Journal of Financial Economics ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Journal of Financial Economics

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

High-frequency quoting, trading, and the efficiency of prices ☆

Jennifer Conrad [a], Sunil Wahal [b,*], Jin Xiang [c]

[a] Kenan Flagler Business School, University of North Carolina at Chapel Hill, USA
[b] WP Carey School of Business, Arizona State University, Tempe, AZ 85287, USA
[c] Discover Financial Services, 2402 W. Beardsley Road, Phoenix, AZ, 85027, USA

ARTICLE INFO

Article history:
Received 14 June 2014
Received in revised form
17 October 2014
Accepted 26 October 2014

JEL classification:
G12
G14

Keywords:
High frequency trading
Market microstructure
Market efficiency

ABSTRACT

We examine the relation between high frequency quotation and the behavior of stock prices between 2009 and 2011 for the full cross section of securities in the US. On average, higher quotation activity is associated with price series that more closely resemble a random walk, and significantly lower cost of trading. We also explore market resiliency during periods of exceptionally high low-latency trading: large liquidity drawdowns in which, within the same millisecond, trading algorithms systematically sweep large volume across multiple trading venues. Although such large drawdowns incur trading costs, they do not appear to degrade the price formation process or increase the subsequent cost of trading. In an out-of-sample analysis, we investigate an exogenous technological change to the trading environment on the Tokyo Stock Exchange that dramatically reduces latency and allows co-location of servers. This shock also results in prices more closely resembling a random walk and a sharp decline in the cost of trading.

1. Introduction

The effect of high-frequency trading on market quality is important and has generated strong interest among academics, practitioners, and regulators. Models of the effect of high-frequency trading on markets generate different predictions, depending on their assumptions and their focus. For example, Budish, Cramton, and Shim (2013) build a model in which the ability to continuously update order books generates technical arbitrage opportunities and a wasteful arms race in which fundamental investors bear costs through larger spreads and thinner markets. Similarly, Han, Khapko, and Kyle (2014) argue that because fast market makers can cancel quotes faster than slow traders, this causes a winner's curse resulting in higher spreads. In contrast, in Aït-Sahalia and Saglam (2013), lower latency generates higher profits and higher liquidity provision. In their model, however, high-frequency liquidity provision declines when market volatility increases, which can lead to episodes of market fragility. In Baruch and Glosten (2013), frequent order cancellations are a standard part of liquidity provision and are generated by limit order traders mitigating the risk that their quotes will be undercut (through rapid submissions and cancellations).

The complementary empirical literature has two strands. The first can be broadly characterized as examining the behavior of high-frequency traders (HFTs) and estimating

☆ We thank Alyssa Kerr and Phillip Howard for research assistance and the Wharton Research Data Service for providing a TAQ-CRSP (Trade and Quote and Center for Research in Security Prices) matching algorithm. We thank Gaelle Le Fol, Terry Hendershott, Mark Seasholes and participants at the Fifth Hedge Fund Research Conference (Paris, France), the Instinet Global Quantitative Equity Conference, the Financial Markets Research Center Conference (Vanderbilt University), the FTSE World Investment Forum, and the Stern Microstructure Conference. We thank the Tokyo Stock Exchange for providing proprietary data around the introduction of arrowhead. Sunil Wahal is a consultant to Dimensional Fund Advisors (DFA). DFA and Discover Financial Services provided no funding or data for this research.

* Corresponding author. Tel.: +1 480 965 8755.
E-mail addresses: j_conrad@kenan-flagler.unc.edu (J. Conrad), Sunil.Wahal@asu.edu (S. Wahal).

http://dx.doi.org/10.1016/j.jfineco.2015.02.008
0304-405X/

Please cite this article as: Conrad, J., et al., High-frequency quoting, trading, and the efficiency of prices. Journal of Financial Economics (2015), http://dx.doi.org/10.1016/j.jfineco.2015.02.008

their effect on markets. Researchers use datasets that explicitly identify high-frequency traders (by some definition), explore the trading strategies they use, test whether those strategies are profitable, and consider whether they impede or improve price discovery. For example, Brogaard, Hendershott and Riordan (2014), Carrion (2013), and Hirshey (2013) all use Nasdaq-identified high-frequency traders (the so-called Nasdaq dataset) and collectively find that HFTs are modestly profitable in aggregate, that they both demand and supply liquidity, and that they appear to impose some adverse selection costs on other traders. Similarly, Hagströmer and Norden (2013) use data on 30 stocks from Nasdaq-OMX Stockholm and find evidence indicating that market-making HFTs reduce short-term volatility. The advantage of this first group of studies is that identification of high-frequency traders is relatively clear-cut. The disadvantage is that only activity on that identifying exchange can be precisely measured. In a fragmented market such as the US, with substantial variation in access (make-take) fees across exchanges and dark pools, HFT behavior in one market (e.g., Nasdaq) might not reflect aggregate market behavior or inform overall market prices. The latter is a primary concern. For example, suppose one observes trades from a high-frequency trader from Exchange X that is known to be cheaper for extracting liquidity for a particular group of stocks. Such a high-frequency trader could be providing liquidity in Exchange Y, but a researcher observing only trades on Exchange X would erroneously draw the conclusion that this high-frequency trader is a liquidity extractor. There are a variety of reasons that there could be a nonrandom distribution of trades across trading venues, ranging from concerns about adverse selection to systematic differences in make-take fees.

The second strand of the empirical literature looks at outcomes in a conditional setting: identifying changes in market structure that facilitate high-frequency activity and examining the consequences. The most important of these papers is Hendershott, Jones, and Menkveld (2011), which examines 1,082 stocks between December 2002 and July 2003. Using the start of autoquotes on the NYSE as an exogenous instrument, they find that algorithmic trading improves liquidity. More recent studies that examine changes in the trading environment in smaller markets draw similar conclusions. For example, Menkveld (2012) examines the effect of the introduction of an electronic exchange (including a large HFT) on trading in a sample of 32 Dutch stocks, and Riordan and Storkenmaier (2012) examine the effect of an upgrade of trading systems in 98 stocks on the Deutsche Borse.

These papers provide evidence on the effects of high frequency trading. Very little evidence exists on a critical aspect of current market structure: high-frequency quoting, that is, the speed of the market environment. The relative scarcity of this evidence is surprising because many theoretical papers describe the speed of changes to the supply curve, which is more closely related to quotations than to trades. And, regulators certainly care about high frequency quotations. Although the 2010 Securities and Exchange Commission Concept Release on Equity Market Structure highlights HFTs as “professional traders acting in a proprietary capacity that engage in strategies that generate a large number of trades on a daily basis,” it also recognizes the importance of high-frequency quoting in that it could represent “phantom liquidity [that] disappears when most needed by long-term investors.”[1] That is, high frequency quoting generates execution risk, which has welfare consequences and is an important characteristic of market structure. For example, Hasbrouck (2013) notes that the execution risk caused by high-speed changes in quotes might not be diversifiable, with slower traders always losing to faster traders. Biais, Foucault, and Moinas (2014) investigate policy approaches, including a Pigovian tax, which could mitigate externalities due to differences in traders' ability to process the amount of information generated by the market, such as the volume of high-frequency quotations. Stiglitz (2014) expresses skepticism that high-frequency quotation or trading is welfare improving and makes a case for slower markets.

Our purpose is to provide large sample evidence on the influence of high frequency quoting on market quality. We do not look at the trading strategies of identified high-frequency traders; instead, we examine market outcomes. We conduct two types of tests: (1) unconditional tests designed to provide evidence for a comprehensive cross section of securities over a long and relatively recent time series, and (2) conditional tests that measure the effect of high-frequency traders during different types of market conditions and over changes in market structure. The latter examine both average and stressed market environments and, separately, a change in trading protocols. Each of the tests either provides new evidence on the influence of high-frequency quotations on markets or fills a gap in the understanding of high-speed markets, or both.

Our sample comes from the two largest equity markets in the world: the full cross section of securities in the US, and the largest three hundred stocks on the Tokyo Stock Exchange (TSE). The sample period is 2009–2011 for the US and 2010, 2011 for Japan. The breadth of the cross section allows for general conclusions, and recent data are important because significant changes have been made in market structure in the past few decades. It is also critical that the time series be long enough to generate statistical power, particularly because cross-sectional independence is likely to be low. Finally, the Japanese data allow for an out-of-sample test in which we can estimate how an exogenous change in the speed of the market changes price discovery and the average cost of trading.

Our measure of high-frequency quoting, which we refer to as quote updates, is any change in the best bid or offer (BBO) quote <u>or</u> size across all quote reporting venues. Each such change can be triggered by the addition of liquidity to the limit order book at the BBO, the cancellation of existing unexecuted orders at the BBO, or the extraction of liquidity

[1] Filings with regulatory bodies, exchanges, trade groups and press accounts, as well as some academic papers, contain numerous suggestions to slow the pace of quotation and trading to what is determined to be a reasonable pace. See, for example, the testimony of the Investment Company Institute to the US House of Representatives, Financial Services Subcommittee on Capital Markets and Government-Sponsored Enterprises, in which the testifier argues for meaningful fees on canceled orders as a mechanism to prevent high-frequency changes in the supply curve (http://www.ici.org/pdf/12_house_cap_mkts.pdf).