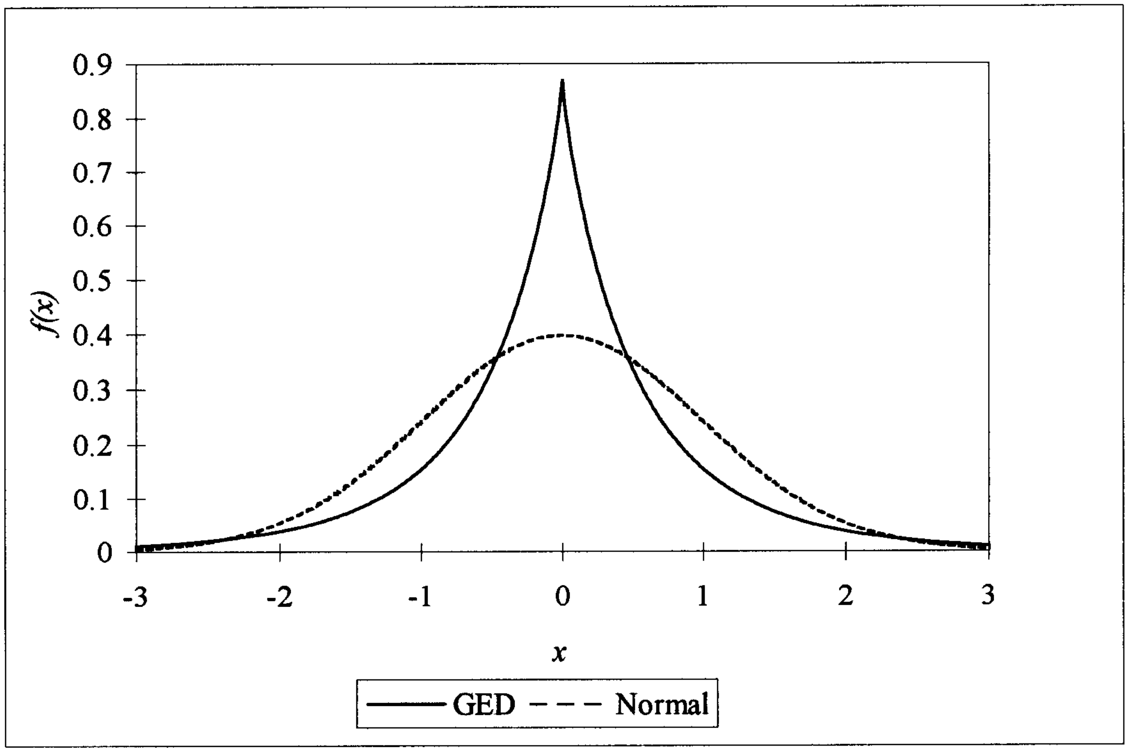


Figure 5

Figure depicts the time of day patterns in the quote exposure cost for ticker symbol AA implied by the model and estimates given in Table 3. The solid line is the 50th percentile of the cost. The upper and lower dashed lines are the 10th and 90th percentiles (respectively). (NB: these are not estimation confidence intervals.)