



Price discovery in tick time[☆]

Bart Frijns^{a,*}, Peter Schotman^b

^a Auckland University of Technology, New Zealand

^b LIFE, Maastricht University and CEPR, the Netherlands

ARTICLE INFO

Article history:

Received 10 September 2007

Received in revised form 25 February 2009

Accepted 7 July 2009

Available online 17 July 2009

JEL classifications:

C32

G15

Keywords:

Price discovery

Tick time models

Nasdaq

Ultra-high frequency data

Microstructure

ABSTRACT

This paper develops a tick time model for the quote setting dynamics on Nasdaq. The model decomposes quotes into an efficient price, asymmetric information and noise. Both the evolution of the efficient price and the information contents of quotes depend on quote durations. New measures for the contribution to price discovery are defined within this model. When aggregated to fixed calendar time intervals, they relate closely to Hasbrouck [Hasbrouck, Joel, 1995, One security, many markets: determining the contribution to price discovery, *Journal of Finance* 50, 1175–1199] information shares. Empirical results for 20 Nasdaq stocks indicate that ECNs, in particular Island, contribute most to price discovery for active stocks. For less active stocks, wholesale market makers contribute most.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

An important measure for price discovery in fragmented markets where information comes from multiple sources (markets or dealers) is the information share introduced by Hasbrouck (1995). This measure has been used extensively studying price discovery for shares listed in various US markets (e.g. Hasbrouck, 1995), among different dealer types trading at Nasdaq (Huang, 2002), among internationally cross-listed shares (e.g. Grammig et al., 2005), and among floor- and screen-based trading systems in the US index futures market (e.g. Hasbrouck, 2003).

The information share, however, typically is not a unique measure when prices are contemporaneously correlated, in which case we can only derive a range for the information share. This range is generally small when the number of information sources is small and the differences among them are large, but in empirical applications with more than two information sources this range is often found to be large. A typical example, one we revisit in this paper, is Huang (2002), who uses the information share to assess the contribution to price discovery for different ECNs (Electronic Communication Networks) and traditional dealers at Nasdaq. Sampling at a one minute frequency, he finds that for some stocks the information share of an important ECN, Island, ranges between 25% and 85%. Baillie et al. (2002) analytically show that upper and lower bounds can differ substantially when

[☆] We thank the editor (Richard Baillie), an anonymous referee, Rob Engle, Joel Hasbrouck, Ronald Mahieu, Bert Menkveld, Franz Palm, Paul Schultz, Avi Wohl, Christian Wolff, Günther Wuijts, seminar participants at Erasmus University, Free University Amsterdam, Auckland University of Technology and New York University, and the participants of the workshop on the Financial Econometrics of Microstructure (Tilburg), and the annual conferences of the European Finance Association (Maastricht), Econometric Society European Meetings (Madrid) and American Finance Association (Philadelphia) for their useful comments and suggestions. All errors are of course our own.

* Corresponding author. Department of Finance, Auckland University of Technology, Private Bag 92006, 1142 Auckland, New Zealand. Tel.: +64 9 921 5706; fax: +64 9 921 9940.

E-mail addresses: bfrijns@aut.ac.nz (B. Frijns), p.schotman@finance.unimaas.nl (P. Schotman).

contemporaneous correlations between price innovations are high, and empirically show that the information share of ECNs for a liquid stock like Yahoo is between 20% and 97%.

A driving factor behind the contemporaneous correlation between prices is data aggregation. When sampling at relatively long intervals it is difficult to observe any informational asymmetries, as each market or dealer has had the time to update their prices. Sampling at higher frequencies minimizes this problem, but introduces stale prices, which may affect measures of price discovery. Another alternative is to sample in tick time, i.e., sample at the frequency at which prices arrive. This minimizes contemporaneous correlations between prices, yet does not introduce stale prices. The aim of this paper is to introduce a price discovery model for tick time data.

This paper makes three contributions to the price discovery literature. First, since tick time data is difficult to model with a vector autoregression, we propose an unobserved components model. This model is a multivariate extension of the structural time series model of Hasbrouck (1993). Within this model the Kalman filter can easily accommodate the irregular timing of tick time observations.¹ Second, when using tick time data, we observe the time between observations. Time is an important factor in microstructure models (e.g. Engle and Patton, 2004 and Furfine, 2007) and is shown to affect the volatility of the efficient price (Engle, 2000) and the information content of trades (Dufour and Engle, 2000). Both effects are included in our model. Third, based on our model, we define new measures for price discovery. We first define measures in tick time, which are duration dependent, and subsequently compute their calendar time equivalents. In tick time, we obtain three measures that address different aspects of the price discovery process. In calendar time we find only one informative measure which resembles Hasbrouck's (1995) information share.

In our empirical part we examine price discovery between two ECNs (Island and Instinet) and three wholesale market makers at Nasdaq for 20 actively traded stocks. We find that the price process does not evolve in calendar time, but in tick time as predicted by Clark (1973). We also find that the information flow to the efficient price is in general less at longer durations, which is in line with Easley and O'Hara (1992). Price discovery measures in tick time are strongly dependent on durations, where some dealers reveal more information when durations are short (active markets) whereas others reveal more when durations are long (inactive markets). Aggregating to calendar time we can often clearly identify the dominant dealer. In terms of price discovery we find that Island tends to dominate the liquid stocks, whereas market makers dominate the less liquid stocks in the sample.

2. A model for quotes in tick time

This section introduces a structural time series model for quote data in tick time. The model is an extension to Hasbrouck's (1993) unobserved components model and is theoretically motivated by Glosten and Harris' (1988) asymmetric information model.

In tick time we need to define which events are informative and should be considered as observations. We consider each change in a dealer quote as an observation and use only these changes to model the price discovery process. This avoids a potential bias from using stale quotes. Because observations are spaced irregularly, we observe the durations between quote innovations. We can therefore study quote setting behavior conditional on durations.

To formalize, consider a dealer market where M dealers issue bid and ask quotes, which arrive at times t_ℓ ($\ell = 1, \dots, L$). Let q_ℓ be the $(2M \times 1)$ vector of all standing quotes at time t_ℓ , where the bid (ask) of dealer i corresponds to element $2i - 1$ ($2i$) of q_ℓ . The duration between two consecutive quote arrivals is measured by $\tau_\ell = t_\ell - t_{\ell-1}$.

Since our interest is in modeling the dynamics of quote updates, we need to select those elements of q_ℓ that are updated at t_ℓ . To do this we define the $(k_\ell \times 2M)$ selection matrix J_ℓ , which contains the rows of the identity matrix that correspond to the elements of q_ℓ that are updated at t_ℓ . The vector of updated quotes is $J_\ell q_\ell$.

We assume that the time series of updated quotes can be described by an unobserved components model,

$$J_\ell q_\ell = J_\ell(c + \mu_\ell + u_\ell), \quad (1)$$

where μ_ℓ is the permanent price component common to all dealers and will be referred to as the efficient price; ι , c , and u_ℓ are $2M$ -vectors of ones, constants and temporary deviations from μ_ℓ , respectively.

The efficient price is assumed to follow a random walk,

$$\mu_\ell = \mu_{\ell-1} + \sigma_\ell r_\ell, \quad (2)$$

with unit variance innovation term r_ℓ , and duration dependent volatility, σ_ℓ . Clark (1973) argues that price processes evolve at a rate by which new information arrives to the market and proxies this arrival rate by the volume of shares traded. Harris (1987) and Ané and Geman (2000) show that it is not so much volume that drives the price process but the arrival of orders. To test for the time scale in which the price process evolves we specify volatility as a function of duration,

$$\sigma_\ell = \sigma \tau_\ell^{\delta_1}, \quad (3)$$

¹ Menkveld et al. (2007) also suggest using an unobserved components model to deal with missing observations. However, their study is distinctly different from ours, as it studies the contribution to price discovery of the overlapping, non-overlapping and overnight periods for several Dutch-US cross-listed stocks.

where δ_1 measures the impact of durations on volatility. If $\delta_1 = \frac{1}{2}$ the variance of the efficient price is proportional to the duration between quote updates and the efficient price is said to evolve in calendar time. If $\delta_1 = 0$ the variance of the efficient price is not affected by quote durations and the efficient price is said to evolve in tick time. In this case the calendar time variance is proportional to the number of quote updates.

The constant c measures the average half-spread and captures e.g., order processing costs. To reduce the number of parameters we assume that bid and ask deviations from m_{ℓ} are symmetric on average, i.e., the elements of c are equal but of opposite sign for bid and ask of the same dealer ($c_{2i-1} = -c_{2i}$ for $i = 1, \dots, M$).

The transitory component in Eq. (1), u_{ℓ} , relates to informational asymmetries among dealers, dealer specific inventory costs and other sources of microstructure noise such as price discreteness. [Glosten and Harris \(1988\)](#) argue that the prices of informed market participants are correlated with changes in the efficient price. We argue that some dealers may also be informed, and apply their rationale to quote updates. As we only look at quotes that are updated at time t_{ℓ} , the quote series $J_{\ell} q_{\ell}$ are information events just like transactions prices. We therefore split u_{ℓ} into two components,

$$u_{\ell} = \alpha_{\ell} r_{\ell} + e_{\ell}, \quad (4)$$

where the $2M$ -vector α_{ℓ} measures the asymmetric information among dealers, and e_{ℓ} is idiosyncratic noise. As for the evolution of the efficient price, we allow this asymmetric information to depend on the time between quote innovations,

$$\alpha_{\ell} = \alpha \tau_{\ell}^{\delta_2} \sigma, \quad (5)$$

where δ_2 measures the impact of quote durations on the asymmetric information.² We expect $\delta_2 < 0$, implying that quotes are less informative at long durations. This would be consistent with [Easley and O'Hara \(1992\)](#) who hypothesize that long durations convey no information and with [Dufour and Engle \(2000\)](#) and [Engle and Patton \(2004\)](#) who find that the price impact of trades is largest at short durations.

The idiosyncratic quote noise, e_{ℓ} , is assumed to be uncorrelated with r_{ℓ} and τ_{ℓ} , and has covariance matrix Ω with block-diagonal structure,

$$\Omega = \begin{pmatrix} \Omega_1 & & \\ & \ddots & \\ & & \Omega_M \end{pmatrix}, \quad (6)$$

where each Ω_i is a (2×2) matrix. The diagonal elements in Ω_i measure the variance of quote innovations not attributable to the efficient price. A low idiosyncratic noise variance therefore indicates that a dealer tracks the efficient price closely. Bids and asks are allowed to correlate for the same dealer, but not for different dealers. The correlation between bid and ask of the same dealer is motivated by inventory cost models, which suggest that a dealer will move both quotes when an inventory shock occurs. Simultaneous changes in the quotes of different dealers can be due to asymmetric information, which is captured by α . The block-diagonal structure of Ω is motivated by the empirical observation that most quote updates are from a single dealer.³ If we do not impose this restriction only a small fraction of the data is available for estimation of the off-diagonal blocks of 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

$$\begin{aligned} J_{\ell} q_{\ell} &= J_{\ell} (c + \iota m_{\ell} + u_{\ell}), \\ m_{\ell} &= m_{\ell-1} + \sigma \tau_{\ell}^{\delta_1} r_{\ell}, \\ u_{\ell} &= \alpha \tau_{\ell}^{\delta_2} \sigma r_{\ell} + e_{\ell}. \end{aligned} \quad (7)$$

Estimation of Eq. (7) is done by Quasi Maximum Likelihood using the Kalman Filter. Because the underlying process is a random walk, we use a diffuse prior and exclude the first 50 observations in the calculation of the likelihood function. To account for the large price change overnight the model is re-initialized everyday with a diffuse prior. However, the parameters are estimated over the whole sample period.

² Although durations between quote innovations of a particular dealer would also be interesting, our primary focus is on durations between any quote update. These durations provide a proxy for the activity of the market and we can test whether asymmetric information is larger during active or inactive periods. In this respect we can capture part of the intra-daily differences between open, mid-day and close, as durations tend to be smaller near the open and close and larger in the mid-day.

³ This assumption could be too restrictive if there is a high degree of collusion among dealers who adjust their quotes simultaneously. However, this would be unlikely in the present case as implicit collusion has decreased considerably since May 1994 ([Christie et al., 1994](#)) and colluding dealers who adjust the same quote at the same second would be highly visible to the market.

Table 1
Summary statistics.

Symbol	Company name	# quote updates	% single quotes	Average duration	Std. dev. duration
AAPL	Apple Computer Inc.	29,787	89.63	15.71	28.73
AMAT	Applied Materials Inc.	105,090	83.58	4.77	5.95
AMGN	Amgen Inc.	40,279	83.97	12.32	22.24
AMZN	Amazon.com, Inc.	150,710	80.52	3.44	4.39
ATHM	At Home Corporation	76,435	86.50	6.34	9.40
CMGI	CMGI, Inc.	90,401	87.06	5.34	7.75
COMS	3Com Corporation	61,049	89.56	7.68	11.34
CPWR	Compuware Corporation	33,301	90.55	13.92	31.03
CSCO	Cisco Systems Inc.	164,480	80.85	3.13	3.52
DELL	Dell Computer Corporation	177,850	77.39	3.02	3.44
INTC	Intel Corporation	171,260	76.92	3.15	3.56
MSFT	Microsoft Corporation	151,110	80.82	3.42	3.94
NOVL	Novell Inc.	18,909	87.88	25.08	43.93
NXTL	Nextel Communications CL-A	19,556	91.63	23.42	42.23
ORCL	Oracle Corporation	87,774	85.81	5.56	7.46
PSFT	Peoplesoft Inc.	24,601	91.12	18.74	32.02
QWST	Qwest Communications Intl Inc.	44,459	88.33	10.68	18.12
SBUX	Starbucks Corporation	14,320	90.06	32.43	64.92
SUNW	Sun Microsystems Inc.	128,370	82.43	4.20	5.15
WCOM	MCI WorldCom Inc.	88,550	83.33	5.66	7.41

This table lists the twenty Nasdaq stocks included in the sample. The number of quote updates is the total number of times a dealer either changes a bid or an ask. The number of single quotes is the frequency of observations for which only one dealer issues a new quote. The last two columns provide the average and standard deviation of the durations between quote updates.

3. Price discovery

In this section we define price discovery measures based on the previously proposed model. We consider three measures to summarize the quote setting behavior of dealers. The first measure considers how dealers innovate their quotes based on the information in the efficient price. The second measure considers the contribution of quote innovations of each dealer to the evolution of the efficient price. The last measure combines the two measures and results in an information share of each dealer similar to [Hasbrouck \(1995\)](#). We first define these measures in tick time and subsequently aggregate them to calendar time equivalents.

3.1. Price discovery in tick time

Suppose that at t_{ℓ} each dealer issues bid and ask quotes and that the previous efficient price, $m_{\ell-1}$, is known to all dealers. Quote updates, v_{ℓ} , reflect the change in efficient price, $m_{\ell} - m_{\ell-1}$, and the dealer-specific noise e_{ℓ} . Substituting Eqs. (2) and (4) in Eq. (1), the innovation of the dealer quotes is equal to

$$v_{\ell} = q_{\ell} - E[q_{\ell} | m_{\ell-1}, q_{\ell-1}] = (\iota\sigma_{\ell} + \alpha_{\ell})r_{\ell} + e_{\ell} = \left(\iota + \frac{\alpha_{\ell}}{\sigma_{\ell}}\right)(m_{\ell} - m_{\ell-1}) + e_{\ell}. \quad (8)$$

From the decomposition (Eq. (8)) we obtain the first measure of the price discovery process. The duration dependent regression coefficient,

$$\beta_{\ell} = \iota + \frac{\alpha_{\ell}}{\sigma_{\ell}}, \quad (9)$$

shows how much of the change in the efficient price is immediately reflected in the quote innovation and we refer to this measure as *dealer liquidity*. The more a dealer incorporates the innovation in the random walk into her quotes, the more liquid she is. Using Eqs. (5) and (3) for the dependence of α_{ℓ} and σ_{ℓ} on duration τ_{ℓ} we find

$$\beta(\tau) = \iota + \alpha\tau^{(\delta_2 - \delta_1)}. \quad (10)$$

At long durations $\beta(\tau)$ will converge to ι when $\delta_1 > \delta_2$. This is theoretically most likely because we expect $0 < \delta_1 < \frac{1}{2}$ and $\delta_2 < 0$. When α is positive, $\beta(\tau)$ is larger than one, meaning that this quote incorporates more information of the random walk and vice versa.

The second measure, referred to as *price discovery*, considers the reverse regression,

$$\Delta m_{\ell} = \gamma'_{\ell} v_{\ell} + \eta_{\ell}, \quad (11)$$

where γ_{ℓ} measures the impact of quote innovations on the efficient price change. Contrary to the reduced form variance decomposition of Hasbrouck (1995), the change in the efficient price is not a deterministic function of the quote innovations. The structural model 1 leads to a remainder term η_{ℓ} in Eq. (11). Being regression parameters, γ_{ℓ} is defined as

$$\gamma_{\ell} = \text{Var}(v_{\ell})^{-1} \text{Cov}(v_{\ell}, \Delta m_{\ell}), \quad (12)$$

with

$$\text{Var}(v_{\ell}) = \Sigma_{\ell} = \sigma_{\ell}^2 \beta_{\ell} \beta_{\ell}' + \Omega, \text{Cov}(v_{\ell}, \Delta m_{\ell}) = \sigma_{\ell}^2 \beta_{\ell}, \quad (13)$$

where the covariance follows from Eq. (8). To solve for γ_{ℓ} , note that the structure of Σ_{ℓ} is that of a full matrix Ω plus a symmetric rank one correction $\sigma_{\ell}^2 \beta_{\ell} \beta_{\ell}'$. We can guess a solution $\gamma_{\ell} = a \Omega^{-1} \beta_{\ell}$ and solve for a using the equality $\Sigma_{\ell} \gamma_{\ell} = \sigma_{\ell}^2 \beta_{\ell}$, leading to

$$\gamma_{\ell} = \frac{\sigma_{\ell}^2}{1 + \sigma_{\ell}^2 \beta_{\ell}' \Omega^{-1} \beta_{\ell}} \Omega^{-1} \beta_{\ell}. \quad (14)$$

As with $\beta(\tau)$ we can make the functional dependence on τ explicit through

$$\gamma(\tau) = \frac{\sigma^2 \tau^{2\delta_1} \Omega^{-1} \beta(\tau)}{1 + \sigma^2 \tau^{2\delta_1} \beta(\tau)' \Omega^{-1} \beta(\tau)}. \quad (15)$$

Contrary to $\beta(\tau)$ this second measure depends strongly on the idiosyncratic noise Ω^{-1} . Assuming all other factors constant, this measure increases for a particular dealer when the corresponding elements in Ω^{-1} increase (i.e., the particular dealer has low idiosyncratic noise) and vice versa.

The last measure resembles Hasbrouck's (1995) *information share*. This measure is defined as the fraction of the variance of the efficient price change that can be attributed to innovations in a dealer's quote updates and can be computed using the R^2 of Eq. (11). From elementary linear regression analysis we can write the R^2 of Eq. (11) as

$$R_{\ell}^2 = \frac{\gamma_{\ell}' \Sigma_{\ell} \gamma_{\ell}}{\sigma_{\ell}^2} = \left(\frac{\sigma_{\ell}^2}{1 + \sigma_{\ell}^2 \beta_{\ell}' \Omega^{-1} \beta_{\ell}} \right) \beta_{\ell}' \Omega^{-1} \beta_{\ell} = \beta_{\ell}' \gamma_{\ell}. \quad (16)$$

The last equality in Eq. (16) immediately suggests how to define information shares of the individual dealers. Let $\beta_{i,\ell} = (\beta_{i,\ell}^b, \beta_{i,\ell}^a)'$ and $\gamma_{i,\ell} = (\gamma_{i,\ell}^b, \gamma_{i,\ell}^a)'$ be the (2×1) subvectors of β_{ℓ} and γ_{ℓ} that correspond to the bid and ask quotes of dealer i . Then we define the information shares following De Jong and Schotman (2008) as

$$\text{IS}_{i,\ell} = \beta_{i,\ell}' \gamma_{i,\ell} = \left(\frac{\sigma_{\ell}^2}{1 + \sigma_{\ell}^2 \beta_{\ell}' \Omega^{-1} \beta_{\ell}} \right) \beta_{i,\ell}' \Omega_i^{-1} \beta_{i,\ell}. \quad (17)$$

Due to the block-diagonal structure of Ω we obtain a unique variance decomposition. The information shares sum to R_{ℓ}^2 . Since all elements of β_{ℓ} depend on τ_{ℓ} , information shares are duration dependent.

Table 2
Duration parameters.

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

This table presents the estimates for the duration parameters δ_1 and δ_2 in Eq. (7). Standard errors are in parentheses. Stocks are referred to by the ticker symbols explained in Table 1. δ_1 measures the impact of time on the innovation in the efficient price process and has values of interest of $\delta_1 = \frac{1}{2}$, which represents a process that evolves in calendar time and $\delta_1 = 0$, which represents a process that evolves in tick time.

Table 3

Asymmetric information.

Estimate	ISLD		INCA		MM1		MM2		MM3	
	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask
Positive	2	1	11	8	13	9	7	7	2	2
Zero	9	6	7	11	1	6	6	4	11	11
Negative	9	13	2	1	6	5	7	9	7	7

This table reports a summary of the estimates for the vector α in Eq. (7). Dealers include the ECNs Island (ISLD) and Instinet (INCA) plus the three most active individual wholesale dealers. Parameters differ for bid and ask quotes. Entries report the number of estimates that are significantly positive, not significantly different from zero, or significantly negative. Estimates are insignificant if the t -statistic is between -2 and 2 .

3.2. Calendar time aggregation

To obtain price discovery measures in calendar time, we aggregate the tick time measures to fixed time intervals. In calendar time we refer to the present time with the suffix (t) . To link calendar time with tick time, let $\mathcal{I}(t)$ represent the closest observation preceding time t , so that $q(t) = q_{\mathcal{I}(t)}$ and $q_{\mathcal{I}} = q(t_{\mathcal{I}})$. Between $t-1$ and t there are $\mathcal{I}(t) - \mathcal{I}(t-1)$ quote updates.

In calendar time innovations in dealer quotes are decomposed as

$$v(t) = q(t) - E[q(t)|m(t-1), q(t-1)] = \iota \Delta m(t) + u_{\mathcal{I}(t)} \quad (18)$$

with $\Delta m(t) = m(t) - m(t-1)$. The change in the efficient price from $\mathcal{I}(t-1)$ to $\mathcal{I}(t)$ is the sum of all the interjacent changes. Only the most recent change is part of both $\Delta m(t)$ and $u_{\mathcal{I}(t)}$. We can write Eq. (18) as

$$v(t) = \iota \sum_{j=\mathcal{I}(t-1)+1}^{\mathcal{I}(t)-1} \sigma_j r_j + \sigma_{\mathcal{I}(t)} \beta_{\mathcal{I}(t)} r_{\mathcal{I}(t)} + e_{\mathcal{I}(t)}. \quad (19)$$

From Eq. (19) we can compute the calendar time price discovery measures $\beta(t)$, $\gamma(t)$ and $\Sigma(t)$. Start by defining the moments

$$\begin{aligned} E[\Delta m(t)^2] &= \sigma(t)^2 = \sum_{j=\mathcal{I}(t-1)+1}^{\mathcal{I}(t)} \sigma_j^2, \\ E[v(t)\Delta m(t)] &= \iota \tilde{\sigma}(t)^2 + \sigma_{\mathcal{I}(t)}^2 \beta_{\mathcal{I}(t)}, \\ E[v(t)v(t)'] &= \Sigma(t) = \tilde{\sigma}(t)^2 \mathbf{u}' + \sigma_{\mathcal{I}(t)}^2 \beta_{\mathcal{I}(t)} \beta_{\mathcal{I}(t)}' + \Omega, \end{aligned} \quad (20)$$

where $\tilde{\sigma}(t)^2 = \sigma(t)^2 - \sigma_{\mathcal{I}(t)}^2$. The volatility, $\sigma(t)$, increases with the length of the interval and is a function of all interjacent durations. This implies that $\Sigma(t)$ converges to a matrix of rank one proportional to $\mathbf{u}'$, which is a consequence of cointegration with a single common trend (i.e., over long enough intervals quote updates of all dealers will be perfectly correlated).

Table 4

Idiosyncratic dealer noise.

	ISLD	INCA	MM1	MM2	MM3
Average	0.36	0.41	0.43	0.43	0.40
Std. dev.	0.47	0.23	0.25	0.29	0.21
Minimum (stock)	0.01	0.09	0.17	0.10	0.10
	(CMGI, DELL)	(MSFT)	(MSFT)	(AMZN, INTC)	(MSFT)
Maximum (stock)	1.80	0.95	1.24	1.38	0.78
	(NXTL)	(AAPL)	(PSFT)	(PSFT)	(CPWR)
Highest	4	2	4	2	8
Lowest	14	0	2	1	3
<i>Correlations</i>					
ISLD		0.66	0.43	0.49	0.20
INCA			0.63	0.59	0.61
MM1				0.85	0.54
MM2					0.51
Duration	0.86	0.75	0.47	0.62	0.28

This table reports a summary of the variance of the idiosyncratic dealer noise components as estimated in Eq. (7). Dealers include the ECNs Island (ISLD) and Instinet (INCA) plus the three most active individual wholesale dealers (MM1, MM2, MM3). The “average” refers to the cross sectional average over all twenty stocks of the diagonal elements of Ω_i . “Std. dev.”, “minimum” and “maximum” are defined analogously with respect to the cross section of stocks. The stock symbol below minimum (maximum) refers to the stock at which the minimum (maximum) occurs. “Highest” is the number of cases for which that dealer has the largest idiosyncratic noise variance among the five dealers. Analogously, “lowest” counts the number of stocks for which that dealer has the lowest noise variance. The *Correlations* refer to the correlation between estimates of idiosyncratic variances over all twenty stocks. The final row reports the correlation between estimates of diagonal elements of Ω_i and the average duration of the quotes for a stock.

Dealer liquidity, $\beta(t)$, is the regression coefficient of quote innovations, $v(t)$, on the change in the efficient price, $\Delta m(t)$. From Eq. (20) we obtain

$$\beta(t) = \frac{E[v(t)\Delta m(t)]}{E[\Delta m(t)^2]} = \frac{\tilde{\sigma}(t)^2}{\sigma(t)^2} \iota + \frac{\sigma_{\epsilon(t)}^2}{\sigma(t)^2} \beta_{\epsilon(t)}, \quad (21)$$

which is a function of all durations between $t-1$ and t . Given the definition of β_{ϵ} in Eq. (9), $\beta(t)$ is linear in α . Elements of $\beta(t)$ have a lower bound of one when the corresponding element of α is positive and vice versa. Since $\sigma(t)^2$ (and $\tilde{\sigma}(t)^2$) increase with the length of the interval, $\beta(t)$ converges to the unit vector ι when intervals increase. The rate of convergence towards ι depends on the duration parameters δ_1 and δ_2 . Thus when calendar time intervals are large, information asymmetries between dealers disappear and all dealers will follow the movements in the efficient price perfectly.

For the price discovery parameters $\gamma(t)$ we again consider the reverse regression,

$$\Delta m(t) = \gamma(t)'v(t) + \eta(t), \quad (22)$$

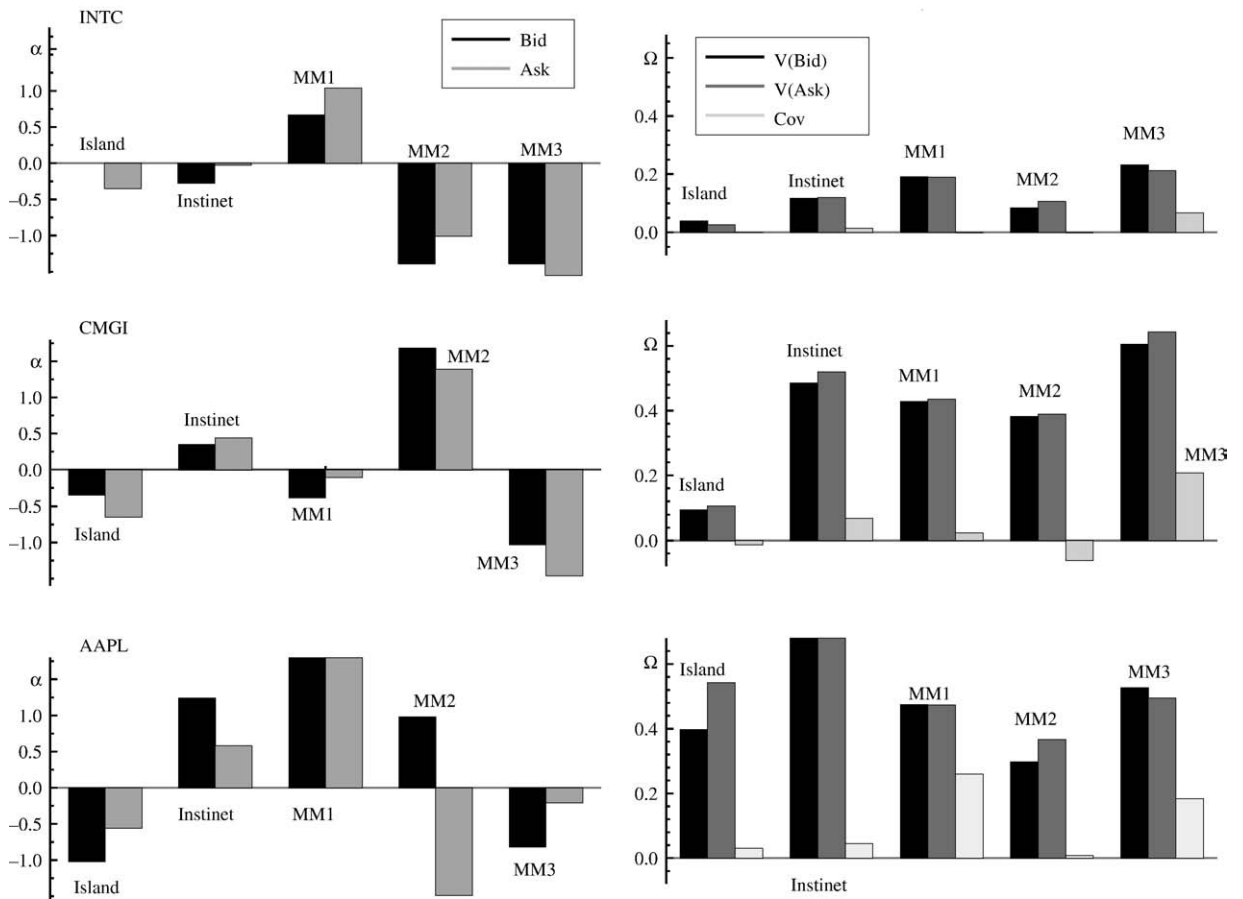
which has regression coefficients

$$\gamma(t) = \Sigma(t)^{-1} \beta(t) \sigma(t)^2. \quad (23)$$

The derivation of $\gamma(t)$ is slightly more involved than it was for the tick time equivalent γ_{ϵ} , since $\Sigma(t)$ is now a symmetric rank two correction on Ω . We guess a solution

$$\gamma(t) = c_{\iota}(t) \Omega^{-1} \iota + c_{\beta}(t) \Omega^{-1} \beta_{\epsilon(t)}, \quad (24)$$

with scalars $c_{\iota}(t)$ and $c_{\beta}(t)$ for which we provide the explicit solution in the [Appendix](#).



Note: These graphs show parameter estimates for asymmetric information (α) and idiosyncratic dealer noise (Ω) for three stocks, Intel, CMGI and Apple. The left part of each graphs shows α , the right part Ω .

Fig. 1. Parameter estimates for three selected stocks.

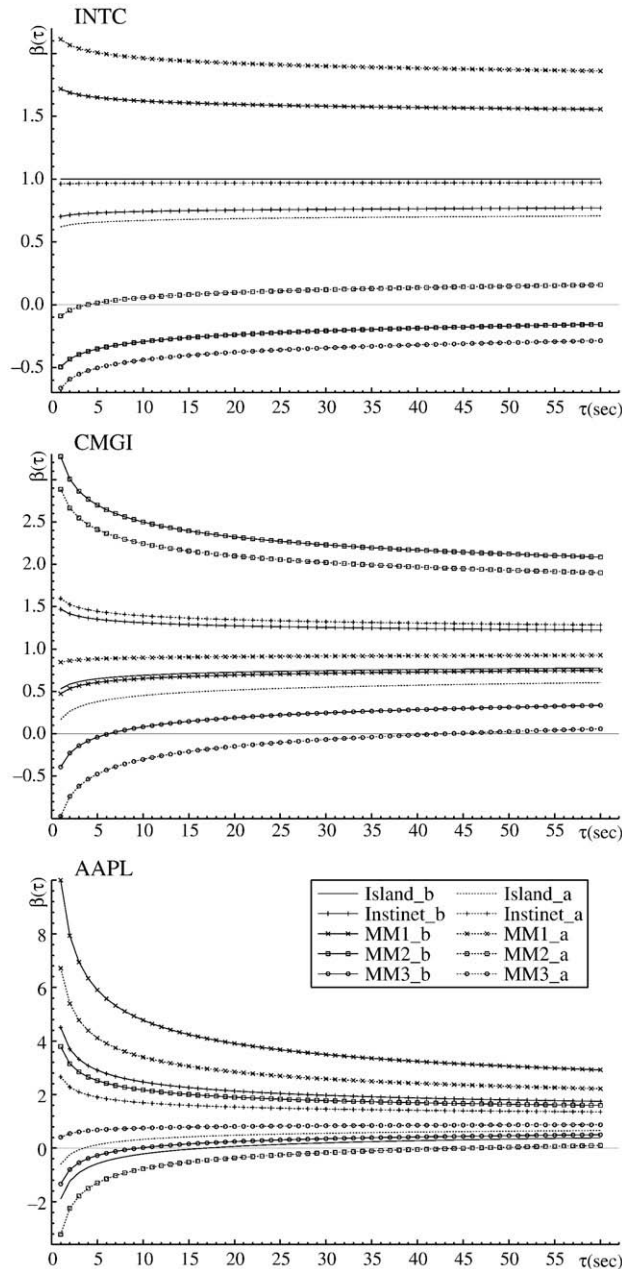
Finally, for the information shares in calendar time we consider the R^2 of Eq. (22),

$$R^2(t) = \gamma(t)' \beta(t), \quad (25)$$

and decompose it into information shares per dealer

$$IS_i(t) = \beta_i(t)' \gamma_i(t).$$

The explicit solution of $IS_i(t)$ is again in the [Appendix](#). For large time intervals the change in dealer quotes will mimic the change in the efficient price. Consequently, the R^2 in Eq. (22) will converge to one. The convergence of R^2 to one together with the convergence of $\beta(t)$ to ι implies that the elements of $\gamma(t)$ and the information shares $IS_i(t)$ will sum to one at large intervals. We therefore conclude that information shares are the only relevant measure when time intervals are large.



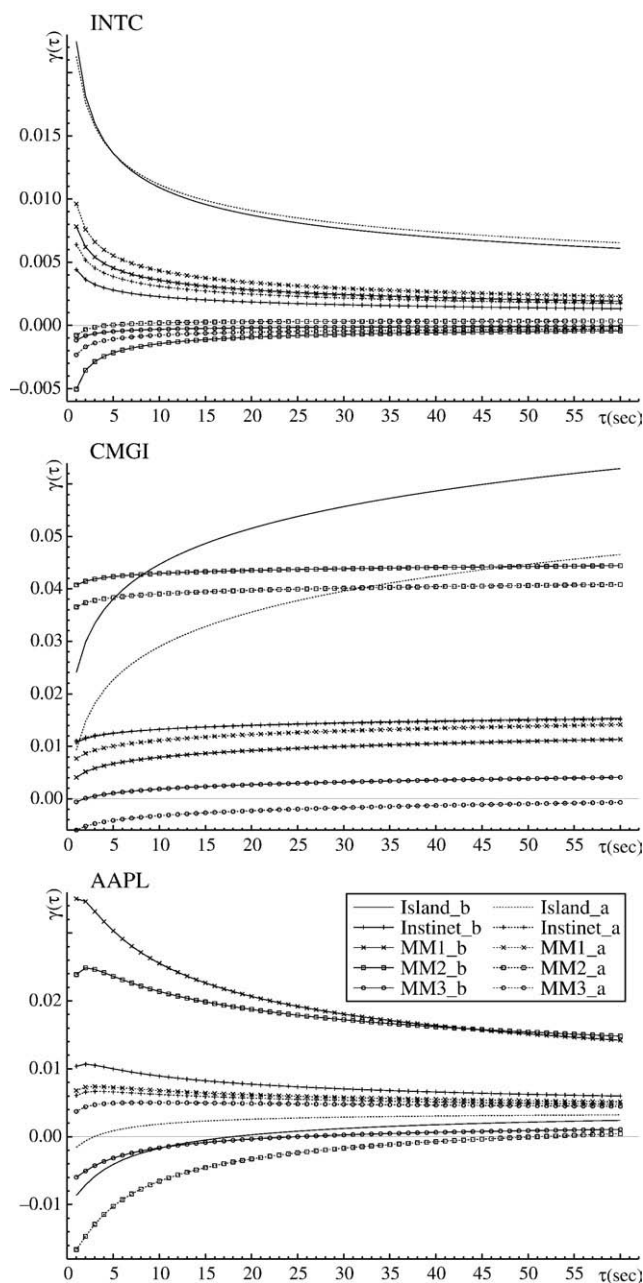
Note: These graphs show the dealer liquidity $\beta(\tau)$ defined in (10) for INTC, CMGI and AAPL as a function of duration.

Fig. 2. Tick time dealer liquidity ($\beta(\tau)$).

4. Data

We use quote data provided by Nasdaq, a database containing intraday quotes and trades at Nasdaq. The quote data contain all quotes issued within normal trading hour time stamped to the nearest second. Most important, they contain the identity of the dealer issuing the quote. We select 20 actively traded companies with varying liquidity for February 1999. Selected stocks and their symbols are reported in Table 1.

Our data are similar to those used by Huang (2002) but from a different month in 1999. Contrasting to Huang (2002), who creates categories of different dealer types, we consider individual dealer quotes as quoting behavior is heterogenous even within categories (see Schultz, 2003). In addition, our model is more suited to describe individual quoting behavior. We consider quotes of the five largest dealers in terms of quoting frequency. This leads to the selection of two ECNs, Island and Instinet, and three wholesale market makers, which differ per stock.



Note: These graphs show the price discovery parameter $\gamma(\tau)$ defined in (15) for INTC, CMGI and AAPL as a function of duration.

Fig. 3. Tick time price discovery ($\gamma(\tau)$).

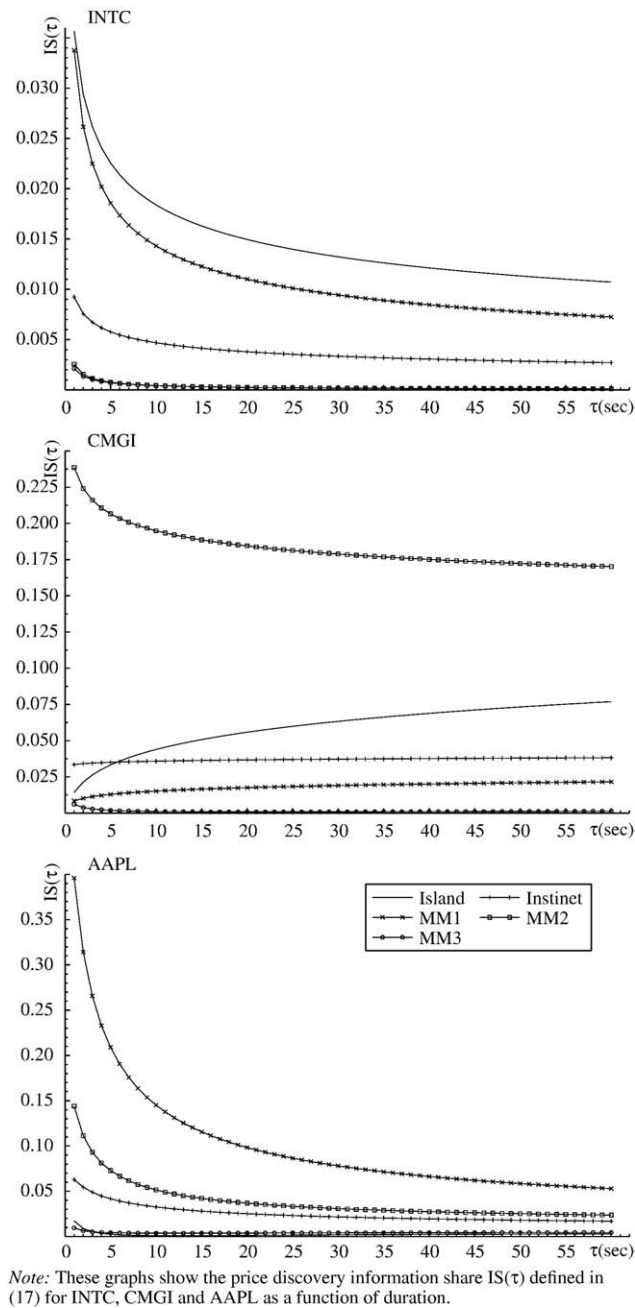


Fig. 4. Tick time information shares $IS(\tau)$.

Since we are interested in modeling innovations in dealer quotes, we remove all stale quotes (e.g., when a dealer only innovates her ask, her bid quote is removed). When a dealer issues multiple quotes at the same time, only the last quote is selected.⁴

We also remove outliers before estimation. Outliers often occur when a dealer is unwilling to trade and issues a quote far away from the inside (best quote in the market). Also, ECNs are not obliged to provide quotes on both sides of the market, and in some cases do not issue a quote on one side of the market. In this case a zero-quote is recorded. We define a quote as an outlier when it is more than \$2 away from the average of the past 50 quotes.

Table 1 provides summary statistics for the filtered data. Our sample contains very actively traded stocks like Dell and less actively traded stocks like Starbucks. The table reveals that many quote updates are by a single dealer, on average about 85%. Average quote durations reported in the fifth column are the inverse of the total number of observations. They range from about three seconds for the

⁴ This often occurs for Island, where many small trades are matched within Island. When such a trade occurs, Island sends new quotes to Nasdaq, which often only reflect a change in quote depth.

most active stocks to half a minute for the less active stocks. For full identification of α , δ_1 and δ_2 we need sufficient time series variation in durations. The standard deviations of durations reported in the last column of Table 1 are reassuring in this respect.

5. Results

In this section we discuss the results. We first provide a summary of the parameter estimates and next address the price discovery measures.

5.1. Parameter estimates

Table 2 presents the estimates of δ_1 and δ_2 . On average δ_1 is slightly negative (-0.06), and in most cases we find δ_1 significantly negative. This rejects the notion that the price discovery process evolves in calendar time and provides evidence for the process evolving in tick time as suggested by Clark (1973) and Ané and Geman (2000). A negative value for δ_1 can arise when the information flow is slower than the arrival rate of quote updates. Quote updates are more frequent than transactions and possibly not all quote updates are equally informative. Especially the large number of short-lived changes on Island could cause a small negative estimate of δ_1 (e.g. Hasbrouck and Saar (2004) show that many orders on Island are canceled after two seconds).

For δ_2 we find a significantly negative value in most cases, indicating that long durations lead to less asymmetric information. These results are in line with Easley and O'Hara (1992), who show that long durations convey little or no information.

The results for δ_1 and δ_2 imply that periods with short durations are periods with higher (calendar time) volatility and more asymmetric information. Consequently, volatile periods are periods where asymmetries are large and vice versa. Since durations tend to be shorter near the open and close of the market, these are the more volatile periods with larger asymmetries.

Table 3 reports summary statistics for the asymmetric information component, α . Although the values of α 's are not that informative on their own (as scaling depends on δ_2 and σ), they do indicate strong heterogeneity among the different dealers. If a dealer has no private information the parameter should be zero. This is often observed for the least active market maker (MM3). It is rarely the case for the most active dealer (MM1), who has the largest number of significant positive α 's. The estimates also show a contrast between the two ECNs: for Instinet estimates tend to be positive, while for Island the majority of estimates are negative.

Table 4 summarizes the estimates of the idiosyncratic dealer noise variance Ω . Average estimates are similar for Instinet and the three individual dealers, while Island differs from the others. For most of the active stocks in the sample (CSCO, DELL, MSFT and INTC) Island has the lowest idiosyncratic dealer variance. For the less active stocks other dealers have lower variances (e.g. AAPL, CPWR and SBUX).

The bottom part of Table 4 shows the correlations between the idiosyncratic noise of the various dealers. The positive correlation suggests that there is much more heterogeneity between different stocks than between different dealers. The idiosyncratic noise also correlates strongly with the average duration (last row of the table). For the less active stocks the noise component tends to be larger. This is most pronounced for Island, which is consistent with the observed spread between the maximum and minimum idiosyncratic noise in the table.

Table 5

Price discovery measures.

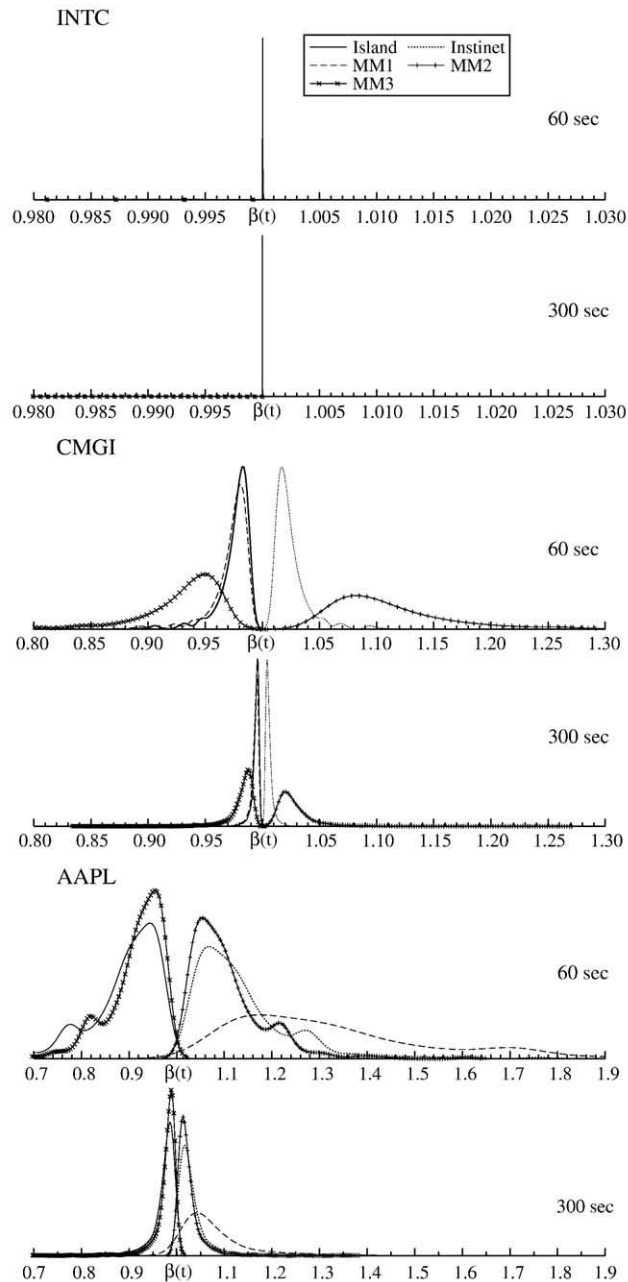
Stock	$\bar{B}$	$s(\beta)$	$\bar{\gamma}$	$s(\gamma)$	Information shares IS(τ)					
					ISLD	INCA	MM1	MM2	MM3	R ²
AAPL	1.39	1.48	0.64	0.85	0.1	2.8	11.2	4.1	0.4	18.6
AMAT	1.04	0.32	1.13	1.06	5.6	1.9	3.4	0.9	0.4	12.2
AMGN	0.22	0.80	0.07	0.74	0.9	1.4	0.2	0.7	1.9	5.1
AMZN	0.51	1.80	− 4.17	8.91	0.0	0.0	0.0	100.0	0.0	100.0
ATHM	1.00	0.20	1.14	0.74	4.3	1.7	3.9	1.1	0.7	11.7
CMGI	0.99	1.00	1.82	1.66	3.4	3.5	1.3	20.6	0.2	29.0
COMS	1.10	0.88	0.91	0.73	3.5	5.2	0.0	0.9	4.3	13.8
CPWR	1.19	0.45	1.01	0.74	0.7	1.2	6.5	5.1	0.8	14.3
CSCO	0.63	1.43	0.82	1.64	4.3	1.9	1.0	1.1	3.7	12.0
DELL	1.47	2.62	1.32	2.44	6.2	1.1	9.5	0.4	2.1	19.4
INTC	0.57	0.89	0.46	0.66	2.6	0.7	2.2	0.1	0.1	5.7
MSFT	0.08	1.09	0.36	0.91	4.0	0.7	0.2	0.1	1.9	6.4
NOVL	1.35	1.24	0.74	0.85	1.9	3.2	6.6	5.0	1.4	18.3
NXTL	0.88	1.30	0.65	1.00	0.1	0.9	9.1	4.6	2.2	16.8
ORCL	0.88	1.16	0.71	0.72	1.4	1.3	8.2	1.3	0.6	12.8
PSFT	5.35	4.78	0.86	1.81	0.0	0.0	99.9	0.0	0.0	99.9
QWST	2.83	2.07	0.99	0.60	1.4	21.7	0.7	8.8	4.1	36.7
SBUX	1.26	0.92	3.57	7.39	0.1	0.5	90.9	0.0	0.1	91.6
SUNW	0.60	0.56	0.46	0.53	2.1	1.2	0.2	0.2	0.8	4.5
WCOM	1.06	0.32	0.51	0.24	1.7	1.2	1.5	0.5	0.5	5.4

This table summarizes the price discovery measures in tick time. Stocks are referred to as detailed in Table 1. Included dealers are the ECNs Island (ISLD) and Instinet (INCA) and the three most active individual wholesale dealers. The first two columns report the average of $\beta(\tau)$ over all dealers (bid and ask) and its standard deviation. The next two columns report average and standard deviations of $\gamma(\tau)$. For each dealer we also report the information shares $\beta_i(\tau)\gamma_i(\tau)$. All measures are evaluated at the average duration of each stock (see Table 1). Parameters $\gamma(\tau)$ and information shares IS(τ) are multiplied by 100.

5.2. Price discovery per quote update

In this section we present the results for the three price discovery measures in tick time. We present detailed results for three representative stocks (INTC, CMGI and AAPL) and show summary statistics for all stocks.

We first present the parameter estimates for α and Ω in Fig. 1 for the three selected stocks, as these help with the interpretation of the price discovery measures. For each stock, an individual market maker has the most positive α . For Island, α is either insignificant or significantly negative. For all three stocks, we find both positive and negative α_i 's.

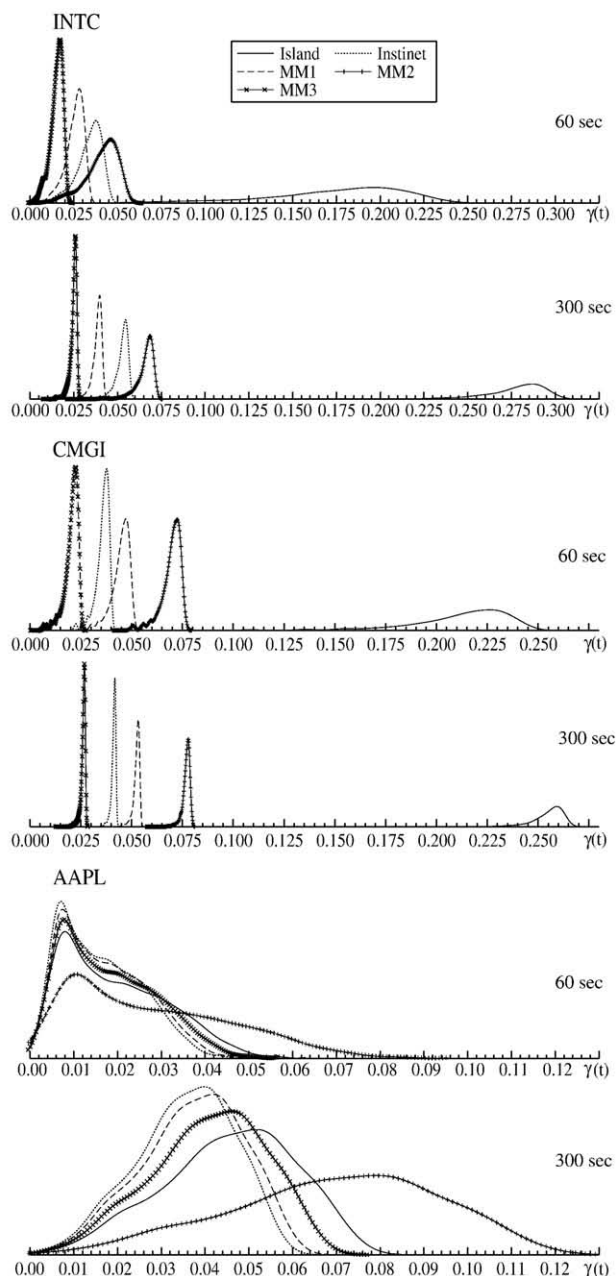


Note: These graphs show the distributions of the calendar time measures for dealer liquidity. These distributions are shown for the bids of INTC, CMGI and AAPL. The aggregates are shown at 60 and 300 second intervals.

Fig. 5. Dealer liquidity in calendar time.

Just as with α , estimates for Ω differ substantially per dealer. For INTC and CMGI, Island has the lowest idiosyncratic noise. For AAPL, MM2 has the lowest. There are also substantial differences in Ω across the stocks, with INTC having a lower idiosyncratic noise than CMGI and AAPL. In addition we find that covariances are small, hence bid and ask quotes of the same dealer are nearly uncorrelated.

In Fig. 2 we show the dealer liquidity measure, $\beta(\tau)$, as a function of quote durations. We observe the dispersion in liquidity, which is largest at short durations, and convergence to unity at longer durations. However, convergence is slow and asymmetries persist at long durations. Furthermore, we observe that the dispersion among dealers is larger for the less active stock (AAPL), suggesting that asymmetric information is more pronounced in less active stocks. A typical pattern is that for each stock at least one dealer always has a $\beta_i(\tau)$ above unity. For INTC and AAPL this is MM1; for CMGI this is MM2. This “dominance” is completely attributable to the α of the particular dealers (see Fig. 1).

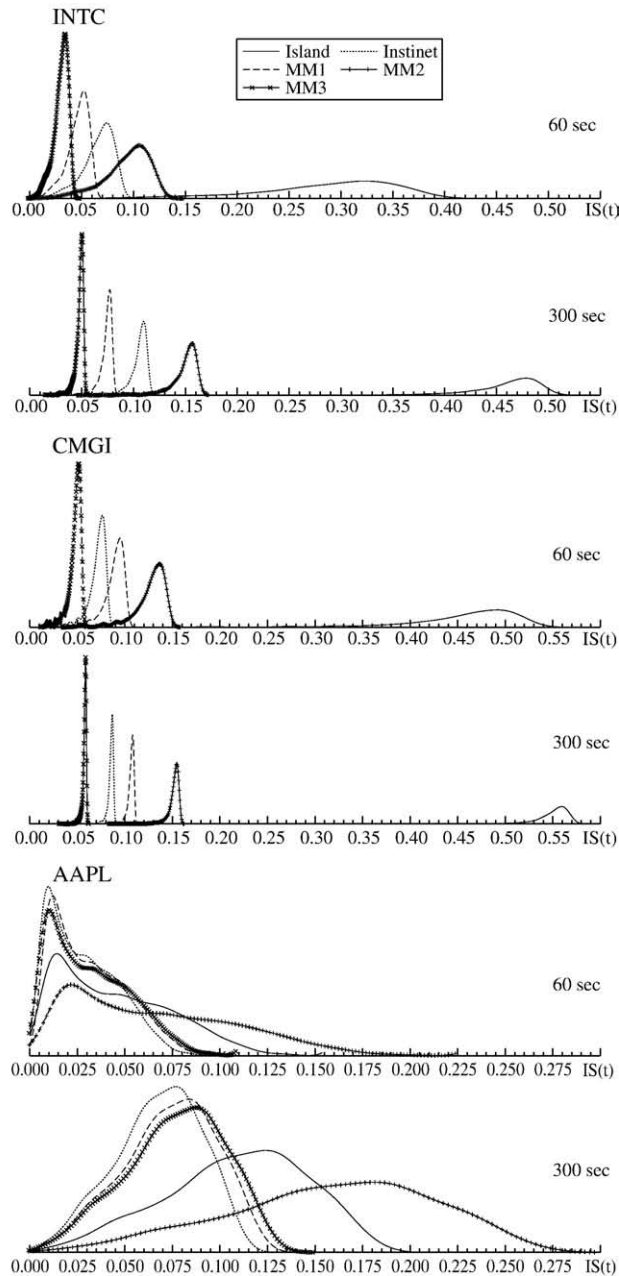


Note: These graphs show the distributions of the calendar time measures for price discovery. These distributions are shown for the asks of INTC, CMGI and AAPL. The aggregates are shown at 60 and 300 second intervals.

Fig. 6. Price discovery in calendar time.

In Fig. 3 we show the price discovery measure, $\gamma(\tau)$. For INTC we observe a clear and persistent dominance of Island at all durations; for AAPL we observe a persistent dominance for MM2. These results are mainly driven by the low idiosyncratic noise of both dealers (see Fig. 1). For CMGI the results are mixed: at short durations MM2 dominates, at long durations Island dominates. The dominance of MM2 at short durations is due to its high $\beta(\tau)$. At long durations $\beta(\tau)$ decreases for MM2 and the low idiosyncratic noise of Island dominates over this asymmetry effect.

Finally, Fig. 4 shows the information shares, $IS(\tau)$. Similar to the other measures, information shares are also highest at short durations. For INTC we observe that Island dominates and is followed by MM1, giving the same results as for $\gamma(\tau)$. Island's dominance is again due to the low idiosyncratic variance. For CMGI we find the highest information share for MM2, which can be attributed to the high $\beta(\tau)$. Island increases in importance at longer durations, due to the steep increase in $\gamma(\tau)$ at longer



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.

Fig. 7. Information shares in calendar time per dealer.

durations. For AAPL, MM1 dominates, since both $\beta(\tau)$ and $\gamma(\tau)$ are relatively high. However, the information share decreases quickly at longer durations.

The results so far provide us with three general conclusions. First, asymmetries between price discovery measures among dealers are largest at short durations, i.e., when markets are active. Second, the dominant dealer has her highest information share at short durations. As durations are shorter near the open and close of the market, more asymmetries should be observed during these periods. Third, since δ_1 is close to zero, calendar time volatility is higher in periods where durations are short. Together with the results for price discovery this indicates that price discovery is larger when volatility is high, which is in line with [Martens \(1998\)](#).

[Table 5](#) provides a summary for $\beta(\tau)$, $\gamma(\tau)$ and $IS(\tau)$ at average durations for all stocks. Similar to the large differences in α we also find large differences in $\beta(\tau)$, with extremely large values for PSFT and QWST. Information shares per quote innovation are quite low and we find that Island dominates for most of the active stocks, while other dealers dominate for less active stocks. For three stocks (AMZN, PSFT, SBUX), $IS(\tau)$ is 100% for a single market maker. For AMZN this concentration is due to the very low idiosyncratic variance for MM2. For SBUX and PSFT this is due to a large negative covariance between the idiosyncratic noise of the ask and bid. This negative covariance is unusual, and for most stocks and dealers we find small positive correlations, consistent with theories of inventory models. The R^2 (the sum of the information shares) is small in tick time. Excluding the three stocks with a single dominant dealer, on average 14.3% of the variance of the efficient price can be explained by the quotes of the top five dealers.

Table 6

Averages for information shares per dealer.

Stock	ISLD	INCA	MM1	MM2	MM3	R^2
<i>A: 60 s intervals</i>						
AAPL	0.04	0.03	0.03	0.07	0.03	0.20
AMAT	0.35	0.07	0.10	0.06	0.04	0.62
AMGN	0.13	0.04	0.11	0.06	0.07	0.41
AMZN	0.00	0.00	0.00	1.00	0.00	1.00
ATHM	0.26	0.07	0.11	0.08	0.05	0.57
CMGI	0.46	0.07	0.09	0.13	0.05	0.80
COMS	0.17	0.05	0.05	0.05	0.06	0.38
CPWR	0.05	0.06	0.13	0.07	0.04	0.35
CSCO	0.61	0.04	0.03	0.02	0.05	0.75
DELL	0.74	0.03	0.01	0.01	0.03	0.82
INTC	0.29	0.07	0.05	0.09	0.03	0.53
MSFT	0.28	0.05	0.03	0.05	0.04	0.45
NOVL	0.01	0.02	0.01	0.04	0.03	0.12
NXTL	0.01	0.04	0.04	0.03	0.06	0.18
ORCL	0.23	0.07	0.05	0.11	0.08	0.54
PSFT	0.00	0.00	0.98	0.00	0.00	0.98
QWST	0.03	0.03	0.04	0.03	0.09	0.22
SBUX	0.00	0.01	0.74	0.01	0.01	0.77
SUNW	0.23	0.08	0.06	0.07	0.05	0.49
WCOM	0.12	0.07	0.04	0.07	0.03	0.33
<i>B: 300 s intervals</i>						
AAPL	0.11	0.07	0.07	0.16	0.08	0.49
AMAT	0.50	0.10	0.14	0.09	0.06	0.89
AMGN	0.22	0.08	0.20	0.11	0.12	0.73
AMZN	0.00	0.00	0.00	1.00	0.00	1.00
ATHM	0.39	0.11	0.16	0.12	0.07	0.85
CMGI	0.55	0.09	0.11	0.15	0.06	0.96
COMS	0.32	0.10	0.09	0.09	0.11	0.71
CPWR	0.10	0.10	0.24	0.12	0.07	0.63
CSCO	0.77	0.05	0.03	0.02	0.06	0.93
DELL	0.87	0.03	0.01	0.01	0.03	0.95
INTC	0.46	0.11	0.07	0.15	0.05	0.84
MSFT	0.49	0.09	0.05	0.09	0.08	0.80
NOVL	0.03	0.04	0.03	0.09	0.06	0.25
NXTL	0.03	0.11	0.09	0.08	0.14	0.45
ORCL	0.36	0.10	0.08	0.17	0.12	0.83
PSFT	0.00	0.00	1.00	0.00	0.00	1.00
QWST	0.07	0.07	0.09	0.06	0.21	0.50
SBUX	0.01	0.01	0.88	0.01	0.01	0.92
SUNW	0.39	0.14	0.09	0.12	0.08	0.82
WCOM	0.25	0.14	0.08	0.14	0.06	0.67

This table presents the average results for information shares. This measure is obtained as the inner product of $\beta(t)$ and $\gamma(t)$. The first five columns present the information shares per dealer $IS_i(t)$, the last column reports the sum of the information shares (R^2). Stocks are referred to by the ticker symbols explained in [Table 1](#). Panels A and B have different levels of time aggregation.

5.3. Calendar time aggregation

To discuss price discovery measures in calendar time, we aggregate to 60 and 300 s intervals, as these are frequencies often used in other studies. For both intervals we compute $\beta(t)$, $\gamma(t)$ and $IS(t)$ given the durations τ_j ($j = \ell(t-1) + 1, \dots, \ell(t)$) observed in period t . This provides us with a time series for each of the three measures.

In Fig. 5 we show the distributions of $\beta(t)$ at the 60 and 300 s aggregation intervals (only bids are shown, asks are similar). Depending on the sign of α these measures are above or below one for all t . The plots clearly show $\beta(t)$'s convergence to unity when the time interval increases. The speed of convergence differs per stocks and generally depends on the stock's activity. For INTC, dealers incorporate most of the information in the efficient price into their quotes within the 60-second interval. For CMGI we observe that there is still some dispersion at the 60-second interval, but most dealers have converged to unity at the five minute level. For AAPL, measures for liquidity have not converged even at the five-minute level.

Results for $\gamma(t)$ are shown in Fig. 6 (only asks are shown, bids are similar). For all stocks, we observe that $\gamma(t)$ increases with the aggregation interval, and that the distribution of $\gamma(t)$ narrows at the longer interval. Island clearly dominates for both INTC and CMGI. For AAPL results are less clear as distributions are very wide and have long right tails. MM2, however, tends to have the highest values for $\gamma(t)$. These findings are in line with the measures in tick time, where Island dominated for both INTC and CMGI, and MM2 dominated for AAPL.

In Fig. 7 we show the distributions of $IS(t)$ per dealer. These graphs are similar to $\gamma(t)$ because $IS_i(t)$ is the inner product of $\beta_i(t)$ and $\gamma_i(t)$, and as $\beta_i(t)$ converges to unity at long intervals. $IS_i(t)$ is therefore approximately equal to the sum of $\gamma_i(t)$ for bid and ask. For CMGI and AAPL calendar time results for the information shares differ substantially from the tick time equivalents. For CMGI, $IS(\tau)$ was highest for MM2 and for AAPL, $IS(\tau)$ was highest for MM1, which was due to the large values for $\beta(\tau)$. $IS(t)$ is highest for Island and MM2 for CMGI and AAPL, respectively. This difference is attributable to the convergence of $\beta_i(t)$.

Table 6 reports average information shares at 60 and 300 s aggregation intervals for all stocks and dealers. Generally, information shares increase with the aggregation interval. For twelve stocks (generally the more active ones) Island dominates, for the other stocks a market maker dominates.

The last column of Table 6 reports the total R^2 for each stock. The R^2 increases considerably when the aggregation interval increases. However, the total amount of variance attributable to all dealers per stock differs. In general, R^2 is higher for more active stocks (e.g., CSCO and DELL) and lower for less active stocks (e.g., NOVL). These results may have implications for e.g. realized volatility when selecting the optimal sampling frequency and where the remaining microstructure noise is an important feature.

6. Concluding remarks

This paper introduces a model for dealer quoting behavior in tick time, which incorporates the effect of quote durations on the volatility and asymmetric information among dealers. Given the model we define three price discovery measures, relating to dealer liquidity, price discovery and information shares both in tick time and in calendar time. By constructing the price discovery measures from tick time data we can avoid many of the simultaneity issues related to the vector autoregressive models.

We find that volatility is higher and that asymmetric information is larger when quote durations are short. Combining these two findings leads to the conclusion that there is more asymmetric information when volatility is high. In tick time, the price discovery measures describe different aspects of the price discovery process. The price discovery measure is mainly driven by a dealer's idiosyncratic noise component, while the information share is driven by a combination of idiosyncratic noise and information asymmetry. In calendar time, informational asymmetries disappear and the information share and the price discovery measure converge. Finally, the sum of the information shares can be used as an indication of how much microstructure noise remains in aggregated dealer quotes and is informative about the speed of price discovery. This information can be used for studies on e.g. realized volatility.

Appendix A. Price discovery in calendar time

This appendix provides the derivation of $\gamma(t)$ and $IS_i(t)$ in calendar time. We start by repeating the definition of $\gamma(t)$ in Eq. (23)

$$\gamma(t) = \Sigma(t)^{-1} \beta(t) \sigma(t)^2 \quad (A1)$$

and the earlier derived moments

$$\begin{aligned} \sigma(t)^2 &= \sum_{j=\ell(t-1)+1}^{\ell(t)} \sigma_j^2 \\ \Sigma(t) &= \tilde{\sigma}(t)^2 \mathbf{u}' + \sigma_{\ell(t)}^2 \beta_{\ell(t)} \beta_{\ell(t)}' + \Omega \\ \tilde{\sigma}(t)^2 &= \sigma(t)^2 - \sigma_{\ell(t)}^2, \\ \beta(t) &= \frac{\tilde{\sigma}(t)^2}{\sigma(t)^2} \mathbf{1} + \frac{\sigma_{\ell(t)}^2}{\sigma(t)^2} \beta_{\ell(t)}, \end{aligned} \quad (A2)$$

To simplify notation we will use the shorthand ℓ for $\ell(t)$, when there can be no confusion that we mean the last quote update of period t , from time $t-1$ to t . As another useful shorthand notation we define the (2×2) matrix

$$V = \begin{pmatrix} V_u & V_{u\beta} \\ V_{u\beta} & V_{\beta\beta} \end{pmatrix} = \begin{pmatrix} \nu' \Omega^{-1} \iota & \nu' \Omega^{-1} \beta_{\ell} \\ \nu' \Omega^{-1} \beta_{\ell} & \beta_{\ell}' \Omega^{-1} \beta_{\ell} \end{pmatrix} \quad (\text{A3})$$

Except V_u the elements of V depend on the duration $\tau_{\ell}(t)$.

To solve for $\gamma(t)$ we guess the solution

$$\gamma(t) = c_{\ell} \Omega^{-1} \iota + c_{\beta} \Omega^{-1} \beta_{\ell} \quad (\text{A4})$$

and substitute this in

$$\Sigma(t) \gamma(t) = \sigma(t)^2 \beta(t) \quad (\text{A5})$$

using the definitions in Eq. (A2) to obtain

$$\begin{aligned} (\tilde{\sigma}(t)^2 u' + \sigma_{\ell}^2 \beta_{\ell}' \beta_{\ell}' + \Omega) (c_{\ell} \Omega^{-1} \iota + c_{\beta} \Omega^{-1} \beta_{\ell}) &= \sigma(t)^2 \left(\frac{\tilde{\sigma}(t)^2}{\sigma(t)^2} \iota + \frac{\sigma_{\ell}^2}{\sigma(t)^2} \beta_{\ell} \right) \\ \Downarrow \end{aligned} \quad (\text{A6})$$

$$c_{\ell} (\tilde{\sigma}(t)^2 V_u + 1) \iota + c_{\ell} \sigma_{\ell}^2 V_{u\beta} \beta_{\ell} + c_{\beta} \tilde{\sigma}(t)^2 V_{u\beta} \iota + c_{\beta} (\sigma_{\ell}^2 V_{\beta\beta} + 1) \beta_{\ell} = \sigma(t)^2 \iota + \sigma_{\ell}^2 \beta_{\ell}$$

Equating coefficients on ι and β_{ℓ} to determine the constants c_{ℓ} and c_{β} gives a system of 2 linear equations in two unknowns

$$\begin{pmatrix} \tilde{\sigma}(t)^2 V_u + 1 & \tilde{\sigma}(t)^2 V_{u\beta} \\ \sigma_{\ell}^2 V_{u\beta} & \sigma_{\ell}^2 V_{\beta\beta} + 1 \end{pmatrix} \begin{pmatrix} c_{\ell} \\ c_{\beta} \end{pmatrix} = \begin{pmatrix} \tilde{\sigma}(t)^2 \\ \sigma_{\ell}^2 \end{pmatrix} \quad (\text{A7})$$

with solution

$$\begin{pmatrix} c_{\ell} \\ c_{\beta} \end{pmatrix} = \frac{1}{D} \begin{pmatrix} V_{\beta\beta} - V_{u\beta} + 1 / \sigma_{\ell}^2 \\ V_u - V_{u\beta} + 1 / \tilde{\sigma}(t)^2 \end{pmatrix} \quad (\text{A8})$$

in which

$$D = (V_u + 1 / \tilde{\sigma}(t)^2) (V_{\beta\beta} + 1 / \sigma_{\ell}^2) - V_{u\beta}^2 \quad (\text{A9})$$

Of particular interest are the expressions for $\gamma(t)$ for small and large time intervals.

First, if the interval is reduced to a single quote update, Eq. (A8) reduces to the tick time solution (Eq. (14)) with $c_{\ell} = 0$ and $c_{\beta} = 1 / (V_{\beta\beta} + 1 / \sigma_{\ell}^2)$,

$$\lim_{\tilde{\sigma}(t)^2 \rightarrow 0} \gamma(t) = \gamma_{\ell} = \frac{\sigma_{\ell}^2}{1 + \sigma_{\ell}^2 V_{\beta\beta}} \beta_{\ell} \quad (\text{A10})$$

Second, for a very large interval, we take the limit as $\tilde{\sigma}(t)^2 \rightarrow \infty$,

$$\bar{\gamma} = \lim_{\tilde{\sigma}(t)^2 \rightarrow \infty} \gamma(t) = \frac{(V_{\beta\beta} - V_{u\beta} + 1 / \sigma_{\ell}^2) \Omega^{-1} \iota + (V_u - V_{u\beta}) \Omega^{-1} \beta_{\ell}}{V_u (V_{\beta\beta} + 1 / \sigma_{\ell}^2) - V_{u\beta}^2} \quad (\text{A11})$$

Premultiplying $\bar{\gamma}$ by ν' and using the definitions of $V_{u\beta}$ and $V_{\beta\beta}$ establishes that the elements of $\bar{\gamma}$ sum to one. To gain more insight in $\bar{\gamma}$ we go back to the original parameterization with α_{ℓ} . Recalling Eq. (9),

$$\beta_{\ell} = \iota + \frac{\alpha_{\ell}}{\sigma_{\ell}}, \quad (\text{A12})$$

we have

$$\bar{\gamma} = \frac{(V_{\beta\beta} - 2V_{u\beta} + V_u + 1 / \sigma_{\ell}^2) \Omega^{-1} \iota + (V_u - V_{u\beta}) \Omega^{-1} \alpha_{\ell} / \sigma_{\ell}}{V_u (V_{\beta\beta} + 1 / \sigma_{\ell}^2) - V_{u\beta}^2} \quad (\text{A13})$$

Also expressing $V_{u\beta}$ and $V_{\beta\beta}$ in terms of α_{ℓ} gives

$$\begin{aligned} V_u - V_{u\beta} &= \iota' \Omega^{-1} \iota - \iota' \Omega^{-1} \beta_{\ell} = -\iota' \Omega^{-1} \alpha_{\ell} / \sigma_{\ell} = -V_{u\alpha} / \sigma_{\ell} \\ V_{\beta\beta} - 2V_{u\beta} + V_u &= \beta'_{\ell} \Omega^{-1} \beta_{\ell} - 2\iota' \Omega^{-1} \beta_{\ell} + \iota' \Omega^{-1} \iota = \alpha'_{\ell} \Omega^{-1} \alpha_{\ell} / \sigma_{\ell} = V_{\alpha\alpha} / \sigma_{\ell}^2 \\ V_u V_{\beta\beta} - V_{u\beta}^2 &= V_u (\iota + \alpha_{\ell} / \sigma_{\ell})' \Omega^{-1} (\iota + \alpha_{\ell} / \sigma_{\ell}) - (V_u + V_{u\alpha} / \sigma_{\ell})^2 = \frac{1}{\sigma_{\ell}^2} (V_u V_{\alpha\alpha} - V_{u\alpha}^2) \end{aligned} \quad (\text{A14})$$

Substituting all coefficients of Eqs. (A14) in (A13) we finally obtain

$$\bar{\gamma} = \frac{(1 + V_{\alpha\alpha}) \Omega^{-1} \iota - V_{u\alpha} \Omega^{-1} \alpha_{\ell}}{V_u + V_u V_{\alpha\alpha} - V_{u\alpha}^2} \quad (\text{A15})$$

The limiting value $\bar{\gamma}$ does not depend on σ or δ_1 , meaning that the price discovery parameters are independent of the process of the efficient price. Since α_{ℓ} depends on τ_{ℓ} and δ_2 , the limiting information shares still depend on the most recent duration.

For the information shares we need the (2×1) sub-vectors $\beta_i(t)$ and $\gamma_i(t)$ corresponding to the bid and ask quotes of dealer i . Combining Eqs. (A2), (A4), (A8) and (A9) we get

$$\begin{aligned} \text{IS}_i(t) &= \gamma_i(t)' \beta_i(t) \\ &= (c_i \Omega^{-1} \iota_2 + c_{\beta} \Omega^{-1} \beta_{i,\ell}) \left(\frac{\tilde{\sigma}(t)^2}{\sigma(t)^2} \iota_2 + \frac{\sigma_{\ell}^2}{\sigma(t)^2} \beta_{i,\ell} \right) \\ &= d_u (\iota_2' \Omega_i^{-1} \iota_2) + d_{u\beta} (\iota_2' \Omega_i^{-1} \beta_{i,\ell}) + d_{\beta\beta} (\beta_{i,\ell}' \Omega_i^{-1} \beta_{i,\ell}), \end{aligned} \quad (\text{A16})$$

with

$$\begin{aligned} d_u &= c_i \frac{\tilde{\sigma}(t)^2}{\sigma(t)^2} \\ d_{u\beta} &= c_i \frac{\sigma_{\ell}^2}{\sigma(t)^2} + c_{\beta} \frac{\tilde{\sigma}(t)^2}{\sigma(t)^2} \\ d_{\beta\beta} &= c_{\beta} \frac{\sigma_{\ell}^2}{\sigma(t)^2} \end{aligned} \quad (\text{A17})$$

References

- Ané, Thierry, Geman, Hélyette, 2000. Order flow, transaction clock, and normality of asset returns. *Journal of Finance* 55, 2259–2284.
- Baillie, Richard T., Booth, Geoffrey, Tse, Yiuman, Zobotina, Tatyana, 2002. Price discovery and common factor models. *Journal of Financial Markets* 5, 309–321.
- Christie, William G., Harris, Jeffrey H., Schultz, Paul H., 1994. Why did Nasdaq market makers stop avoiding odd-eighth quotes? *Journal of Finance* 49, 1841–1860.
- Clark, P.K., 1973. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* 41, 135–156.
- De Jong, Frank, Schotman, Peter C., 2008. Price discovery in fragmented markets, Maastricht University. Working Paper.
- Dufour, Alfonso, Engle, Robert E., 2000. Time and the price impact of a trade. *Journal of Finance* 55, 2467–2498.
- Easley, David, O'Hara, Maureen, 1992. Time and the process of security price adjustment. *Journal of Finance* 47, 905–927.
- Engle, Robert, 2000. The econometrics of ultra-high frequency data. *Econometrica* 68, 1–22.
- Engle, Robert F., Patton, Andrew J., 2004. Impacts of trades in an error-correction model of quote prices. *Journal of Financial Markets* 7, 1–25.
- Furfine, Craig, 2007. When is inter-transaction time informative? *Journal of Empirical Finance* 14, 310–332.
- Glosten, L.R., Harris, L.E., 1988. Estimating the components of the bid/ask spread. *Journal of Financial Economics* 21, 123–142.
- Gammig, J., Melvin, M., Schlag, C., 2005. Internationally cross-listed stock prices during overlapping trading hours: price discovery and exchange rate effects. *Journal of Empirical Finance* 12, 139–164.
- Harris, Lawrence, 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, Joel, 1993. Assessing the quality of a security market: a new approach to transaction-cost measurement. *Review of Financial Studies* 6, 191–212.
- Hasbrouck, Joel, 1995. One security, many markets: determining the contribution to price discovery. *Journal of Finance* 50, 1175–1199.
- Hasbrouck, Joel, 2003. Intraday price formation in U.S. equity index markets. *Journal of Finance* 58, 2375–2399.
- Hasbrouck, Joel, Saar, Gideon, 2004. Limit orders and volatility in hybrid markets: the Island ECN. Working Paper. In New York University.
- Huang, Roger D., 2002. The quality of ECN and Nasdaq market maker quotes. *Journal of Finance* 57, 1285–1319.
- Martens, Martin, 1998. Price discovery in high and low volatility periods: open outcry versus electronic trading. *Journal of International Financial Markets, Institutions and Money* 8, 243–260.
- Menkveld, Albert J., Koopman, Siem Jan, Lucas, Andre, 2007. Modeling around-the-clock price discovery for cross-listed stocks using state space models. *Journal of Business and Economic Statistics* 25, 213–225.
- Schultz, Paul H., 2003. Who makes markets. *Journal of Financial Markets* 6, 49–72.