Price Discovery in Tick Time*

Bart Frijns†
Auckland University of Technology

Peter Schotman‡
Limburg Institute of Financial Economics and CEPR

October 12, 2004

*We like to thank Ronald Mahieu, Franz Palm, seminar participants at Erasmus University and the Vrije Universiteit of Amsterdam, and participants at the Common Features in Maastricht Conference, the workshop on the Financial Econometrics of Microstructure workshop, the European Finance Association and the Econometric Society European Meetings for their useful comments and suggestions.

†Corresponding Author: Bart Frijns, Department of Finance, Auckland University of Technology, Private Bag 92006, 1020 Auckland, New Zealand. Phone: +64-9-917 9999 etxn 5713, Fax: +64-9-917 9629, E-mail: bart.frijns@aut.ac.nz.

‡Peter Schotman, LIFE Maastricht University, E-mail: p.schotman@berfin.unimaas.nl.

Price Discovery in Tick Time

Abstract

In this paper we propose a tick time model for dealer quote interactions using ultra-high-frequency data. This model includes duration functions to measure the time dependence of volatility as well as information asymmetry. In order to asses price discovery we define several measures in tick time. These measures can be aggregated to calendar time and we define a comparable measure to Hasbrouck (1995) information shares. In our empirical part we examine the Island and Instinet Electronic Communication Networks, and three wholesale market makers for 20 actively traded stocks with varying liquidity at Nasdaq. Our results include that volatility does not increase with the duration between quote updates, and that longer quote durations lead to lower price discovery. In terms of price discovery we find that ECNs tend to dominate the liquid stocks, whereas market makers dominate the less liquid stocks.

Keywords: Price Discovery, Tick Time models, Nasdaq, Ultra-high frequency data, Microstructure.

JEL Classifications: C32, G15.

1 Introduction

Price discovery is the process of how different information sources contribute to the evolution of the underlying value of an asset. In a fragmented market with multiple dealers, like the Nasdaq, each dealer contributes to the price discovery process. Interesting questions are which dealer contributes most, how quick the discovery process works, and how it depends on market circumstances like liquidity, volatility and trading intensity. An important measure for the price discovery contribution of a dealer is the information share defined by Hasbrouck (1995). Huang (2002) provides an application to a variety of stocks on the Nasdaq to study the relationship between ECNs (Electronic Communication Networks) and traditional dealers.

Hasbrouck (1995) defines information shares as the part of the variance of the random walk component of returns that can be attributed to a particular market or dealer. But when quote updates of various dealers are contemporaneously correlated, the variance decomposition will not be unique. Hasbrouck (1995) suggests alternative Choleski decompositions to establish upper and lower bounds. For a particular stock at Nasdaq, Baillie, Booth, Tse, and Zabotina (2002) show, however, that upper and lower bounds can differ substantially, even if returns are observed at one minute intervals. When the sampling frequency is very high this contemporaneous correlation is minimal. For this reason, tick time data would be the preferred observation frequency. The use of tick time data for estimating price discovery among Nasdaq dealers is the main topic of this paper.

This paper provides three main contributions. First, we suggest a different model for the dynamics of quote changes. The traditional vector autoregressive of Hasbrouck (1995) is not suited for tick time data, as it always requires some time aggregation to construct a vector of quote updates for every time interval. An unobserved components specification is more suitable for ultra-high frequency data, since it does not require a complete vector of dealer quotes in each time interval. The model we propose is an extension of Hasbrouck (1993) to a setting with multiple dealers where quotes arrive in tick time.

Our second contribution is that we allow for duration effects on the quote dynamics. Time is an important factor in microstructure (Engle and Patton (2004)). It affects the volatility of the efficient price (Engle (2000)) and it also has an impact on the information content of dealer quote updates (Dufour and Engle (2000)). Both effects will be included in the model.

Our third contribution is the definition of measures for price discovery. These measures are extensions of the calendar time measures developed by De De Jong and Schotman (2003) for an unobserved components model. We define these measures in tick time, where they are a function of the time between quote innovations. Additionally, we integrate these measures over time to reflect their calendar time equivalents. One of these calendar time measures resembles the Hasbrouck (1995) information shares. The structural interpretation of the unobserved components model provides a decomposition that leads to unambiguous information shares.

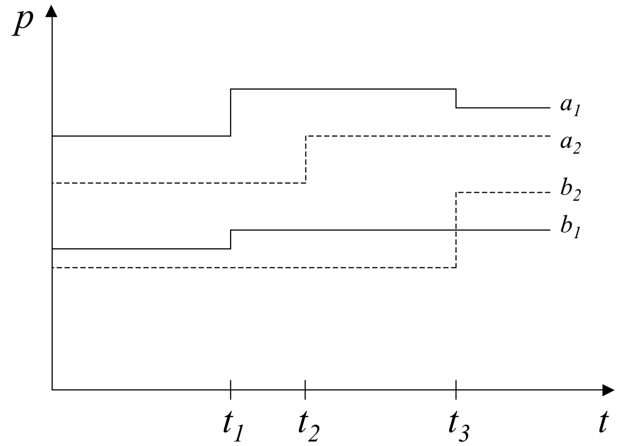
In our empirical part we examine the quotes of two ECN's (Island and Instinet) and three wholesale market makers at Nasdaq for 20 actively traded stocks with varying liquidity. These market makers are selected as the most active in terms of quoting frequency. For a tick time model it is more natural to consider the quote setting behavior of individual dealers instead of classes of dealers as in Huang (2002). Considering individual dealers is also in line with Schultz (2003) who finds a lot of heterogeneity among individual market makers.

As a preview of the results, we find that volatility does not evolve in calendar time, but either in tick time as mentioned by Clark (1973) or even less than tick time (see Dufour and Engle (2000)). The information flow to the efficient price is in general less at longer durations. We confirm the hypothesis of Easley and O'Hara (1992), which states that long durations convey no information. Similar results were found by Dufour and Engle (2000), and Engle and Patton (2004).

Price discovery measures in tick time appear strongly dependent on durations. Some dealers reveal information when durations are short whereas others reveal information when durations are long. Aggregating to calendar time we can often clearly identify the dominant dealer. In terms of price discovery we find that ECNs tend to dominate the liquid stocks, whereas market makers dominate the less liquid stocks. When the interval over which we aggregate increases, all these measures converge to single point estimates.

The remainder of this paper is structured as follows. Section 2 discusses specification and estimation of the model. Section 3 defines the measures for price discovery in tick time and derives the calendar time aggregation of these measures. In section 4 we discuss the data. Section 5 presents the results of the model and the price discovery measures. Finally section 6 concludes.

Figure 1: Quote Arrivals



2 A Model for Quotes in Tick Time

In this section we introduce a structural time series model for data in tick time and is an extension of the unobserved components model as proposed by Hasbrouck (1993). This model also fits the theoretical framework as proposed by Glosten and Harris (1988).

Consider a dealer market where M dealers issue bid and ask quotes. Quotes arrive at times t_ℓ ($\ell = 1, \ldots, L$). The time between two consecutive quote arrivals is the quote duration. We are interested in modelling the dynamics of quote updates conditional on durations and assume that these durations are exogenous.

When sampling in tick time all quote changes are included. These can be changes in the bid and/or ask quotes of one or more dealers. In figure 1 we illustrate the quote arrival process. At t_1 dealer 1 increases her bid and ask quote simultaneously. Next at t_2, dealer 2 increases only her ask. At time t_3 both dealers change one of their quotes.

With M dealers in the market we have $2M$ different time series of quote updates. Let q_ℓ be the $2M$-vector of all standing quotes at time t_ℓ. The bid (ask) of dealer i corresponds to element $2i-1$ ($2i$) of q_ℓ. To describe the quote dynamics in a model, we define I_ℓ as the vector of identities of the quotes that have changed at t_ℓ. This vector I_ℓ has dimensions $k_\ell \times 1$. When, e.g. dealer i changes her bid at t_ℓ, I_ℓ has a single element, $I_\ell = 2i-1$. When dealer i changes both bid and ask, $I_\ell = (2i-1, 2i)$. The vector of quotes that are updated at t_ℓ can be written as $q_{I_\ell,\ell}$. The time series properties of updated quotes are modelled as an unobserved components model. These components are a deterministic part (c), which is a $2M$-vector due to the average spread between a quote series and the efficient price.

This constant deviation captures e.g. the order processing costs the market maker faces. For a bid we expect negative elements in c, for an ask we expects these elements positive. The second part is a permanent part (m_ℓ), which is common to all dealers. The last part is a transitory part (u_ℓ), which is a $2M$-vector of factors not captured by c and m_ℓ. This part captures the informational asymmetries among dealers, but also inventory costs and other kinds of noise. The decomposition for quotes follows as

$$q_{I_\ell,\ell} = c_{I_\ell} + \iota m_\ell + u_{I_\ell,\ell}, \tag{1}$$

where ι is a unit vector.[1] The permanent part, also called the efficient price follows a random walk, with time dependent volatility (σ_ℓ). The impact of time on the volatility of the price process was questioned by Clark (1973), Harris (1987) and Ané and Geman (2000), among others.

$$m_\ell = m_{\ell-1} + \sigma_\ell r_\ell. \tag{2}$$

The innovation term r_ℓ has unit variance. We specify volatility as

$$\sigma_\ell = \sigma \tau_\ell^{\delta_1}, \tag{3}$$

where τ_ℓ is the duration between quotes, normalized by dividing it by the average quote duration over the whole sample.[2] The parameter δ_1 measures the impact of quote durations on the volatility of the random walk. If $\delta_1 = \frac{1}{2}$ the random walk is said to evolve in calendar time. In this case the variance of $(m_p - m_s)$ is equal to $(\sum_{\ell=s+1}^{p} \tau_\ell)\sigma^2$, and thus proportional to the length of the calendar time interval. When $\delta_1 = 0$ then the variance of the random walk is not affected by the time between quote updates. In this case the random walk evolves in tick time. The calendar time variance is proportional to the number of quote updates (see e.g. Clark (1973)).

The last term in (1) is the transitory term u_ℓ, which measures the temporary deviations from the efficient price and includes asymmetric information, inventory effects and other sources of noise. To capture the asymmetric information we allow the transitory component for every dealer, to correlate with the innovation in the efficient price. This is in line with the theoretical framework of Glosten and Harris (1988), who argue that the prices of

[1] The length of this unit vector is suppressed at the length of the vector of quote updates.

[2] This normalization only affects the σ parameter, which now refers to the volatility per tick, instead of the volatility per second.

informed market participants are correlated with the innovation in the efficient price. The decomposition for the transitory component therefore reads

$$u_{I_\ell,\ell} = \alpha_{I_\ell} \tau_\ell^{\delta_2} \sigma r_\ell + e_{I_\ell,\ell}. \tag{4}$$

where the asymmetric information is captured by α. We allow this information component to depend on the duration between quotes. The impact time has on the asymmetric information is measured by δ_2. In case $\delta_2 > 0$, quotes become more informative at longer durations. If $\delta_2 < 0$ the opposite holds. The idiosyncratic quote noise ($e_{I_\ell,\ell}$) is uncorrelated with r_ℓ and τ_ℓ and has covariance matrix Ω.

To estimate the model we put it in state space form for a time series process with missing observations. Following Harvey (1989) we write

$$\begin{aligned} J_\ell' q_\ell &= J_\ell' c + J_\ell' \iota m_\ell + J_\ell' u_\ell, \\ m_\ell &= m_{\ell-1} + \sigma_\ell r_\ell, \\ J_\ell' u_\ell &= J_\ell' \alpha \tau_\ell^{\delta_2} \sigma r_\ell + J_\ell' e_\ell, \end{aligned} \tag{5}$$

where J_ℓ is a matrix that selects the elements from the quote vector that are updated. This matrix has dimension $2M \times k_\ell$. In this case $J_\ell' q_\ell$ is a $k_\ell \times 1$ vector of updated quotes. This pre-multiplication is applied to all the components that we discussed.

An advantage of putting the model in a state space form is that with some additional restrictions the parameters c, α, δ_1, δ_2, σ and Ω can be identified. The first restriction is bid and ask deviations from the efficient price are symmetric. This restriction is merely imposed to reduce the number of parameters in the model. The second restriction is on the idiosyncratic quote noise (e_ℓ). We assume the covariance matrix of this noise term (Ω) to have a block diagonal structure. This structure only allows bid and ask quotes of the same dealer to be correlated. We motivate this structure by the fact that the quote noise includes inventory effects and a remaining microstructure noise. Theoretical models for inventory costs adhere to the notion that when a dealer receives inventory she will lower her ask to induce a trade at the opposite side, but also lowers her bid to avoid receiving additional inventory. Therefore, a specific dealer may wish to alter both quotes simultaneously due to the inventory position she has. Simultaneous changes in the quotes of different dealers can be due to asymmetric information both dealers have, which is captured by α or occurs by mere chance.

De Jong and Schotman (2003) show that for a model in calendar time, imposing structure on Ω is necessary for the identification of α. In tick time there are a few additional effects that hinder this identification. When the dimension of the quote vector is one at each observation, the identification of α is not guaranteed. Without duration functions α can only be identified up to a translation along the unit vector. See Appendix A for a full derivation. Duration functions help to make identification possible, but only if $\delta_1 \neq 0$ and $\delta_2 \neq 0$. When these measures are close to zero identification will be troublesome.

Estimation of (5) is done using the Kalman Filter. As the underlying process is a random walk, the system cannot be initialized using the long run mean of the underlying process. We use a diffuse prior to initialize the random walk process. To incorporate the potentially large price change overnight the same approach is followed. Everyday the model is re-initialized with a diffuse prior. The first 50 observations on each day are excluded in the maximization of the Likelihood function as the variance of the prediction error has not converged to normal levels.

As the model considers quote updates, the dimension of the quote vector will frequently be one. This makes computations in the Kalman Filter recursions straightforward as determinants and inverses are given. However since this dimension changes over time the model will not converge to a steady state and the full recursion will have to be computed for each observation. For this reason the filtering routine is computationally demanding.

3 Price Discovery

In this section we define price discovery measures in tick time using the unobserved components model of the previous section. These tick time measures are subsequently aggregated to calendar time equivalents.

3.1 Price Discovery in Tick time

We consider three measures to summarize the quote setting behavior of dealers. The first measure considers how dealers incorporate information in the efficient price into their quote innovations. The second measure considers the contribution of quote innovations to the evolution of the efficient price. The last measure combines both measures in one single measure, which considers the total information share of dealers.

To explore the implications of the model we consider the following scenario. Suppose at t_ℓ each dealer issues bid and ask quotes, and suppose that the previous efficient price

$m_{\ell-1}$ is known to all dealers. The quote updates reflect both the change in efficient price, $m_\ell - m_{\ell-1}$ between $t_{\ell-1}$ and t_ℓ and the dealer noise e_ℓ. The innovation of the dealer quotes is then equal to

$$\begin{aligned} v_\ell = q_\ell - \mathrm{E}_{\ell-1}[q_\ell] &= \iota(m_\ell - m_{\ell-1}) + u_\ell \\ &= (\iota + \alpha\tau_\ell^{\delta_2-\delta_1})(m_\ell - m_{\ell-1}) + e_\ell. \end{aligned} \tag{6}$$

The amount of information in the efficient price that is incorporated into quotes is referred to as *dealer efficiency*. The more of the innovation in the random walk that is incorporated in dealer quotes, the more efficient they are. To obtain a measure for dealer efficiency we consider the covariance between quote updates as a fraction of the total variance of the random walk

$$\beta(\tau) = Cov(v_\ell \Delta m_\ell)/Var(\Delta m_\ell) = \iota + \alpha\tau^{(\delta_2-\delta_1)}, \tag{7}$$

which is a function of the duration of quote innovations. The crucial parameters are α, δ_1 and δ_2. At long durations this measure will converge to 1 when $\delta_1 > \delta_2$, and diverge in the opposite case. The sign of α determines whether 1 is an upper or a lower bound for the respective element in $\beta(\tau)$. When α is larger than zero, $\beta(\tau)$ will be larger than one. This indicates that this quote incorporates more information of the random walk and vice versa. Hence whether a dealer is more or less efficient depends solely on the sign of α. Whether this efficiency persists over long or short durations is dependent on the values of δ_1 and δ_2.

For the second measure we consider a variance decomposition of the efficient price similar to Hasbrouck's (1995) analysis of *price discovery*. It relates the change in the efficient price to innovations in dealer quotes,

$$\Delta m_\ell = \gamma(\tau)' v_\ell + \varepsilon_\ell, \tag{8}$$

where the regression coefficients $\gamma(\tau)$ are defined as

$$\gamma(\tau) = Var(v_\ell)^{-1} Cov(v_\ell \Delta m_\ell). \tag{9}$$

For the covariance matrix $\Sigma(\tau) = Var(v_\ell)$ we find

$$\Sigma(\tau) = \Omega + \sigma^2\tau^{2\delta_1}(\iota + \alpha\tau^{\delta_2-\delta_1})(\iota + \alpha\tau^{\delta_2-\delta_1})'. \tag{10}$$

The covariance follows as

$$Cov(v_\ell \Delta m_\ell) = \sigma^2\tau^{2\delta_1}(\iota + \alpha\tau^{\delta_2-\delta_1}). \tag{11}$$

To arrive at an expression for $\gamma(\tau)$ we apply the matrix inversion lemma on $\Sigma(\tau)$,

$$\Sigma(\tau)^{-1} = \Omega^{-1} - \frac{\sigma^2\tau^{2\delta_1}\Omega^{-1}(\iota + \alpha\tau^{\delta_2-\delta_1})(\iota + \alpha\tau^{\delta_2-\delta_1})'\Omega^{-1}}{\sigma^2\tau^{2\delta_1}(\iota + \alpha\tau^{\delta_2-\delta_1})'\Omega^{-1}(\iota + \alpha\tau^{\delta_2-\delta_1})}, \tag{12}$$

and therefore find

$$\gamma(\tau) = \frac{\sigma^2\tau^{2\delta_1}\Omega^{-1}(\iota + \alpha\tau^{\delta_2-\delta_1})}{1 + \sigma^2\tau^{2\delta_1}(\iota + \alpha\tau^{\delta_2-\delta_1})'\Omega^{-1}(\iota + \alpha\tau^{\delta_2-\delta_1})}. \tag{13}$$

This expression of $\gamma(\tau)$ can also be written as a function of $\beta(\tau)$

$$\gamma(\tau) = \frac{\sigma^2\tau^{2\delta_1}\Omega^{-1}\beta(\tau)}{1 + \sigma^2\tau^{2\delta_1}\beta(\tau)'\Omega^{-1}\beta(\tau)}. \tag{14}$$

Hence this measure combines the dealer efficiency measure and the amount of idiosyncratic noise each dealer quote has.

The last measure resembles the information shares of Hasbrouck (1995). These are defined as the proportional contribution of dealer quote innovations to the innovation in the efficient price. This is determined by considered how much of the variance of the efficient price can be explained by the variance in quote innovations. This can be expressed in terms of the R^2 of regression (8). From standard regression theory it is known that this is the inner product of $\gamma(\tau)$ and $\beta(\tau)$

$$R^2(\tau) = \gamma(\tau)'\beta(\tau). \tag{15}$$

Hasbrouck (1995) information shares have the characteristic that they cannot be assigned to dealers uniquely. Given that the matrix inversion lemma can be applied and the block diagonal structure of Ω, $R^2(\tau)$ can be decomposed as the sum of the information shares per dealer ($IS_i(\tau)$).

3.2 Calendar Time Aggregation

In the previous section we discussed price discovery and efficiency of single quote updates. This section defines similar measures of quote setting behavior over time. To obtain these measures we aggregate the tick time measures to fixed time intervals. Let us first establish the relationship between tick time and calendar time. In calendar time we refer to t as the present time and $t+1$ as the time at the next time interval. Let $Q(t)$ be the vector of quotes at time t. Assume that deviations from the efficient price are included in the information set of all dealers. In calendar time innovations in dealer quotes are decomposed as

$$V(t) = Q(t) - \mathrm{E}_{t-1}[Q(t)] = \iota(m(t) - m(t-1)) + u(t). \tag{16}$$

The innovation in quotes is equal to the change in the efficient price plus the noise around this efficient price. To establish the link between calendar time and tick time, let ℓ_t represent the observation closest to time t. The tick time equivalent (16) is

$$v_{\ell_t} = \iota(m_{\ell_t} - m_{\ell_{t-1}}) + u_{\ell_t}. \tag{17}$$

The change in the efficient price from ℓ_{t-1} to ℓ_t is the sum of all interjacent changes. We write (17) as the sum of these changes

$$v_{\ell_t} = \iota(m_{\ell_t} - m_{\ell_{t-1}}) + u_{\ell_t} = \iota \sum_{j=\ell_{t-1}+1}^{\ell_t} (m_j - m_{j-1}) + u_{\ell_t}. \tag{18}$$

Dealer efficiency is considered as the change in the dealer quotes that can be attributed to the change in the efficient price. This is the regression coefficient of quote innovations regressed on the change in the efficient price. We obtain this measure by considering the covariance of quote innovations with the change in the efficient price as a fraction of the total variance caused by the efficient price. The covariance follows as

$$\begin{aligned} Cov(v_{\ell_t}, (m_{\ell_t} - m_{\ell_{t-1}})) &= \mathrm{E}[(\iota \sum_{j=\ell_{t-1}+1}^{\ell_t} \tau_j^{\delta_1} \sigma r_j + \alpha \tau_{\ell_t}^{\delta_2} \sigma r_{\ell_t})(\sum_{j=\ell_{t-1}+1}^{\ell_t} \tau_j^{\delta_1} \sigma r_j)] \\ &= \sigma^2(\iota \sum_{j=\ell_{t-1}+1}^{\ell_t} \tau_j^{2\delta_1} + \alpha \tau_{\ell_t}^{\delta_1+\delta_2}). \end{aligned} \tag{19}$$

The variance of the random walk is

$$Var(m_{\ell_t} - m_{\ell_{t-1}}) = \sigma^2 \sum_{j=\ell_{t-1}+1}^{\ell_t} \tau_j^{2\delta_1} \tag{20}$$

The measure for dealer efficiency follows from this,

$$\beta(t) = \iota + \alpha \frac{\tau_{\ell_t}^{\delta_1+\delta_2}}{\sum_{j=\ell_{t-1}+1}^{\ell_t} \tau_j^{2\delta_1}}. \tag{21}$$

Again $\beta(t)$ is a re-scaling of α. As the interval between ℓ_t and ℓ_{t-1} increases, $\beta(t)$ converges to unity. Thus when intervals are larger, dealers incorporate information about the true price more efficiently into their quotes. The rate of convergence is dependent on α, δ_1 and δ_2. Depending on the sign of α this convergence will occur from above ($\alpha > 0$) or below ($\alpha < 0$). When convergence in $\beta(t)$ occurs quickly then dealers quote efficiently in this stock.

For price discovery we consider the reverse regression as for dealer efficiency. We consider the change in the efficient price that can be attributed to innovations in dealer quotes

$$\gamma(t)' = Cov(V_{\ell_t}, (m_{\ell_t} - m_{\ell_{t-1}}))\Sigma(t)^{-1}. \tag{22}$$

The variance of quote innovations follows as

$$\begin{aligned} \Sigma(t) &= Var(\iota(m_{\ell_t} - m_{\ell_{t-1}} + u_{\ell_t})) \\ &= \sigma^2(\iota\iota' \sum_{t_\ell=\ell_{t-1}+1}^{\ell_t-1} \tau_{t_\ell^{2\delta_1}} + (\iota\tau_{\ell_t}^{\delta_1} + \alpha\tau_{\ell_t}^{\delta_2})(\iota\tau_{\ell_t}^{\delta_1} + \alpha\tau_{\ell_t}^{\delta_2})') + \Omega. \end{aligned} \tag{23}$$

This matrix depends on all the interjacent durations from ℓ_{t-1} to ℓ_t. Therefore the inverse of this matrix is not straightforward. To extract the duration dependence, the matrix inversion lemma needs to be applied twice. For the sake of notational comfort the inverse is not displayed.

As mentioned before information shares are determined by the R^2 of the regression of dealer quotes on the efficient price. Again this measure is the inner product of dealer efficiency and price discovery

$$R^2(t) = \gamma(t)'\beta(t). \tag{24}$$

Here the inversion lemma does not only extract the durations form the calculation of the inverse. For information shares this lemma is needed to assign them to dealers individually. As the inverse of $\Sigma(t)$ is a function of the inverse of Ω, which was known to have a block diagonal structure, these information shares can be assigned uniquely.

When the sampling interval increases the change in dealer quotes will to a large extent represent the change in the efficient price. As a consequence $R^2(t)$ converges to one. This convergence of $R^2(t)$ will be related to the convergence of $\beta(t)$ to a unit vector. Although stated without proof, if $R^2(t)$ and $\beta(t)$ converge to unity, then the sum of the elements in $\gamma(t)$ add up to one. As the covariance between quote innovations and the change in the random walk is positive, as well as the variance of the random walk, elements in $\gamma(t)$, will converge to single point estimates between zero and one.

The total value of $R^2(t)$ indicates how much microstructure noise there is left in dealer quotes. Results from this could have consequences for e.g. the sampling intervals used for realized volatility as described by Andersen, Bollerslev, Diebold, and Labys (2001). When $R^2(t)$ is low there is still a lot of microstructure noise in the data and longer sampling intervals should be used.

4 Data

The data used in this study is provided by Nastraq. This data set contains all trades and quotes that occur within normal trading hours at Nasdaq. From this data set we consider dealer quote data. This provides all quotes issued within trading hours, time stamped to the nearest second. Most important, it contains the identity of the dealer that issues the quote. For our study we select 20 highly traded companies with different liquidity listed at Nasdaq for February 1999. The selected stock and their ticker symbols are reported in table 1. Huang (2002) uses the same data set as ours in 1999 but uses different months in his study. He creates categories of different types of dealers. Schultz (2003) argues that dealer quoting behavior is heterogeneously, hence we consider individual dealers. We consider the dealer quotes of the 5 largest dealers, in terms of quoting frequency. This leads to the selection of 2 ECNs, Island and Instinet, and 3 market makers, which change depending on the stock considered.

The data provided needs to be filtered before it can be used in the model. As we consider the innovation in a quote (either bid or ask quote) the other quotes are removed from the data. First, when multiple quotes are issued at the same second, the last quote in this sequence is selected.[3] Second, a change in the depth of the quote is not considered. All quotes that do not change the bid or ask are removed. When a dealer only updates a bid this is the only quote that we will consider in the model. Due to this selection criteria the number of newly issued quotes is often 1.

INSERT TABLE 1 HERE

Another issue is the treatment of outliers. When a dealer is unwilling to trade she

[3]Especially in the case of Island this occurs frequently. This is caused since there are many small trades matched within Island itself. Every time this happens Island will send the new best standing quote to Nasdaq, which in many cases in merely a change in the depth of the quote

will issue a quote far away from the inside (best quote in the market). This can happen regularly on one side of the market when a dealer is unwilling to take more inventory. Although these quotes send a very strong signal to the other marker participants, this is something not considered here as we address the issues of price discovery. We define an outlier as a quote that is more than $ 2 away from the average of the past 50 quotes. These outliers are deleted.

Table 2 reports some summary statistics of the data after filtering. The first column reports the total number of quotes given by dealers in each stock. Our data set incorporates very liquid stocks like DELL as well as illiquid stocks like SBUX. The second column report the fraction of single quote updates. In section 2 we discussed the identification problems that could occur. The number of quotes is inversely related to the duration between quotes. Further note that there is a weak relationship between the percentage of single quotes and the total number of quotes issued. The other issue that can lead to identification problems is the variance in durations. These seem to be large enough not to cause any identification problems.

INSERT TABLE 2 HERE

5 Results

This section provides the results of the model in section 2 and the measures of price discovery discussed in section 3.

5.1 Parameter Estimates

In this part we discuss the parameter estimates from model (5). In table 3 we report the estimates for the duration parameters, including their standard errors. The parameter δ_1 measures the duration effect on the innovation in the efficient price. This parameter is significantly negative in 50% of the cases or zero in most other cases. For most stocks the random walk evolves in tick time or we find that long durations between quotes decreases the volatility of the random walk, which is consistent with the findings of Dufour and Engle (2000). The value of $\delta_1 = \frac{1}{2}$, which represents a random walk in calendar time can easily be rejected.

INSERT TABLE 3 HERE

The parameter δ_2 measures the duration dependence of the asymmetric information. In most cases this parameter is negative and more negative than δ_1. Longer durations between quote updates result in less asymmetric information about the efficient price. Hence at short durations quotes are more informative. These results are similar to the findings of Dufour and Engle (2000) and Engle and Patton (2004). They find that the price impact of trades is larger at short durations. Our results also confirm the reasoning of Easley and O'Hara (1992). Their model hypothesizes that long durations convey no information. The combined results of the duration parameters indicate that short durations means higher volatility, and more asymmetric information. As a consequence volatile periods are periods where asymmetries are large and vice versa.

INSERT TABLE 4 HERE

In table 4 we report the α's from (5) including their standard errors. An α significantly larger/smaller than zero means that this dealer has more/less exposure the innovation in the efficient price than the average of dealers. Although these α's are not that informative on their own, they do indicate large heterogeneity among the different dealers. Moreover it also indicates the heterogeneity present in the bid and ask quotes themselves. This heterogeneity among stocks will cause different results when discussing price discovery. It also indicates that dealer cannot just be grouped in categories as their characteristics are very different among individual dealers.

5.2 Price Discovery in Quote Updates

In this section we discuss the results of these measure per quote update.

INSERT FIGURE 2 HERE

To discuss dealer efficiency we consider the innovations in quotes that are due to a change in the efficient price. This measure is a re-scaling of the α's reported before, and is dependent on the duration estimates for their convergence. In most of the cases δ_1 is larger than δ_2, which means that for long durations the estimate for $\beta(\tau)$ will converge to one. This is expected, as when there is a long time between quote updates, the change in quotes is likely to be caused by a change in the efficient price. As the measures for $\beta(\tau)$ are duration dependent, they cannot be shown in a table. In figure 2 we draw $\beta(\tau)$ for three specific stocks in the sample as a function of duration. These stocks were selected to represent liquid stocks (INTC), less liquid stocks (CMGI) and illiquid stocks (AAPL). For all stocks $\beta(\tau)$ converges to one, as $\delta_2 < \delta_1$. Even when durations are as long as one minute, the change in quotes can still not be explained fully by the change in the underlying price. As there was heterogeneity among the α's of the different stock there is the same heterogeneity in $\beta(\tau)$.

INSERT FIGURE 3 HERE

For measures of price discovery in tick time we perform the reverse regression of the efficient price on quote innovations. We consider to what extend a single innovation in a dealer quote can explain the change in the efficient price. Again this measure is duration dependent and should be considered relatively with respect to Island. In figure 3 we show the same stocks as before. The dynamics of these estimates cannot be derived straightforward from the parameters as for $\beta(\tau)$. In the case of INTC we see a clear dominance of the ECNs over the market makers. The price discovery is highest at very short durations, and declines rapidly when durations get longer. However for INTC durations most often are very short. For CMGI, Island dominates persistently. The price discovery for Island does not decrease as the durations increase. An explanation for this is that the α's for Island are close to zero and the duration parameter δ_2 is also close to zero. For AAPL we see a similar pattern as for INTC. Interesting is the crossing line for MM1 and MM2 at very short durations.

INSERT FIGURE 4 HERE

The information shares we consider are determined by the R^2 of the regression of the change in the efficient price on the innovation in dealer quotes. This is the inner product of $\beta(\tau)$ and $\gamma(\tau)$. This $R^2(\tau)$ is decomposed to dealer specific parts using the structure of the model. In figure 4 we present the graphs for the three stocks. As a general results information shares are highest at short durations. For INTC the dominance of Island is obvious. For CMGI MM2 dominates and for AAPL Instinet leads in terms of information shares. Another general result is that for CMGI the information share of a tick is much higher than that of the other two stocks.

5.3 Calendar time aggregated Price Discovery

In the previous part we discussed some results for the information content of quote innovations. When these measures are aggregated over time, these provide useful information on the total efficiency of dealer as well as stocks and allow us to determine the total price contribution of dealers. Another benefit of time aggregation is that the measure we find can be compared to traditional measures, like Hasbrouck (1995) information shares. In this part we aggregate to one-minute and 5-minute intervals, as these intervals are most commonly used in other research.

INSERT TABLE 5 HERE

INSERT FIGURE 5 HERE

We again start by discussing dealer efficiency. In table 5 we report these measures in 60 second intervals and 300 second intervals. Depending on the sign of α, these measures are higher or lower than one. As mentioned when the time period over which we aggregate grows, these measures should converge to one. This convergence is clearly seen for all of the stocks in the sample. Many stocks have achieved this convergence after five-minutes, like INTC, but for the less liquid stocks there are still some inefficiencies (see e.g. QWST). In figure 5 we show the distributions of $\beta(t)$ for the three stocks that we evaluate at 60 and 300 seconds. The results from these graphs again proof the heterogeneity among dealers. It also displays the skewed distribution of this measure.

INSERT TABLE 6 HERE

INSERT FIGURE 6 HERE

The same is done for the measure of price discovery in table 6. A general result that can be seen is that price discovery increases when the sampling interval gets larger. At both sampling intervals we see that Island has very high measures for price discovery compared to the other dealers, especially for the more liquid stocks in the sample. These measures can again be shown in distribution plots in figure 6. We show these distributions for three stocks that we have selected and only show the asks. Clearly when the sampling interval gets larger the distributions converge to a point estimate. In contrast to information shares, our measure benefits from time aggregation. The longer the sampling interval the more precisely we can assign information shares uniquely. In the case of INTC and CMGI Island clearly dominates the price discovery process. For AAPL after five minutes there is still no convergence to fixed points.

INSERT TABLE 7 HERE

INSERT FIGURE 7 HERE

Finally we consider the $R^2(t)$ of the regression when we aggregate over time. By applying the inversion lemma twice, this $R^2(t)$ can be decomposed and information shares can be assigned uniquely. Results for information shares per dealer are shown in table 7. Again comparing the 60 and 300 second interval we see that information shares increase. Interesting is the total dominance of specific market makers for the stocks PSFT and SBUX. Another interesting result is that we never find any dominance Instinet. This is in contrast to the findings of Huang (2002). In figure 7 we show these $R^2(t)$ per dealer for the three stocks. These show the same pattern as the graphs for $\gamma(t)$. This is because the measure for $\beta(t)$ converge to one and the $R^2(t)$ is the sum of the elements in $\gamma(t)$.

INSERT FIGURE 8 HERE

In figure 8 we show the total sum of the $R^2(t)$ for a one-, two- and five-minute interval. With these graphs we can discuss the efficiency of the stock as a whole. When we move to longer sampling intervals more information is incorporated in the efficient price of the stock. In the case of INTC and CMGI, this information is almost fully incorporated after five minutes. In the case of AAPL we see that after five minutes there is still a lot of specific dealer noise.

Implications of this also run into a different field of research. When measuring realized volatility of Andersen, Bollerslev, Diebold, and Labys (2001) one usually samples at 5-minute intervals. For the most liquid stocks this seems to be a good sampling interval as most information is incorporated in the stock. However for the less liquid stock microstructure effects can still dominate the movement of the stock price and thus the assumed convergence to the total integrated volatility does not hold.

6 Conclusion

This paper introduced a model for dealer quoting behavior in tick time. Quote innovations are modelled as they arrive and the model can be estimated using a Kalman Filter. From the model we derived measures for dealer efficiency and price discovery in tick time. Consequently these measures were aggregated to calendar time equivalents. We showed that when these measures are aggregated over a sufficiently long period of time, these will converge to single point estimates for price discovery.

In our empirical results we find that more volatility is generated at shorter durations. We also find that dealer quotes tend to be more informative when durations are short. Using the measures for price discovery we can identify the information share of the dealer examined.

Finally the results for the calendar time aggregation, indicate that for some stocks, the microstructure noise in the price process is still relatively large. This implies that at a sampling interval of five minutes quotes are still noisy. Hence measures like realized volatility, for some stocks might still be noise approximations of the real volatility.

A Derivation of the Moment conditions

This appendix discusses the identification problems that can arise. The assumption that there will be only one quote update leads to the restriction that $J_\ell'\iota = 1$. To make this point consider no variation in durations. Hence the duration functions can be left out for the moment. Constants are left out for the sake of notational clarity. Consider model (5) in section 2.

$$\begin{aligned} J_\ell' q_\ell &= J_\ell' \iota m_\ell + J_\ell' u_\ell, \\ m_\ell &= m_{\ell-1} + \sigma r_\ell, \\ J_\ell' u_\ell &= J_\ell' \alpha \sigma r_\ell + J_\ell' e_\ell, \end{aligned} \tag{25}$$

Given the structure of this model we can derive the variance and the first order auto covariance,

$$
\begin{aligned}
E[\Delta q_\ell^2] = \gamma_0(t_\ell) = & \\
& (1 + 2J_\ell'\alpha + J_\ell'\alpha\alpha' J_\ell + J_{\ell-1}'\alpha\alpha' J_{\ell-1})\sigma^2 \\
& + J_\ell'\Omega J_\ell + J_{\ell-1}'\Omega J_{\ell-1},
\end{aligned} \tag{26}
$$

$$
\begin{aligned}
E[\Delta q_\ell, \Delta q_{\ell-1})] = \gamma_1(t_{\ell-1}) = & \\
& - \sigma^2(J_{\ell-1}'\alpha + J_{\ell-1}'\alpha\alpha' J_{\ell-1}) \\
& - J_{\ell-1}'\Omega J_{\ell-1}.
\end{aligned} \tag{27}
$$

Note that the variance is driven only by the information in the last observation and that the first order auto covariance is driven by the lagged information. To show that α cannot be identified uniquely, define

$$\dot{\alpha} \equiv \alpha + x\iota \tag{28}$$

The moment conditions can be rewritten as

$$
\begin{aligned}
\gamma_0(t_\ell) = & \sigma^2(1 + 2J_\ell'\dot{\alpha} + J_\ell'\dot{\alpha}\dot{\alpha}' J_\ell + J_{\ell-1}'\dot{\alpha}\dot{\alpha}' J_{\ell-1}) \\
& \sigma^2(2x + x\dot{\alpha}' J_\ell + xJ_\ell'\dot{\alpha} + x\dot{\alpha}' J_{\ell-1} + xJ_{\ell-1}'\dot{\alpha} + 2x^2) \\
& + J_\ell'\dot{\Omega} J_\ell + J_{\ell-1}'\dot{\Omega} J_{\ell-1}
\end{aligned} \tag{29}
$$

$$
\begin{aligned}
\gamma_1(t_{\ell-1}) = & - \sigma^2(J_{\ell-1}'\dot{\alpha} + J_{\ell-1}'\dot{\alpha}\dot{\alpha}' J_{\ell-1}) \\
& - \sigma^2(x + xJ_{\ell-1}'\dot{\alpha} + x\dot{\alpha}' J_{\ell-1} + x^2) \\
& - J_{\ell-1}'\dot{\Omega} J_{\ell-1}.
\end{aligned} \tag{30}
$$

Which can be solved when

$$\dot{\Omega} = \Omega + \sigma^2 x\iota\iota'(\iota\iota' + \dot{\alpha}\iota' + \iota\dot{\alpha}' + x\iota\iota'). \tag{31}$$

Hence the model stated cannot identify the parameters in the α vector uniquely. When durations are included in the model the model looks like

$$
\begin{aligned}
J_\ell' q_\ell &= J_\ell'\iota m_\ell + J_\ell' u_\ell, \\
m_\ell &= m_{\ell-1} + \sigma_\ell r_\ell, \\
J_\ell' u_\ell &= J_\ell'\alpha\tau_\ell^{\delta_2}\sigma r_\ell + J_\ell' e_\ell,
\end{aligned} \tag{32}
$$

where

$$\sigma_\ell = \sigma \tau_\ell^{\delta_1}. \tag{33}$$

Again moments can be derived for this model:

$$\begin{aligned} \gamma_0(t_\ell) = {} & \sigma^2(\tau_\ell^{2\delta_1} + 2\tau_\ell^{\delta_1+\delta_2} J_\ell' \alpha + \tau_\ell^{2\delta_2} J_\ell' \alpha\alpha' J_\ell + \tau_{\ell-1}^{2\delta_2} J_{\ell-1}' \alpha\alpha' J_{\ell-1}) \\ & + J_\ell' \Omega J_\ell + J_{\ell-1}' \Omega J_{\ell-1} \end{aligned} \tag{34}$$

$$\begin{aligned} \gamma_1(t_{\ell-1}) = {} & -\sigma^2(\tau_{\ell-1}^{\delta_1+\delta_2} J_{\ell-1}' \alpha + \tau_{\ell-1}^{2\delta_2} J_{\ell-1}' \alpha\alpha' J_{\ell-1}) \\ & - J_{\ell-1}' \Omega J_{\ell-1}. \end{aligned} \tag{35}$$

The vector α can now be identified uniquely iff durations change over time. When these durations are fixed a linear substitute can be found in the way shown before and the vector α cannot be identified.

References

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, 2001, "The distribution of realized exchange rate volatility," *Journal of the American Statistical Association*, 96, 42–55.

Ané, T., and H. Geman, 2000, "Order Flow, Transaction Clock, and Normality of Asset Returns," *Journal of Finance*, 55, 2259–2284.

Baillie, R. T., G. Booth, Y. Tse, and T. Zabotina, 2002, "Price discovery and common factor models," *Journal of Financial Markets*, 5, 309–321.

Clark, P., 1973, "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41, 135–156.

De Jong, F., and P. Schotman, 2003, "Price Discovery in Fragmented Markets," Working Paper.

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

Durbin, J., and S. J. Koopman, 2001, *Time Series Analysis by State Space Methods*, Oxford University Press, Oxford.

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

Engle, R., 2000, "The Econometrics of Ultra-High frequency Data," *Econometrica*, 68, 1–22.

Engle, R. F., and A. J. Patton, 2004, "Impacts of Trades in an Error-Correction Model of quote prices," *Journal of Financial Markets*, 7, 1–25.

Glosten, L. R., and L. Harris, 1988, "Estimating the Components of the Bid/Ask Spread," *Journal of Financial Economics*, 21, 123 – 142.

Harris, L., 1987, "Transaction Data Tests of the Mixture of Distributions Hypothesis," *Journal of Financial and Quantitative Analysis*, 22, 127–141.

Harvey, A., 1989, *Forecasting, structural time series models and the Kalman Filter*, Cambridge University Press, Cambridge.

Hasbrouck, J., 1993, "Assessing the Quality of a Security Market: A New Approach to Transaction-Cost Measurement," *Review of Financial Studies*, 6, 191–212.

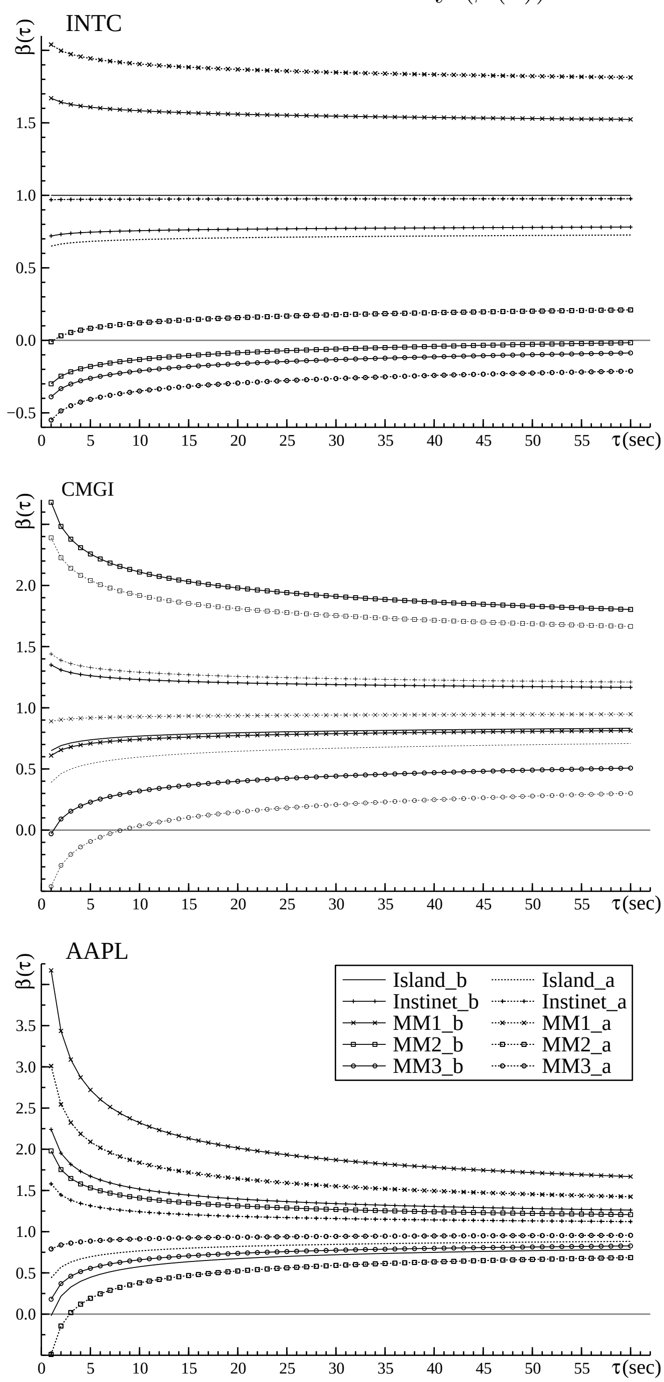
Hasbrouck, J., 1995, "One Security, Many Markets: Determining the Contribution to Price Discovery," *Journal of Finance*, 50, 1175–1199.

Huang, R. D., 2002, "The Quality of ECN and Nasdaq Market Maker Quotes," *Journal of Finance*, 57, 1285–1319.

O'Hara, M., 1995, *Market Microstructure Theory*, Blackwell Publishers Inc., Massachusetts.

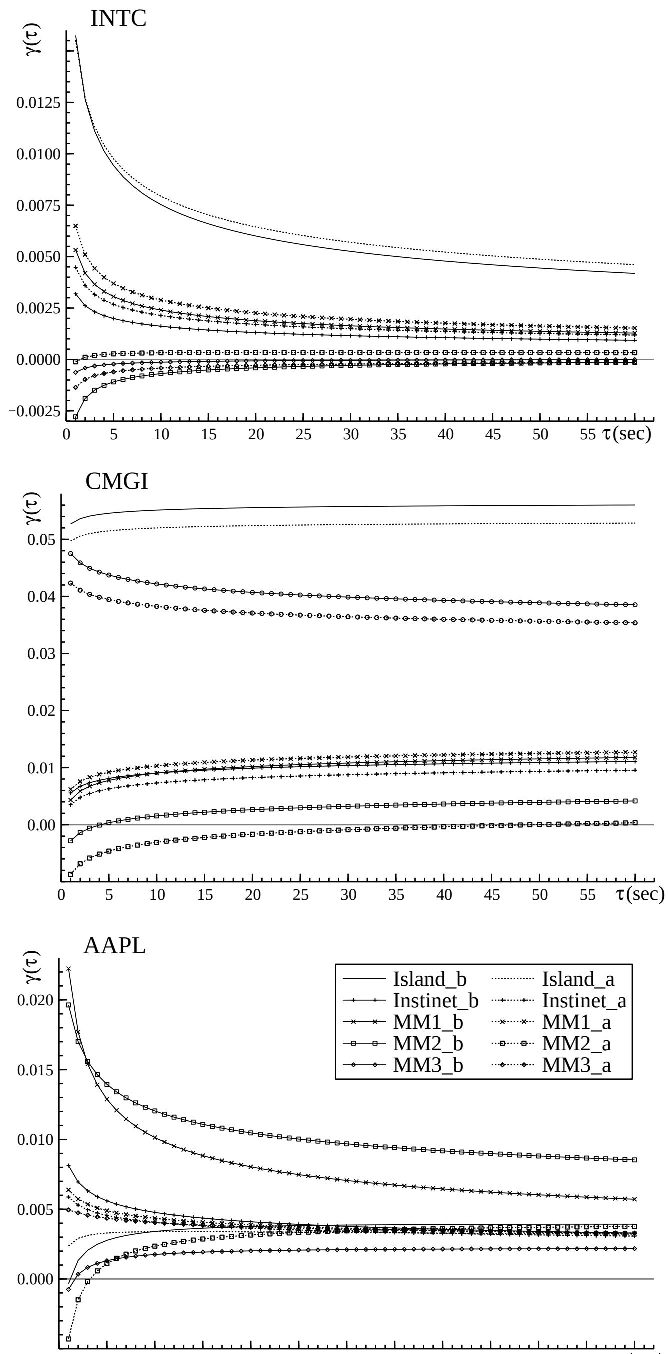
Schultz, P. H., 2003, "Who makes markets," *Journal of Financial Markets*, 6, 49–72.

Figure 2: Tick time measure for dealer efficiency ($\beta(\tau)$) as a function of duration



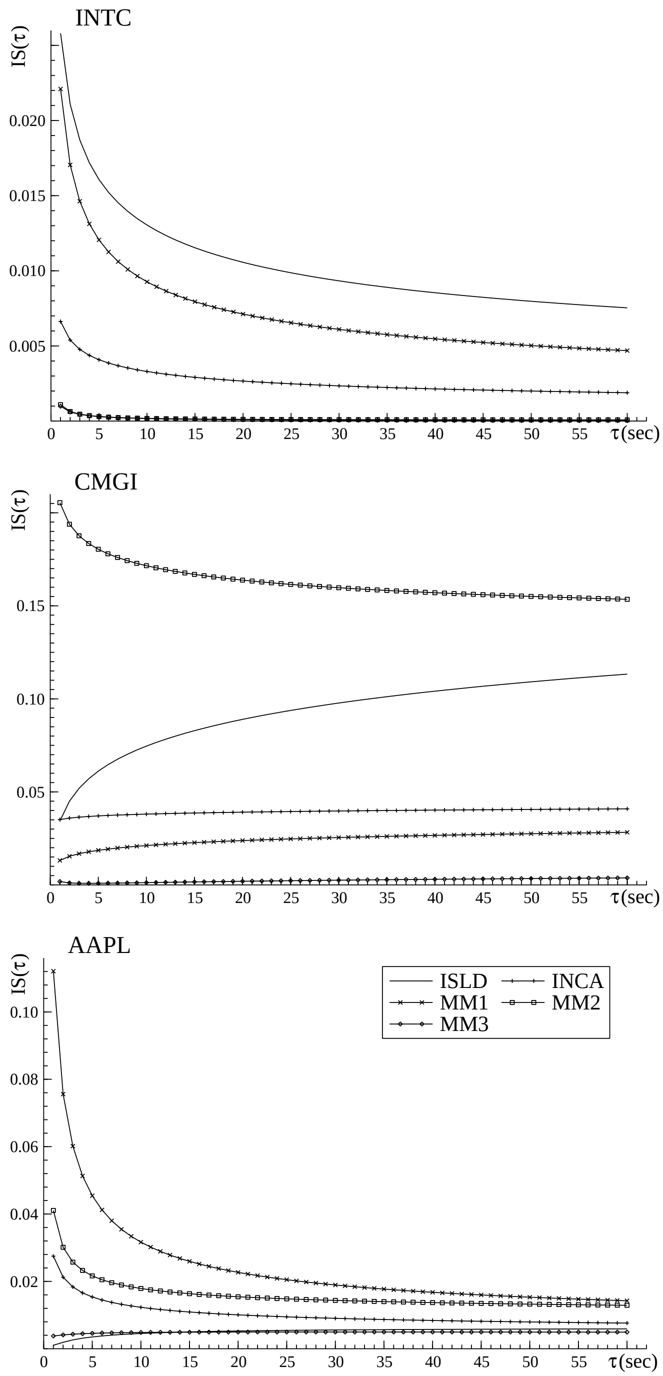
Note: These graphs show the dealer efficiency measures in tick time for INTC, CMGI and AAPL. These measures are plotted as a function of duration.

Figure 3: Tick time measure for price discovery ($\gamma(\tau)$) as a function of duration



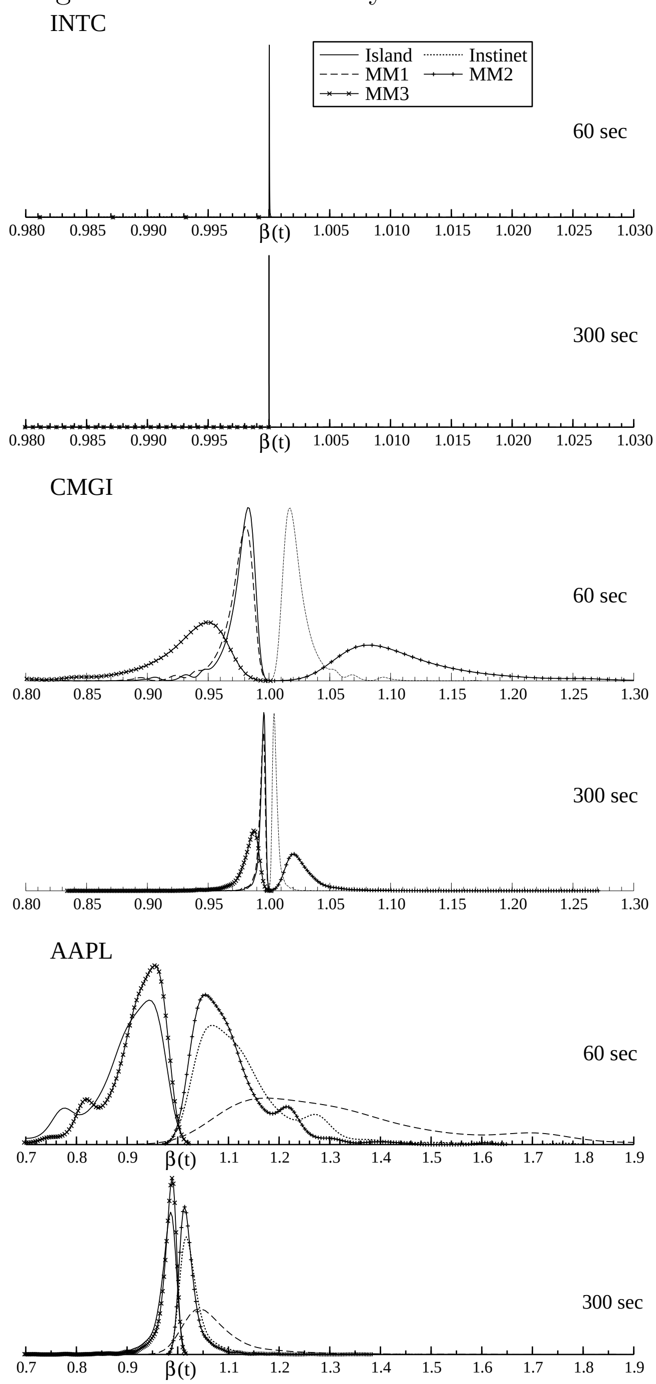
Note: These graphs show the price discovery measures in tick time for INTC, CMGI and AAPL. These measures are plotted as a function of duration.

Figure 4: Tick time measure for Information Shares as a function of duration



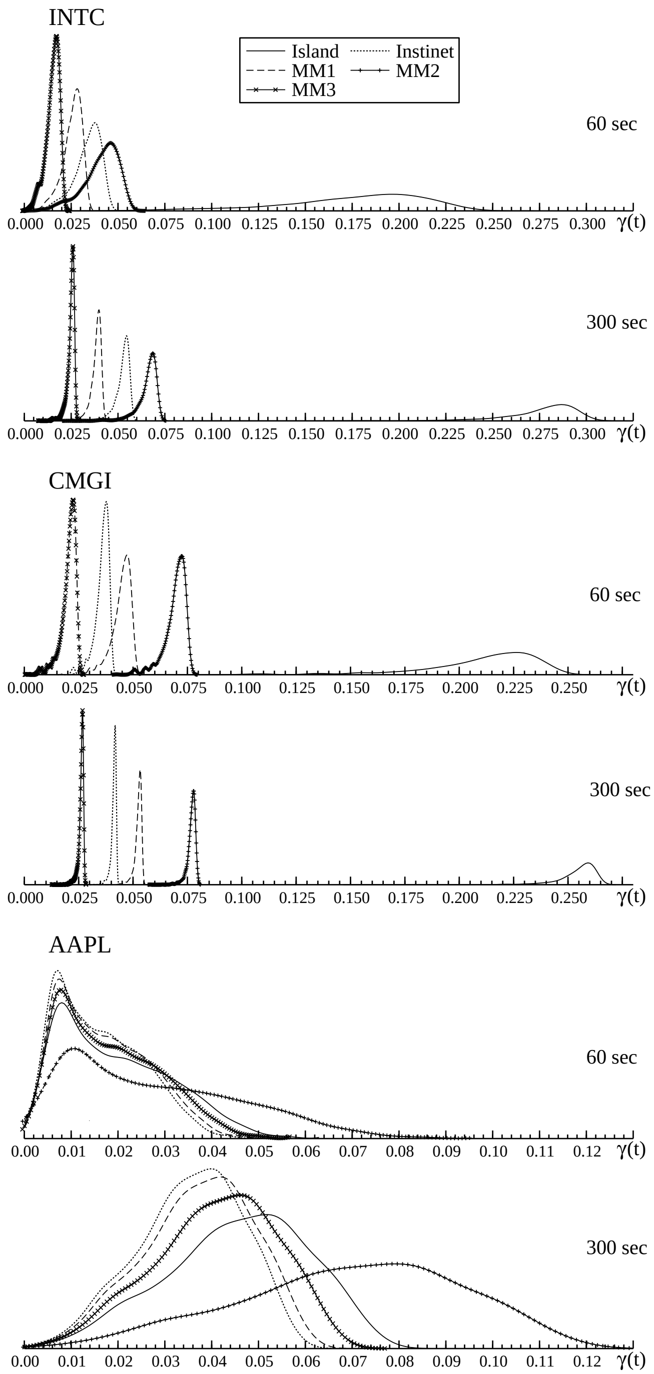
Note: These graphs show information shares in tick time for INTC, CMGI and AAPL per dealer. These measures are plotted as a function of duration.

Figure 5: Dealer efficiency in Calendar time



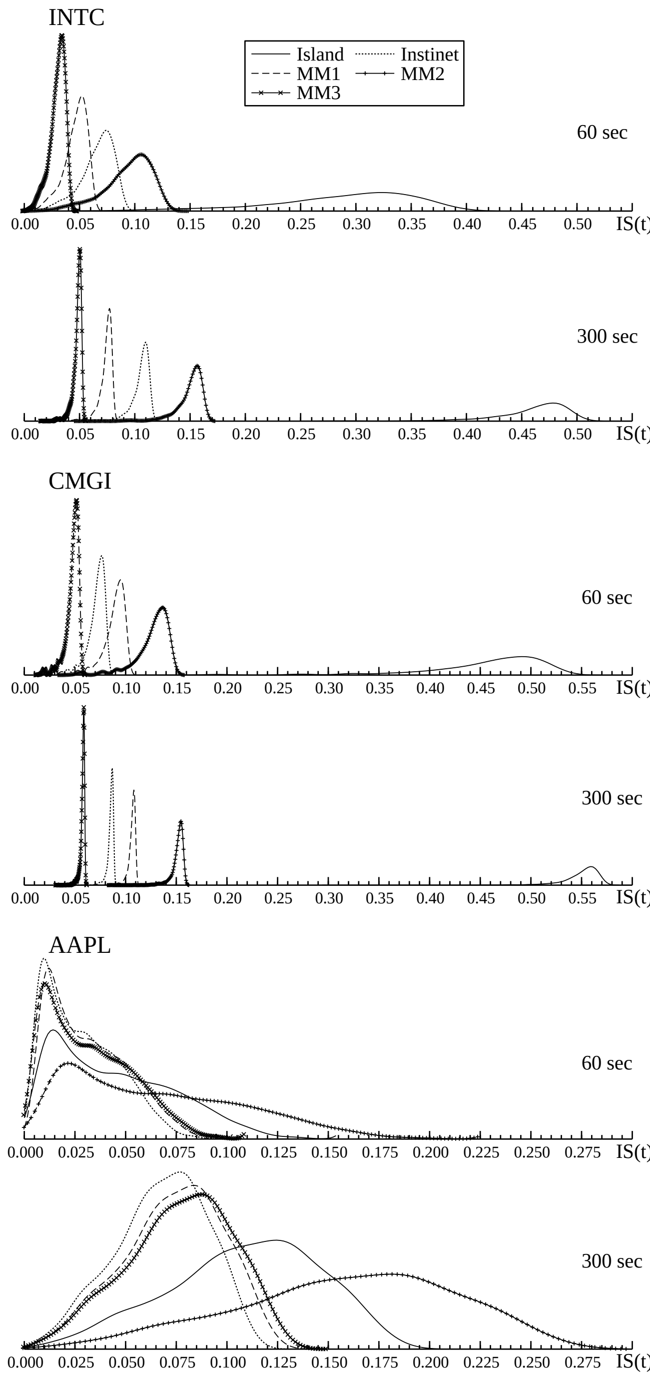
Note: These graphs show the distributions of the calendar time measures for dealer efficiency. These distributions are shown for the bids of INTC, CMGI and AAPL. Island is not shown as its value is always 1. The aggregates are shown at 60 and 300 second intervals.

Figure 6: Price Discovery in Calendar time



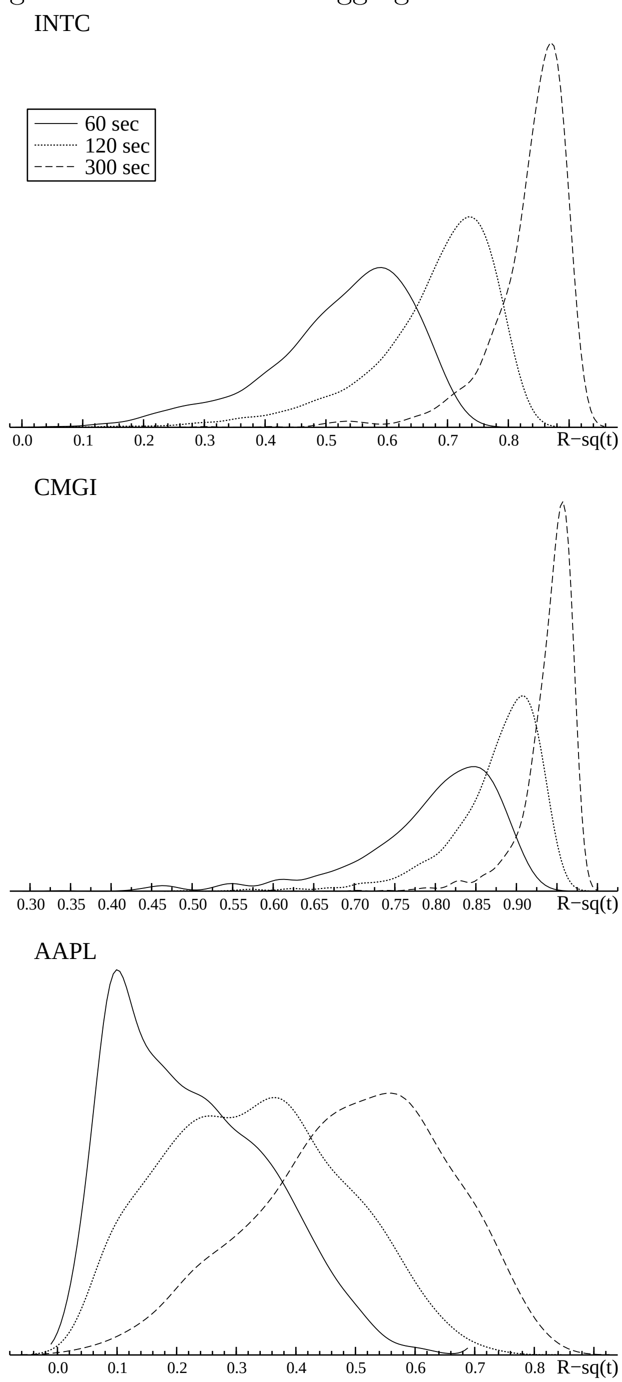
Note: These graphs show the distributions of the calendar time measures for price discovery. These distributions are shown for the asks of INTC, CMGI and AAPL. The aggregates are shown at 60 and 300 second intervals.

Figure 7: Information Shares in Calendar time per dealer



Note: These graphs show the distributions of the calendar time measures for Information Shares. These distributions are shown for INTC, CMGI and AAPL. The aggregates are shown at 60 and 300 second intervals. This measure is obtained by taking the inner product of the measure for dealer efficiency and price discovery. Using the specific structure of the model and by applying the matrix inversion lemma, the inner product can be decomposed to dealers.

Figure 8: Calendar time aggregated Information



Note: These graphs show the distributions of the total information incorporated in dealer quotes over a specific time interval. These distributions are shown for INTC, CMGI and AAPL. The aggregates are shown at 60, 120 and 300 second intervals.

Table 1: List of Company Ticker symbols and Company names

Ticker Symbol	Company name
AAPL	Apple Computer Inc.
AMAT	Applied Materials Inc.
AMGN	Amgen Inc.
AMZN	Amazon.com, Inc.
ATHM	At Home Corporation
CMGI	CMGI, Inc.
COMS	3Com Corporation
CPWR	Compuware Corporation
CSCO	Cisco Systems Inc.
DELL	Dell Computer Corporation
INTC	Intel Corporation
MSFT	Microsoft Corporation
NOVL	Novell Inc.
NXTL	Nextel Communications CL-A
ORCL	Oracle Corporation
PSFT	Peoplesoft Inc.
QWST	Qwest Communications Intl Inc.
SBUX	Starbucks Corporation
SUNW	Sun Microsystems Inc.
WCOM	MCI WorldCom Inc.

Table 2: Summary statistics

Stock	# of Obs	% of single quotes	Average Duration	Std of Duration
AAPL	29,787	89.63	15.71	28.73
AMAT	105,090	83.58	4.77	5.95
AMGN	40,279	83.97	12.32	22.24
AMZN	150,710	80.52	3.44	4.39
ATHM	76,435	86.50	6.34	9.40
CMGI	90,401	87.06	5.34	7.75
COMS	61,049	89.56	7.68	11.34
CPWR	33,301	90.55	13.92	31.03
CSCO	164,480	80.85	3.13	3.52
DELL	177,850	77.39	3.02	3.44
INTC	171,260	76.92	3.15	3.56
MSFT	151,110	80.82	3.42	3.94
NOVL	18,909	87.88	25.08	43.93
NXTL	19,556	91.63	23.42	42.23
ORCL	87,774	85.81	5.56	7.46
PSFT	24,601	91.12	18.74	32.02
QWST	44,459	88.331	10.68	18.12
SBUX	14,320	90.056	32.43	64.92
SUNW	128,370	82.43	4.20	5.15
WCOM	88,550	83.33	5.66	7.41

Note: This table presents some summary statistics of the data after being filtered. Column 1 represents the total number of quotes issued for each stock. Column 2 indicates in how many cases there is only 1 quote innovation in the observation matrix. Column 3 gives the average duration between any quote updates and column 4 gives the respective standard deviation.

Table 3: Duration Parameters

Stock	δ_1	δ_2
AAPL	-0.06 (0.03)	-0.44 (0.05)
AMAT	0.02 (0.02)	-1.20 (0.17)
AMGN	0.04 (0.02)	-0.21 (0.17)
AMZN	0.03 (0.01)	0.50 (0.02)
ATHM	0.01 (0.02)	-1.26 (0.18)
CMGI	0.07 (0.01)	-0.11 (0.04)
COMS	-0.11 (0.02)	-0.51 (0.06)
CPWR	0.00 (0.02)	-0.72 (0.13)
CSCO	-0.07 (0.01)	-0.24 (0.04)
DELL	-0.13 (0.01)	-0.30 (0.03)
INTC	-0.17 (0.02)	-0.23 (0.04)
MSFT	-0.20 (0.02)	-0.13 (0.03)
NOVL	-0.23 (0.04)	-0.59 (0.05)
NXTL	-0.05 (0.03)	-0.44 (0.08)
ORCL	0.02 (0.02)	-0.82 (0.05)
PSFT	0.00 (0.02)	-0.23 (0.03)
QWST	-0.14 (0.02)	0.23 (0.03)
SBUX	-0.01 (0.03)	-0.32 (0.06)
SUNW	-0.06 (0.02)	-0.96 (0.11)
WCOM	-0.11 (0.02)	-1.17 (0.14)

Note: This table reports the estimates and standard errors of the duration parameters in the model,

$$\begin{aligned} J'_\ell q_\ell &= J'_\ell c + J'_\ell \iota m_\ell + J'_\ell u_\ell, \\ m_\ell &= m_{\ell-1} + \tau_\ell^{\delta_1} r_\ell, \\ J'_\ell u_\ell &= J'_\ell \alpha \tau_\ell^{\delta_2} r_\ell + J'_\ell e_\ell, \end{aligned}$$

Column 1 gives the duration parameters on the random walk component, column 2 reports the duration parameters of the correlation between dealer quotes and the random walk.

Table 4: Quote specific relations to the efficient price

Stock	$ISLD_b$	$ISLD_a$	$INCA_b$	$INCA_a$	$MM1_b$	$MM1_a$	$MM2_b$	$MM2_a$	$MM3_b$	$MM3_a$
AAPL	-1.02	-0.56	1.24	0.58	3.17	2.01	0.98	-1.49	-0.82	-0.21
	(0.20)	(0.21)	(0.37)	(0.26)	(0.40)	(0.35)	(0.19)	(0.25)	(0.36)	(0.28)
AMAT	-0.03	-0.06	0.07	0.38	0.49	0.33	-0.08	-0.12	0.02	-0.63
	(0.02)	(0.02)	(0.05)	(0.10)	(0.13)	(0.11)	(0.05)	(0.06)	(0.14)	(0.19)
AMGN	-1.59	-0.69	0.13	0.14	-1.44	-0.97	-1.70	-1.70	0.17	-0.13
	(0.43)	(0.42)	(0.20)	(0.28)	(0.69)	(0.60)	(0.93)	(1.09)	(0.41)	(0.43)
AMZN	0.00	0.12	0.93	1.38	-0.33	-0.66	-0.97	-5.21	-0.12	-0.05
	(0.05)	(0.05)	(0.08)	(0.08)	(0.12)	(0.13)	(0.16)	(0.13)	(0.20)	(0.18)
ATHM	-0.08	-0.11	0.08	0.04	0.38	0.28	-0.15	-0.16	-0.13	-0.17
	(0.03)	(0.04)	(0.04)	(0.03)	(0.13)	(0.10)	(0.06)	(0.07)	(0.08)	(0.09)
CMGI	-0.35	-0.61	0.35	0.44	-0.39	-0.11	1.68	1.39	-1.03	-1.46
	(0.08)	(0.09)	(0.15)	(0.18)	(0.17)	(0.19)	(0.15)	(0.15)	(0.21)	(0.23)
COMS	-0.07	-0.02	1.27	0.79	-1.21	-1.02	0.00	-0.24	0.00	1.49
	(0.08)	(0.07)	(0.17)	(0.18)	(0.28)	(0.30)	(0.14)	(0.22)	(0.17)	(0.28)
CPWR	-0.22	-0.31	-0.12	0.04	0.59	0.37	0.79	0.89	0.12	-0.26
	(0.09)	(0.11)	(0.08)	(0.08)	(0.21)	(0.19)	(0.32)	(0.36)	(0.18)	(0.18)
CSCO	-0.41	-0.42	0.44	0.48	0.55	0.03	0.76	0.80	-3.00	-2.92
	(0.07)	(0.08)	(0.10)	(0.09)	(0.18)	(0.25)	(0.20)	(0.25)	(0.22)	(0.23)
DELL	-0.47	-0.47	0.12	0.26	4.54	5.34	0.01	0.60	-2.39	-2.84
	(0.04)	(0.04)	(0.08)	(0.07)	(0.34)	(0.37)	(0.12)	(0.16)	(0.11)	(0.18)
INTC	0.00	-0.35	-0.28	-0.03	0.67	1.04	-1.39	-1.01	-1.39	-1.55
	(0.09)	(0.08)	(0.09)	(0.08)	(0.15)	(0.16)	(0.10)	(0.11)	(0.20)	(0.19)
MSFT	-0.07	-0.07	-0.10	0.05	-0.81	-0.18	-1.31	-1.22	-2.80	-2.73
	(0.09)	(0.09)	(0.10)	(0.09)	(0.19)	(0.23)	(0.12)	(0.18)	(0.17)	(0.18)
NOVL	1.16	-0.07	1.69	0.16	1.73	-2.58	0.30	1.00	0.15	-0.07
	(0.32)	(0.12)	(0.28)	(0.11)	(0.32)	(0.44)	(0.15)	(0.20)	(0.16)	(0.16)
NXTL	-1.44	-1.25	-0.49	-0.01	1.89	1.23	1.51	-1.32	0.16	-1.50
	(0.39)	(0.42)	(0.23)	(0.34)	(0.45)	(0.42)	(0.48)	(0.62)	(0.20)	(0.38)
ORCL	-0.34	-0.46	0.20	-0.04	2.62	-2.15	-0.27	-0.05	-0.42	-0.26
	(0.06)	(0.06)	(0.07)	(0.09)	(0.22)	(0.18)	(0.07)	(0.08)	(0.12)	(0.15)
PSFT	0.93	-0.43	1.90	0.18	11.98	9.51	8.09	9.20	-0.24	2.33
	(0.36)	(0.21)	(0.30)	(0.26)	(1.44)	(1.28)	(1.03)	(1.14)	(0.54)	(0.56)
QWST	0.99	0.22	4.57	5.46	-0.45	0.23	2.79	3.44	0.70	0.37
	(0.69)	(0.43)	(0.47)	(0.42)	(0.42)	(0.63)	(0.71)	(1.11)	(0.19)	(0.21)
SBUX	0.14	-0.08	1.21	0.22	1.87	1.28	-0.52	-1.10	-0.27	-0.16
	(0.19)	(0.32)	(0.34)	(0.23)	(0.41)	(0.31)	(0.34)	(0.43)	(0.24)	(0.19)
SUNW	-0.08	-0.31	-0.04	0.00	-0.55	-0.50	-1.07	-1.61	0.10	0.10
	(0.06)	(0.08)	(0.07)	(0.06)	(0.11)	(0.14)	(0.19)	(0.27)	(0.11)	(0.13)
WCOM	-0.02	-0.12	0.24	-0.13	0.65	0.49	-0.26	-0.36	0.17	-0.05
	(0.04)	(0.05)	(0.06)	(0.05)	(0.17)	(0.15)	(0.08)	(0.09)	(0.08)	(0.06)

Note: This table presents the estimates for α and their standard errors. These $\alpha's$ are estimated in

$$J_\ell' q_\ell = J_\ell' c + J_\ell' \iota m_\ell + J_\ell' u_\ell,$$
$$m_\ell = m_{\ell-1} + \tau_\ell^{\delta_1} r_\ell,$$
$$J_\ell' u_\ell = J_\ell' \alpha \tau_\ell^{\delta_2} r_\ell + J_\ell' e_\ell.$$

The subscripts b and a represent the respective bid or ask side of a dealer.

Table 5: Average values for dealer efficiency ($\beta(t)$) per dealer quote

Stock	$ISLD_b$	$ISLD_a$	$INCA_b$	$INCA_a$	$MM1_b$	$MM1_a$	$MM2_b$	$MM2_a$	$MM3_b$	$MM3_a$
60 Second intervals										
AAPL	0.89	0.94	1.13	1.06	1.34	1.22	1.11	0.84	0.91	0.98
AMAT	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99
AMGN	0.81	0.92	1.02	1.02	0.83	0.89	0.80	0.80	1.02	0.98
AMZN	1.00	1.02	1.14	1.21	0.95	0.90	0.85	0.22	0.98	0.99
ATHM	1.00	1.00	1.00	1.00	1.01	1.01	1.00	1.00	1.00	1.00
CMGI	0.98	0.96	1.02	1.03	0.97	0.99	1.12	1.10	0.93	0.90
COMS	1.00	1.00	1.08	1.05	0.92	0.94	1.00	0.99	1.00	1.10
CPWR	0.99	0.99	0.99	1.00	1.03	1.02	1.04	1.04	1.01	0.99
CSCO	0.98	0.98	1.02	1.02	1.02	1.00	1.03	1.03	0.87	0.87
DELL	0.98	0.98	1.01	1.01	1.19	1.23	1.00	1.03	0.90	0.88
INTC	1.00	0.98	0.99	1.00	1.03	1.05	0.93	0.95	0.93	0.93
MSFT	1.00	1.00	0.99	1.00	0.95	0.99	0.92	0.93	0.83	0.84
NOVL	1.17	0.99	1.25	1.02	1.26	0.62	1.04	1.15	1.02	0.99
NXTL	0.82	0.84	0.94	1.00	1.24	1.16	1.19	0.83	1.02	0.81
ORCL	0.99	0.99	1.01	1.00	1.07	0.94	0.99	1.00	0.99	0.99
PSFT	1.16	0.93	1.33	1.03	3.05	2.63	2.39	2.58	0.96	1.40
QWST	1.52	1.11	3.39	3.85	0.76	1.12	2.46	2.80	1.37	1.19
SBUX	1.02	0.99	1.21	1.04	1.33	1.23	0.91	0.80	0.95	0.97
SUNW	1.00	0.99	1.00	1.00	0.99	0.99	0.97	0.96	1.00	1.00
WCOM	1.00	1.00	1.01	1.00	1.02	1.01	0.99	0.99	1.00	1.00
300 Second intervals										
AAPL	0.97	0.99	1.03	1.02	1.08	1.05	1.03	0.96	0.98	0.99
AMAT	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AMGN	0.96	0.98	1.00	1.00	0.96	0.97	0.95	0.95	1.00	1.00
AMZN	1.00	1.00	1.03	1.05	0.99	0.98	0.97	0.81	1.00	1.00
ATHM	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
CMGI	0.99	0.99	1.01	1.01	0.99	1.00	1.03	1.02	0.98	0.98
COMS	1.00	1.00	1.02	1.01	0.98	0.99	1.00	1.00	1.00	1.02
CPWR	1.00	1.00	1.00	1.00	1.01	1.00	1.01	1.01	1.00	1.00
CSCO	1.00	1.00	1.00	1.00	1.00	1.00	1.01	1.01	0.97	0.97
DELL	1.00	1.00	1.00	1.00	1.04	1.05	1.00	1.01	0.98	0.98
INTC	1.00	1.00	1.00	1.00	1.01	1.01	0.99	0.99	0.99	0.99
MSFT	1.00	1.00	1.00	1.00	0.99	1.00	0.98	0.99	0.97	0.97
NOVL	1.04	1.00	1.06	1.01	1.07	0.90	1.01	1.04	1.01	1.00
NXTL	0.96	0.96	0.98	1.00	1.06	1.04	1.05	0.96	1.01	0.95
ORCL	1.00	1.00	1.00	1.00	1.02	0.99	1.00	1.00	1.00	1.00
PSFT	1.04	0.98	1.08	1.01	1.47	1.38	1.32	1.36	0.99	1.09
QWST	1.12	1.03	1.54	1.65	0.95	1.03	1.33	1.41	1.08	1.04
SBUX	1.01	1.00	1.06	1.01	1.10	1.07	0.97	0.94	0.99	0.99
SUNW	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.99	1.00	1.00
WCOM	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Note: This table reports the averages for dealer efficiency ($\beta(t)$) per dealer quote. This measure is defined as the covariance between quote innovations and the change in the efficient price, divided by the variance of the efficient price change. Panel A shows this measure aggregated up to 60 seconds, panel B up to 300 seconds. All measures over time converge to 1.

Table 6: Average values for price discovery ($\gamma(t)$) per dealer quote

Stock	$ISLD_b$	$ISLD_a$	$INCA_b$	$INCA_a$	$MM1_b$	$MM1_a$	$MM2_b$	$MM2_a$	$MM3_b$	$MM3_a$
					60 Second intervals					
AAPL	0.03	0.02	0.02	0.02	0.02	0.02	0.04	0.03	0.01	0.02
AMAT	0.14	0.22	0.04	0.04	0.05	0.05	0.03	0.03	0.02	0.02
AMGN	0.06	0.06	0.03	0.02	0.06	0.06	0.03	0.02	0.05	0.04
AMZN	0.30	0.38	0.02	0.03	0.01	0.01	0.07	0.04	0.01	0.01
ATHM	0.13	0.13	0.04	0.04	0.06	0.05	0.04	0.04	0.02	0.03
CMGI	0.24	0.21	0.04	0.04	0.04	0.04	0.07	0.07	0.02	0.02
COMS	0.10	0.08	0.03	0.03	0.02	0.02	0.03	0.02	0.03	0.03
CPWR	0.03	0.03	0.03	0.03	0.07	0.06	0.04	0.04	0.02	0.02
CSCO	0.30	0.32	0.02	0.02	0.02	0.01	0.01	0.01	0.02	0.01
DELL	0.34	0.41	0.01	0.02	0.01	0.01	0.01	0.00	0.01	0.01
INTC	0.12	0.18	0.03	0.03	0.02	0.03	0.05	0.04	0.01	0.02
MSFT	0.16	0.14	0.03	0.03	0.02	0.02	0.03	0.02	0.02	0.02
NOVL	0.01	0.01	0.01	0.01	0.01	0.00	0.02	0.02	0.01	0.01
NXTL	0.01	0.00	0.03	0.02	0.03	0.02	0.02	0.01	0.04	0.02
ORCL	0.12	0.12	0.03	0.03	0.03	0.03	0.06	0.05	0.03	0.05
PSFT	0.02	0.03	0.02	0.02	0.08	0.09	0.00	0.00	0.02	0.01
QWST	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.05	0.05
SBUX	0.00	0.00	0.00	0.00	0.32	0.30	0.00	0.00	0.01	0.01
SUNW	0.11	0.13	0.04	0.04	0.03	0.02	0.04	0.04	0.03	0.02
WCOM	0.05	0.07	0.03	0.04	0.02	0.02	0.04	0.03	0.01	0.02
					300 Second intervals					
AAPL	0.06	0.05	0.03	0.04	0.04	0.04	0.09	0.07	0.04	0.04
AMAT	0.19	0.30	0.05	0.05	0.07	0.07	0.04	0.05	0.02	0.03
AMGN	0.12	0.10	0.04	0.04	0.10	0.10	0.06	0.04	0.07	0.06
AMZN	0.32	0.41	0.03	0.03	0.02	0.02	0.08	0.05	0.01	0.01
ATHM	0.20	0.19	0.05	0.06	0.08	0.08	0.06	0.06	0.03	0.04
CMGI	0.29	0.25	0.04	0.04	0.05	0.05	0.08	0.08	0.03	0.03
COMS	0.17	0.15	0.06	0.04	0.04	0.04	0.05	0.04	0.06	0.05
CPWR	0.05	0.05	0.05	0.06	0.13	0.11	0.06	0.06	0.04	0.04
CSCO	0.37	0.40	0.03	0.03	0.02	0.02	0.01	0.01	0.03	0.02
DELL	0.39	0.48	0.01	0.02	0.01	0.01	0.01	0.00	0.02	0.01
INTC	0.19	0.28	0.05	0.05	0.04	0.04	0.08	0.07	0.02	0.02
MSFT	0.26	0.24	0.05	0.05	0.03	0.03	0.05	0.04	0.04	0.03
NOVL	0.02	0.02	0.02	0.02	0.02	0.01	0.04	0.04	0.03	0.04
NXTL	0.02	0.01	0.06	0.05	0.05	0.05	0.04	0.04	0.08	0.06
ORCL	0.18	0.18	0.05	0.05	0.05	0.04	0.09	0.08	0.05	0.07
PSFT	0.06	0.08	0.04	0.04	0.12	0.14	0.01	0.00	0.04	0.03
QWST	0.03	0.04	0.02	0.02	0.06	0.05	0.03	0.02	0.10	0.11
SBUX	0.00	0.00	0.01	0.01	0.41	0.38	0.01	0.01	0.01	0.01
SUNW	0.17	0.21	0.07	0.07	0.06	0.04	0.06	0.06	0.04	0.04
WCOM	0.11	0.14	0.06	0.07	0.04	0.04	0.07	0.07	0.03	0.03

*Note:*This table reports the averages for price discovery ($\gamma(t)$) per dealer quote. This measure is obtained from the regression of the efficient price change on quote innovations. Panel A shows this measure aggregated up to 60 seconds, panel B up to 300 seconds.

Table 7: Averages for Information shares per dealer

	60 seconds					300 seconds				
Stock	Island	Instinet	MM1	MM2	MM3	Island	Instinet	MM1	MM2	MM3
AAPL	0.04	0.03	0.03	0.07	0.03	0.11	0.07	0.07	0.16	0.08
AMAT	0.35	0.07	0.10	0.06	0.04	0.50	0.10	0.14	0.09	0.06
AMGN	0.13	0.04	0.11	0.06	0.07	0.22	0.08	0.20	0.11	0.12
AMZN	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00
ATHM	0.26	0.07	0.11	0.08	0.05	0.39	0.11	0.16	0.12	0.07
CMGI	0.46	0.07	0.09	0.13	0.05	0.55	0.09	0.11	0.15	0.06
COMS	0.17	0.05	0.05	0.05	0.06	0.32	0.10	0.09	0.09	0.11
CPWR	0.05	0.06	0.13	0.07	0.04	0.10	0.10	0.24	0.12	0.07
CSCO	0.61	0.04	0.03	0.02	0.05	0.77	0.05	0.03	0.02	0.06
DELL	0.74	0.03	0.01	0.01	0.03	0.87	0.03	0.01	0.01	0.03
INTC	0.29	0.07	0.05	0.09	0.03	0.46	0.11	0.07	0.15	0.05
MSFT	0.28	0.05	0.03	0.05	0.04	0.49	0.09	0.05	0.09	0.08
NOVL	0.01	0.02	0.01	0.04	0.03	0.03	0.04	0.03	0.09	0.06
NXTL	0.01	0.04	0.04	0.03	0.06	0.03	0.11	0.09	0.08	0.14
ORCL	0.23	0.07	0.05	0.11	0.08	0.36	0.10	0.08	0.17	0.12
PSFT	0.00	0.00	0.98	0.00	0.00	0.00	0.00	1.00	0.00	0.00
QWST	0.03	0.03	0.04	0.03	0.09	0.07	0.07	0.09	0.06	0.21
SBUX	0.00	0.01	0.74	0.01	0.01	0.01	0.01	0.88	0.01	0.01
SUNW	0.23	0.08	0.06	0.07	0.05	0.39	0.14	0.09	0.12	0.08
WCOM	0.12	0.07	0.04	0.07	0.03	0.25	0.14	0.08	0.14	0.06

Note: This table reports the averages for information shares per dealer. This measure is obtained as the inner product of $\lambda(t)$ and $\theta(t)$. The matrix inversion lemma is used to assign the information shares to separate dealers. This is possible as we impose structure on the idiosyncratic dealer noise. We reports the measure aggregated up to 60 seconds and 300 seconds.