Trades and Quotes: A Bivariate Point Process

ROBERT F. ENGLE
New York University

ASGER LUNDE
Aarhus School of Business

ABSTRACT

This article formulates a bivariate point process to jointly analyze trade and quote arrivals. In microstructure models, trades may reveal private information that is then incorporated into new price quotes. This article examines the speed of this information flow and the circumstances that govern it. A joint likelihood function for trade and quote arrivals is specified in a way that recognizes that an intervening trade sometimes censors the time between a trade and the subsequent quote. Models of trades and quotes are estimated for eight stocks using Trade and Quote database (TAQ) data. The essential finding for the arrival of price quotes is that information flow variables, such as high trade arrival rates, large volume per trade, and wide bid–ask spreads, all predict more rapid price revisions. This means prices respond more quickly to trades when information is flowing so that the price impacts of trades and ultimately the volatility of prices are high in such circumstances.

KEYWORDS: duration analysis, market microstructure, transaction data.

Financial markets are designed to rapidly match buyers and sellers of assets at mutually agreeable prices. When this process is examined in detail, there are two types of events, which are observable at most financial exchanges. Traders buy and sell assets and specialists post quotes. Traders observe the posted prices and previous trades to determine their strategies, and similarly the specialists observe past trades and prices to decide what quotes to post. Since trades and quote revisions do not occur simultaneously, the times of each event presumably represent some optimization and potentially convey information.

The authors would like to thank the National Science Foundation for support (grant no. SBR-9422575). The second author thanks the Center for Nonlinear Modeling in Economics, University of Aarhus. Valuable comments have been received from Ken Nyholm, seminar participants at the University of Aarhus, the Center for Analytical Finance in Aarhus, University of California–San Diego, and Camp Econometrics IV at Catalina Island. Moreover, the suggestions of two anonymous referees and the editors have led to significant improvements in this article. All errors remain the authors responsibility. Address correspondence to Asger Lunde, Department of Information Science, Aarhus School of Business, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark, or e-mail: alunde@asb.dk.

Journal of Financial Econometrics, Vol. 1, No. 2, pp. 159–188

DOI: 10.1093/jjfinec/nbg011

In analyzing the information content of a trade, it is now common not only to examine the impact on prices of the direction of the trade [i.e., Hasbrouck (1996) and his many references], but also the timing of the trade [as in Easley and O'Hara (1992), Ghysels and Jasiak (1998), Engle (2000), Dufour and Engle (2000), and Grammig and Wellner (2002)]. The timing of the quote response, however, has not been examined. *How long do specialists wait after executing a trade until they post new quotes?* This is relevant in determining the speed of the price response to trades and ultimately to the rapidity of information absorption and market clearing.

This article analyzes these two time scales and estimates a model that relates them. We treat the arrival of trades and succeeding quotes as a bivariate, dependent point process. The arrival of each type of event is influenced by the past history of both processes and also by information such as the bid–ask spread, volume of trades, and other predetermined variables. The essential finding is that information flow variables, such as high trade arrival rates, large volume per trade, and wide bid/ask spreads, all predict more rapid price revisions. This means prices respond more quickly to trades when information is flowing, so that the price impacts of trades and ultimately the volatility of prices are high in such circumstances.

The model incorporates the fact that detailed records of financial transactions include two prices with different interpretations. A quote reflects one market participant's willingness to trade. It is firm only up to a given size and may be improved, both in terms of price and/or quantity. It may well reflect limit orders that are known to the specialist but often not to other participants. Transaction prices are agreed prices between counterparties; however, they do not reflect opportunities to trade. For example, a trade that occurs at the ask price is likely to be between a buyer and the specialist or a limit order. This price is not available to a seller. Similarly a trade price for a small-volume transaction is not generally available to a large transaction trader.

Some alternative models have recently been suggested by Rydberg and Shephard (1999b), Kamionka (2000), Russell (2001), and Bowsher (2002). These models are built in the multivariate point processes framework treated in, for example, Cox and Lewis (1972). In this setting, trades and quotes would be treated as the marginal event processes. Kamionka (2000) and Bowsher (2002) suggest looking at the so-called pooled process of all trades and quotes. In these formulations, the time from a trade to a quote is not modeled, as there may be intervening trades that censor this observation. The expected time could not be estimated without an adjustment for the arrival of intervening trades.

Finally, the importance of the questions addressed in this article is highlighted by the connection to the questions first addressed in Hasbrouck (1991) and further elaborated in Dufour and Engle (2000) and Chordia, Roll, and Subrahmanyam (2001). How long does it take for information to be incorporated into prices?

In the next section, the economic background is sketched. Section 2 presents the statistical framework. Section 3 presents the data and discusses institutional surroundings in which these emerged. In Section 4 we present the estimation results and a discussion. Section 5 gives a summary and some concluding remarks.

1 ECONOMIC MOTIVATION

Empirical findings of recent studies such as Engle and Russell (1998), Dufour and Engle (2000), and Engle (2000) are consistent with the predictions of the literature studying the market microstructure of financial markets [see O'Hara (1995)]. In the early microstructure models, time does not matter per se. In Kyle (1985), orders are batched together and cleared at predetermined points in time. Hence the arrival times of individual orders are of no relevance to the market marker. The sequential trade framework suggested by Glosten and Milgrom (1985) has orders arriving according to some stochastic fashion independent of any time parameters. Thus the timing of trades is also irrelevant in this model. If, however, time can be correlated with any factor related to the asset price, then the rate of trade arrivals conveys information to the agents. And as the agents learn from watching the flow of trades, the adjustment of prices to information will also depend on time [O'Hara (1995: 168)].

The notion of time was introduced into economic models by Diamond and Verrecchia (1987) and Easley and O'Hara (1992). Put very briefly, the first model shows that observing a low rate of trade arrival implies the presence of *bad news*. This result is derived from short-sell constrains. Easley and O'Hara (1992) introduce event uncertainty into the sequential trade framework. The uncertainty is whether informed traders received a signal about the value of the asset. Their model implies that a low trading intensity means *no news*, because the informed traders only trade when they get a signal.

The empirical studies mentioned above seem to favor the Easley and O'Hara model. Dufour and Engle (2000) found that time durations are negatively correlated with the absolute value of the following quote revision, and that the spread is negatively correlated with lagged durations. Engle (2000) and Engle and Russell (1998) derived a relationship between arrival rates and volatility. Engle (2000) modeled both the arrival times of trades and characteristics of these events, sometimes called marks. He modeled time according to the autoregressive conditional duration (ACD) model and the marks are modeled conditional on the times. Thus the estimated expected durations are included in the volatility equation of an autoregressive conditionally heteroscedastic (ARCH)-type model. It was found that longer durations were associated with lower volatilities, and interpreted as no news reduces volatility; this supports Easley and O'Hara.

In an asymmetric information framework, the specialist quotes bid and ask prices to offset the expected losses from trading with informed traders. Once a trade has occurred, the specialist can reevaluate his quotes. If the trade was a buy, then there is a slightly increased probability that the information possessed by a fraction of the traders was positive for the asset. The specialist will increase both bid and ask prices at this time and possibly change the spread. The amount by which the specialist moves the quotes depends on the information he has from trades thus far and the assessment of the fraction of traders who are informed. The higher the fraction, the greater the response to the trade.

A central question in market microstructure is how fast and how completely new information is incorporated into prices. A key but unnoticed part of this question is the timing of quote changes in response to trades. The timing of the specialist's response is assumed to be immediate in models such as Glosten and Milgrom (1985). However, in Easley and O'Hara (1992), the calendar time between his revisions can change. If there is no information event, then trading will slow down and consequently the time between quote revisions will become longer. However, there is still no delay from a trade to a quote revision. Only in the case where trades have no information, or where the discreteness of quotes is greater than the size of the desired revision, will there be a delay between trades and quote revisions. Thus there is a prediction that in a market with fewer information traders and slower trade rates, the time to revise quotes should be longer.

A deeper analysis of the timing of quote setting must be tied to the supply of limit orders. Since on the New York Stock Exchange (NYSE), the specialist participates in a relatively small number of trades, his quotes reflect the tightness of the limit order book. If limit buy orders are all above his asking quote, he may increase the quote to execute new market orders against the book. Similarly if there are limit buy orders within his spread, he may reduce the ask to again reflect the prices at which trades can be executed. In this interpretation, quotes may change in the absence of trades or other news simply because of changes in the order book. Similarly trades may not result in a quote change if the limit order book is unchanged.

A quote consists of four numbers, a bid and an ask price, and a bid and an ask quantity, called the quoted depth. The specialist guarantees to transact at least the quoted depth at the quoted prices. The specialist not only can post a quote to signify new prices, but also new depth. While both reflect the response to trades and to the limit order book, the economic incentives to adjust the quotes may be different for prices and quantities. In the empirical work, quotes which alter the prices are called midquotes, which is the average of the bid and ask prices, and are analyzed separately from all quotes.

2 METHODOLOGY

In trade or quote datasets, only one type of event can occur, namely a trade or the post of a quote, respectively. When combining trade and quote data, a more complicated situation arises. Now two types of events will be occurring as time passes and the associated marks may be different variables for the two types. There are very few general accounts on multivariate point processes in the literature. A comprehensive treatment was given by Cox and Lewis (1972), from whom we adopted some of the terminology and notation.

Denote by $t_1, \ldots, t_{i-1}, t_i$ the sequence of clock times at which a trade of a given asset occurred, and by $q_1, \ldots, q_{i-1}, q_i$ the timing of bid-ask quote revisions for this particular asset. A general bivariate model for these processes would involve associating a counting process,

$$N(s_1, s_2) = \{(N^t(s_1), N^q(s_2)\},$$

with the bivariate point process. Here $N^t(s_1)$ and $N^q(s_2)$ are the number of trades and quotes in $(0, s_1]$ and $(0, s_2]$, respectively. Further, a bivariate sequence of intervals $\{T_i, Q_j\}$ is defined. Here T_i is the time between the $(i-1)$th and the ith trade; and the $\{Q_j\}$ sequence is defined similarly. This specification might suggest constructing a bivariate model from $\{(T_i, Q_i)\}_{i=1}^N$. This could be fruitful in some situations, but in general it is not a useful approach because, in the present case, events in the two processes with a common serial number will be far apart in real time. This leaves the specification of the dependence between pairs of T_i and Q_i very difficult. Later it will be seen how our model circumvents this problem of asynchronous starting points of duration pairs.

2.1 The Model

Let $t_1, \ldots, t_{i-1}, t_i$ be defined as above. Using these trade arrival times, we define the sequence $t_1^{\text{fq}}, \ldots, t_{i-1}^{\text{fq}}, t_i^{\text{fq}}$, with t_i^{fq} denoting the clock time of the first quote arriving after the trade at t_{i-1}. Given these point processes, two types of durations are defined. Both types start with a trade occurring at time t_{i-1}. Hence, define by $X_i = t_i - t_{i-1}$ the forward recurrence time to the next trade, and denote by $Y_i = t_i^{\text{fq}} - t_{i-1}$ the forward recurrence time to the next quote. We call these random variables a *forward trade duration* and a *forward quote duration*, respectively. Together X_i and Y_i constitute a bivariate duration process that eliminates the synchronization problem mentioned above. The trade times each initiate a waiting time for the next quote to occur. The structure does not comprise the assumption that the process of forward trade durations is independent of the timing of quotes or information associated with the quote, such as spreads. This is merely a possibility that can be tested. Giving the model this structure has economic intuition, as in most transaction datasets the trade times are the only observable events that make the specialist reflect on his current bid price, ask price, or the depth at these prices. The series constructed in this manner now have the property that durations with a common serial number have the same time origin. While it is clearly possible to model the time from a quote to the following trade, this would be an entirely different model, designed to answer questions different from the ones we are addressing.

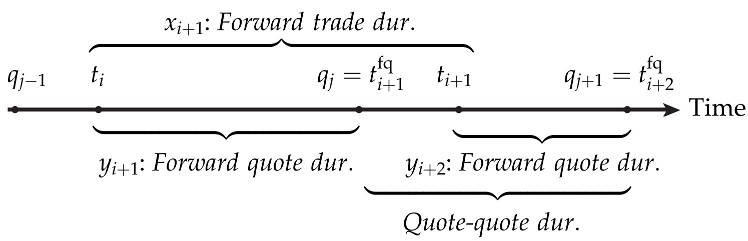
It will often be the case for very frequently traded stocks that a new trade occurs before the next quote, that is, X_i might be less than Y_i. The trade conveys information that is likely to change our beliefs about when the next quote will occur, and especially our beliefs about what will happen at the next quote. This means that at time t_i, our expectation of the arrival of the next quote will change even though the initiated quote spell was not completed. To embed this feature in the model, cases of $Y_i > X_i$ are treated as *censored* forward quote durations. This is done by defining

$$\tilde{Y}_i \equiv (1 - d_i)Y_i + d_i X_i, \quad \text{where } d_i \equiv I_{\{Y_i > X_i\}}.$$

We call $\tilde{Y}_i$ the *observed forward quote duration*. The associated indicator variable d_i takes the value 1 if $\tilde{Y}_i$ was censored, and in such cases we only know that the ith forward quote duration was longer than $\tilde{Y}_i$.

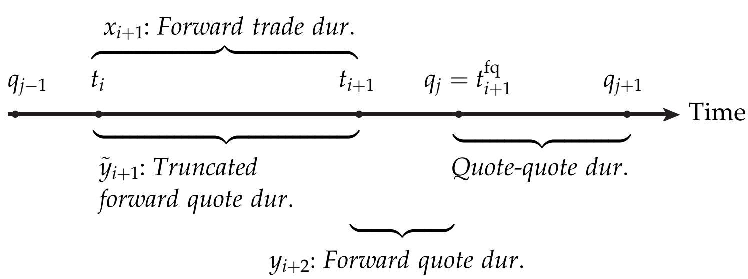
By construction, this framework will give rise to one of two outcomes, an uncensored observation or a censored observation. To envision the process it is helpful to look at two illustrations.

Scenario 1: No censoring



In this scenario, the forward quote duration is observed as q_j occurs before t_{i+1}. This is not the case for the second scenario.

Scenario 2: Censoring



As pointed out by the referee, quotes occurring like q_{j+1} are ignored in the construction of forward quote durations. This is not an important issue as long as we focus the article on modeling the specialist's reaction time from executing a trade until new quotes are posted. It is the time until the specialists start the updating process, which manifests itself into one or a series of rapidly following quotes. Apart from this, it is important to note that such quotes are only ignored when constructing the forward quote duration. They are very much retained in the model, through explanatory variables, such as the spread, interquote durations, and backward quote durations. We return to these matters in detail after finishing the model specification.

The statistical model can now be built by specifying the parameterization of the bivariate duration process given by $\{(X_i, \tilde{Y}_i)\}_{i=1}^N$. Assume that the *i*th observation has a joint density conditional on all relevant and available information as of time t_{i-1}. Modeling this joint distribution directly would be a very complex matter, but fortunately a simpler expression can easily be obtained. Without loss of generality, the joint density can be written as the product of the conditional density and the resulting marginal density. Thus we write this as

$$p(x_i, \tilde{y}_i|\mathcal{F}_{i-1}; \boldsymbol{\omega}) = g(x_i \,|\, \mathcal{F}_{i-1}; \boldsymbol{\omega}_1) f(\tilde{y}_i \,|\, x_i, \mathcal{F}_{i-1}; \boldsymbol{\omega}_2) \tag{1}$$

and call $g(\cdot \,|\, \mathcal{F}_{i-1}; \boldsymbol{\omega}_1)$ the trade density and $f(\cdot \,|\, x_i, \mathcal{F}_{i-1}; \boldsymbol{\omega}_2)$ the quote density. Here $\boldsymbol{\omega}$ is a vector containing all the parameters in the model. $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$ are subvectors of parameters pertaining to the trade and quote density. These may or may not be distinct parameters. $\mathcal{F}_{i-1}$ is the history of the bivariate duration process given by $\{(X_{j-1}, \tilde{Y}_{j-1})\}_{j=1}^{i-1}$ and any explanatory variables observable at time t_{i-1}.

Before we turn to the actual parameterization, a few words about the model defined so far are required. The process for the forward trade duration is assumed to be of the ACD type suggested by Engle and Russell (1998).[1] However, the observed forward quote durations are censored and this must be modeled carefully. The process of censoring times is in fact the forward trade duration process, and hence the censoring times will be a dependent process. Unlike conventional duration and competing risk models [see, e.g., Cox and Oakes (1984)], we observe the censoring threshold for each observation and can model it directly. In summary, we are *not* modeling the observed forward quote durations $\tilde{Y}_i$, but the forward quote durations Y_i, which will be clear below.

We now return to the parametric specification of the model. Let $\psi_i(\boldsymbol{\omega}_1) = E(X_i \,|\, \mathcal{F}_{i-1}; \boldsymbol{\omega}_1)$, then using a conditional exponential distribution, the trade density is specified as

$$g(x_i \,|\, \mathcal{F}_{i-1}; \boldsymbol{\omega}_1) = \psi_i(\boldsymbol{\omega}_1)^{-1} \exp\left\{\frac{-x_i}{\psi_i(\boldsymbol{\omega}_1)}\right\}, \tag{2}$$

[1] Over the last half of the previous decade the interest in modeling interevent durations has produced a large number of interesting articles. Some of these are Lunde (1997), Jasiak (1998), Bauwens and Veredas (1999), Hautsch (1999), Rydberg and Shephard (1999a), Bauwens and Giot (2000a), Fernandes and Grammig (2000), Gerhard and Hautsch (2000), Grammig and Maurer (2000), Golia (2001), Zhang, Russell, and Tay (2001), Ghysels, Gourieroux and Jasiak (2002). Some of these models are compared in Bauwens et al. (2000).

where the expected duration follows an exponential linked ACD-type model.[2] It is given by

$$\psi_i(\boldsymbol{\omega}_1) \equiv \psi_i = \exp\left\{\alpha + \delta \ln(\psi_{i-1}) + \gamma \frac{x_{i-1}}{\psi_{i-1}} + \boldsymbol{\beta}'\mathbf{Z}_{i-1}\right\}, \tag{3}$$

with $\mathbf{Z}_{i-1}$ being a vector of explanatory variables known at time t_{i-1}. Equation (3) will be referred to as the trade equation. Note that ψ^{-1}, the inverse of the expected duration, is the trading intensity. It gives the instantaneous rate at which trades arrive.

The quote density takes into account that some of the observations are censored. This is done in the usual way for models with censored observations. Thus we have

$$\begin{aligned} f(\tilde{y}_i \,|\, x_i, \mathcal{F}_{i-1};\ \boldsymbol{\omega}_2) &= h_Y(\tilde{y}_i \,|\, x_i,\ \mathcal{F}_{i-1};\ \boldsymbol{\omega}_2)^{1-d_i} S_Y(\tilde{y}_i \,|\, x_i,\ \mathcal{F}_{i-1};\ \boldsymbol{\omega}_2)^{d_i} \\ &= h_Y(y_i \,|\, x_i,\ \mathcal{F}_{i-1};\ \boldsymbol{\omega}_2)^{1-d_i} S_Y(x_i \,|\, x_i,\ \mathcal{F}_{i-1};\ \boldsymbol{\omega}_2)^{d_i}, \end{aligned} \tag{4}$$

where h_Y and S_Y are the density function and the survivor function,[3] respectively, for the forward quote duration. $h_Y(\cdot \,|\, x_i,\ \mathcal{F}_{i-1};\ \boldsymbol{\omega}_2)$ is the actual forward quote density and will be termed so. The intuition is that an uncensored observation will contribute to the likelihood with the actual forward quote density, and a censored observation contributes with the probability that the duration will be longer than the censoring threshold. The density is similar to the density for the forward trade durations, except for the important feature that it is conditional on x_i. Let $\varphi_i(\boldsymbol{\omega}_2) = E(Y_i \,|\, x_i,\ \mathcal{F}_{i-1};\ \boldsymbol{\omega}_2)$, then

$$h_Y(y_i \,|\, x_i,\ \mathcal{F}_{i-1};\ \boldsymbol{\omega}_2) = \varphi_i(\boldsymbol{\omega}_2)^{-1} \exp\left\{\frac{-y_i}{\varphi_i(\boldsymbol{\omega}_2)}\right\}. \tag{5}$$

Note, as $\varphi_i(\boldsymbol{\omega}_2)$ is the conditional expectation of Y_i (and not $\tilde{Y}_i$), we are actually building a model for the expected forward quote duration. This expected duration also follows an exponential ACD-type model, with the quote equation given by

$$\begin{aligned} \varphi_i(\boldsymbol{\omega}_2) &\equiv \varphi_i \\ &= \exp\left\{\mu + \rho \ln(\varphi_{i-1}) + \delta_1 \frac{\tilde{y}_{i-1}}{\varphi_{i-1}} + \delta_2 \frac{\tilde{y}_{i-1}}{\varphi_{i-1}} d_{i-1} + \tau \frac{x_i}{\psi_i} + \boldsymbol{\eta}'\mathbf{V}_{i-1}\right\}, \end{aligned} \tag{6}$$

where $\mathbf{V}_{i-1}$ typically contains some of the variables of $\mathbf{Z}_{i-1}$. Here φ^{-1} is the quoting intensity, which is the rate at which the specialist posts his quotes. Note as we can only condition on lagged *observed* forward quote durations, we include $(\tilde{y}_{i-1}/\varphi_{i-1})d_{i-1}$, which allows us to assess the impact of having some observations censored.

---

[2] The main reason for choosing this form is it's computational convenience, as it implies that the parameter space is unrestricted. This contrasts with the ACD model, which requires parameter restrictions to keep the expected duration positive. It closely resemblance the exponential generalized autoregressive conditionally heteroscedastic (EGARCH) of Nelson (1991), with the key feature that x_i/ψ_i is an i.i.d. *excess duration*. This specification is also advocated by Bauwens and Giot (2000a), who call it a *log-ACD*$_2$ model.

[3] The survivor function, $S_Y(t)$, is one minus the cumulative distribution function, $H_Y(t) = \int_0^t h_Y(u)du$, which is the probability that the duration will last longer than t.

The definitions given above imply that the expected forward trade duration $\{\ln(\psi_i)\}$ and the expected forward quote duration $\{\ln(\varphi_i)\}$ both follow ARMA models with exogenous explanatory variables [often called ARMAX models, see, e.g., Harvey (1990, p. 264)], to which the usual ARMA stationarity condition applies. Specifically, X_i and Y_i (and hence also $\tilde{Y}_i$) are stationary if δ and ρ are strictly less than one, respectively. This result is given in Bauwens and Giot (2002b) and Bauwens, Galli, and Giot (2002).

2.2 Estimation and Inference

Under Equation (1), the log-likelihood can be expressed as

$$\begin{aligned}\mathcal{L}(\boldsymbol{\omega};\mathbf{X},\tilde{\mathbf{Y}}) &= \sum_{i=1}^{N} l_i(\boldsymbol{\omega}) = \sum_{i=1}^{N} \ln p(x_i, \tilde{y}_i \mid \mathcal{F}_{i-1};\boldsymbol{\omega}) \\ &= \sum_{i=1}^{N} [\ln g(x_i \mid \mathcal{F}_{i-1};\boldsymbol{\omega}_1) + \ln f(\tilde{y}_i \mid x_i, \mathcal{F}_{i-1};\boldsymbol{\omega}_2)] \\ &= \sum_{i=1}^{N} l_i^g(\boldsymbol{\omega}_1) + \sum_{i=1}^{N} l_i^f(\boldsymbol{\omega}_2), \end{aligned} \tag{7}$$

which has to be maximized with respect to the parameters $(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2)$. To get good starting values for this problem, we take the approach of maximizing $\sum_{i=1}^{N} l_i^g(\boldsymbol{\omega}_1)$ first and then, conditional on this, $\sum_{i=1}^{N} l_i^f(\boldsymbol{\omega}_2)$ is maximized. Following this, we start the joint maximization of $\sum_{i=1}^{N} l_i(\boldsymbol{\omega})$ at the optimum found in the first step, delivering the maximum-likelihood estimates $\hat{\boldsymbol{\omega}} = (\hat{\boldsymbol{\omega}}_1,\hat{\boldsymbol{\omega}}_2)$. This two-step approach reduces the estimation time considerably, which is very important considering the huge datasets analyzed in this article. If $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$ are completely separate parameters, then maximization of the joint likelihood will be equivialent to maximization of the two separate likelihoods. If not, separate maximization will lead to estimates slightly different from maximum-likelihood estimates.

The estimation approach is semiparametric, in that we do not assume that the true densities of g and h are exponential, as stated in Equations (2) and (5).[4] The log-likelihood function is called a quasi-likelihood function. This method only requires specifying the conditional mean of the distribution. Then quasi maximum likelihood (QML) estimators can be obtained which are consistent for $\boldsymbol{\omega}$ and have a well-defined asymptotic covariance matrix. QML methods were introduced into econometrics by White (1982), and the results that justify the present application are analogous to Bollerslev and Wooldridge (1992), as shown by Engle and Russell (1998). The robust covariance matrix for $\boldsymbol{\omega}$ is calculated as

$$\mathbf{A}(\hat{\boldsymbol{\omega}})^{-1}\mathbf{J}(\hat{\boldsymbol{\omega}})\mathbf{A}(\hat{\boldsymbol{\omega}})^{-1} = \left[\sum_{i=1}^{N}\mathbf{H}_i(\hat{\boldsymbol{\omega}})\right]^{-1}\left[\sum_{i=1}^{N}\mathbf{s}_i(\hat{\boldsymbol{\omega}})\mathbf{s}_i(\hat{\boldsymbol{\omega}})'\right]\left[\sum_{i=1}^{N}\mathbf{H}_i(\hat{\boldsymbol{\omega}})\right]^{-1}, \tag{8}$$

[4] It is clearly possible to use a general density such as the generalized gamma distribution as suggested by Lunde (1998). But we have no economic theory suggesting what the shapes of these densities should be. It is not a question we are going to address in this article.

with

$$\mathbf{s}_i(\hat{\boldsymbol{\omega}}) = \frac{\partial l_i}{\partial \boldsymbol{\omega}}(\hat{\boldsymbol{\omega}}) \quad \text{and} \quad \mathbf{H}_i(\hat{\boldsymbol{\omega}}) = \frac{\partial^2 l_i}{\partial \boldsymbol{\omega} \partial \boldsymbol{\omega}'}(\hat{\boldsymbol{\omega}})$$

being the score vector and the Hessian matrix, respectively.

It seems appropriate to note that the censoring connection in Equation (4) is derived from the exponential density, and consequently QML interpretations of the quote equation are only in the sense of White (1982), such that convergence is to the maximizer of the asymptotic likelihood.[5]

2.3 Residuals and Specification Tests

The residual analysis assesses the validity of the exponential duration distributions used in the QML approach and the amount of remaining autocorrelation not explained by the specified model. Thus, under the null that the model is truly exponential and that the conditional expected duration is correctly specified, the residuals should be identical and independent unit exponentially distributed.

Generally residuals with a unit exponential distribution are defined as follows. If T has survivor function $S(t)$, then $S(T)$ is uniformly distributed and $-\ln S(T)$ has a unit exponential distribution. Thus for the trade part we define the residual to be

$$\begin{aligned} \xi_i &= -\ln(S(x_i \,|\, \mathcal{F}_{i-1}; \hat{\boldsymbol{\omega}}_1)) \\ &= x_i \psi_i(\mathcal{F}_{i-1}; \hat{\boldsymbol{\omega}}_1)^{-1}, \end{aligned} \tag{9}$$

which is identical to the residual defined in Engle and Russell (1998). These residuals are often called Cox–Snell residuals, as they are derived from the general definition of residuals given by Cox and Snell (1968).[6]

If the ith individual is censored, so too is the corresponding residual, and thus in general we obtain a set of uncensored and a set of censored residuals which cannot be regarded on the same footing. We may therefore seek to modify the Cox–Snell residuals, taking explicit account of the censoring. Suppose that $\tilde{y}_i$ is censored. The Cox–Snell residual for this observation is then given by

$$r_i = \tilde{y}_i \varphi_i(x_i, \mathcal{F}_{i-1}; \hat{\boldsymbol{\omega}}_2)^{-1}.$$

If the fitted model is correct, then r_i can be taken to have a unit exponential distribution. The cumulative hazard function of this distribution increases linearly with time, hence the greater the value of the duration, the greater the value of that residual. It hereby follows that the residual for this duration at the actual unobserved termination time will be greater than the residual evaluated at the observed censoring time. To account for this, the Cox–Snell residual can be modified by adding a positive constant, called the excess residual. Using the lack of memory

---

[5] A possible misspecification may be resolved by using a more general distribution, as suggested in Lunde (1998). Still this is beyond of the aim of this article, as the model is quite complex already.

[6] A very readable exposition of residuals for survival models may be found in Collett (1994).

property of the exponential distribution, we know that because r_i has a unit exponential distribution, the excess residual will also have a unit exponential distribution. The expected value of the excess residual is therefore one. This suggests defining the residuals for the quote part as

$$\varepsilon_i = \tilde{y}_i \varphi_i(x_i, \mathcal{F}_{i-1}; \hat{\omega}_2)^{-1} + d_i.$$

As above, if the fitted model is correct, then ε_i can be taken to have a unit exponential distribution.

In the empirical sections we report a number of Lagrange multiplier (LM) tests. These tests are defined as given in Wooldridge (1994: section 4.6), but for convenience we reproduce it in the present setting. Hence, let the objective be to test the r nonlinear restrictions on the parameter vector and write this as

$$\mathsf{H}_0\colon \mathbf{c}(\boldsymbol{\omega}) = \mathbf{0},$$

where $\mathbf{c}(\boldsymbol{\omega})$ is an $r \times 1$ vector function of $\boldsymbol{\omega}$, which we assume to be $p \times 1$, and $r \leq p$. Let $\mathbf{C}(\boldsymbol{\omega}) \equiv \nabla_{\boldsymbol{\omega}} \mathbf{c}(\boldsymbol{\omega})$ be the $r \times p$ gradient of $\mathbf{c}(\omega)$. Then using the heteroscedasticity-consistent covariance matrix of Equation (8), with the notation of the previous section, the LM test is defined as

$$Q_{\mathrm{LM}} = N^{-1} \left(\sum_{i=1}^{N} \mathbf{s}_i(\tilde{\boldsymbol{\omega}}) \right)' \mathbf{A}(\tilde{\boldsymbol{\omega}})^{-1} \mathbf{C}(\tilde{\boldsymbol{\omega}})'$$
$$\times \left(\mathbf{C}(\tilde{\boldsymbol{\omega}}) \mathbf{A}(\tilde{\boldsymbol{\omega}})^{-1} \mathbf{J}(\tilde{\boldsymbol{\omega}}) \mathbf{A}(\tilde{\boldsymbol{\omega}})^{-1} \mathbf{C}(\tilde{\boldsymbol{\omega}})' \right)^{-1} \mathbf{C}(\tilde{\boldsymbol{\omega}}) \mathbf{A}(\tilde{\boldsymbol{\omega}})^{-1} \left(\sum_{i=1}^{N} \mathbf{s}_i(\tilde{\boldsymbol{\omega}}) \right),$$

where all quantities are evaluated at the estimates, $\tilde{\boldsymbol{\omega}}$, obtained after imposing the restrictions. This statistic is asymptotically $\chi^2(r)$ distributed under H_0.

3 DATA DESCRIPTION

The data are extracted from the Trade and Quote (TAQ) database. The TAQ database is a collection of all trades and quotes in NYSE, American Stock Exchange (AMEX), and National Association of Securities Dealers Automated Quotation (NASDAQ) securities. We only consider trades and quotes on the NYSE. Schwartz (1993) and Hasbrouck, Sofianos, and Sosebee (1993) document NYSE trading and quoting procedures.

Among the 50 stocks with the highest capitalization value on December 31, 1996, eight stocks were randomly selected. The names and some summary statistics are given in Tables 1 and 2. Trades reported within the same second were treated as one trade, with the volumes of the multiple trades aggregated. For the quotes we also filtered out multiple occurrences of quotes. The sample period is the two months from August 4, 1997, to September 30, 1997, which gives a total of 42 trading days. All stocks in the period had more quotes than trades. By comparing the last two columns of Table 1, it appears that less than half of the quotes

Table 1 Selected NYSE stocks.

Company	Symbol	Shares	Value	Trades	All quotes	Changing midquotes
Procter & Gamble Company	PG	694	74,595	46932	59658	28087 (71.7%)
Disney (Walt) Company (The)	DIS	682	47,496	28390	39299	15349 (70.8%)
Federal National Mortgage Ass.	FNM	1129	42,053	24910	34527	10827 (75.6%)
General Motors Corporation	GM	757	42,182	32618	39007	14087 (75.3%)
Bank-American Corporation	BAC	387	38,632	34764	47515	19172 (70.5%)
McDonald's Corporation	MCD	830	37,572	24720	27693	9528 (86.1%)
Monsanto Company	MTC	822	31,954	25324	30208	11607 (76.0%)
Schlumberger Limited	SLB	309	30,848	27787	40363	18177 (66.2%)

This table presents 8 randomly selected stocks from the 50 leading NYSE stocks in market value as of December 31, 1996. Shares and values are in millions. Columns 4 and 5 report the number of trades and quotes after cleaning the data. The last column gives the number of quotes associated with a change in the midquote. The percentages in parentheses are the amount of the quotes in question actually used in defining the forward quote durations.

comprised revisions (changes) of the midquote,[7] the rest of the revisions being pure depth revisions. In Table 3, midquote durations are systematically negatively correlated with lagged volume, measured as shares transacted on the previous trade, negatively correlated with the lagged spread, and positively correlated with trade durations.

3.1 The Bivariate Duration Process

To construct the bivariate duration process of forward trade and quote durations, we begin as outlined in Section 1 by calculating the forward trade durations simply as the intertrade arrival times. The first duration every day is the duration from the second to the third trade that day. Thus the NYSE opening and the high volume associated with this are excluded from the analysis. This is important because the opening trade is fundamentally different from all other trades during the trading day. The procedure is that the market for a stock opens when the specialist finds a price that balances the buy and sell orders for the issue. The specialist does this by matching market orders that come in through the Opening Automated Report Service (OARS), a feature that accepts preopening market orders, public limit orders, and non-OARS-eligible market orders that come into the electronic display book or printers, and orders from the trading crowd. In this way the NYSE opening resembles a call market where orders are batched together for execution at a predetermined time at a single price, in contrast with the rest of the day, which is a continuous market where market orders are executed

[7] It is of no importance whether one consider midquote revisions or spread revisions as these coincide. In the data considered, the specialist almost exclusively changed the bid or the ask, and very rarely both at the same revision.

Table 2 Summary statistics.

Firm	Average forward trade duration	Average price	Average volume	Time S.L.Q.	Sells	Inside	Buys
Panel A: Trade statistics (all durations are in seconds)							
PG	20.7 (24.9)	129.4 (24.7)	1148.1 (2234.7)	39.3 (56.2)	38.0	12.2	49.8
DIS	34.3 (42.9)	78.7 (1.3)	1473.1 (3228.6)	79.7 (107.2)	37.5	17.6	44.9
FNM	38.9 (46.2)	45.5 (1.4)	2915.6 (6672.5)	136.2 (207.0)	39.3	18.3	42.3
GM	29.9 (37.2)	64.9 (2.5)	2459.7 (6899.6)	92.3 (138.8)	36.8	19.5	43.7
BAC	28.0 (36.4)	71.3 (2.8)	1840.9 (4145.9)	65.9 (103.0)	39.1	12.5	48.4
MCD	39.3 (44.0)	48.6 (2.1)	2623.4 (6943.2)	79.7 (107.2)	43.8	14.9	41.3
MTC	38.0 (47.7)	42.8 (3.5)	2950.7 (7749.0)	112.2 (174.1)	35.4	15.7	48.9
SLB	35.0 (44.4)	77.7 (3.7)	1658.6 (3850.7)	70.1 (105.6)	38.2	13.5	48.3

Firm	Average observed forward quote duration	Average forward quote duration	Average quote-quote duration	Average spread (%)	Censored (%)
Panel B: Quote statistics (all durations are in seconds)					
PG	12.9 (29.5)	29.0 (52.9)	34.5 (42.5)	0.142 (0.08)	56.4
DIS	21.5 (44.0)	59.2 (100)	63.2 (85.4)	0.169 (0.08)	61.2
FNM	28.5 (56.3)	100.8 (186)	89.0 (134)	0.234 (0.10)	66.3
GM	21.5 (44.2)	70.2 (122)	68.9 (99.4)	0.173 (0.08)	66.8
BAC	17.0 (31.9)	46.4 (90.3)	50.6 (73.2)	0.203 (0.11)	60.5
MCD	30.5 (64.8)	127.9 (255)	101.3 (161)	0.222 (0.10)	68.2
MTC	27.0 (63.5)	84.7 (169)	83.1 (119)	0.314 (0.18)	64.3
SLB	19.2 (34.1)	47.8 (89.6)	53.3 (74.5)	0.197 (0.11)	56.3

This table gives summary statistics for the datasets. Panel A presents statistics related to trades. Average forward trade duration is the average length of the durations between successive trades. The next column gives the average transaction price, followed by the average size of the amount of shares traded. The fifth column reports the average time since the most recent quote. Columns six through eight report the amount of transactions classified as sells, at midquote, and buys, respectively. Panel B relates to quotes. Average observed forward quote duration is the average length of the truncated forward quote durations. The next column is the average length of the observed forward quote durations, followed by the average length of the quote-quote durations. The last column reports the amount of the observed forward quote durations that were censored. Numbers in parentheses are standard deviations.

Table 3 Simple correlations with forward quote durations.

	$tr.dur_t$		$tr.dur_{t-1}$		$volume_{t-1}$		$spread_{t-1}$		$\Delta spread_{t-1}$	
Firm	r	p	r	p	r	p	r	p	r	p
Panel A: Observed forward quote duration (only changing midquotes included)										
PG	.294	<.001	.060	<.001	−.017	<.001	−.038	<.001	−.003	<.001
DIS	.424	<.001	.066	<.001	−.038	<.001	−.068	<.001	−.017	.002
FNM	.451	<.001	.075	<.001	−.069	<.001	−.077	<.001	−.023	<.001
GM	.431	<.001	.069	<.001	−.023	<.001	−.065	<.001	−.022	<.001
BAC	.500	<.001	.108	<.001	−.042	<.001	−.084	<.001	−.025	<.001
MCD	.406	<.001	.053	<.001	−.052	<.001	−.084	<.001	−.027	<.001
MTC	.346	<.001	.057	<.001	−.048	<.001	−.054	<.001	−.014	.015
SLB	.486	<.001	.092	<.001	−.044	<.001	−.060	<.001	−.015	.005
Panel B: Forward quote duration (only changing midquotes included)										
PG	.144	<.001	.096	<.001	−.024	<.001	−.094	<.001	−.024	<.001
DIS	.172	<.001	.097	<.001	−.048	<.001	−.178	<.001	−.044	<.001
FNM	.164	<.001	.106	<.001	−.082	<.001	−.192	<.001	−.057	<.001
GM	.188	<.001	.137	<.001	−.054	<.001	−.161	<.001	−.049	<.001
BAC	.196	<.001	.130	<.001	−.051	<.001	−.162	<.001	−.044	<.001
MCD	.149	<.001	.102	<.001	−.068	<.001	−.193	<.001	−.050	<.001
MTC	.144	<.001	.087	<.001	−.057	<.001	−.165	<.001	−.027	<.001
SLB	.180	<.001	.108	<.001	−.054	<.001	−.144	<.001	−.033	<.001

This table gives some simple correlations of the forward quote durations with trade durations, lagged trade durations, lagged volume, lagged spread, and the change in the lagged spread. The p values are calculated using that $r(\sqrt{N-2})/(\sqrt{1-r^2}) \sim N(0, 1)$ holds asymptotically.

immediately upon submission. This motivates why the opening requires special attention and modeling, a task that is beyond the scope of this article.

The relation between the size of a trade and its impact on market liquidity is complex. Large orders are usually broken into smaller trades, although sometimes they are traded directly through the upstairs market. We reduce the impact of large trades by using a square root transformation. This is a compromise between excluding them entirely and modeling them separately. This has been used by many authors [see Hasbrouck (1996b)].

Overnight durations were also omitted from the sample. Now for every trade, the prevailing quote is the most recent quote that occurred at least five seconds before the trade. As on the NYSE floor, where posting new quotes is given priority over recording completed trades, a quote revision will often precede the trade from which it was instigated. Thus to compute the forward quote durations we delay every quote time five seconds and then the forward quote duration is the time from the present trade to the next quote. Matching trades with quotes in this way overcomes the concern for mistimed recordings. This "five-second rule" was suggested by Lee and Ready (1991). To build the observed forward quote durations, every pair of forward trade and quote durations are compared. If the forward quote is longer than the forward trade duration, the observed forward quote duration is censored. The magnitude of censoring is reported in panel B of Table 2. About 62.5% of the forward quote durations are censored.

As discussed earlier in the article the forward quote duration series ignores some quotes by construction. Table 1 shows that about 75% of the quotes are preserved. This means about 25% of the midquote revisions had a new midquote revision *without* an intervening trade. To assess the importance of this we calculated summary statistics for these quote-quote durations. The calculations showed that the average of the median durations is 10.5 seconds and that 75% of these durations are *less* than 26.1 seconds. This is vastly different from the statistics for quote-quote durations *with* intervening trades. For these durations, the average of the medians is 58.75 seconds, and 75% are more than 29.75 seconds. This gives fair support for our intuition that a quote following a trade starts the specialists updating process manifesting itself into one or a series of rapidly following quotes.

3.2 Explanatory Variables

With the dependent process defined, we need to specify the explanatory variables to be put into $\mathbf{Z}_{i-1}$ and $\mathbf{V}_{i-1}$. There are many possibilities: more lagged values of the dependent process, the time since the most recent quote, the spread, the volume, etc. Of course, it is preferable to include variables that have economic interpretations. There is a huge amount of literature on the importance of information contained in spreads and trading volume which is nicely reviewed in Karpoff (1987), Hasbrouck (1996), Goodhart and O'Hara (1997), Coughenour and Shastri (1999), and Madhavan (2000). The recurrent theme is that high trading volume and high spreads are associated with periods with a high rate of information flow to the market. In Easley O'Hara (1992), a high intensity of trading and price revisions is again associated with high information flow. We test this connection using the explanatory variables defined below. We give further interpretations of the included variables when we discuss the results of the estimation.

Hereafter we denote observations coming at time t_i simply with subscript i. Hence we define

$$\mathsf{p}_i \equiv \text{price of shares traded at time } t_i,$$

and

$$\mathsf{vol}_i \equiv \text{number of shares traded at time } t_i,$$

of which $\sqrt{\mathsf{vol}_i}$ is used to down-weight very large trades. For the quoting information we have the relative spread, defined as

$$\mathsf{spr}_i \equiv 100(\ln \mathsf{ask}_i - \ln \mathsf{bid}_i),$$

where the indices refer to the quote prevailing at time t_i. In the above definition, the relative spread is preferred to the nominal spread ($\mathsf{ask}_i - \mathsf{bid}_i$) because it is dimensionless and can therefore be directly compared between different stocks.[8] To smooth out the effect of very-short-term bouncing, we compute the following equally weighted moving average:

$$\mathsf{lev.spr}_i \equiv \text{average of the 10 most recent spreads.}$$

As the average quote-quote durations range from 34.5 to 101.3 seconds, $\mathsf{lev.spr}_i$ ranges from an average of 5 to 15 minutes. To capture short-run moments in the spread, the change in the spread from trade to trade was included as well, this is $\Delta\mathsf{spr}_i \equiv \mathsf{spr}_i - \mathsf{spr}_{i-1}$.

Trades may be classified into buys and sells using the technique of Lee and Ready (1991). Trades at prices above the midquote are associated with buys (initiated by a buyer) and are given a value of 1; trades below the midquote are called sells (initiated by a seller) and are given a value of -1. This variable is often referred to as a buy-sell trade indicator. The rational for this classification is that trades originating from buyers are most likely to be executed at or near the ask, while sell orders trade at or near the bid. This scheme classifies all trades except those that occur at the midquote. We do not apply the *tick* rule; trades at the midquote are given a value of zero. This results in the following variable

$$\mathsf{sign}_i = \begin{cases} 1 & \text{if } \mathsf{p}_i > (\mathsf{ask}_i + \mathsf{bid}_i)/2 \\ 0 & \text{if } \mathsf{p}_i = (\mathsf{ask}_i + \mathsf{bid}_i)/2 \\ -1 & \text{if } \mathsf{p}_i < (\mathsf{ask}_i + \mathsf{bid}_i)/2. \end{cases}$$

On average, 15.9% of the volumes traded are given a value of zero. Using these values, we compute the accumulated signed volume. This is calculated using a moving window of 10 trades. We include the absolute value of this variable as an explanatory variable, that is,

$$\mathsf{abs}(\mathsf{s.vol})_i \equiv \left| \sum_{j=1}^{10} \mathsf{sign}_{i-j} \mathsf{vol}_{i-j} \right|.$$

This variable is a measure of the imbalance of trades over approximately 5-minute intervals. It is related to the depth measure VNET introduced by Engle and Lange (2001). Giving the sizes corresponding to trades executed at the mid-quote zero weight excludes these sizes from the constructed variable. As the trades at the midquote often correspond to crossed orders, these do not contribute to imbalance of the specialists inventory, because he is not trading on his own account.

To incorporate the effects of periods of infrequent price revisions, we include the backward time to the most recent quote, denoted $\mathsf{Back.Q}_i$, and

$$\mathsf{lev.QQ}_{i-1} \equiv \text{average of the 10 most recent quote-quote durations.}$$

[8] It is important to note that variables are lagged with respect to the trade time. Hence the third lagged spread would be the prevailing spread three trades ago.

Table 4 Summary of dependent and explanatory variables.

Coefficient	Name	Description	Units
Panel A: Dependent			
—	ψ_i	Expected forward trade duration (time between trades)	30 seconds
—	φ	Expected forward quote duration (time from trade to quote)	30 seconds
Panel B: Explanatory			
β_1, η_3	lev.QQ_{i-1}	Mean of the 10 most recent QQ durations	30 seconds
β_2, η_4	ΔSpr_{i-1}	Change in the relative spread, from the trade which initialized the present duration and the most recent trade	Percent
β_3, η_5	lev.Spr_{i-1}	Mean of the 10 most recent relative spreads	Percent
β_4, η_6	$\sqrt{\text{vol}}_{i-1}$	Square root of the size of the trade which initialized the present duration	100 shares
β_5, η_7	$\text{abs}(\text{s.vol})_{i-1}$	Absolute value of accumulated signed size of the 10 previous trades	100 shares
δ_1	$\tilde{y}_{i-1}/\varphi_{i-1}$	Observed forward quote duration relative to expected	Fraction
δ_2	$d_{i-1}\tilde{y}_{i-1}/\varphi_{i-1}$	Censored observed forward quote duration relative to expected	Fraction
τ	x_i/ψ_i	Forward trade duration relative to expected	Fraction
η_1, γ	x_{i-1}/ψ_{i-1}	Lagged forward trade duration relative to expected	Fraction
η_2	ψ_{i-1}	Lagged expected forward trade duration	30 seconds
η_8	Back.Q_{i-1}	Time from the most recent quote to the trade initializing the present duration	30 seconds
$\kappa_{j=1,\ldots,6}$	$\text{D}_{i-1}^{j=1,\ldots,6}$	Six hourly time-of-day dummies	—

This table summarizes the units and definitions of the dependent and the explanatory variables included in the analysis. The β and the η coefficients refers to lagged information sets, $\boldsymbol{Z}_{i-1}$ and $\boldsymbol{V}_{i-1}$, respectively.

Like lev.spr_i this gives an average over approximately 5 to 15 minutes, depending on the speed of the market.

In the definitions of lev.spr, abs(s.vol), and lev.QQ we have used a fixed number of lagged observations; 10 in each case. One could also take the approach of using the observations in a fixed time window, but this would cause these variables to be based on a very different number of observations across the day and across stocks. The specific choice of 10 lagged observations was selected to avoiding oversmoothing microstructure effects. Table 4 summarizes the computation and explanation of the explanatory variables associated with the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\eta}$ of Equations (3) and (6).

Typically the market exhibits high activity in the morning and before closure. Around lunchtime, the activity is mostly lower. The daily pattern in the dependent variable must be reflected in patterns in the independent variables. We take the approach of introducing six hourly dummy variables into the trade and the quote equations. This is a particularly convenient solution, since the censoring of quotes by intervening trades could be seriously distorted by attempting to adjust each separately. Specifically the dummies are denoted $\mathsf{D}_i^1, \ldots, \mathsf{D}_i^6$ and take the value one if the duration initiated at time t_i started between 10:00–11:00, …, 15:00–16:00, respectively.[9]

The estimated functional forms are given as follows:

$$\begin{aligned}\ln(\psi_i) = \alpha &+ \delta\ln(\psi_{i-1}) + \gamma\frac{x_{i-1}}{\psi_{i-1}} + \beta_1\mathsf{lev.QQ}_{i-1} + \beta_2\Delta\mathsf{Spr}_{i-1}\\ &+ \beta_3\mathsf{lev.Spr}_{i-1} + \beta_4\sqrt{\mathsf{vol}}_{i-1} + \beta_5\mathrm{abs}(\mathsf{s.vol})_{i-1} + \beta_6\mathsf{Back.Q}_{i-1}\\ &+ \kappa_1\mathsf{D}_i^1 + \kappa_2\mathsf{D}_i^2 + \kappa_3\mathsf{D}_i^3 + \kappa_4\mathsf{D}_i^4 + \kappa_5\mathsf{D}_i^5 + \kappa_6\mathsf{D}_i^6 \end{aligned} \tag{10}$$

and

$$\begin{aligned}\ln(\varphi_i) = \mu &+ \rho\ln(\varphi_{i-1}) + \delta_1\frac{\tilde{y}_{i-1}}{\varphi_{i-1}} + \delta_2\frac{\tilde{y}_{i-1}}{\varphi_{i-1}}d_{i-1} + \tau\frac{x_i}{\psi_i} + \eta_1\frac{x_{i-1}}{\psi_{i-1}}\\ &+ \eta_2\ln(\psi_{i-1}) + \eta_3\mathsf{lev.QQ}_{i-1} + \eta_4\Delta\mathsf{Spr}_{i-1} + \eta_5\mathsf{lev.Spr}_{i-1}\\ &+ \eta_6\sqrt{\mathsf{vol}}_{i-1} + \eta_7\mathrm{abs}(\mathsf{s.vol})_{i-1} + \eta_8\mathsf{Back.Q}_{i-1}\\ &+ \varkappa_1\mathsf{D}_i^1 + \varkappa_2\mathsf{D}_i^2 + \varkappa_3\mathsf{D}_i^3 + \varkappa_4\mathsf{D}_i^4 + \varkappa_5\mathsf{D}_i^5 + \varkappa_6\mathsf{D}_i^6 \end{aligned} \tag{11}$$

(see Table 4 for the definition of the explanatory variables). Table 3 gives some simple correlations of the forward quote durations and the explanatory variables in the quote equations. Note that strong correlations are present for the actual uncensored forward quote duration. We will see these correlations reflected in the estimated models.

4 ESTIMATION AND RESULTS

Maximization of the log-likelihood as outlined in Section 1.2 was performed in C++ using the simplex method as found in Press et al. (1992). We often formulate the discussion in terms of intensities, which are the reciprocal of the expected durations, since these may have particularly clear interpretations. In the next two sections, the model estimated using the midquote arrivals is presented and discussed.

[9] Another approach is often adopted from Engle and Russell (1998). They suggested making a multiplicative diurnal adjustment using a cubic spline. They estimate their model both in two steps and simultaneously, and reported that the results remained stable. Recently a highly sophisticated semiparametric method has been suggested by Veredas, Rodriguez-Poo and Espasa (2001).

4.1 The Trade Equation

The estimates of the trade equation are presented in Table 5. We present the equation (suppressing the dummies), stressing the signs of the coefficients and putting in boldface type the ones that are generally significant at the 5% level:

$$\ln(\psi_i) = \overset{-}{\boldsymbol{\alpha}} + \overset{+}{\boldsymbol{\delta}} \ln(\psi_{i-1}) + \overset{+}{\boldsymbol{\gamma}} \frac{x_{i-1}}{\psi_{i-1}} + \overset{+}{\beta_1} \text{lev.QQ}_{i-1} + \overset{+}{\boldsymbol{\beta}_2} \Delta\text{Spr}_{i-1}$$
$$+ \overset{-}{\beta_3} \text{lev.Spr}_{i-1} + \overset{-}{\boldsymbol{\beta}_4} \sqrt{\text{vol}}_{i-1} + \overset{+}{\beta_5} \text{abs(s.vol)}_{i-1} + \overset{-}{\beta_6} \text{Back.Q}_{i-1}.$$

The trading intensity shows a very high degree of persistence, with δ greater than 0.95 for most stocks. All stocks have γ, the coefficient on the surprise term, positive and significant. Hence these two parameters are as expected from earlier studies.

The most conspicuous of the explanatory variables is the square root of the volume. The systematically and significantly negative coefficients reveal that large trades initiate shorter durations than small trades so that the arrival intensity rises after a big trade. Presumably this is because large trades are more likely to be information based and therefore signal the presence of information trading [as, e.g., in Easley and O'Hara (1987)]. Since informed traders will be in a hurry to exploit deteriorating informational advantages, information flow leads to fast trading.

A high bid–ask spread is another sign of informational traders, and it has a predominantly negative but not significant sign in the level, as would be expected. However, it has a positive and generally significant coefficient in the change. Thus rising spreads initially discourage trades, but once the spread has risen permanently to a new higher level, trades can increase again.

Lagged quote durations have a mixed positive impact on trade durations and the time since the last quote has a negative and systematic effect on trade arrivals; both are of mixed significance. When quotes have not been revised for a long time, trade intensities rise. When demand and supply are out of balance, measured by abs(s.vol)_{i-1}, there is a positive but not significant effect on the duration to the next trade.

The diagnostic tests in panel B of Table 5 are very well behaved. The Ljung–Box tests for no autocorrelation in the residuals, $LB(\xi)$, are dramatically reduced from the corresponding Ljung–Box statistic, $LB(X)$, for the raw durations. They are potentially clean, depending upon the appropriate p value for such large sample sizes. Further, only a first-order model is estimated, and probably higher-order models would be even more satisfactory from this point of view.

Some excess dispersion is still present, as the standard deviation of the residuals exceeds 1. To assess the significance of this excess dispersion we apply a simple test suggested by Engle and Russell (1998). The null of no excess dispersion is based on the statistics $\sqrt{N}((\hat{\sigma}_\xi^2 - 1)/\sigma_\nu)$, where $\hat{\sigma}_\xi^2$ is the sample variance of $\hat{\xi}$, which should be one under the null hypothesis. σ_v is the standard deviation of $(\xi - 1)^2$, which equals $\sqrt{8}$ under the null of a unit exponential distribution. This statistic has a limiting normal distribution under the null with a 5% critical value of 1.645. Performing this test on our samples reveals that excess dispersion is still

Table 5 Estimates for the trade equation.

Firm	$\hat{\alpha}$ $(t_{\alpha=0})$	$\hat{\delta}$ $(t_{\delta=1})$	$\hat{\gamma}$ $(t_{\gamma=0})$	$\hat{\beta}_1$ $(t_{\beta_1=0})$	$\hat{\beta}_2$ $(t_{\beta_2=0})$	$\hat{\beta}_3$ $(t_{\beta_3=0})$	$(\times 10^{-1})$ $\hat{\beta}_4$ $(t_{\beta_4=0})$	$(\times 10^{-3})$ $\hat{\beta}_5$ $(t_{\beta_5=0})$	$(\times 10^{-2})$ $\hat{\beta}_6$ $(t_{\beta_6=0})$
Panel A: *Parameter estimates*									
PG	**−.0439** (−6.90)	**.972** (5.60)	**.0442** (11.7)	**.5618** (3.86)	**.0695** (4.19)	−.0191 (−1.25)	**−.0295** (−6.68)	.0304 (1.86)	−.0072 (−0.14)
DIS	−.0072 (−1.57)	**.981** (3.23)	**.0340** (6.82)	.0443 (0.57)	.0230 (1.55)	−.0564 (−2.42)	**−.0183** (−3.67)	.0076 (0.77)	−.0817 (−2.31)
FNM	.0047 (0.60)	**.950** (3.94)	**.0374** (8.13)	**.2550** (2.62)	**.0498** (3.22)	−.0431 (−1.66)	**−.0255** (−5.27)	.0144 (2.20)	−.0259 (−1.42)
GM	**−.0182** (−5.80)	**.993** (3.74)	**.0250** (9.54)	−.0124 (−0.29)	−.0078 (−0.91)	−.0018 (−0.14)	**−.0043** (−2.88)	.0011 (0.35)	−.0303 (−2.05)
BAC	**−.0293** (−4.85)	**.978** (3.74)	**.0393** (7.76)	.0575 (0.63)	.0055 (0.85)	−.0056 (−0.47)	**−.0165** (−4.27)	.0058 (0.78)	−.0291 (−1.07)
MCD	.0002 (0.04)	**.985** (3.68)	**.0279** (7.09)	.0225 (0.84)	**.0333** (3.34)	**−.0528** (−3.31)	**−.0138** (−4.96)	**.0112** (2.63)	−.0015 (−0.22)
MTC	−.0050 (−1.73)	**.992** (4.82)	**.0256** (8.11)	.0269 (−1.01)	.0108 (2.28)	−.0083 (−1.46)	**−.0116** (−5.79)	.0025 (0.75)	−.0024 (−0.21)
SLB	**−.0209** (−6.59)	**.993** (3.97)	**.0255** (8.83)	.0213 (0.42)	.0048 (0.52)	−.0122 (−1.15)	−.0024 (−1.25)	−.0143 (−2.14)	−.0341 (−1.52)

							Panel B: Diagnostics			
Firm	$(\times 10^{-2})$ $\hat{\kappa}_1$ $(t_{\kappa_1=0})$	$(\times 10^{-2})$ $\hat{\kappa}_2$ $(t_{\kappa_2=0})$	$(\times 10^{-2})$ $\hat{\kappa}_3$ $(t_{\kappa_3=0})$	$(\times 10^{-2})$ $\hat{\kappa}_4$ $(t_{\kappa_4=0})$	$(\times 10^{-2})$ $\hat{\kappa}_5$ $(t_{\kappa_5=0})$	$(\times 10^{-2})$ $\hat{\kappa}_6$ $(t_{\kappa_6=0})$	$LB(X)$	$LB(\xi)$	s_{ξ}	E-R EDT
PG	.1244 (0.71)	**.7032** (3.09)	**1.1666** (4.09)	.6218 (2.39)	−.0058 (−0.04)	**−.4526** (−2.65)	**7185.0**	28.9	1.0716	**11.35**
DIS	.2701 (1.81)	**.5870** (3.18)	**.8764** (3.01)	.3642 (1.53)	−.1500 (−1.07)	−.3458 (−2.43)	**2160.6**	9.2	1.1764	**22.86**
FNM	.3243 (1.35)	**1.1172** (5.71)	**1.7546** (5.29)	**1.0728** (2.97)	−.1105 (−0.62)	−.4046 (−1.93)	**1625.0**	11.4	1.1162	**13.71**
GM	.1433 (1.54)	.1858 (1.96)	.2404 (2.17)	−.0119 (−0.11)	−.2173 (−2.23)	**−.3355** (−3.71)	**3276.7**	**38.7**	1.1340	**18.25**
BAC	**.5546** (2.84)	**1.4134** (4.43)	**1.5759** (3.84)	.8488 (2.30)	.3478 (1.45)	−.0540 (−0.27)	**4810.0**	29.5	1.1669	**23.83**
MCD	.0715 (0.41)	.2183 (1.17)	.4828 (2.56)	.0370 (0.16)	−.2328 (−1.40)	−.3851 (−2.54)	**2113.1**	**36.0**	1.0435	**4.94**
MTC	.1394 (0.91)	.1371 (0.95)	.1745 (1.03)	.0224 (0.12)	−.3363 (−2.36)	**−.6080** (−4.31)	**3892.3**	30.4	1.1117	**13.27**
SLB	.3147 (2.19)	**.5417** (3.87)	**.6108** (3.26)	.2231 (1.38)	.0119 (0.08)	−.0142 (−0.10)	**3021.0**	**51.2**	1.1555	**19.75**

This table reports estimation results for the trade equation defined in Equation (10). Panel A gives the parameter estimates with t-statistics reported in parentheses. The parameters $\kappa_1, \ldots, \kappa_6$ represent the possible diurnal effect of six hourly dummies, 10:00–11:00, 11:00–12:00, 12:00–13:00, 13:00–14:00, 14:00–15:00, and 15:00–16:00. In panel B some diagnostics are given. In the first column $LB(X)$ is short for the Ljung–Box test with 15 lags on X, the second column is likewise for ξ. E-R EDT is short for Engle–Russell excess dispersion test. Numbers in boldface are significant at the 99% level, numbers in normal type are significant at the 95% level. The numbers typed with very small types are insignificant.

left in the residuals. In Lunde (1998), a generalized gamma version of the trade equation is estimated. This model is able to remove the excess dispersion completely. But our model is complex enough as it is, and our experiments showed that it did not change any of the conclusions.

In Table 6 we present some diagnostics based on the joint estimation. A LM test, as described in Section 2.3, for a single lagged quote duration proved insignificant. This shows that quote timing, apart from the time since the last quote, is of little importance for the trading intensity. In Table 6 we also report the results of the likelihood ratio test for the joint significance of the diurnal dummies, $\kappa_1, \ldots, \kappa_6$. It appears that they are highly significant in all cases. The piecewise linear time-of-day splines based on the coefficients $\mu, \kappa_1, \ldots, \kappa_6$ are plotted in Figure 1. The inverse U-shaped pattern confirms the findings of previous studies.

Overall the trade equation is consistent with our economic models and quite stable across stocks. It reveals the importance of volume and spreads in predicting trade intensities, and finds that quote revisions are less important.

Before we turn our focus to the quote equation, it seems appropriate to try to quantify the economic significance of the estimated relationship. Using the units of the explanatory variables presented in Table 4, we can ask for the effect of an increase in, for example, volume on the expected duration. Since the dependent variable is the log of the expected waiting time, all coefficients can be interpreted as the percentage change due to a unit change in the regressor. For volume, the estimated coefficient for FNM is -0.00255. Hence if $\sqrt{\text{vol}_i}$ quadruples from one trade to the next (moving from a trade of 100 shares to 1600 is not an usual event for the stocks under study), then the expected duration will decrease with $4 \cdot 0.00255$ (or 1.02%). On average this amounts to $38.9 \cdot 0.0102 = 0.4$ second within

Table 6 Joint diagnostics.

Firm	Max like	LM-trade	LM-quote	LR-n.d.tr.	LR-n.d.qu.	LR-n.d.both
PG	−44408.61	*.09*	*2.16*	***156.42***	***79.02***	***243.38***
DIS	−46804.91	*.03*	*1.57*	***127.20***	***31.76***	***155.04***
FNM	−45191.54	*1.00*	6.17	***126.98***	***82.24***	***192.04***
GM	−47903.60	*1.53*	4.78	***104.82***	***20.28***	***127.48***
BAC	−46222.24	*1.23*	*.11*	***203.84***	***47.06***	***255.50***
MCD	−44908.55	*.37*	4.97	***74.56***	***44.88***	***120.60***
MTC	−44545.75	*.13*	*.01*	***142.50***	***51.54***	***220.98***
SLB	−45583.72	*.24*	***9.43***	***105.40***	***32.26***	***137.20***

For the joint model presented in Tables 5 and 7.

Tables 5 and 7 give several diagnostics for the quote equation corresponding to the model defined in this table. In the first column, $LB(\tilde{Y})$ is short for the Ljung–Box test with 15 lags on $\tilde{Y}$, the second column is likewise for ε. LM is the Lagrange multiplier test for the validity of excluding the most recent quote-quote duration, one test for the trade equation and one for the quote equation. LR is a likelihood ratio test for the validity of excluding diurnal dummies. LR-n.d.tr. excludes dummies from the trade equation, LR-n.d.qu. from the quote equation, and LR-n.d.both excludes dummies from both equations. Numbers in boldface are significant at the 99% level, numbers in normal type are significant at the 95% level. The numbers in italic are insignificant.

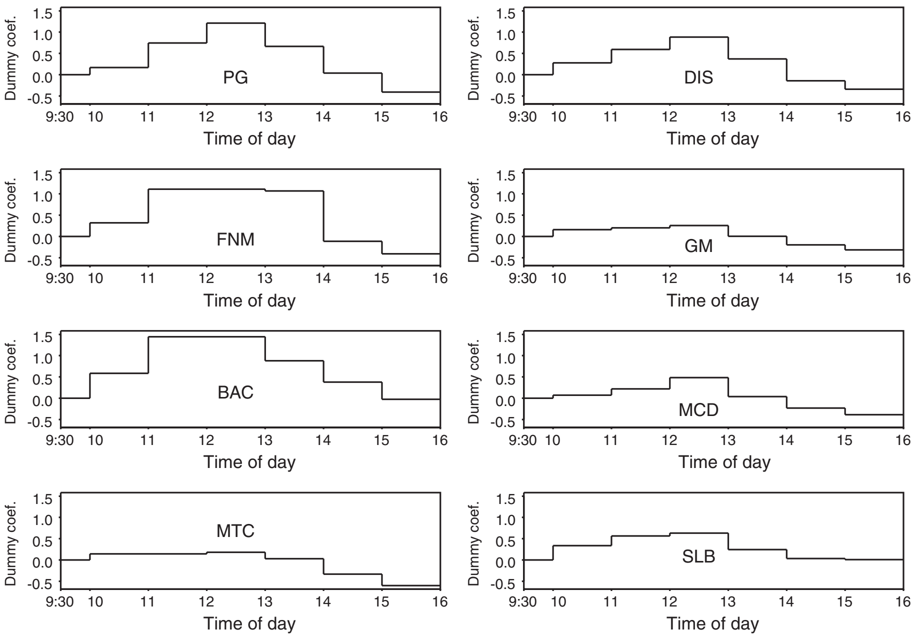


Figure 1 Piecewise linear time-of-day spline for the trade equation.

one trade. However, because of the lagged dependent variable, the long-run effect will be $\beta_4/(1-\rho)$, and as $\hat{\rho} = 0.95$, the long-run effect would be $4 \cdot 0.00255/0.05$ (i.e., 20.4%). On average this amounts to 8 seconds, and will be even more in periods where the trading intensity is below average. Carrying out these calculations for $\Delta\mathsf{Spr}_i$, just increasing it by 50%, shows that the short-run effect is an increase of $0.5 \cdot 0.0498$ (i.e., 2.5%) amounting to about 1 second on average. By consulting panel B of Table 2, it appears that a change of more than 50% in the relative spread is unlikely. We will now turn our attention to the quote equation, where effects of much larger magnitude will appear.

4.2 The Quote Equation

Table 7 reports the estimates of the quote equation estimated for midquote arrivals. The dependent variable can be interpreted as the time from a trade to a midquote revision, taking in to account the censoring effect of a trade that arrives before the next quote. The equation is reproduced here (with the dummies suppressed) with the typical sign.

$$\begin{aligned}\ln(\varphi_i) = {} & \overset{+}{\boldsymbol{\mu}} + \overset{+}{\boldsymbol{\rho}} \ln(\varphi_{i-1}) + \overset{+}{\boldsymbol{\delta}}_1 \frac{\tilde{y}_{i-1}}{\varphi_{i-1}} + \overset{-}{\boldsymbol{\delta}}_2 \frac{\tilde{y}_{i-1}}{\varphi_{i-1}} d_{i-1} + \overset{+}{\boldsymbol{\tau}} \frac{x_i}{\hat{\psi}_i} + \overset{\pm}{\eta}_1 \frac{x_{i-1}}{\hat{\psi}_{i-1}} \\ & + \overset{+}{\eta}_2 \ln(\hat{\psi}_{i-1}) + \overset{+}{\boldsymbol{\eta}}_3 \mathsf{lev.QQ}_{i-1} + \overset{-}{\boldsymbol{\eta}}_4 \Delta\mathsf{Spr}_{i-1} + \overset{-}{\boldsymbol{\eta}}_5 \mathsf{lev.Spr}_{i-1} \\ & + \overset{-}{\boldsymbol{\eta}}_6 \sqrt{\mathsf{vol}}_{i-1} + \overset{+}{\boldsymbol{\eta}}_7 \mathrm{abs}(\mathsf{s.vol})_{i-1} + \overset{+}{\boldsymbol{\eta}}_8 \mathsf{Back.Q}_{i-1}\end{aligned}$$

Table 7 Estimates for the quote equation.

Firm	$\hat{\mu}$ $(t_{\mu=0})$	$\hat{\rho}$ (SE)	$\hat{\delta}_1$ $(t_{\delta_1=0})$	$\hat{\delta}_2$ $(t_{\delta_2=0})$	$\hat{\tau}$ $(t_{\tau=0})$	$\hat{\eta}_1$ $(t_{\eta_1=0})$	$\hat{\eta}_2$ $(t_{\eta_2=0})$	$\hat{\eta}_3$ $(t_{\eta_3=0})$	$\hat{\eta}_4$ $(t_{\eta_4=0})$	$\hat{\eta}_5$ $(t_{\eta_5=0})$
PG	**−.322** (−6.92)	.616 (0.06)	**.270** (8.66)	**−.097** (−2.90)	**.125** (17.0)	.013 (0.73)	−.017 (−0.59)	**.099** (5.53)	**−.901** (−4.70)	−.251 (−1.93)
DIS	**.207** (4.78)	.589 (0.05)	**.253** (10.0)	−.044 (−1.23)	**.209** (21.6)	−.042 (−1.97)	.009 (0.25)	**.028** (3.63)	**−1.49** (−7.04)	**−1.69** (−7.99)
FNM	**.487** (8.96)	.505 (0.04)	**.349** (9.63)	−.076 (−1.64)	**.249** (20.9)	.025 (1.20)	−.051 (−0.75)	**.029** (4.68)	**−1.38** (−8.10)	**−1.83** (−10.9)
GM	.060 (1.38)	.550 (0.03)	**.367** (11.5)	−.108 (−2.42)	**.184** (19.9)	**.044** (2.80)	.062 (1.64)	**.029** (4.31)	**−1.72** (−10.3)	**−.747** (−4.88)
BAC	.054 (1.18)	.517 (0.04)	**.373** (9.08)	**−.130** (−3.10)	**.209** (24.2)	.017 (1.10)	.087 (2.32)	**.053** (5.87)	**−1.29** (−9.96)	**−1.14** (−8.15)
MCD	**.423** (4.94)	.481 (0.04)	**.338** (9.70)	−.139 (−2.55)	**.303** (22.3)	−.034 (−1.51)	.012 (0.23)	**.025** (4.28)	**−2.05** (−10.3)	**−1.62** (−5.01)
MTC	.088 (1.87)	.528 (0.05)	**.316** (9.22)	−.082 (−1.76)	**.225** (19.7)	−.018 (−0.79)	−.050 (−1.70)	**.045** (6.46)	**−.703** (−5.44)	**−.539** (−5.26)
SLB	−.077 (−1.58)	.396 (0.04)	**.260** (10.9)	−.077 (−2.45)	**.172** (19.0)	−.001 (−0.01)	**.139** (3.46)	**.048** (4.59)	**−.783** (−4.83)	**−.693** (−4.27)

Firm	$(\times 10^{-1})$ $\hat{\eta}_6$ $(t_{\eta_6=0})$	$(\times 10^{-3})$ $\hat{\eta}_7$ $(t_{\eta_7=0})$	$(\times 10^{-1})$ $\hat{\eta}_8$ $(t_{\eta_8=0})$	$\hat{\varkappa}_1$ $(t_{\varkappa_1=0})$	$\hat{\varkappa}_2$ $(t_{\varkappa_2=0})$	$\hat{\varkappa}_3$ $(t_{\varkappa_3=0})$	$\hat{\varkappa}_4$ $(t_{\varkappa_4=0})$	$\hat{\varkappa}_5$ $(t_{\varkappa_5=0})$	$\hat{\varkappa}_6$ $(t_{\varkappa_6=0})$
PG	**−.119** (−6.53)	.113 (1.62)	**.328** (5.27)	−.028 (−1.72)	−.019 (−0.95)	.035 (1.57)	−.008 (−0.36)	**−.064** (−3.27)	.018 (1.02)
DIS	**−.215** (−12.9)	**.393** (4.38)	**.177** (5.02)	.041 (1.86)	.056 (2.20)	**.096** (3.27)	**.071** (2.62)	.018 (0.80)	**.073** (2.69)
FNM	**−.263** (−22.9)	**.138** (2.74)	**.188** (6.91)	**.105** (3.21)	**.159** (4.31)	**.219** (5.21)	.101 (2.50)	.027 (0.81)	**.140** (3.80)
GM	**−.170** (−18.9)	**.206** (4.09)	**.215** (7.90)	**.061** (2.65)	.049 (1.82)	.061 (1.99)	.074 (2.50)	.019 (0.75)	**.071** (2.75)
BAC	**−.193** (−16.8)	.108 (1.94)	**.239** (7.14)	−.059 (−2.48)	−.035 (−1.15)	.032 (0.93)	−.048 (−1.51)	**−.084** (−3.21)	−.004 (−0.17)
MCD	**−.242** (−16.2)	**.250** (4.11)	**.212** (7.72)	**.088** (2.68)	**.121** (3.31)	**.161** (4.00)	**.127** (3.12)	.065 (1.90)	**.152** (3.76)
MTC	**−.172** (−15.3)	.054 (1.35)	**.199** (6.50)	**.090** (2.87)	**.117** (3.43)	**.114** (3.21)	**.132** (3.52)	.080 (2.37)	**.170** (4.28)
SLB	**−.204** (−15.1)	**.380** (4.18)	**.362** (8.81)	.067 (2.32)	.061 (1.88)	**.099** (2.58)	.074 (2.02)	.030 (0.92)	**.122** (3.47)

This table reports estimation results for the quote equation defined in Equation (11). The parameters $\varkappa_1, \ldots, \varkappa_6$ represent the possible diurnal effect of six hourly dummies, 10:00–11:00, 11:00–12:00, 12:00–13:00, 13:00–14:00, 14:00–15:00, and 15:00–16:00. t-statistics are reported in parentheses. Numbers in boldface are significant at the 99% level, numbers in normal type are significant at the 95% level. The numbers typed with very small types are insignificant. This does not apply to ρ.

(the variable definitions are given in Table 4). The conditional intensity is again persistent, but much less so than in trades as judged by the autoregressive coefficient. The lagged standardized forward quote duration is very significant, and there is some evidence that censored durations need to be treated differently from uncensored durations.

Trade timing is very important in explaining quote timing. The current standardized forward trade duration is probably the most significant variable, which is not particularly surprising. Not because it is also the censoring threshold (remember that φ_i is the conditional expectation of Y and not $\tilde{Y}$), but because we should expect both the forward trade and the forward quote duration to be driven by the same informational (common) factor, and hence highly correlated.[10]

Although the lagged value of the standardized forward trade duration is negative for some stocks, the total effect of lagged trade durations is also positive. Recognizing that this is a distributed lagged relation between quote durations and standardized trade durations, the lag coefficients are readily calculated. The impact coefficient is simply τ, the first lag effect is $\rho\tau + \eta_1$, which is positive for all stocks, and the long-run effect is $(\tau + \eta_1)/(1-\rho)$, which is also clearly positive for all stocks. Thus short trade durations predict short quote durations. Similarly, because η_2 is positive, short expected trade durations also lead to short quote durations. In contrast with the trade equation, the level of quote durations is also very significantly positive.

The spread variables enter in a very interesting fashion. Both whenever the spread has recently increased or when the level of the spread is high, quotes are more rapidly adjusted. Conversely, recent decreases in the spread lead to slow adjustment, as well as permanent decreases. Presumably the specialist prefers a wide spread in order to profit from his monopoly position. Hence he may choose to leave the spread wide as long as possible until either demand falls or competition from the limit order book rises to erode this position. We see, from the trade equation, that a temporary increase in spreads slows trades, but we don't see the response of the limit order book. Lange (1997) shows that limit orders arrive quickly when spreads temporarily widen, but slow down for permanently wide spreads.

In summary, these effects reveal that rising and permanently high spreads are associated with fast price adjustment. This is consistent with both theoretical and empirical models. For example, the Easley and O'Hara (1992) model, as discussed by Dufour and Engle (2000), finds that when the proportion of informed traders is high, then the spread will be high and the price impact of a trade will be high. This will appear as large price revisions following trades when spreads are high. Since only price revisions that exceed the tick size will appear as new quotes, this matches the empirical finding. Similarly, using different empirical methods, Dufour and Engle (2000) and Engle and Lange (2001) show that market liquidity, whether measured by price impact or depth, is generally reduced when markets are faster. Engle (2000) finds that price volatility is greater when time between trades is shorter and when spreads are wider.

---

[10] We are greatful to Larry Glosten for making this point.

The strong negative volume effect is again consistent with both theoretical and empirical models. Large volume is typically a signal of informational trading and consequently should imply high price impacts and price volatility, which in this continuous-time framework means short quote durations. It is also consistent with an inventory model in which a large-volume trade is more likely to push the specialist away from his preferred inventory position and induce rapid quote revisions. Finally, large volume is consistent with a sticky limit order book model, since a large volume is more likely to eliminate the innermost price and lead to an immediate change in quotes.

The final two variables, the buy-sell imbalance and time since last quote, are both very significantly positive, although there appear to be theoretical arguments to support either sign. Trade imbalance can be interpreted as evidence that liquidity suppliers are willing to tolerate one-sided demands and continue to take the other side. Hence it is an indicator that asymmetric information is not a current fear and consequently price changes can be postponed. Similarly, long times since prices were revised can be interpreted as evidence that there is little information flow and therefore even longer durations can be expected. This is an implication of a declining hazard function which has frequently been found in this literature [see, e.g., Engle and Russell (1998)]. There is presumed to be some latent variable that we might call liquidity, that is only measured by its consequences, such as the spread or order imbalance. Of course, both of these variables can alternatively be interpreted as evidence that prices need to be changed.[11]

In Table 8, several diagnostic tests are presented for the quote equation. The extent of autocorrelation in quote durations in the first column is quite extraordinary; even though the second column reveals that the final model does not pass at a 1% level, the improvement is dramatic. To reduce the residual autocorrelation even further, we investigated several specifications with more lags of the expected and observed duration variables. In some cases these reduced the correlations, and had minimal effects on the other parameter estimates. However, for simplicity we keep the reported common specification. All stocks showed significant excess dispersion. Again we could resolve this using a more flexible duration distribution.

Table 6 presents some diagnostics based on the joint estimation. The LM test for a single lagged quote duration gave mixed results, proving its importance in some cases. The results of the likelihood ratio test for the joint significance of the diurnal dummies $\varkappa_1, \ldots, \varkappa_6$ showed significance in all cases. The dummy coefficients are plotted in Figure 2. As the trading intensity enters the quote equation, the diurnal effects of the trading intensity are transferred to the quote equations by

[11] When quote arrivals that modify only depth are included as well, there are approximately twice as many observations. With an identical specification, apart from the extra quotes, a somewhat different answer is found. Now the information variables such as volume and trade arrival rate and level of spread have a slowing or ambiguous effect on quote arrivals. This result is an indication that depth revisions are relatively more frequent when the market is slower. A possible explanation is that depth quotes are derived directly from the limit order book and limit orders are submitted more cautiously when the market shows informational trading.

Table 8 Diagnostics for the quote equation.

Firm	$LB(\tilde{Y})$	$LB(\varepsilon)$	s_{ε}
PG	***76517.6***	***103.2***	0.8345
DIS	***17055.5***	***65.4***	0.7865
FNM	***22167.4***	***43.3***	0.7204
GM	***68727.9***	***73.8***	0.6703
BAC	***10945.4***	***52.5***	0.7409
MCD	***49841.1***	***46.3***	0.7120
MTC	***36250.7***	***57.0***	0.7803
SLB	***3295.9***	***50.1***	0.8210

Diagnostics for quote equation of Table 7.
This table gives some diagnostics for the quote equation corresponding to the model defined in Table 8. The details are as in panel B of Table 5.

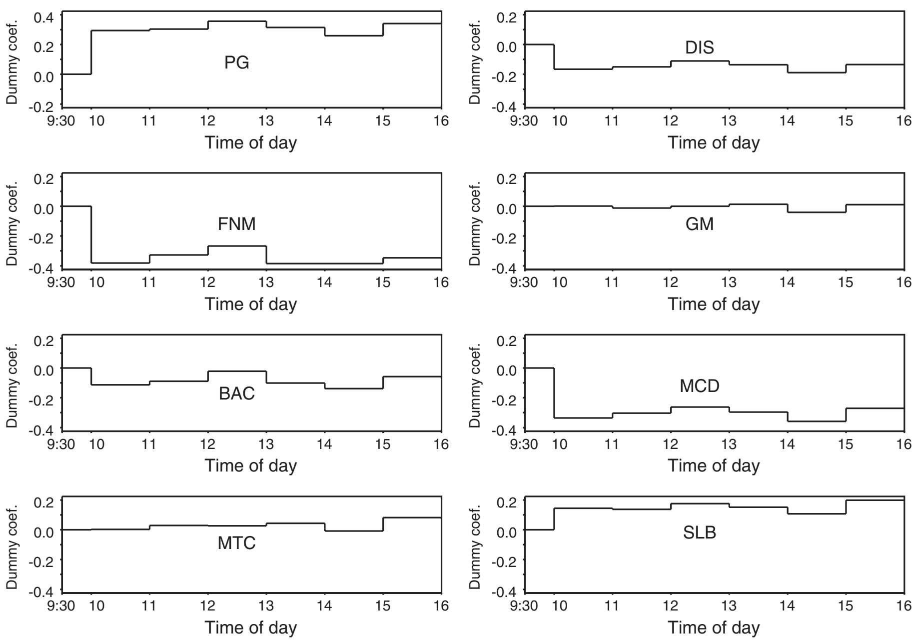


Figure 2 Dummy coefficients for the quote equation.

$\tau(x_i/\psi_i) + \eta_1(x_{i-1}/\psi_{i-1}) + \eta_2 \ln(\psi_{i-1})$. Hence the spline in Figure 2 represents small adjustments of the induced effect of the trade dummies.

The quote equation nevertheless tells a very consistent story, and reflects the unconditional correlations of Table 3. Evidence of information flows, whether it is past short quote durations, current and past high trade intensities, high levels of spread or volume, or low levels of market liquidity, all lead to short quote revision times. This is the underlying model of price dynamics in response to trading and news, which gives us a microstructure view of price volatility and its correlates.

Before we turn to the conclusion we must quantify the economic significance of the estimated relationship, as done for the trade equation. The procedure is just like above, and for this example we use BAC, because its coefficients are in the middle of the overall range of the stocks. For volume, the estimated coefficient for BAC is -0.01927. Hence if $\sqrt{\mathsf{vol}}_i$ quadruples from one trade to the following, then the expected duration will decrease 7.7%. On average, this amounts to $46.4 \cdot 0.077 = 3.6$ seconds within one trade. However, the long-run effect would be $4 \cdot 0.01927/0.52$, (i.e., 14.8%), or about 7 seconds. Carrying out these calculations for $\Delta\mathsf{Spr}_i$, just increasing it by 50%, shows that the short-run effect is a decrease of $0.5 \cdot 1.2934$ (or 64.7%), amounting to 30 seconds on average. The long-run effects through the level of the spread will be $0.5 \cdot 1.1387/0.52$ (i.e., 109%), amounting to 50.6 seconds on average. The effects of changes in abs(s.vol)$_i$ are very small.

5 CONCLUSION

This article has developed a bivariate model of the arrival of trades and quotes for stocks traded on the NYSE. The time between a trade and a new price quote is argued to be an interesting measure of the speed at which information is incorporated into prices. It is also related to the price impact of the trade, the liquidity of the stock, and ultimately the volatility of the stock. The model seeks to determine under what circumstances price adjustment is rapid and when it is slow.

Since trade quote durations are not observed when there is an intervening trade, the bivariate model must allow censoring of the quote durations by the trade durations. A bivariate likelihood function is formulated with censoring of one variable by the other. This likelihood function is a generalization of the ACD model of Engle and Russell (1998). It is maximized under a variety of specifications for a trade equation and for a quote equation conditional on the trade duration. Specification tests and robustness checks are carried out.

The essential finding for the arrival of quotes that revise prices is that information flow variables, such as high trade arrival rates, large volume per trade, and wide bid–ask spreads, all predict more rapid price revisions. This means prices respond more quickly to trades when information is flowing so that the price impacts of trades and ultimately the volatility of prices, are high in such circumstances.

Received October 23, 2001; revised October 12, 2002; accepted March 17, 2003

REFERENCES

Bauwens, L., F. Galli, and P. Giot. (2002). "The Moments of log-ACD Models." unpublished manuscript, CORE, Louvain-La-Neuve.

Bauwens, L., and P. Giot. (2000a). "The Logartihmic ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks." *Annales d'Economie et de Statistique 60*, 117–149.

Bauwens, L., and P. Giot. (2000b). "The Moments of First-Order log-ACD." in *Modeling Seasonality and Periodicity*, Proceedings of the 3rd International Symposium on Frontiers of Time Series Modeling, 213–237.

Bauwens, L., P. Giot, J. Grammig, and D. Veredas. (2000). "A Comparison of Financial Duration Models via Density Forecasts." CORE Discussion Paper 2000/60, Louvain-La-Neuve.

Bauwens, L., and D. Veredas. (1999). "The Stochastic Conditional Duration Model: A Latent Factor Model for the Analysis of Financial Durations." CORE Discussion Paper 9958, Université Catholique de Louvain; forthcoming in *Journal of Econometrics*.

Bollerslev, T., and J. M. Wooldridge. (1992). "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariance." *Econometric Reviews 11*, 143–172.

Bowsher, C. G. (2002). "Modelling Security Market Events in Continuous Time: Intensity Based, Multivariate Point Process Models." Unpublished manuscript, Nuffield College, Oxford.

Chordia, T., R. Roll, and A. Subrahmanyam. (2001). "Evidence on the Speed of Convergence to Market Efficiency." Finance Paper 11-01, UCLA.

Collett, D. (1994). *Modelling Survival Data in Medical Research*. London: Chapman & Hall.

Coughenour, J., and K. Shastri. (1999). "Symposium on Market Microstructure: A Review of Empirical Evidence." *Financial Review 34*, 1–28.

Cox, D. R., and E. J. Snell. (1968). "A General Definition of Residuals." *Journal of the Royal Statistical Society B30*, 248–275.

Cox, D. R., and P. A. W. Lewis. (1972). "Multivariate Point Processes." In L. M. Le Cam, J. Neyman, and E. L. Scott, (eds.), Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. III, Berkeley: University of California Press, pp. 401–448.

Cox, D. R., and D. Oakes. (1984). *Analysis of Survival Data*. London: Chapman & Hall.

Diamond, D. W., and R. E. Verrecchia. (1987). "Constraints on Short-Selling and Asset Price Adjustments to Private Information." *Journal of Financial Economics 18*, 277–311.

Dufour, A., and R. F. Engle. (2000). "Time and the Impact of a Trade." *Journal of Finance 55*, 2467–2498.

Easley, D., and M. O'Hara. (1987). "Price, Trade Size, and Information in Securities Markets." *Journal of Financial Economics 16*, 69–90.

Easley, D., and M. O'Hara. (1992). "Time and the Process of Security Price Adjustment." *Journal of Finance 47*, 905–927.

Engle, R. F. (2000). "The Econometrics of Ultra-High Frequency Data." *Econometrica 68*, 1–22.

Engle, R. F., and J. Lange. (2001). "Predicting VNET: A Model of the Dynamics of Market Depth." *Journal of Financial Markets 2*, 113–142.

Engle, R. F., and J. R. Russell. (1998). "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data." *Econometrica 66*, 1127–1163.

Fernandes, M., and J. Grammig. (2000). "Non-Parametric Specification Tests for Conditional Duration Models." Manuscript, European University Institute and University of Frankfurt.

Gerhard, F., and N. Hautsch. (2000). "Determinants of Inter-Trade Durations and Hazards Rates Using Proportional Hazard ARMA models." Discussion Paper 00/02, CoFE, Konstanz.

Ghysels, E., C. Gourieroux, and J. Jasiak. (2002). "Stochastic Volatility Duration Models." Forthcoming in *Journal of Econometrics*.
Ghysels, E., and J. Jasiak. (1998). "GARCH for Irregularly Spaced Financial Data: The ACD-GARCH Model." *Studies in Nonlinear Dynamics and Econometrics 2*, 133–149.
Glosten, L., and P. Milgrom. (1985). "Bid, Ask, and Transactions Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics 13*, 71–100.
Golia, S. (2001). "Long Memory Effect in Ultra-High Frequency Data." *Quaderni di Statistica 3*, 43–52.
Goodhart, C. A. E., and M. O'Hara. (1997). "High Frequency Data in Financial Markets: Issues and Applications." *Journal of Empirical Finance 4*, 73–114.
Grammig, J., and K. O. Maurer. (2000). "Non-Monotone Hazard Functions and the Autoregressive Conditional Duration Model." *Econometrics Journal 3*, 16–38.
Grammig, J., and M. Wellner. (2002). "Modelling the Interdependence of Volatility and Intertransaction Duration Processes." *Journal of Econometrics 106*, 369–400.
Harvey, A. (1990). *The Econometric Analysis of Time Series*, 2nd ed. Deddington: Philip Alan.
Hasbrouck, J. (1991). "Measuring the Information Content of Stock Trades." *Journal of Finance 46*, 179–207.
Hasbrouck, J. (1996). "Modelling Market Microstructure Time Series." In G. S. Maddala and C. R. Rao, (eds.), *Handbook of Statistics*, vol. 14, North-Holland, Amsterdam, pp. 647–692.
Hasbrouck, J., G. Sofianos, and D. Sosebee. (1993). "Orders, Trades, Reports and Quotes at the New York Stock Exchange." Technical report, NYSE, Research and Planning Section.
Hautsch, N. (1999). "Analyzing the Time Between Trades with a Gamma Compounded Hazard Model." Discussion Paper 99/03, CoFE, Konstanz.
Jasiak, J., 1998). "Persistence in Intertrade Durations." *Finance 19*, 166–195.
Kamionka, T. (2000). "La Modélisation des Données Haute Fréquence." *Journal de la Société Française de Statistique 141*, 167–211.
Karpoff, J. M. (1987). "The Relation Between Price Changes and Trading Volume: A Survey." *Journal of Financial and Quantitative Analysis 22*, 109–126.
Kyle, A. S. (1985). "Continuous Auctions and Insider Trading." *Econometrica 53*, 1315–1335.
Lange, J. (1997). The Supply of Liquidity and the Order Submission Process, University of California–San Diego.
Lee, C. M., and M. J. Ready. (1991). "Inferring Trade Direction from Intraday Data." *Journal of Finance 46*, 733–746.
Lunde, A. (1997). "A Conjugate Gamma Model for Durations in Transactions." Working paper, University of Aarhus.
Lunde, A. (1998). "A Generalized Gamma Autoregressive Conditional Duration Model." Working paper, University of Aarhus.
Madhavan, A. (2000). "Market Microstructure: A Survey." *Journal of Financial Markets 3*, 205–258.
Nelson, D. B. (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica 59*, 347–370.
O'Hara, M. (1995). *Market Microstructure Theory*, Cambridge, MA: Blackwell.
Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannary. (1992). *Numerical Recipes in C*, 2nd ed. Cambridge: Cambridge University Press.

Russell, J. R. (2001). "Econometric Modelling of Multivariate Irregulary-Spaced High-Frequency Data." Unpublished manuscript, University of Chicago.

Rydberg, T. H., and N. Shephard. (1999b). "Modelling Trade-by-Trade Price Movements of Multiple Assets Using Multivariate Compound Poisson Processes." Oxford: Manus Nuffield College.

Rydberg, T. H., and N. Shephard. (1999a). "BIN Models for Trade-by-Trade Data: Modelling the Number of Trades in a Fixed Interval of Time." Oxford: Manus Nuffield College.

Schwartz, R. A. (1993). *Reshaping the Equity Markets*. Homewood, IL: Business One Irwin.

Veredas, D., J. Rodriguez-Poo, and A. Espasa. (2001). "On the (Intradaily) Seasonality and Dynamics of a Financial Point Proces: A Semiparametric Approach." CORE Discussion Paper 2001/XX, Université Catholique de Louvain.

White, H. (1982). "Maximum Likelihood Estimation of Misspecified Models." *Econometrica 50*, 1–26.

Wooldridge, J. M. (1994). "Estimation and Inference for Dependent Processes." In R. F. Engle and D. L. McFadden (eds.), *Handbook of Econometrics*, vol. IV. Amsterdam, The Netherlands: Elsevier Science B.V., pp. 2640–2738.

Zhang, M. Y., J. R. Russell, and R. S. Tay. (2001). "A Nonlinear Autoregressive Conditional Duration Model with Applications to Financial Transaction Data." *Journal of Econometrics 104*, 179–201.