Journal of Empirical Finance 16 (2009) 759–776

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

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.

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).

0927-5398/$ – see front matter
doi:10.1016/j.jempfin.2009.07.002

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 ($\ell=1,\ldots,L$). 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,

$$J_\ell q_\ell = J_\ell(c+\iota m_\ell + u_\ell), \tag{1}$$

where m_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 m_ℓ, respectively.

The efficient price is assumed to follow a random walk,

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

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,

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

[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.