2017

THREE ESSAYS ON LIQUIDITY IN MODERN MARKETS

Konstantin Sokolov
kbsokolov@gmail.com

THREE ESSAYS ON LIQUIDITY IN MODERN MARKETS

Author:

Konstantin Sokolov

DISSERTATION

Submitted to the Department of Finance

in partial fulfilment of the requirements for

Doctor of Philosophy in Management – Finance

Lazaridis School of Business and Economics

Wilfrid Laurier University

DECLARATION OF CO-AUTHORSHIP/PREVIOUS PUBLICATION

I, Konstantin Sokolov, hereby declare that this thesis incorporates material that is a result of joint research as follows: This thesis incorporates the outcome of joint research undertaken in collaboration with Jonathan Brogaard, Allen Carrion, Thibaut Moyaert, Ryan Riordan and Andriy Shkilko. The collaboration is covered in Chapter 2 of the thesis. In addition, this thesis incorporates the outcome of joint research undertaken in collaboration with Andriy Shkilko. The collaboration is covered in Chapter 3 of the thesis. The key ideas, primary contributions and test designs are those of the author, and the data analysis was performed by the author.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

This thesis includes one original paper that has been accepted for publication to the *Journal of Financial Economics*. Contained in Chapter 2 of this thesis, the publication title of the paper is "HFT and Extreme Price Movements".

ABSTRACT

Recent technological advancements have challenged financial markets. Academic researchers, regulators and market participants voice concerns that modern markets bear the negative externalities of such advancements. Specifically, they are concerned that today's markets are becoming more fragile and unfair to less sophisticated traders. This work employs empirical methodology to test whether these concerns are justified. This thesis contains three essays:

The first essay studies whether modern markets become less liquid during intraday extreme price movements (EPMs). When a price moves in a certain direction, liquidity providers face two opposing incentives. The first incentive is to stay in the market to accumulate more inventory in anticipation of a price reversal. The second incentive is to withdraw due to capital constraints, inventory and adverse selection risks. Using data from Canadian and U.S. markets, I find that the former incentive is stronger during intraday EPMs. This finding alleviates concerns that prices are subject to periods of extreme volatility due to systematic liquidity withdrawals. Contrary to these concerns, liquidity providers appear sufficiently incentivized to dampen intraday volatility.

The second essay examines the activity of a specific type of modern liquidity providers – high frequency traders (HFTs) – around EPMs. I find that, on average, HFTs provide liquidity during EPMs by absorbing imbalances created by non-high frequency traders (nHFTs). Yet HFT liquidity provision is limited to EPMs in single stocks. When several stocks experience simultaneous EPMs, HFT liquidity demand dominates their supply. There is little evidence of HFTs causing EPMs.

The third essay studies whether recent technological advancements result in higher costs for less sophisticated traders. In modern markets, trading firms spend generously to gain a speed advantage over their rivals. The marketplace that results from this rivalry is characterized by speed differentials whereby some traders are faster than others. Is such a marketplace optimal? To answer this question, I study a series of exogenous weather-related episodes that temporarily remove the speed advantages of the fastest traders by disrupting their microwave networks. During these episodes, adverse selection declines accompanied by improved liquidity and reduced volatility. Liquidity improvement is larger than the decline in adverse selection consistent with the emergence of latent liquidity and enhanced competition among liquidity suppliers. The results are confirmed in an event-study setting, whereby a new business model adopted by one of the technology providers reduces speed differentials among traders, which results in liquidity improvements.

ACKNOWLEDGEMENTS

I express sincere gratitude to my advisor and dissertation committee chair, Dr. Andriy Shkilko, whose patience, passion and continued support during my doctoral studies further motivated me to complete this PhD. Thank you, Andriy, for being a great mentor and for providing constructive and insightful feedback.

I also thank my committee members, Dr. Phelim Boyle, Dr. Fabricio Perez and Dr. Brian Smith, for their support and guidance over the dissertation process. I am grateful to Dr. Andreas Park and Dr. Mark Reesor, who generously accepted the role of external examiners, for their insightful recommendations and sincere encouragement. I am also thankful for the support that I received from the Finance Area and the PhD program office. I thank the Ontario Graduate Scholarship program for financially supporting my research. I would like to acknowledge the support and help of my wife, Yana.

I am grateful for feedback on my research from Amber Anand, Robert Battalio, Hendrik Bessembinder, Michael Brolley, Jonathan Brogaard, Sabrina Buti, Allen Carrion, Jean-Edouard Colliard, Erich and Ruth Engler, Sean Foley, Kingsley Fong, Thierry Foucault, Austin Gerig, Emmanuel Gobet, Michael Goldstein, Nathan Halmrast, Terrence Hendershott, Stuart Hinson, Laurence Lescourret, Katya Malinova, Sophie Moinas, Thibaut Moyaert, Olena Nikolsko-Rzhevska, Maureen O'Hara, Michael Pagano, Roberto Pascual, Jireh Ray, Ryan Riordan, Wing Wah Tham, Jun Uno, Kumar Venkataraman, René Wells, Petter Wiberg, Bart Yueshen, Yelin Zhang and Haoxiang Zhu.

STATEMENT OF ORIGINALITY

I, Konstantin Sokolov, hereby declare that the material contained in this dissertation is original work performed primarily by me under the supervision of my dissertation committee. This includes the development of ideas and research models, data analysis and writing of the dissertation. During the process of conducting this research, my committee members were involved in reading my work, commenting on it and helping me improve it. My supervisor, Dr. Andriy Shkilko, provided oversight for all aspects of my research and focused on helping me refine my ideas and writing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. DOES LIQUIDITY EVAPORATE? INTRADAY EVIDENCE

1.1. Introduction

Modern markets are often claimed to bear alarming levels of short-term and long-term volatility. Violent price movements pose especially high concern when they are not related to changes in fundamentals. According to Greenwood and Thesmar (2011) and Khan, Kogan and Serafeim (2012), uninformed liquidity demand can lead to non-fundamental price pressures that persist for a considerable time. Such price pressures can be observed not only in the long run. When liquidity demand is not absorbed by liquidity supply within a short time interval, it has the potential to trigger intraday price movements of large magnitude. Easley, López de Prado and O'Hara (2011, 2012) and Kirilenko, Kyle, Samadi and Tuzun (2016) show that, during one such price pressure episode – the flash crash of May 6, 2010 – electronic liquidity providers reduced their activity. The joint agency report[1] on the October 15, 2014 U.S. Treasury Flash Crash also discusses a reduction in liquidity supply. Echoing an industry-wide concern, CFTC chair Timothy Massad recently noted that, during large intraday price movements, liquidity providers tend to cancel standing orders at the top of the book and then place the orders at deeper book layers.[2] Because of this behavior, liquidity tends to evaporate when it is most needed.

This concern often finds support in academic studies on evaporating liquidity. According to these studies, when prices are under pressure, market makers reposition liquidity to deeper layers of the book, as they look to benefit from the reversals that follow (e.g., Nagel, 2012; Hendershott and Menkveld, 2014; So and Wang, 2014). Due to such repositioning, liquidity demanders quickly consume thin layers of the book, leading to price overshooting and subsequent reversals.

[1] https://www.treasury.gov/press-center/press-releases/Documents/Joint_Staff_Report_Treasury_10-15-2015.pdf
[2] http://www.cftc.gov/PressRoom/SpeechesTestimony/opamassad-30

The literature on contrarian trading suggests, however, that large price movements should be associated with improved liquidity supply. Instead of repositioning the limit orders to the deeper layers of the book, liquidity providers should quote more aggressively, because on average large price movements are likely to be driven by uninformed order flow leading to large reversals (Lehmann, 1990; Lo and MacKinlay, 1990; Gromb and Vayanos, 2002; Avramov, Chordia and Goyal, 2006). According to this view, liquidity demanders should find sufficient liquidity supply throughout the book during price pressure episodes.

This paper tests whether limit order book depth and the price impacts of marketable orders point to the dominance of evaporating liquidity or competitive contrarian liquidity provision during intraday price movements. I design tests assuming the following model of liquidity provision: First, competitive liquidity providers choose how much depth to place at every level of the limit order book. Second, liquidity demanders begin submitting marketable orders that consume some of the depth thus moving the price. Third, liquidity providers choose to either keep the remaining limit orders or cancel them. If they cancel, liquidity evaporates as in Nagel (2012), and prices overshoot. Finally, prices reverse once liquidity demand has been exhausted and the book bounces back.

My findings suggest that for intraday price pressures evaporating liquidity does not dominate modern markets. On average, the per share impact of marketable orders decreases and the limit order book depth increases in the magnitude of price movements.

Although the data generally suggest that liquidity does not evaporate, I find some evidence for the existence of market making strategies suggested by Nagel (2012), Hendershott and Menkveld (2014) and So and Wang (2014). Specifically, I observe that limit order book depth is net cancelled in the way of price movements and repositioned deeper in the book. The economic effect of this repositioning is however small and affects less than 6% of total depth.

The literature on evaporating liquidity suggests that short-term price inefficiencies serve as a profit source for market makers. To support the existence of the contrarian trading incentive, I show that the profitability of limit order trading increases in the magnitude of intraday price inefficiencies. The aggregate profits from liquidity provision increase by $9-92 per price movement with one standard deviation increase in the transitory component of price volatility. This confirms that liquidity providers are willing to increase contrarian limit order placement when the expected transitory price component is large.

The remainder of the paper is as follows: Section 1.2 describes how the paper fits into the extant literature, Section 1.3 describes the data and methodology, Section 1.4 reports summary statistics, Section 1.5 contains empirical findings on limit order book dynamics and contrarian profits, and Section 1.6 concludes.

1.2. Literature review

The paper builds on three strands of literature. The first strand emphasizes the risks and constraints of liquidity provision, which implies that market makers should withdraw during price pressures. The second strand of literature discusses the profitability of contrarian strategies, which implies that contrarian traders should accumulate inventory against the direction of price pressures. The third strand combines the first two by suggesting that the size and frequency of price reversals are increasing when market makers scale back to earn higher compensation for the risks of liquidity provision through contrarian profits. The goal of this paper is to test whether large intraday price pressures create significant risks so that liquidity providers will supply less liquidity consistent with the behaviour suggested by the third strand.

The size of inventory that liquidity providers are willing to accumulate depends on multiple incentives and constraints. In the theory literature, severe instances of mispricing, such as stock market bubbles (Shleifer and Vishny, 1997) and downward liquidity spirals (Brunnermeier and Pedersen, 2009; Huang and Wang, 2009), are often attributed to the lack of funding liquidity available to contrarian traders. Comerton-Forde et al. (2010) support the abovementioned theoretical predictions by showing that NYSE market makers provide less liquidity when their balance sheet revenues decrease. Furthermore, the significant co-movement of stocks can drain the capital of liquidity providers and lead to illiquidity and large price reversals (Andrade, Chang and Seasholes, 2008; Hameed, Kang and Viswanathan, 2010; Hameed and Mian, 2015).

Market makers may choose to limit their inventory exposure even if capital constraints are not binding. For example, it is in a liquidity provider's interest to withdraw quotes in anticipation of incoming informed market orders. Bonart and Gould (2015) show that limit order traders successfully adjust their strategies to market order arrivals. Hendershott, Jones and Menkveld (2011) observe that, with the proliferation of algorithmic liquidity provision, quote cancellation and repositioning account for a higher proportion of price discovery than trades. Liquidity shortages due to increased anticipated informed trading may occasionally cause excessive price fluctuations (Easley, Lopez de Prado and O'Hara, 2011, 2012). Panayides (2007) shows that liquidity can be temporarily improved by affirmative obligations forcing market makers to accumulate informed order flow. When the obligations are not binding, liquidity deteriorates because market makers have to compensate for the position losses incurred during mandatory participation.

Finally, liquidity providers may fail to accommodate order flow if it creates significant dislocations in their preferred holding portfolios (Amihud and Mendelson, 1980). Large order flows may impede portfolio diversification and set the liquidity provider's inventory risk above the acceptable level. According to Madhavan and Sofianos (1998), NYSE specialists maintain target inventory levels by selectively withdrawing from liquidity provision on either the buy or the sell side of the bid ask spread. Malinova and Park (2016) show that modern low-latency market makers continue to manage their inventory through quote withdrawal, which leads to lower liquidity and higher volatility. Market makers may be especially unwilling to hold highly risky portfolios because the stochastic nature of liquidity demand makes the portfolio holding period uncertain (Ho and Stoll, 1981). This effect may further aggravate the depletion of the limit order book during price pressures.

However, price pressures incentivise liquidity providers and arbitrageurs to increase liquidity supply in anticipation of higher contrarian profits. Due to the execution uncertainty of non-marketable limit orders, liquidity providers are often assumed to be less informed than traders who require immediacy. Although liquidity providers may not know if a specific price pressure will result in a permanent or a transitory price dislocation, they may observe that most price movements revert over time and engage in contrarian trading (Gromb and Vayanos, 2002; Avramov, Chordia and Goyal, 2006; Hendershott and Seasholes, 2007; Biais, Declerck and Moinas, 2016). According to Lehmann (1990) and Lo and MacKinlay (1990), the expected reversible price component increases in the magnitude of the price movements. This finding implies that large price movements should lead to high expected profits from contrarian trading and create strong incentives for contrarian liquidity provision.

The literature suggesting that liquidity providers should withdraw during large price movements and the literature on the profitability of contrarian trading are generalized by papers proposing the risk-return trade-off of contrarian liquidity provision. The trade-off between the risks and contrarian profits of market making in volatile markets has been studied by Nagel (2012). He estimates that contrarian trading becomes more profitable during volatile daily and multiple-day intervals. He suggests that, when uncertainty is high, liquidity providers scale back and extract additional contrarian profits to compensate for risks and constraints. This leads to the evaporating liquidity effect when uninformed order flow creates large price movements that reverse over time.

So and Wang (2014) show that, consistent with Nagel (2012), price reversals become more prevalent and contrarian trading strategies become more profitable during intervals of high uncertainty around earnings announcements. Hendershott and Menkveld (2014) develop a model predicting that the daily reversible price component may become larger when risk-averse market makers rebalance their inventory. Following Weill (2007), constrained market makers may refrain from liquidity provision at certain points during price reversals, which will lead to jumps and aggravate the magnitude of these reversals. As such, the literature on evaporating liquidity suggests that the market is often in equilibrium where market makers have incentives to withdraw from liquidity provision with the purpose of benefiting from resulting price reversals.

A burgeoning literature on algorithmic and high-frequency trading takes the debate on whether the risk-return trade-off of liquidity provision results in an equilibrium with evaporating liquidity to the intraday level. This literature divides microstructure researchers into two opposing camps.

On the one hand, several papers find that algorithmic liquidity providers may mitigate intraday volatility spikes while making contrarian profits (Anand and Venkataraman, 2016; Subrahmanyam and Zheng, 2015; Brogaard et al., 2016). The mechanism for such mitigation has been discussed by Ahn, Bae and Chan (2001), who find that liquidity providers step in after periods of high intraday volatility and mitigate the magnitude of price movements. The finding that intraday algorithmic traders act as contrarians during intraday price movements is supported by the literature suggesting that algorithmic trading leads to more efficient intraday prices (Hendershott, Jones and Menkveld, 2011; Brogaard, Hendershott and Riordan, 2014; Conrad, Wahal and Xiang, 2015). Colliard (2015) shows that, although contrarian traders and market makers counteract short-term price reversals and make prices more efficient in the short run, they may impede price discovery in the long run. This camp, as a whole, suggests that the technological advancement of liquidity provision pushes markets to the equilibrium where intraday price movements are dominated by competitive contrarian liquidity provision.

On the other hand, some studies show that algorithmic traders have incentive to reposition liquidity to the deeper layers of the limit order book and increase the magnitude of intraday price movements (Kirilenko, Kyle, Samadi and Tuzun, 2016; van Kervel, 2015). The finding that algorithmic traders have the potential to exacerbate price fluctuations is consistent with the results of Korajczyk and Murphy (2015) and Menkveld and van Kervel (2015) who show that high frequency traders switch from liquidity supply to liquidity demand during the execution of large institutional trades. As a whole, this camp as a whole suggests that the technological advancement of liquidity provision pushes markets to the equilibrium where intraday price movements are dominated by evaporating liquidity.

This paper takes one step further and tests whether intraday price movements in the modern Canadian and US equity markets are in equilibrium dominated by competitive contrarian liquidity provision or evaporating liquidity. Although the paper does not attempt to identify the impact of algorithmic and high-frequency trading activity per se, it provides empirical evidence that is generally consistent with the former equilibrium.

1.3. Data and methodology

The main results are obtained from two datasets. The first is order data from the Toronto Stock Exchange (TSX) for the full year of 2006, and the second is millisecond trade and quote DTAQ data for the full year of 2014. The TSX dataset is the same as that used by Anand and Venkataraman (2016). Reliable estimation of the reversible price component requires a sufficient number of observations, which are often not available in small stocks. To ensure reliability, I use the TSX60 index constituents, which are the most active stocks in the Canadian market.

Although the Canadian market circa 2006 lacks some features of today's markets, such as the dominance of high frequency trader (HFT) algorithms and the proliferation of dark pools, the TSX data have two advantages. First, the data contain all limit order placements and cancellations: thus, allowing an examination of the limit order book beyond the best quotes. Second, the TSX data allows exact identification regarding whether a particular trade was triggered by marketable ask or bid order while the trade direction is not identified in the DTAQ data. I use Lee and Ready (1991) trade classification algorithm to identify trade direction in the DTAQ sample.

To address the possibility that the 2006 sample is not representative of modern markets, I test my findings on a 2014 sample of U.S. stocks from the DTAQ database. Although the DTAQ data do not report orders, they are available for a recent time period when high frequency trading accounts for a substantial share of the trading volume. Moreover, the data allow me to confirm that the results are not unique to the Canadian market. Unlike the Canadian market circa 2006, trading activity in the modern U.S. market is significant outside of the major indices, allowing for proper estimation of a reversible price component for smaller stocks. A broad sample of stocks with different capitalizations will allow me to observe possible differences in limit order book depth dynamics around price reversals. I construct the sample of US stocks using a procedure similar to Chakrabarty, Jain, Shkilko and Sokolov (2016). This procedure involves two steps: First, I assign market capitalization and trading volume rank to all stocks in the CRSP database, and second, I select three groups of thirty stocks each as large, medium and small based on the sum of the market capitalization and trading volume ranks.

The sample is constructed in tick time, where tick is an update of the midquote price. Calendar time sampling would lead to an incomparable number of intervals with zero activity across different samples. Another alternative could be trade time sampling. Trade time sampling does not have the shortcomings of calendar time sampling but, as shown by Aït-Sahalia, Mykland and Zhang (2011), leads to the overestimation of the reversible price component due to bid-ask bounce. As such, tick time sampling does not have either of the above shortcomings and appears to be the best alternative for the purpose of current analysis.

The primary set of results is focused on the dynamics of depth and expected reversible price component over the series of consecutive same-directional midquote updates forming the price movements. All price movements during the trading day are included in the sample. The typical price movement can be presented as follows:

Return sign	+	+	+	+	−	−	−	+	+	+	…
Number of quote updates before pivot	4	3	2	1	3	2	1	3	2	1	
Pivot point	no	no	no	yes	no	no	yes	no	no	yes	

I estimate the expected reversible price component as the signed difference between the current pivot price[3] and forecasted pivot price at t+1. Price forecast could be obtained using an autoregressive model similar to the one used in Hasbrouck (1993). Although I confirm my findings with Hasbrouck methodology on a daily level, I choose a simplified approach for intraday estimation of the expected reversible price component.

The intraday reversible price components are estimated using the following steps: First, I rank all pivot-to-pivot log returns for the same stock into twenty buckets by magnitude. Then I run the following AR(1) model for each stock:

$$R_{i,t} = \alpha_i + \beta_i R_{i,t-1} + \varepsilon_{it}$$

where $R_{i,t}$ and $R_{i,t-1}$ are the current and lagged midquote returns for the stock i.

The coefficients are used to obtain the expected future return estimate $\hat{R}_t$. This estimate is the reversible price component at time $t-1$. For each interval, it shows the percentage of price that is expected to be corrected in the future conditioned on past return. This variable can be compared to the return at time $t-1$ in sign and magnitude. Conditional on execution at the pivot points, the expected compensation for contrarian liquidity providers will be a sum of effective spread and reversible price component.

[3] Since most of the returns within the sequence of midquote-to-midquote updates have a magnitude bounded by the minimum tick size, it is more appropriate to estimate the reversible component of the total pivot-to-pivot return. The results, however, are robust to estimation with midquote-to-midquote returns and are available upon request.

1.4. Summary statistics

To capture the differences in liquidity provision during price movements of different magnitude, I split the sample into two subsamples. The first one includes only price movements with the magnitude below the 90th percentile while the second one includes the price movements with the magnitude above 90th percentile. Table 1.1 reports descriptive statistics. The average time between the two midquote updates ranges between 3 and 10 seconds for non-extreme price movements (non-EPMs) and between 22 and 33 seconds for extreme price movements (EPMs). Since modern quote-to-trade ratios tend to be high, the majority of non-EPM price movements happen without trading, with the median number of shares traded per price movement being as low as zero. This, however, does not imply that volume is trivial when trading takes place. Even for a subsample of non-EPMs, several hundreds of shares are demanded on average in the direction of return. The EPM sample contains intervals with much more intense trading activity ranging from a few thousand to more than ten thousand shares traded during the average EPM.

When more shares are demanded on one side of the quotes than on the other, volume imbalance is created. Consistent with the literature on market making inventory management, prices yield to such volume imbalances. The ratio of shares demanded in the direction of return to those demanded against the direction of the return is about two.

This volume imbalance ratio is consistent across the TSX and US samples. Unlike the TSX data, the data on US markets does not report whether each particular trade is initiated by buy or sell marketable order, and I use Lee and Ready (1991) trade classification algorithm to identify trade direction in the US data. The consistency of the volume imbalance ratio for both the TSX and the US samples suggests that Lee and Ready (1991) algorithm provides a reasonable estimation of trade direction.

Despite the differences in sample periods, the TSX sample is comparable to the US sample. Trading activity in TSX stocks most closely resembles trading activity in US medium-sized stocks. Nonetheless, consistent with lower levels of algorithmic and high-frequency trading activity, price movements last longer in the TSX sample.

In modern markets, liquidity is provided by a variety of market participants. Among them, high-frequency traders have the technological capacity to react to order flow exceptionally quickly. Therefore, they can withdraw limit orders with ease as EPMs develop. I note that high-frequency traders are present only in the US sample.

The behavior of liquidity demanders during negative price movements may be affected by short-selling restrictions. The TSX sample may therefore contain price declines of smaller magnitude. This said, the restrictions are unlikely to affect the US sample because they are triggered only after a 10% price decline. It is also possible that the magnitude of price movements is reduced by the routing software. Since order routers observe the state of the book in real time, they may opt out of demanding too much if the book is thin thereby mitigating the price movements.

1.5. Results

1.5.1. The dominance of the contrarian trading incentive

The success of contrarian trading implies that the magnitude of the pricing error is positively related to the probability of a price reversal. If it were not the case, then pricing errors would systematically increase over time, leading to losses for contrarian traders, and eventually to market crashes. Pricing errors cannot be observed, which makes studies that use various techniques for pricing error estimation subject to criticism that the results depend on the choice of these techniques. However, price movements are observed, and according to Lehmann (1990) and Lo and MacKinlay (1990), larger price movements should contain a larger reversible component.

I use probit analysis to test if the relation between the probability of a reversal and the magnitude of the pricing error proxied by the return magnitude is consistent with the profitability of contrarian trading. Specifically, I estimate the probability that the current midquote update is a pivot point conditional on accumulated return. According to the coefficients reported in the Table 1.2, the return magnitude is positively related to the probability of reversal. The result is consistent across all samples. On average, the probability of a price reversal increases by 14% with one basis point increase in cumulative EPM return for the most active stocks. Pricing errors do not systematically accumulate over time. This shows that contrarian traders have a stronger incentive to trade against the direction of return when the pricing error increases.

Contrarian trading incentive defines the shape of the limit order book before a price movement begins. Specifically, deeper layers of the book typically contain more shares. As such, available contrarian depth should increase as market orders consume first layers of the book unless substantial amount of depth is cancelled.

I estimate contrarian depth as a share depth of the best bid (ask) quote for intervals with a negative (positive) midquote return weighted by the proportion of time the quote depth remained unchanged in the entire duration of the price movement. This variable represents instant liquidity that limit order traders are willing to provide at the best contrarian quote. Figure 1.1 plots contrarian depth and the reversible price component by return magnitude percentiles. According to the figure, consistent with the dominance of contrarian trading incentive, contrarian limit order depth increases with an increase in return magnitude and the reversible price component.

The depth of the best contrarian quotes suggests that intraday price movements do not typically accelerate because of a lack of contrarian liquidity provision. Notably, however, not all of the limit orders placed at the best quotes end up being executed. Some limit orders can be cancelled and repositioned, as the literature on constrained liquidity provision suggests. For example, according to van Kervel (2015), observed limit order book depth can be misleading as some liquidity providers may cancel limit orders when they observe substantial liquidity demand. To address this possibility, I construct additional measures of the willingness of an average liquidity provider to accumulate inventory against the direction of the price movements.

The first measure is the elasticity of the best contrarian quotes. For a positive (negative) price movement, it captures the number of shares that liquidity providers supply to marketable buy (sell) orders per one-cent update of the best ask (bid) quote:

$$ContrElast = \frac{DirVolume}{\Delta BestContrarianQuote}$$

Although I focus on the elasticity of the contrarian quote in testing whether contrarian trading incentive dominates the incentive to reposition liquidity, I provide estimates of non-contrarian elasticity, which is computed in a similar manner:

$$NonContrElast = \frac{InDirVolume}{\Delta BestNonContrarianQuote}$$

The second measure takes into account that marketable orders traded against the direction of price movements can decrease the inventory risks of liquidity providers. It is computed as the inverse[4] of Amihud (2002) illiquidity measure, except that the data allows me to use the net volume traded in the direction of return instead of the total volume:

$$InvAmihud = \frac{DirVolume - InDirVolume}{|Ret|}$$

Table 1.3 reports statistics for EPMs and non-EPMs. Consistent with the literature on contrarian trading, an average EPM contains a reversible price component of a magnitude that is several times higher than that of an average non-EPM. In line with Figure 1.1, the best contrarian depth is 12-50% higher during EPMs. This suggests that net liquidity supply during an intraday price movement is dominated by incentives of contrarian trading rather than by risk aversion and capital constraints. Contrarian limit order book depth is significantly higher during EPMs.

The estimates of the elasticity of contrarian quotes point to the same direction as the estimates of the best limit order book depth. Specifically, in the TSX sample, the best contrarian quote yields by one cent after 826 shares are demanded during the EPM, while it yields to demand after 304 shares are demanded during non-EPMs. The result is consistent and statistically significant for all samples. Comparison of inverse Amihud measures for EPM and non-EPM intervals reveals that it takes a two times higher net volume imbalance to move the midquote price by one cent during EPMs than during non-EPMs. For example, in the TSX sample, it takes 913 shares of net volume to move the midquote by one cent during EPMs while it takes only 450 shares to achieve the same result during non-EPMs. The corresponding net volumes for the sample of the most active US stocks are 7,470 and 3,547 shares, respectively.

The fact that both contrarian quote elasticity and inverse Amihud measure point to an increase in liquidity supply during EPMs suggests that evaporating liquidity is not a concern for an average intraday EPM. Even if some of the limit order book depth is repositioned from the

[4] By design, the sample does not contain zero-return intervals. This allows me to invert the original measure for more intuitive economic interpretation.

best quotes at the beginning of a price movement towards the end of a price movement, thus increasing its magnitude, the volume executed per one-cent price update ends up being higher during larger price movements than it is during smaller ones.

1.5.2. Does liquidity evaporate at all on the intraday level?

Contrarian incentive appears to be stronger than the incentive to scale back from liquidity provision on the intraday level. In this section, I use TSX data to examine if there is any evidence of evaporating liquidity. According to Nagel (2012), Hendershott and Menkveld (2014) and So and Wang (2014), liquidity providers have an incentive to reposition limit orders further from the best quotes so that larger pricing errors would compensate for risks and constraints.

Figure 1.2 describes limit order book dynamics around the pivot points of intraday reversals. Intraday limit order book dynamics during price movements has three important features. First, the limit order book depth is net placed at the future best contrarian quote at the pivot point. Second, the depth is net cancelled at the limit order book layers between the current price and the future pivot point. Third, the amount of depth net cancelled is economically small for both EPMs and non-EPMs.

The above observations suggest that the evaporating liquidity effect described by Nagel (2012), Hendershott and Menkveld (2014) and So and Wang (2014) for the daily and multiple-day price reversals also exists for the intraday ones. That said, consistent with my earlier finding that the contrarian limit order book depth increases in the magnitude of the intraday price movements, the evaporating liquidity does not exacerbate the magnitude of these movements.

Table 1.4 contains detailed analysis of the limit order book depth placement and cancellation dynamics during the five ticks before the pivot point.[5] As the price moves in a certain direction, contrarian depth is withdrawn from the limit order book layers in the direction of the price updates and placed at the future best contrarian quote. For example, during non-EPMs, one tick before the stock stops falling (raising) at the best contrarian quote P_{bid} (P_{ask}), there is, on average, 80 shares net cancelled at the quote above P_{bid} (below P_{ask}) and 262 shares net placed at the quote P_{bid} (P_{ask}). One tick before the end of the EPMs, there is, on average, 53 shares net cancelled in the way of the price movements and 83 shares placed at the future best pivot quotes. The number of shares cancelled represents 6.7% and 2.6% of the average best contrarian depth during non-EPMs and EPMs, respectively. As such, despite the existence of the evaporating liquidity effect during intraday price movements, this effect does not increase in the magnitude of price movements and has small economic significance.

The contrarian depth placed at the layers that are less competitive than the future best pivot quote is not executed during the ongoing price movement. The decrease in contrarian depth placement at these layers suggests two non-exclusive explanations. First, the liquidity providers who engage in reversal trading strategies described by the literature on evaporating liquidity have a smaller incentive to assume inventory risk after a certain point. Second, such liquidity providers can successfully predict the pivot points of intraday EPMs and non-EPMs.

The above results point to the existence of limit order repositioning on the intraday level that is consistent with the activity of liquidity providers described by Nagel (2012), Hendershott and Menkveld (2014) and So and Wang (2014) on the daily and multiple-day level. In line with the increasing information content of limit orders documented by Hendershott, Jones and

[5] Limit order book depth dynamics prior to the last five ticks before the pivot point continues the pattern observed during the last five ticks. The results are available upon request.

Menkveld (2011) and O'Hara (2015), the repositioned limit order book depth appears to be highly predictive of the price movement ending point. However, the economic effect of the repositioned liquidity is small, and it does not increase with the magnitude of price movements.

1.5.3. Profitability of contrarian liquidity provision

The positive relation between contrarian liquidity provision and return magnitude is consistent with liquidity providers extracting contrarian profits from price reversals, as suggested by Nagel (2012). However, an alternative explanation is especially pertinent to intraday price movements: the possibility that increased limit order book depth during large intraday price movements is due to limit orders being picked off by fast informed liquidity demanders rather than due to a significant pricing error component of the large price movements. If this is the case, then large intraday price movements and pricing errors would result in losses for aggregated limit order traders on average.

Although the profitability of contrarian strategies and aggregated limit order trading has been documented by earlier studies, there is no empirical evidence for the relation between pricing errors and profits from liquidity provision. The goal of this section is to test whether aggregated profits of limit order trading are increasing consistent with contrarian trading incentive or decreasing consistent with informed liquidity demand by fast traders.

I estimate aggregate profits from a limit order strategy using a methodology developed by Handa and Schwartz (1996). This methodology allows estimating the total daily profit from liquidity provision as a sum of two components:

$$\Pi = (P_a - P_b) * Min(N_a, N_b) + (\bar{P} - P_{EOD}) * (N_a - N_b),$$

where P_a and P_b are the average prices per share executed through non-marketable sell and buy limit orders, N_a and N_b are the numbers of shares traded against ask and bid quotes, P_{EOD} is the end-of-day midquote,[6] $\bar{P} = P_a$ if $N_a > N_b$ and $\bar{P} = P_b$ if $N_b > N_a$.

The first term corresponds to the aggregate realized profit from spread and reversals, and the second term corresponds to the unrealized profit from the end-of-day position. To make the data comparable across stock-days with different numbers of reversals, I scale the profits by the number of pivot points per day. Gross profit estimates are not affected by possible make-or-take fee structure discrepancies. Since liquidity provision rebates are independent of the positioning of non-marketable limit orders in the limit order book, I do not include them into profit calculations.

The main independent variable for the current profit analysis is the pricing error ratio. The larger the ratio, the higher the proportion of a reversible component of the total daily volatility. I estimate the pricing error ratio following the methodology developed by Hasbrouck (1993). First, I estimate the vector autoregression (VAR) system with ten lags:

$$r_t = a_1 r_{t-1} + a_2 r_{t-2} + \cdots + b_1 x_{t-1} + b_2 x_{t-2} + \cdots \ v_{r,t}$$

$$x_t = c_1 r_{t-1} + c_2 r_{t-2} + \cdots + d_1 x_{t-1} + d_2 x_{t-2} + \cdots \ v_{x,t},$$

where r_t is the difference in log midquotes;[7] x_t is the column vector of three signed trade-related variables (a signed trade indicator, signed trading volume, and signed square root of trading volume) that allows for a nonlinear relation between returns and trades, and $v_{r,t}$ and $v_{x,t}$ are

[6] Handa and Schwartz (1996) use trade data and benchmark EOD inventory against closing trade. I follow the approach of Brogaard, Hendershott and Riordan (2014) and use EOD midquote as a benchmark.

[7] Using calendar time for price sampling could result in incomparable estimates of the pricing error, as liquid stocks have higher midquote update frequencies than the less liquid ones. To make empirical tests comparable across different stocks, I compute intraday returns in tick time.

zero-mean serially uncorrelated disturbances. I estimate the VAR on the day-by-day basis. Next, I invert the VAR to obtain the vector moving average (VMA) representation:

$$r_t = a_0^* v_{r,t} + a_1^* v_{r,t-1} + \ldots + b_0^* v_{x,t} + b_1^* v_{x,t-1} + \cdots$$

$$x_t = c_0^* v_{r,t} + c_1^* v_{r,t-1} + \ldots + d_0^* v_{x,t} + d_1^* v_{x,t-1} + \cdots$$

Using the return equation from the VMA model, I estimate the transitory price component:

$$s_t = \alpha_0 v_{r,t} + \alpha_1 v_{r,t-1} + \ldots + \beta_0 v_{x,t} + \beta_1 v_{x,t-1} + \cdots,$$

where $\alpha_j = -\sum_{k=j+1}^{\infty} a_k^*$ and $\beta_j = -\sum_{k=j+1}^{\infty} b_k^*$.

The estimate of s_t shows the sum of expected future price updates given current and past shocks to return and trading.

The pricing error variance can be computed as:

$$\sigma_s^2 = \sum_{j=0}^{\infty} \left[\alpha_j \; \beta_j\right] Cov(v) \begin{bmatrix} \alpha_j \\ \beta_j \end{bmatrix}$$

I estimate the pricing error ratio as a square root of the pricing error variance scaled by the volatility of the log of midquote price:

$$PricingErrorRatio = \frac{\sigma_s}{\sigma_{\log(p)}}$$

It is important to note that the Hasbrouck (1993) methodology estimates the lower bound for the pricing error. It assumes that the public information set includes only past returns, volume and trading direction. I acknowledge that pricing errors of a higher magnitude may be obtained with a more sophisticated methodology.[8] However, underestimating the magnitude of the pricing error is unlikely to lead to a different conclusion about the relation between the pricing error and

[8] For example, Brogaard, Hendershott and Riordan (2014) apply the Kalman filter to estimate the impact of high-frequency traders on mispricing.

liquidity provision, as long as my estimate of the pricing error is increasing in the magnitude of the true pricing error.

Table 1.5 reports the summary statistics for profit components and the pricing error ratio. Most aggregate limit order trading profit comes from reversals and spread. Profit estimates for the Canadian market are significantly larger than those obtained for the US market, which implies that the latter has lower execution costs for marketable orders. Estimated mean pricing error ratios are between 1.2 and 2.5%. This does not mean, however, that price inefficiencies have low economic significance. Hasbrouck's (1993) methodology provides an estimate of the lower bound of the unobserved pricing error and the true economic significance of the pricing error is likely to be greater than estimated.

I estimate the relation between the limit order trading profit and the pricing error ratio with the following linear model:

$$Profit_{it} = \beta_1 PricingErrorRatio_{it} + \beta_2 PESpread_{it} + \beta_3 nTrades_{it} + \beta_4 DailyReturn_{it} + \varepsilon_{it},$$

where $Profit_{it}$ is the total profit or one of the profit components; $PricingErrorRatio_{it}$ is the pricing error ratio; $PESpread_{it}$ is the average percentage effective spread; $nTrades_{it}$ is the average number of trades per price movement; and $DailyReturn_{it}$ is the total daily return magnitude. All variables are estimated on the stock-day level and standardized across stocks.

The coefficient estimates are reported in Table 1.6. Consistent with the predictions of Nagel (2012), one standard deviation increase in the pricing error ratio corresponds to an increase in the total limit order trading profits of 4-17%. This translates into a $92.3 increase in profit per price movement in the Canadian sample and $76.5, $9.6 and $8.8 increases in the U.S. large, medium and small samples.[9] These estimates are adjusted for volatility, volume and

[9] Although the profit estimates appear to be very large, they are associated with substantial risk. For instance, in the TSX sample, the ratio of the profit median to the standard deviation is less than 0.1.

spread. Since the pricing error ratio captures the reversible component of volatility, it is more closely related with the spread and reversal components than the end-of-day position component of the total profit. There is significant correlation between the end-of-day position profit and the pricing error ratio for the Canadian and US small stocks. This suggests that, consistent with the predictions of Hendershott and Menkveld (2014), liquidity providers may take advantage of price pressures to manage their inventories.

1.6. Conclusions

This paper studies contrarian liquidity provision during price fluctuations. Academic researchers and regulators often voice concerns that liquidity providers have incentives to scale back during large intraday price movements; thus, exacerbating them. These concerns are fostered by the literature on evaporating liquidity that suggests that the risks and constraints of market making lead to the equilibrium where liquidity providers scale back during daily and multiple-day price movements to benefit from resulting price reversals.

I alleviate the above concerns for intraday price movements by showing that, during intraday price movements, contrarian incentive dominates the incentive to scale back from liquidity provision. Consistent with competitive contrarian trading, liquidity provision is increasing in the magnitude of intraday price movements. This result implies that the low-liquidity equilibrium suggested by the literature on evaporating liquidity for daily and multiple-day price fluctuations does not typically hold for intraday ones. Liquidity providers typically have enough capacity to intensify contrarian liquidity provision with the magnitude of intraday price movements.

Although contrarian liquidity provision increases in the magnitude of intraday price movements, I find some evidence consistent with the scaling back of liquidity providers proposed by the literature on evaporating liquidity. Specifically, limit order depth is withdrawn as intraday price movements develop and are repositioned to the quotes that are closer to the pivot point. However, the economic significance of such repositioning is small and decreasing in the magnitude of price movements. This result suggests that some intraday traders engage in the reversal strategies described in the literature on evaporating liquidity.

Enhanced liquidity provision during large price movements is consistent with the profitability of reversal trading strategies developed by several studies. These studies assume that price inefficiency serves as compensation for liquidity provision. I support this assumption by finding that the aggregate profits of limit order traders increase with the intraday pricing error. Moreover, I find some evidence that pricing errors assist limit order traders in managing end-of-day positions.

1.7. References

Ahn, H.-J., Bae K.-H. and Chan K., 2001, Limit orders, depth, and volatility: Evidence from the stock exchange of Hong Kong, *The Journal of Finance* 56, 767–788.

Aït-Sahalia, Y., Mykland P. and Zhang L., 2011, Ultra high frequency volatility estimation with dependent microstructure noise, *Journal of Econometrics* 160, 160–175.

Amihud, Y., 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

Amihud, Y. and Mendelson H., 1980, Dealership market: Market-making with inventory, *Journal of Financial Economics* 8, 31–53.

Anand, A. and Venkataraman K., 2016, Market conditions, fragility and the economics of market making, *Journal of Financial Economics*, Forthcoming.

Andrade, S., Chang C. and Seasholes M., 2008, Trading imbalances, predictable reversals, and cross-stock price pressure, *Journal of Financial Economics* 88, 406–423.

Avramov, D., Chordia, T. and Goyal, A., 2006, Liquidity and autocorrelations in individual stock returns, *The Journal of Finance* 61, 2365–2394.

Biais, B., Declerck, F. and Moinas, S., 2016, Who supplies liquidity, how and when? BIS working paper.

Bonart, J. and Gould M., 2015, Strategic liquidity provision in a limit order book, Working paper.

Brogaard, J., Carrion, A., Moyaert, T., Riordan, R., Shkilko, A. and Sokolov, K., 2016, High-frequency trading and extreme price movements, *Journal of Financial Economics*, Forthcoming.

Brogaard, J., Hendershott, T. and Riordan, R., 2014, High frequency trading and price discovery, *Review of Financial Studies* 27, 2267–2306.

Brunnermeier, M. and Pedersen L., 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2238.

Chakrabarty, B., Jain, P., Shkilko, A. and Sokolov, K., 2016, Speed of market access and market quality: Evidence from the SEC naked access ban, Working paper.

Colliard, J.-E., 2015, Catching falling knives: Speculating on market overreaction. Working paper.

Comerton-Forde, C., Hendershott, T., Jones, C., Moulton, P. and Seasholes, M., 2010, Time variation in liquidity: The role of market-maker inventories and revenues, *The Journal of Finance* 65, 295–331.

Conrad, J., Wahal, S. and Xiang, J., 2015, High-frequency quoting, trading, and the efficiency of prices, *Journal of Financial Economics* 116, 271–291.

Easley, D., Lopez De Prado, M. and O'Hara, M., 2011, The microstructure of the flash crash: Flow toxicity, liquidity crashes, and the probability of informed trading, *Journal of Portfolio Management* 37, 118.

Easley, D., Lopez De Prado, M. and O'Hara, M., 2012, Flow toxicity and liquidity in a high-frequency world, *Review of Financial Studies* 25, 1457–1493.

Greenwood, R. and Thesmar, D., 2011, Stock price fragility, *Journal of Financial Economics* 102, 471–490.

Gromb, D. and Vayanos, D. 2002, Equilibrium and welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics* 66, 361–407.

Hameed, A. and Mian, G., 2015, Industries and stock return reversals, *Journal of Financial and Quantitative Analysis* 50, 89–117.

Hameed, A., Kang, W. and Viswanathan, S., 2010, Stock market declines and liquidity, *The Journal of Finance* 65, 257–293.

Handa, P. and Schwartz, R., 1996, Limit order trading, *The Journal of Finance* 51, 1835–1861.

Hasbrouck, J., 1993, Assessing the quality of a security market: A new approach to transaction-cost measurement, *Review of Financial Studies* 6, 191–212.

Hendershott, T., Jones, C. and Menkveld, A., 2011, Does algorithmic trading improve liquidity? *The Journal of Finance* 66, 1–33.

Hendershott, T. and Menkveld, A., 2014, Price pressures, *Journal of Financial Economics* 114, 405–423.

Hendershott, T. and Seasholes, M., 2007, Market maker inventories and stock prices, *American Economic Review* 97, 210–214.

Ho, T. and Stoll, H., 1981, Optimal dealer pricing under transactions and return uncertainty, *Journal of Financial Economics* 9, 47–73.

Huang, J. and Wang, J., 2009, Liquidity and market crashes, *Review of Financial Studies* 22, 2607–2643.

Khan, M., Kogan, L. and Serafeim, G., 2012, Mutual fund trading pressure: Firm-level stock price impact and timing of SEOs, *The Journal of Finance* 67, 1371–1395.

Kirilenko, A., Kyle, A., Samadi, M. and Tuzun, T., 2016, The flash crash: The impact of high-frequency trading on an electronic market, *The Journal of Finance*, forthcoming.

Korajczyk, R. and Murphy, D., 2015, High frequency market making to large institutional trades, Working paper.

Lee, C. and Ready, M., 1991, Inferring trade direction from intraday data, *The Journal of Finance* 46, 733–746.

Lehmann, B., 1990, Fads, martingales, and market efficiency, *Quarterly Journal of Economics* 105, 1–28.

Lo, A. and MacKinlay, C., 1990, When are contrarian profits due to stock market overreaction? *Review of Financial Studies* 2, 175–205.

Madhavan, A. and Sofianos, G., 1998, An empirical analysis of NYSE specialist trading, *Journal of Financial Economics* 48, 189–210.

Malinova, E. and Park, A., 2016, "Modern" market makers, Working paper.

Menkveld, A. and van Kervel, V., 2015, High-frequency trading around large institutional orders, Working paper.

Nagel, S., 2012, Evaporating liquidity, *Review of Financial Studies* 25, 2005–2039.

O'Hara, M., 2015, High frequency market microstructure, *Journal of Financial Economics* 116, 257–270.

Panayides, M., 2007, Affirmative obligations and market making with inventory, *Journal of Financial Economics* 86, 513–542.

Shleifer, A. and Vishny, R., 1997, The limits of arbitrage, *The Journal of Finance* 52, 35–55.

So, E. and Wang, S., 2014, News-driven return reversals: Liquidity provision ahead of earnings announcements, *Journal of Financial Economics* 114, 20–35.

Subrahmanyam, A. and Zheng, H., 2015, Limit order placement by high-frequency traders, Working paper.

van Kervel, V., 2015, Competition for order flow with fast and slow traders, *Review of Financial Studies* 28, 2094–2127.

Weill, P. O., 2007, Leaning against the wind, *The Review of Economic Studies* 74, 1329–1354.

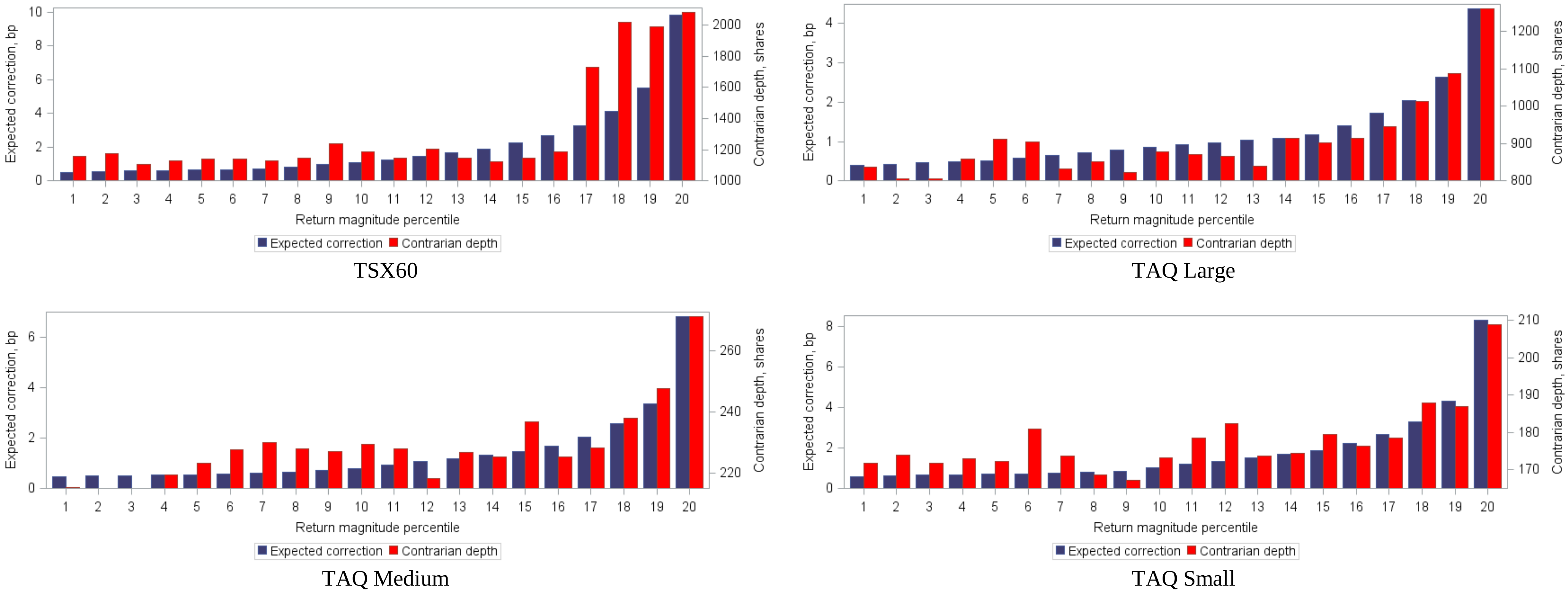
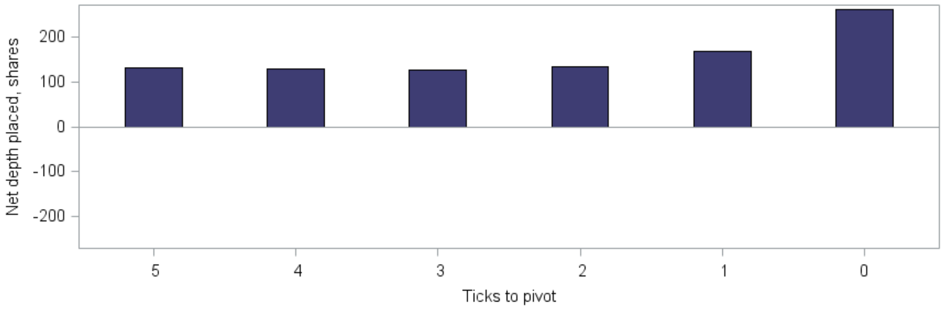
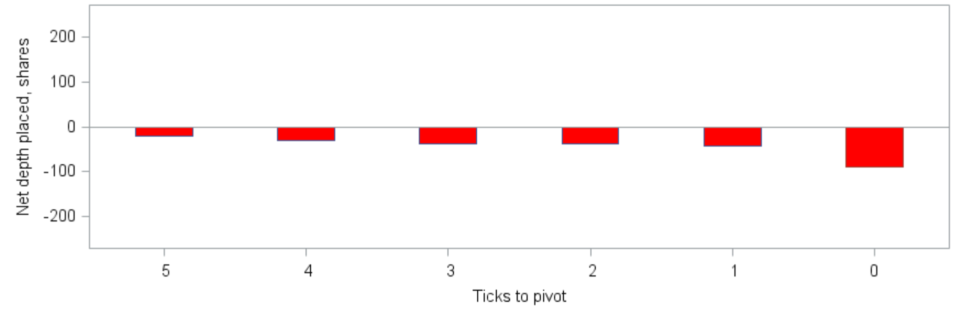


Figure 1.1. Expected price correction and depth

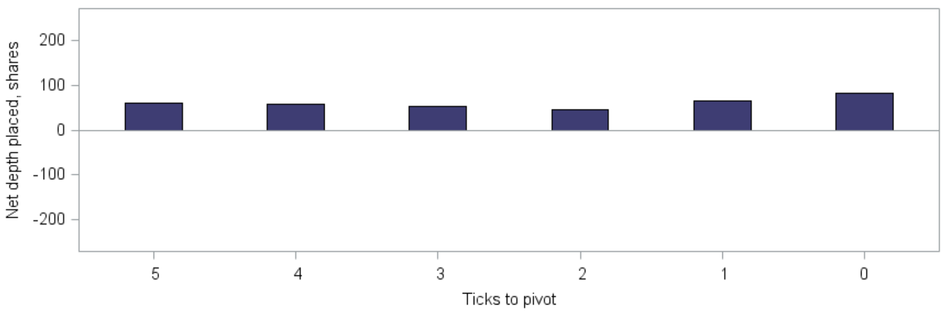
The figure shows the expected price correction and contrarian best quote depth aggregated by return magnitude percentiles. The expected price correction is adjusted for the return direction, so it represents the transitory component of the contemporaneous price. Contrarian depth is the depth of the best quote in the direction of return. The first figure represents the Canadian TSX60 sample, while the second, third and fourth figures represent US large, medium and small stocks, respectively.



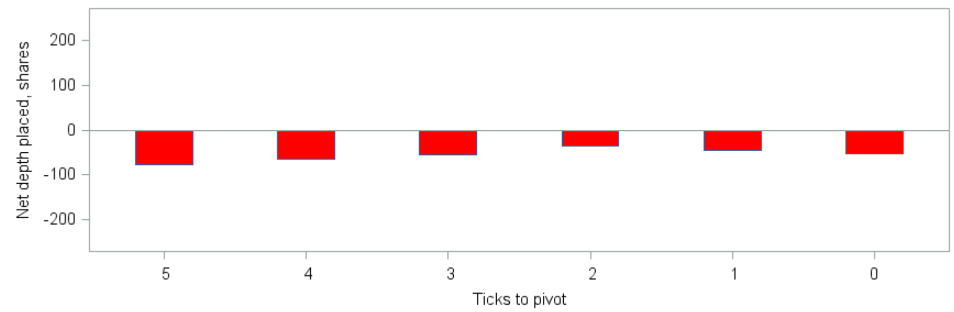
Net depth placed at the future best contrarian pivot quote, non-EPMs



Sum of net depth placed at the limit order book layers better than future best contrarian pivot quote, non-EPMs



Net depth placed at the future best contrarian pivot quote, EPMs



Sum of net depth placed at the limit order book layers better than future best contrarian pivot quote, EPMs

Figure 1.2. Limit order book dynamics during price movements

The figure shows limit order book depth dynamics during five midquote updates before the pivot point. The top plots correspond to intervals with a return magnitude below the 90th percentile, and the bottom plots correspond to the intervals with return above the 90th percentile. The left plots report depth dynamics exactly at the future best contrarian quote, and the right plots report aggregate depth dynamics at all limit order book layers that undercut the future best contrarian quote.

Table 1.1. Intraday summary statistics

The table reports intraday summary statistics for the intraday price movements. $AbsRet$ is the absolute return, $DirVol$, $InDirVol$ is the number of shares traded in the direction and against the direction of return, accordingly. $NetVol$ is the difference between $DirVol$ and $InDirVol$. $Duration$ is the time that the price moves in the same direction. $TicksPerMovement$ is the number of midquote updates in the price movements. Panel A shows the coefficients for the price movements with absolute return below the 90th percentile, and Panel B shows the coefficients for the sample of returns above the 90th percentile.

Panel A: Non-EPMs

TSX60	Mean	Median	Std
$AbsRet, bp$	2.444	1.652	2.427
$DirVol, shares$	308	0	7,295
$InDirVol, shares$	157	0	10,095
$NetVol, shares$	151	0	12,042
$Duration, seconds$	10.159	0.960	42.897
$TicksPerMovement$	1.529	1	0.925
$NPriceMovements$	24,303,780		
TAQ Large	**Mean**	**Median**	**Std**
$AbsRet, bp$	1.370	1.029	1.221
$DirVol, shares$	1,824	0	15,950
$InDirVol, shares$	795	0	10,548
$NetVol, shares$	1,029	0	12,134
$Duration, seconds$	3.469	0.040	22.630
$NTicksPerMovement$	1.572	1	0.895
$NPriceMovements$	28,302,974		
TAQ Medium	**Mean**	**Median**	**Std**
$AbsRet, bp$	1.625	1.160	1.925
$DirVol, shares$	163	0	1,831
$InDirVol, shares$	65	0	1,431
$NetVol, shares$	98	0	1,772
$Duration, seconds$	4.110	0.033	20.134
$TicksPerMovement$	1.534	1	0.980
$NPriceMovements$	25,883,432		
TAQ Small	**Mean**	**Median**	**Std**
$AbsRet, bp$	2.074	1.468	2.480
$DirVol, shares$	100	0	1,607
$InDirVol, shares$	41	0	1,088
$NetVol, shares$	58	0	1,450
$Duration, seconds$	5.369	0.051	35.423
$TicksPerMovement$	1.548	1	0.971
$NPriceMovements$	20,602,427		

Panel B: EPMs			
TSX60	Mean	Median	Std
$AbsRet, bp$	13.057	10.775	9.214
$DirVol, shares$	2,362	300	30,426
$InDirVol, shares$	941	0	25,408
$NetVol, shares$	1,421	200	34,021
$Duration, seconds$	33.549	6.500	117.748
$TicksPerMovement$	3.707	3	2.493
$NPriceMovements$	2,703,496		
TAQ Large	Mean	Median	Std
$AbsRet, bp$	5.255	4.112	3.698
$DirVol, shares$	10,519	2,549	49,768
$InDirVol, shares$	5,047	700	30,699
$NetVol, shares$	5,472	1,400	33,129
$Duration, seconds$	22.242	4.614	73.572
$TicksPerMovement$	5.114	5	2.496
$NPriceMovements$	3,145,945		
TAQ Medium	Mean	Median	Std
$AbsRet, bp$	7.842	5.959	45.318
$DirVol, shares$	1,156	301	6,512
$InDirVol, shares$	424	0	5,873
$NetVol, shares$	732	203	6,680
$Duration, seconds$	23.343	5.060	72.534
$TicksPerMovement$	4.913	4	3.363
$NPriceMovements$	2,877,966		
TAQ Small	Mean	Median	Std
$AbsRet, bp$	9.586	7.110	27.496
$DirVol, shares$	638	100	5,022
$InDirVol, shares$	236	0	2,933
$NetVol, shares$	402	95	3,870
$Duration, seconds$	27.286	4.023	111.422
$TicksPerMovement$	4.632	4	3.247
$NPriceMovements$	2,291,077		

Table 1.2. Probit regressions

The table reports coefficients and marginal effects of a probit model described in Section 5.1. The dependent variable is the probability that the current midquote update is the last one in the sequence of same-directional midquote updates. $AbsCRet$ is the absolute cumulative return accumulated since the beginning of a price movement. $NetVol$ is the net number of shares traded in the direction of return, and $PESpread$ is the percentage effective spread. Panel A shows the coefficients for the price movements with absolute return below 90th percentile and Panel B shows the coefficients for the sample of returns above 90th percentile. P-values are given in parentheses.

Panel A: Non-EPMs

	TSX60	TAQ Large	TAQ Medium	TAQ Small
$Intercept$	0.267	0.015	0.278	0.308
	(0.00)	(0.00)	(0.00)	(0.00)
$AbsCRet, bp$	0.060	0.210	0.054	0.025
	0.022	0.080	0.020	0.009
	(0.00)	(0.00)	(0.00)	(0.00)
$NetVol, 10{,}000\ shares$	-0.002	-0.034	-0.292	-0.340
	-0.001	-0.013	-0.109	-0.127
	(0.00)	(0.00)	(0.00)	(0.00)
$PESpread, bp$	-0.022	-1.709	-0.053	-0.035
	-0.008	-0.653	-0.020	-0.013
	(0.00)	(0.00)	(0.00)	(0.00)
Pseudo R-Squared	0.007	0.020	0.003	0.002

Panel B: EPMs

	TSX60	TAQ Large	TAQ Medium	TAQ Small
$Intercept$	-1.242	-1.287	-0.900	-0.880
	(0.00)	(0.00)	(0.00)	(0.00)
$AbsCRet, bp$	0.073	0.137	0.009	0.009
	0.021	0.033	0.002	0.003
	(0.00)	(0.00)	(0.00)	(0.00)
$NetVol, 10{,}000\ shares$	-0.004	-0.032	-0.071	-0.173
	-0.001	-0.008	-0.020	-0.048
	(0.00)	(0.00)	(0.00)	(0.00)
$PESpread, bp$	-0.175	-4.716	-0.028	-0.068
	-0.050	-1.146	-0.008	-0.019
	(0.00)	(0.00)	(0.00)	(0.00)
Pseudo R-Squared	0.129	0.085	0.008	0.010

Table 1.3. Intraday price movements

The table reports statistics for key variables of interest during price movements with the magnitude above 90th percentile (EPMs) and below 90th percentile (non-EPMs). $RevPriceComponent$ is the deviation of price from the expected underlying efficient price as estimated in the Data and Methodology section. $ContrDepth$ is the best ask(bid) depth if the return is positive(negative) and $NonContrDepth$ is the depth of the opposite best quote. $ContrElast$ is the number of shares traded through marketable buy(sell) orders per one-cent move of best ask(bid) quote during positive (negative) price movement and $NonContrElast$ is estimated accordingly the opposite best quote. $InvAmihud$ is the net volume imbalance divided by return in basis points. **, *** correspond to the statistical significance at 0.05% and 0.01%.

TSX	EPM	non-EPM	difference
$RevPriceComponent, bp$	7.675	1.433	6.242***
$ContrDepth, shares$	2,034	1,238	796***
$NonContrDepth, shares$	1,140	838	303***
$ContrElast, shares$	826	304	523***
$NonContrElast, shares$	402	124	277***
$InvAmihud$, 1000 $shares$	913	450	463***
TAQ Large	EPM	non-EPM	difference
$RevPriceComponent, bp$	3.390	0.889	2.501***
$ContrDepth, shares$	1,216	870	346***
$NonContrDepth, shares$	1,485	1,536	-51***
$ContrElast, shares$	2,654	1,670	984***
$NonContrElast, shares$	1,276	688	588***
$InvAmihud$, 1000 $shares$	7,470	3,547	3,923***
$TAQ\ Medium$	EPM	non-EPM	difference
$RevPriceComponent, bp$	4.809	0.994	3.815***
$ContrDepth, shares$	271	221	50***
$NonContrDepth, shares$	333	319	14***
$ContrElast, shares$	353	175	178***
$NonContrElast, shares$	133	63	70***
$InvAmihud$, 1000 $shares$	832	364	468***
$TAQ\ Small$	EPM	non-EPM	difference
$RevPriceComponent, bp$	5.894	1.276	4.618***
$ContrDepth, shares$	203	178	25***
$NonContrDepth, shares$	239	240	-1**
$ContrElast, shares$	178	106	72***
$NonContrElast, shares$	69	38	30***
$InvAmihud$, 1000 $shares$	315	129	186***

Table 1.4. Dynamics of limit order book depth

The table reports net depth placement dynamics around the future best pivot quotes during five ticks before the price reaches the pivot point. The best contrarian quote at pivot is the best ask (bid) quote at the turning point of the positive (negative) price movement. The best non-contrarian quote at pivot is the best bid (ask) quote at the turning point of the positive (negative) price movement. Limit order book layers that are better or worse than the future best pivot quotes are identified accordingly. The net depth is estimated as non-marketable share volume placed minus cancelled at the given layer of the limit order book. Panel A shows the coefficients for the price movements with absolute return below 90th percentile, and Panel B shows the coefficients for the sample of returns above the 90th percentile. P-values are given in parentheses.

Panel A: non-EPMs

ticks to pivot	best contrarian quote at pivot	1 cent better	2 cents better	3 cents better	4 cents better	5 cents better	6 cents better	7 cents better	8 cents better	9 cents better	10 cents better	10+ cents better
0	261.6	-75.9	-8.0	-2.7	-0.2	-0.1	0.1	-3.3	0.0	0.0	0.1	0.3
	(0.00)	(0.00)	(0.00)	(0.00)	(0.09)	(0.20)	(0.27)	(0.00)	(0.77)	(0.15)	(0.39)	(0.00)
1	168.0	-22.7	-8.8	-6.5	-2.0	-0.4	-0.6	-2.3	-0.1	-0.1	0.0	0.1
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.03)	(0.90)	(0.20)
2	133.4	-17.4	-8.2	-7.0	-2.2	-0.5	-0.6	-2.4	0.0	-0.1	-0.1	0.2
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.63)	(0.14)	(0.03)	(0.23)
3	127.3	-14.1	-11.9	-7.4	-1.9	-0.4	-0.5	-2.1	0.5	0.0	0.0	0.9
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.05)	(0.01)	(0.00)	(0.04)	(0.71)	(0.72)	(0.00)
4	128.4	-9.4	-11.7	-8.3	-1.6	-0.2	0.7	-2.0	0.2	0.0	0.1	2.2
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.53)	(0.12)	(0.00)	(0.54)	(0.96)	(0.78)	(0.00)
5	132.1	-8.9	-6.3	-8.2	-2.1	-0.4	-0.1	-2.8	0.9	0.5	0.5	5.1
	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)	(0.27)	(0.92)	(0.00)	(0.00)	(0.10)	(0.14)	(0.00)
		1 cent worse	2 cents worse	3 cents worse	4 cents worse	5 cents worse	6 cents worse	7 cents worse	8 cents worse	9 cents worse	10 cents worse	10+ cents worse
0		34.3	10.3	0.3	0.3	-1.8	-4.8	-5.9	-2.9	-2.8	-3.5	20.6
		(0.00)	(0.00)	(0.13)	(0.13)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
1		25.3	2.4	-2.7	-1.2	-2.6	-2.8	-4.4	-3.9	-2.9	-3.5	35.2
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
2		17.7	0.7	-2.9	-2.4	-2.5	-1.9	-4.5	-3.5	-2.9	-2.7	45.4
		(0.00)	(0.11)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
3		18.4	-0.2	-2.6	-2.2	-2.2	-2.5	-4.1	-3.5	-3.4	-2.9	54.5
		(0.00)	(0.76)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
4		17.4	0.2	-3.7	-2.6	-1.7	-2.2	-4.4	-2.9	-2.8	-1.7	50.3
		(0.00)	(0.80)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.08)	(0.00)
5		19.4	-0.5	-3.4	-2.1	-2.2	-2.4	-4.5	-3.7	-3.1	0.9	55.2
		(0.00)	(0.68)	(0.00)	(0.03)	(0.04)	(0.01)	(0.00)	(0.00)	(0.00)	(0.68)	(0.00)

ticks to pivot	best non-contrarian quote at pivot	1 cent better	2 cents better	3 cents better	4 cents better	5 cents better	6 cents better	7 cents better	8 cents better	9 cents better	10 cents better	10+ cents better
0	348.5	-27.4	-5.4	-3.6	-1.3	-0.1	-0.5	-1.0	-0.2	-0.1	0.0	-0.5
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.12)	(0.00)	(0.00)	(0.00)	(0.00)	(0.89)	(0.00)
1	184.0	-26.1	-7.6	-4.7	-2.0	-0.5	-1.0	-1.6	-0.4	-0.3	0.1	-1.0
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.66)	(0.00)
2	105.4	-25.5	-8.8	-5.7	-2.0	-0.6	-1.0	-1.5	-0.5	-0.3	-0.2	-1.0
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)
3	75.2	-23.9	-8.2	-5.3	-1.9	-0.8	-0.7	-1.4	-0.5	0.2	-0.1	-0.9
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.42)	(0.43)	(0.00)
4	58.7	-21.7	-7.5	-5.8	-2.5	-0.7	-0.6	-1.7	-0.5	0.1	0.3	-1.8
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.33)	(0.00)	(0.00)	(0.75)	(0.19)	(0.00)
5	44.3	-16.6	-6.0	-5.1	-2.6	-1.0	-1.3	0.2	-0.8	-0.1	0.0	-2.7
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.78)	(0.00)	(0.87)	(0.88)	(0.00)
		1 cent worse	2 cents worse	3 cents worse	4 cents worse	5 cents worse	6 cents worse	7 cents worse	8 cents worse	9 cents worse	10 cents worse	10+ cents worse
0		33.6	-3.2	-5.6	2.5	3.5	1.7	-1.6	1.2	1.7	-1.6	83.6
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
1		130.0	52.2	3.5	2.7	1.8	0.3	-2.3	-2.2	0.4	-1.5	67.8
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.21)	(0.00)	(0.00)	(0.58)	(0.00)	(0.00)
2		82.8	116.6	58.6	8.1	2.8	0.0	-1.5	-2.6	-2.9	-2.1	56.5
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.96)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
3		25.2	96.4	125.1	47.0	6.3	-0.1	-1.9	-2.4	-2.5	-3.5	46.5
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.80)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
4		-6.8	44.6	109.9	90.2	44.1	1.5	-1.9	-1.7	0.0	-4.1	30.6
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.07)	(0.02)	(0.02)	(0.94)	(0.00)	(0.00)
5		-21.0	9.0	60.1	68.4	91.9	52.9	-2.2	-5.1	0.1	-0.2	26.8
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.07)	(0.00)	(0.90)	(0.80)	(0.00)

Panel B: EPMs												
ticks to pivot	best contrarian quote at pivot	1 cent better	2 cents better	3 cents better	4 cents better	5 cents better	6 cents better	7 cents better	8 cents better	9 cents better	10 cents better	10+ cents better
0	83.0	4.2	1.0	-11.4	-5.6	-2.6	-5.0	-1.2	-3.4	-2.7	-2.2	-24.5
	(0.00)	(0.40)	(0.77)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
1	65.2	-24.1	17.2	-3.3	-5.5	-1.3	-3.3	-2.2	-2.1	-1.8	-1.4	-17.9
	(0.00)	(0.00)	(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
2	45.9	-8.3	-11.8	1.7	1.0	-1.0	-1.9	-2.5	-1.7	-0.9	0.6	-11.8
	(0.00)	(0.00)	(0.00)	(0.42)	(0.47)	(0.00)	(0.04)	(0.00)	(0.00)	(0.00)	(0.61)	(0.00)
3	52.6	-2.3	-24.8	-12.8	1.2	0.2	-2.5	-3.1	-1.3	-0.7	-1.7	-8.5
	(0.00)	(0.20)	(0.00)	(0.00)	(0.56)	(0.70)	(0.00)	(0.00)	(0.00)	(0.00)	(0.35)	(0.00)
4	58.6	3.6	-18.6	-26.9	-6.6	0.7	-2.8	-2.9	-0.9	-0.1	-3.8	-6.3
	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.33)	(0.00)	(0.00)	(0.00)	(0.62)	(0.19)	(0.00)
5	59.4	5.5	-15.8	-28.2	-19.5	-1.2	-5.1	-3.5	-0.5	0.2	-2.2	-6.0
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.08)	(0.00)	(0.00)	(0.25)	(0.52)	(0.29)	(0.00)
		1 cent worse	2 cents worse	3 cents worse	4 cents worse	5 cents worse	6 cents worse	7 cents worse	8 cents worse	9 cents worse	10 cents worse	10+ cents worse
0		9.9	33.5	5.9	12.9	5.5	-1.5	1.6	0.4	-0.8	-1.7	50.6
		(0.05)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.40)	(0.04)	(0.00)	(0.00)
1		16.4	31.5	2.8	3.8	0.9	0.9	-2.3	-2.1	-1.7	-3.0	61.7
		(0.00)	(0.00)	(0.01)	(0.00)	(0.18)	(0.12)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
2		22.2	17.1	1.5	-1.4	-0.2	0.2	-3.7	-3.8	-1.8	-3.6	62.8
		(0.00)	(0.00)	(0.15)	(0.24)	(0.75)	(0.76)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
3		20.0	7.4	-2.0	-0.5	-0.4	-0.6	-3.7	-3.4	-2.3	-2.7	71.7
		(0.00)	(0.00)	(0.02)	(0.67)	(0.54)	(0.41)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
4		20.4	2.8	-3.8	-3.7	0.1	-1.7	-3.4	-3.8	-0.7	-1.6	73.9
		(0.00)	(0.06)	(0.00)	(0.00)	(0.88)	(0.01)	(0.00)	(0.00)	(0.26)	(0.01)	(0.00)
5		20.5	2.9	-3.0	-4.0	0.7	-2.1	-3.7	-3.4	-1.6	-1.1	74.1
		(0.00)	(0.16)	(0.11)	(0.00)	(0.45)	(0.01)	(0.00)	(0.00)	(0.04)	(0.16)	(0.00)

ticks to pivot	best non-contrarian quote at pivot	1 cent better	2 cents better	3 cents better	4 cents better	5 cents better	6 cents better	7 cents better	8 cents better	9 cents better	10 cents better	10+ cents better
0	270.1	12.3	5.8	-1.1	-0.8	-0.3	-0.6	-9.1	-0.7	-0.6	-0.6	-10.5
	(0.00)	(0.00)	(0.04)	(0.19)	(0.44)	(0.01)	(0.18)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
1	126.4	2.5	3.9	-0.8	-0.1	-0.5	-1.4	-4.9	-1.0	-0.8	-0.6	-9.5
	(0.00)	(0.27)	(0.07)	(0.28)	(0.86)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
2	38.7	-3.7	-0.8	-1.2	-0.1	-0.5	-1.3	-1.7	-0.7	-0.5	-1.7	-5.6
	(0.00)	(0.01)	(0.67)	(0.13)	(0.95)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.16)	(0.00)
3	13.7	-8.5	-4.0	-2.6	-0.9	-0.6	-0.7	1.8	-0.2	-0.4	-2.2	-3.9
	(0.00)	(0.00)	(0.00)	(0.00)	(0.13)	(0.00)	(0.00)	(0.00)	(0.47)	(0.00)	(0.22)	(0.00)
4	7.6	-14.6	-3.6	-3.1	-0.7	-0.6	-0.6	7.5	-0.1	-0.3	-0.2	-2.9
	(0.00)	(0.00)	(0.00)	(0.00)	(0.36)	(0.00)	(0.00)	(0.00)	(0.84)	(0.00)	(0.03)	(0.00)
5	10.2	-15.4	-2.4	-3.2	-1.2	-0.8	-0.7	12.1	-0.4	-0.3	-0.1	-2.5
	(0.00)	(0.00)	(0.01)	(0.00)	(0.02)	(0.03)	(0.00)	(0.00)	(0.00)	(0.01)	(0.41)	(0.00)
		1 cent worse	2 cents worse	3 cents worse	4 cents worse	5 cents worse	6 cents worse	7 cents worse	8 cents worse	9 cents worse	10 cents worse	10+ cents worse
0		133.8	-4.0	-35.2	-2.9	-8.5	-6.0	-14.7	-9.6	-1.4	-7.2	-33.7
		(0.00)	(0.42)	(0.00)	(0.16)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
1		148.0	180.6	12.6	-13.5	-3.9	-5.1	-8.0	-8.9	-4.0	-3.7	-23.3
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
2		52.3	165.0	96.0	17.2	5.0	-0.3	-2.9	-5.1	-7.7	-2.4	-9.6
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.75)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
3		0.4	116.0	148.3	72.7	21.2	5.1	2.2	-2.0	-7.9	-4.7	-9.2
		(0.88)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
4		-31.3	51.0	146.7	125.7	52.2	19.9	9.6	3.8	-5.8	-5.6	-3.2
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.34)
5		-42.7	14.6	100.9	137.8	83.0	40.7	18.3	9.9	-2.2	-5.1	-1.3
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.05)	(0.00)	(0.78)

Table 1.5. Profits summary statistics

The table reports summary statistics for the estimated profits from liquidity provision and daily pricing errors. $MM\_ret\_spr\_rev$ is the liquidity providers' profit component corresponding to spread and reversals. $MM\_ret\_position$ is the liquidity providers' profit component corresponding to end-of-day position profits. MM_ret_total is the sum of the two components. All profit components are estimated in dollars and scaled by the number of reversals per day. $pricing\_error\_ratio$ is the proportion of intraday price variance attributable to the pricing error.

TSX60	Mean	Median	Std
MM_ret_total	632.4	230.1	2186.4
$MM\_ret\_spr\_rev$	438.1	153.0	1536.4
$MM\_ret\_position$	194.4	30.2	1170.8
$pricing\_error\_ratio$	2.45%	0.99%	5.46%
TAQ Large	Mean	Median	Std
MM_ret_total	65.9	15.6	801.3
$MM\_ret\_spr\_rev$	69.1	15.1	559.8
$MM\_ret\_position$	-4.1	0.0	361.5
$pricing\_error\_ratio$	1.48%	0.88%	2.25%
TAQ Medium	Mean	Median	Std
MM_ret_total	2.7	1.0	61.3
$MM\_ret\_spr\_rev$	3.1	1.1	44.9
$MM\_ret\_position$	-0.5	0.0	21.6
$pricing\_error\_ratio$	1.20%	0.63%	2.66%
TAQ Small	Mean	Median	Std
MM_ret_total	3.3	0.6	54.6
$MM\_ret\_spr\_rev$	3.3	0.7	39.9
$MM\_ret\_position$	-0.5	0.0	18.9
$pricing\_error\_ratio$	2.26%	0.80%	7.87%

Table 1.6. Profits regression

The table reports regression coefficients of the following model:

$$Depvar_{it} = \beta_1 PricingErrorRatio_{it} + \beta_2 PESpread_{it} + \beta_3 nTrades_{it} + \beta_4 DailyReturn_{it} + \varepsilon_{it},$$

where the dependent variable is the total profit of liquidity provision, profits from reversals and spread and profits from accumulated end of day position. $PricingErrorRatio$ is the proportion of intraday price variance attributable to the pricing error, $PESpread$ is the average percentage effective spread, $nTrades$ is the number of trades per day and $DailyReturn$ is the daily return magnitude. Regressions are estimated with fixed effects. All variables are standardized. P-values corresponding to double clustered standard errors are reported in parentheses.

Panel A: Dependent – Total profits of limit order trading

	TSX60	TAQ Large	TAQ Medium	TAQ Small
$PricingErrorRatio$	0.0422	0.0955	0.1566	0.1619
	(0.00)	(0.00)	(0.08)	(0.00)
$PESpread$	-0.0816	-0.0106	0.0108	-0.0218
	(0.01)	(0.60)	(0.68)	(0.46)
$nTrades$	0.0754	0.0381	0.0472	0.0666
	(0.00)	(0.10)	(0.19)	(0.17)
$DailyReturn$	0.0593	-0.0397	-0.1092	-0.0735
	(0.01)	(0.01)	(0.08)	(0.02)
$Adj. R^2$	0.01	0.01	0.02	0.03

Panel B: Dependent – Profits from spread and reversals

	TSX60	TAQ Large	TAQ Medium	TAQ Small
$PricingErrorRatio$	0.0316	0.1528	0.0935	0.1113
	(0.01)	(0.00)	(0.02)	(0.00)
$PESpread$	-0.0786	-0.0339	-0.0113	-0.0582
	(0.00)	(0.11)	(0.63)	(0.03)
$nTrades$	0.0676	0.0739	0.0723	0.1179
	(0.00)	(0.00)	(0.09)	(0.02)
$DailyReturn$	0.0555	-0.0473	-0.0484	-0.0688
	(0.01)	(0.00)	(0.08)	(0.01)
$Adj. R^2$	0.01	0.03	0.01	0.03

Panel C: Dependent – Profits from end-of-day position

	TSX60	TAQ Large	TAQ Medium	TAQ Small
$PricingErrorRatio$	0.0533	-0.0269	0.0405	0.0643
	(0.00)	(0.13)	(0.66)	(0.00)
$PESpread$	-0.0725	0.0298	0.0424	0.0290
	(0.01)	(0.13)	(0.08)	(0.28)
$nTrades$	0.0607	-0.0304	-0.0156	-0.0389
	(0.00)	(0.23)	(0.48)	(0.36)
$DailyReturn$	0.0592	-0.0146	-0.1487	-0.0442
	(0.01)	(0.16)	(0.07)	(0.03)
$Adj. R^2$	0.01	0.01	0.01	0.01

Chapter 2. HIGH FREQUENCY TRADING AND EXTREME PRICE MOVEMENTS

2.1. Introduction

In modern markets, high frequency traders (HFTs) play an important role in providing liquidity (Hasbrouck and Saar, 2013; Menkveld, 2013; Malinova, Park and Riordan, 2014, Conrad, Wahal and Xiang, 2015). Generally, the rise of HFT has been accompanied by a reduction in trading costs (Angel, Harris and Spatt, 2011; Jones, 2013; Harris, 2013) and an increase in price efficiency (Carrion, 2013; Brogaard, Hendershott and Riordan, 2014; Chaboud, Chiquoine, Hjalmarsson and Vega, 2014). Nevertheless, liquidity provision by HFTs is endogenous as they are typically not obligated to stabilize markets in periods of stress. A growing literature finds that endogenous liquidity providers (ELPs) often withdraw from the market during such periods (Raman, Robe and Yadav, 2014; Bongaerts and Van Achter, 2015; Cespa and Vives, 2015; Korajczyk and Murphy, 2015; Anand and Venkataraman, 2016). The focus of this study is HFT behavior during stressful conditions.

We define stressful periods as unexpected and rapidly developing extreme price movements (EPMs) that belong to the 99.9^{th} percentile of the return distribution. While a growing body of work examines HFT activity during normal conditions, less attention has been given to periods of market stress such as EPMs. Our main finding is that, on average, HFTs trade in the opposite direction of EPMs and supply liquidity to non-high frequency traders (nHFTs) by absorbing their trade imbalances. This result holds even during the largest EPMs and during the times when nHFTs demand substantial amounts of liquidity. Notably, HFTs supply liquidity both to the EPMs that eventually reverse and the EPMs that result in permanent price changes. This means that an average HFT trade during extreme price movements provides liquidity to aggressive, occasionally informed, nHFTs.

Even though EPMs occur quickly, they consist of multiple sequential trades. If HFT algorithms are designed to stop providing liquidity during EPMs, technology would allow them to withdraw limit orders as EPMs develop. Yet the results imply that the algorithms are designed to remain in the market, likely because doing so is profitable. Although revenue estimates are noisy, we find evidence that the revenues are greater on days when EPMs occur. Despite the enhanced revenue potential, the data show that HFTs do not cause EPMs. Our results complement those of Bessembinder, Carrion, Tuttle and Venkataraman (2016), who show that liquidity provision increases around large uninformed predictable trades. In our setting EPMs are generally unpredictable and are occasionally informed, yet the incentive to provide liquidity remains. Our findings expand the understanding of resiliency of modern markets in stressful times.

Although HFTs stabilize prices during an average EPM, we find clear limits to HFT liquidity provision. HFT liquidity supply is outstripped by their liquidity demand when more than one stock simultaneously undergoes an EPM (we refer to these instances as co-EPMs). We show that during such periods, HFTs accumulate substantial position risk, which likely triggers risk controls, particularly for their liquidity supplying strategies. Focusing on one exceptionally large co-EPM, the 2010 Flash Crash, Kirilenko, Kyle, Samadi and Tuzun (2016) find that HFTs withdrew from liquidity provision. Reflecting on the Crash, the regulators have expressed concern that incentives to provide liquidity are deficient during market-wide periods of stress (CFTC-SEC, 2011). Our findings generalize these results and deepen our understanding of market-wide liquidity shortages and offer evidence in support of the regulators' view.

Theory suggests that ELPs may choose several ways of reacting to order imbalances. Traders described by Grossman and Miller (1988) choose to supply liquidity during order imbalances. On the contrary, the predatory traders of Brunnermeier and Pedersen (2005) opt to demand liquidity.

The back-runners of Yang and Zhu (2015) supply liquidity until they recognize an institutional trading pattern and then switch to demanding liquidity. In our setting, HFT behavior during an average EPM is more consistent with that described by Grossman and Miller (1988), although the data point to net HFT liquidity demand during co-EPMs and occasional back-running for long EPM sequences.

2.2. Data, EPM detection and summary statistics

2.2.1. HFT data

The HFT data come from NASDAQ and span two years: 2008 and 2009. These data have been previously used by Carrion (2013), Brogaard, Hendershott and Riordan (2014), and O'Hara, Yao and Ye (2014), among others. For each trade the dataset contains an indicator for whether an HFT or an nHFT participates on the liquidity-supplying or the liquidity-demanding side of the trade. When preparing the data NASDAQ identified 26 firms that act as independent HFT proprietary trading firms based on its knowledge of the firm's activity. A firm is identified by NASDAQ as an HFT if it trades frequently, holds small intraday inventory positions, and ends the day with a near zero inventory. HFTs on NASDAQ have no obligation to stabilize prices during stressful times (Bessembinder, Hao and Lemmon, 2011; Clark-Joseph, Ye and Zi, 2016) and so are ideal participants to study liquidity provision by ELPs.

The data allow us to directly observe HFT liquidity provision and demand. We are subject to the same limitations as the abovementioned studies, mainly that we cannot observe individual HFT activity and that we only observe trading on NASDAQ. Although trades on NASDAQ make up 30-40% of all trading activity in the sample stocks, it is a possible that during EPMs HFTs provide liquidity on NASDAQ while taking it from the other markets. We are unable to refute this possibility. Nonetheless, we believe that such liquidity transfer is unlikely as liquidity provision on NASDAQ is not systematically more attractive than it is on other venues during the sample period.

2.2.2. EPM identification

We identify EPMs as extreme changes in the National Best Bid and Offer (NBBO) midquotes. The use of midquotes instead of trade prices allows us to reduce the effect of the bid-ask bounce. In untabulated results we find similar effects when using trade prices. We obtain the midquotes from the NYSE Trade and Quote database (TAQ) after adjusting the data according to the recommendations of Holden and Jacobsen (2014). Specifically, we (i) interpolate the times of trades and the times of NBBO quotes within a second, (ii) adjust for withdrawn quotes, (iii) delete locked and crossed NBBO quotes, and (iv) delete trades reported while the NBBO is locked or crossed. To avoid focusing on price dislocations that may be caused by market opening and closing procedures, we only consider trading activity between 9:35 a.m. and 3:55 p.m.

Using the filtered TAQ midquotes, we compute 10-second absolute midquote returns. The choice of the 10-second sampling frequency is based on two offsetting considerations. On the one hand, detecting EPMs that result from brief liquidity dislocations requires a relatively short sampling interval. On the other hand, a sampling interval that is too short may split an EPM into several price changes that are not large enough to be captured by the identification procedure. The choice of 10-second intervals is a compromise between these two considerations. As a robustness check, we repeat the main analyses for several alternative interval lengths: 1 second, 5 seconds, 30 seconds and 1 minute. The results are qualitatively similar.

The NASDAQ HFT dataset contains 120 stocks divided into three size categories: large, medium and small. There are 40 stocks in each category. Medium and small stocks trade rather infrequently, and there are usually insufficient observations to draw statistically robust conclusions about HFT and nHFT activity. The main analysis therefore focuses on the 40 largest stocks. In a similar application, and driven by similar considerations, Andersen, Bollerslev, Diebold and Ebens (2001) also focus on the largest stocks when detecting EPMs. The sample of 40 largest stocks contains 45.2 million 10-second intervals.

We use three approaches to identify EPMs. The first approach is straightforward and simply labels all intervals that belong to the 99.9th percentile of 10-second absolute midpoint returns for each stock as EPMs. The second approach is more sophisticated and accounts for predictable return correlations in time and across firms. First, for each day we estimate a short-term market model of the following form:

$$Ret_{it} = p_{t-1}SPY_{t-1} + \cdots + p_{t-10}SPY_{t-10} + q_{t-1}Ret_{it-1} + \cdots + q_{t-10}Ret_{it-10} + \varepsilon_{it}, \quad (1)$$

where Ret_{it} is stock i's return over the ten-second interval t, and SPY_t is the return on the S&P 500 ETF (SPRD). Second, we use the coefficients from the previous day's regressions to compute residuals of the current day's model. Third, we label all intervals with residuals that belong to the 99.9th percentile as EPMs. As a robustness check, we use in-sample residuals, with the model estimated over the full sample. The results are similar.

Both approaches select intervals with the largest absolute returns out of 45.2 million 10-second intervals, and define them as EPMs. The intuitive nature of these techniques is appealing, yet they come with two limitations. First, the 99.9 cutoffs are stock-specific and therefore implicitly assume that each stock is equally likely to undergo an EPM. Consequently, the 99.9 technique may (over-) under-sample stocks that are (less) more prone to EPMs. The second limitation is that the techniques (especially the first one) are agnostic to volatility conditions and therefore tend to oversample periods of high volatility. We suggest that understanding HFT behavior is relevant regardless of accompanying volatility. Nevertheless, to formally address this limitation, we repeat the analysis using a third EPM detection technique, the Lee and Mykland's (2012) methodology, which accounts for contemporaneous (local) volatility.

Throughout the main manuscript, we use the results obtained from the second identification technique where EPMs are based on the residuals from Equation (1). A summary of results obtained using the first and the third techniques is reported in the robustness section. The results obtained from these techniques are in line with those reported in the paper.

Finally, in unreported results, we find that the 99.9th percentile returns closely correspond to the 99.9th percentile of trade imbalances. An EPM identification that focuses on the largest imbalances rather than the largest returns produces a similar sample.

2.2.3. Summary statistics

Table 2.1 reports the descriptive statistics for the sample of 45,200 EPMs in Panel A and, for comparison, the full sample of 10-second intervals in Panel B. The statistics expectedly show that returns, trading activity, and bid-ask spreads are considerably larger during the EPMs than during an average 10-second period. The average absolute EPM return is 0.478%, which is more than 17 times (or about 10 standard deviations) larger than the full-sample return. Trading activity is also substantially higher; increasing from 18 trades per 10 seconds to 72 trades. Dollar trading volume increases from $76,076 to $462,950, and share volume increases by a similar magnitude. Finally, the quoted and relative spreads nearly double during EPMs suggesting that liquidity is impaired during these events.

The number of positive EPMs is approximately equal to the number of negative EPMs. In unreported results, we find that EPM characteristics such as the absolute return magnitude, trading volume, and quoted spreads are similar for positive and negative EPMs. HFT and nHFT behavior is also similar across these different types of events. The results reported in the remainder of the paper report combined positive and negative EPMs.

Figures 2.1 and 2.2 report the EPM time series. In both figures, the scale of the vertical axis is logarithmic. Figure 2.1 reports the intraday frequency of EPMs, with 50.3% of the events occurring in the first hour of trading. This pattern is consistent with studies that document high price volatility and information uncertainly in the morning hours (Chan, Christie and Schultz, 1995; Egginton, 2014). The remaining EPMs are distributed relatively evenly throughout the day, with a moderate increase near the end of the day.[10] Figure 2.2 plots the daily frequency of EPMs during the 2008-2009 sample period. Most EPMs (66.3%) occur during the months of September, October and November of 2008, the height of the financial crisis.

2.3. HFT and nHFT activity around EPMs

In this section, we show that HFTs provide liquidity to nHFTs during a typical EPM, even when the EPM is very large and even when the price change is permanent. We also show that HFT liquidity supply is overshadowed by their demand when several stocks undergo simultaneous EPMs and also during long sequences of EPMs. We also show that liquidity provision during an average EPM is profitable, yet we find no evidence that HFTs trigger EPMs to benefit from this profitability.

[10] Aitken, Cumming and Zhan (2015) find that proliferation of HFT has reduced instances of end-of-day price manipulation.

2.3.1. A typical EPM

To measure HFT activity during EPMs, we use directional trade imbalances computed as the difference between trading activity in the direction of the EPM and trading activity in the opposite direction: $HFT^{D} = HFT^{D+} - HFT^{D-}$ and $HFT^{S} = HFT^{S+} - HFT^{S-}$, where HFT^{D} is HFT liquidity demand, HFT^{S} is HFT liquidity supply, and the superscripts + (-) indicate activity in the same (opposite) direction of the EPM return. For example, if HFTs demand 20 shares of liquidity in the direction of the price movement and demand 1 share in the opposite direction, HFT^{D} is +19. Similarly, if HFTs supply 20 shares of liquidity against the direction of the EPM and supply 4 shares in the direction of the EPM, HFT^{S} is -16. We compute similar metrics for nHFTs.

In addition, we introduce two net imbalance metrics, HFT^{NET} ($nHFT^{NET}$) computed as the sum of HFT^{D} and HFT^{S} ($nHFT^{D}$ and $nHFT^{S}$). Since liquidity is typically provided against the direction of return, $(n)HFT^{S}$ usually has a negative value, and the sum of $(n)HFT^{D}$ and $(n)HFT^{S}$ is in effect the difference between liquidity demanding and liquidity providing volume. Net imbalances indicate the direction in which net trading activity by a particular trader type is occurring relative to the EPM direction. For example, a positive (negative) net HFT imbalance indicates overall trading in the direction (opposite) of the EPM.

To begin the discussion of HFT and nHFT activity around EPMs Figure 2.3 reports the cumulative return (CRET) as well as HFT^{D}, $nHFT^{D}$ and HFT^{NET} starting 100 seconds prior to an average EPM and up to 100 seconds afterwards. We make the following expositional choices. First, the figure includes both positive and negative EPMs, and we invert the statistics for the latter. Second, we benchmark the signs for HFT and nHFT activity to the EPM return. For example, if the EPM return is positive, a negative HFT^{D} ten seconds after the EPM means that HFTs sell the stock via liquidity demanding orders, effectively counteracting the effects of the positive EPM that occurred ten seconds earlier.

Figure 2.3 shows that prices change significantly during the EPM interval, and then revert somewhat during the remaining 100 seconds (10 intervals).[11] There is a large increase in $nHFT^{D}$ during the EPM, with a share imbalance of more than 5,300. In the meantime, HFT^{D} is about 2,300 shares. More importantly, HFT^{NET} is negative, indicating that HFT liquidity supply offsets HFT liquidity demand and that HFTs absorb volume imbalances created by nHFTs.[12]

The results in Figure 2.3 provide first evidence on HFT and nHFT behavior around EPMs. In Table 2.2, we examine EPM event windows in more detail. Specifically, we focus on event windows that span 20 seconds before and after the EPM interval and report liquidity demand and supply statistics for HFTs and nHFTs. We find that HFT^{NET} is statistically significant in the direction opposite of returns during interval t (the EPM interval) and the two following intervals. Statistical significance is preserved when we cluster the standard errors in time. Further, upon splitting HFT activity into demand and supply, we observe that HFTs trade in the direction of the EPM with their liquidity demanding trades (HFT^{D} is 2,296 shares) and in the opposite direction with their liquidity supplying trades (HFT^{S} is 2,539 shares). HFTs provide 243 shares of net liquidity against the direction of an average EPM. This finding is contrary to the belief held by some market observers that HFTs trade large amounts in the direction of EPMs.

Is 243 shares too small a quantity to claim that HFTs stabilize prices? The results in Table 2.2 are simple averages and therefore do not suggest that HFT liquidity provision is limited to 243 shares per EPM. Rather, 243 is the number of shares that nHFTs demand during an average EPM.

[11] Our reliance on quote midpoints aims to focus the analysis on the permanent component of the security price. This said, some non-permanent components remain. Specifically in Figure 2.3, the return slopes downward up to t=0, then jumps, and then partly reverses after the EPM. This process is best described as a combination of random walk and stationary noise processes.

[12] The net imbalance metrics are designed so that HFT^{NET}=-$nHFT^{NET}$.

Beyond being liquidity providers during EPMs, do HFTs trigger EPMs? In the 10-second interval starting 20 seconds prior to an EPM (t-20), HFT^{NET} and $nHFT^{NET}$ do not show any directionality. However, in t-10 HFTs trade against the direction of the future EPM return.[13] As such, it appears that HFTs do not trigger EPMs. We examine this issue in more detail in a subsequent section.

Following an EPM, HFTs continue to trade in the opposite direction of the EPM return, but unlike in interval t they primarily use liquidity demanding trades. Specifically, HFTs demand a net of 113 shares against the direction of the preceding EPM return in interval t+10. This suggests that HFTs may speed up the reversal process. We study reversals in more detail in the following section.

2.3.2. EPM types: reversals and permanent price changes

The literature suggests that large price movements can be triggered by at least two types of events: information arrival and trade imbalances. A news arrival, for instance, often results in prices adjusting rapidly to incorporate information. In an efficient market, such price movements will be permanent. Alternatively, trade imbalances usually arise because impatient traders submit large volumes of buy or sell orders and push prices away from the fundamental values. Price movements arising from such pressures are transitory and are followed by reversals. Figure 2.4 presents an illustration.

Do HFTs provide liquidity to both EPM types? To answer this question, we divide the sample into transitory and permanent EPMs. The former are characterized by significant, yet temporary, price changes followed by reversals. We identify these as EPMs that revert by more than 2/3 by the end of a 30-minute period. The latter, permanent, EPMs do not revert by more than 1/3

[13] In Table 2.2, as in Figure 2.3, we benchmark the signs of HFT and nHFT volume to the EPM return.

by the end of this period. To allow for a clean separation of the two EPM types, we exclude the EPMs that revert by more than 1/3 and less than 2/3; these are 14.2% of the sample. The results are robust to using alternative reversal thresholds (e.g., reversals of more than 1/2 of the EPM return), time thresholds of 1, 10 and 20 minutes, and allowing reversals to occur by the end of the trading day.

In Table 2.3, we examine the characteristics of the two EPM types and HFT activity around them. Despite a significant difference in post-EPM price patterns, other EPM characteristics (i.e., returns, trading activity, HFT participation and spreads) are similar across the two types (Panel A). For instance, the average absolute return is 0.481% during both a typical transitory and a typical permanent EPM. In Panel B, we describe HFT activity around the two EPM types. Consistent with the full sample results, HFTs provide liquidity to both types during interval t.

2.3.3. EPM magnitude

Although the EPMs in the sample represent the 99.9^{th} percentile of all price movements, the setup may obscure the picture for the largest EPMs, during which HFT activity may differ from what has been discussed so far. Kirilenko, Kyle, Samadi and Tuzun (2016) show that when prices reached extraordinary lows during the 2010 flash crash, HFTs withdrew from liquidity provision. So far, the results suggest that EPMs are not accompanied by similar withdrawals. But what about the largest EPMs? In Table 2.4, we examine if HFT liquidity provision varies in EPM magnitude, and particularly if HFTs provide liquidity to the largest of the extreme price movements.

Table 2.4 reports summary statistics and HFT^{NET} results for EPMs divided into four magnitude quartiles, from the relatively small (Q1) to the largest (Q4). As expected, trading volume and spreads increase in return magnitude (Panel A). HFT liquidity provision also increases, going from 116 shares in Q1 to 546 shares in Q4 (Panel B). Insofar as these results are generalizable to events like the 2010 Flash Crash, they suggest that it was probably not the magnitude of the crash that triggered HFT withdrawal.

2.3.4. EPM types: standalone and co-EPMs

The 2010 Flash Crash was characterized not only by the magnitude of price movements, but also by the large number of stocks that were affected. It is possible that liquidity withdrawals during the crash were due to the HFT firms' risk controls that were triggered when accumulated inventories reached high levels. The Flash Crash was a uniquely large and rare event, and it is not clear if it should be viewed as typical of HFT behavior in instances of multi-stock price movements. To examine this issue, we define co-EPMs as those that occur in two or more stocks during the same 10-second time interval and repeat earlier analyses.

Panel A of Table 2.5 reports that co-EPMs comprise 57% of the sample. The prevalence of co-EPMs should not be surprising given the exceptionally high EPM occurrence during the 2008 financial crisis when prices of multiple assets experienced large simultaneous movements (Figure 2.2). An average co-EPM includes 3.5 stocks. The average return is 0.487% during a standalone EPM and 0.471% during a co-EPM. Trading activity metrics are noticeably different between the two types, with dollar volume during the standalone EPMs being about 74% higher than that during the co-EPMs.

Panel B shows that HFTs supply 1,296 shares of net liquidity during the standalone EPMs. In the meantime, they demand 549 shares of net liquidity during the co-EPMs. This reversal in HFT behavior is striking. In Figure 2.5, we examine its evolution. To do so, we plot HFT^D and HFT^S for standalone and co-EPMs. As previously, the metrics are computed on a per-stock basis. HFT^D and

HFT^S decline rapidly when more than one stock undergoes simultaneous EPMs. Notably, the decline is more pronounced for supply than demand. As such, the risk controls triggered during co-EPMs appear to affect liquidity supplying strategies more than they affect liquidity demanding strategies, giving rise to positive HFT^{NET} during co-EPMs.

Note that even though HFT activity per stock declines, total inventory accumulated during co-EPMs may increase. For instance, inventory accumulated during the 10-stock co-EPM is 7,960 (= 796 $\times$ 10) shares, more than 6 times the inventory accumulated during the standalone EPMs. As such, even though HFTs reduce activity on the per-share basis, the risk of their total positions may increase. This risk may be further exacerbated if, in addition to the co-EPM stocks, other stocks in the HFT portfolios are experiencing large price movements. Such movements, even if they do not qualify as EPMs, may affect total HFT position risk. We examine this issue next.

2.3.5. Co-EPMs and position risk

To gain a better understanding of the risks assumed by HFTs during co-EPMs, we turn to the concept of value at risk (VaR). We caution that our data do not contain capital positions or inventories of individual HFTs, so we are unable to estimate the true VaR. Rather, we follow the general intuition of VaR analyses and refer to the results as quasi-VaR (QVaR). Specifically, we rely on the non-parametric method of Allen, Bali and Tang (2012) and begin by estimating, on a daily basis, the 99th percentile of the 10-second absolute returns for the portfolio of sample stocks. Note that constituent returns vary during instances of tail portfolio returns. To account for this, we estimate average stock returns for each sample stock i during the instances of portfolio tail returns on each day d, Ret_{id}^{tail}. The contribution of individual stocks to the portfolio tail return varies slowly. As such, the composition of portfolio tail returns on day $d-1$ is a sufficient proxy for the expected composition on day d. With this in mind, we compute intraday QVaR as follows:

$$QVaR_t = -\min\left(\sum_{i}^{40} Ret_{id-1}^{tail} \times DINV_{it}, \sum_{i}^{40} -Ret_{id-1}^{tail} \times DINV_{it}\right) \quad (2)$$

where $DINV_{it}$ is the dollar inventory in stock i accumulated by the HFTs during the interval t valued at the last midquote of the interval. The first term captures potential portfolio losses if the EPM is followed by a positive tail return, and the second term captures potential losses if the EPM is followed by a negative tail return. We then select the minimum of the two terms to estimate the maximum loss. In a nutshell, QVaR estimates the expected dollar loss during the following 10-second interval if the HFT portfolio experiences an unfavorable 99th percentile return.

Figure 2.6 shows that QVaR increases steadily in the number of stocks experiencing a co-EPM. Specifically, it increases from \$287 during intervals without EPMs to \$936 for standalone EPMs, and further to over \$5,000 for intervals with more than 10 EPMs. This increase is driven by the inventory accumulation in both the stocks undergoing EPMs and in the rest of the portfolio and likely explains the HFTs' tendency to reduce risk exposure on a per-stock basis.

2.3.6. EPM sequences

Earlier results show that HFTs provide substantial liquidity to the standalone EPMs, yet demand liquidity during co-EPMs. In Panel A of Table 2.6, which serves as a companion to Figure 2.5, we examine HFT sensitivity to the number of stocks in a co-EPM. The sensitivity is high; HFT^{NET} switches from being negative during the standalone EPMs to zero for the two-stock co-EPMs and to being positive for co-EPMs that involve three and more stocks. The results suggest

HFT liquidity supply is sensitive to inventory risk. This is consistent with Amihud and Mendelson (1980) and Comerton-Forde et al. (2010), who show that market maker strategies depend on inventories.

Given this fragility, it is possible that HFTs also remain on the sidelines on days with long sequences of EPMs, especially if these EPMs have the same return direction. In Panel B of Table 2.6, we examine if this occurs. The data show that HFTs usually provide net liquidity to the first four EPMs in the sequence and reduce net liquidity provision to zero if the sequence continues. There is some evidence of positive HFT^{NET} for very long sequences.

2.3.7. Does HFT activity during EPMs differ from their usual behavior?

Research shows that HFTs usually demand liquidity in the direction of returns (e.g., Brogaard, Hendershott and Riordan, 2014). If this pattern persisted during EPMs, we would observe significantly positive and large HFT^{NET}. On the contrary, we find that the pattern reverses for standalone EPMs. Although the pattern does not reverse for co-EPMs, it is possible that the positive HFT-return relation is reduced even for these EPMs. Accounting for return magnitude, HFTs may demand less liquidity during the times when multiple stocks undergo EPMs than they normally would. To examine this issue, we turn to the following multivariate setting:

$$HFT^{NET}{}_{it} = \alpha + \beta_1 1_{EPM_{it}} + \beta_2 Ret_{it} + \beta_3 Vol_{it} + \beta_4 Spr_{it} + \boldsymbol{Lags_{kit-\sigma}\gamma_{k\sigma}} + \varepsilon_{it}, \quad (3)$$

where HFT^{NET} is the difference between HFT^{D} and HFT^{S} as discussed earlier; $1_{EPM\ it}$ is a dummy variable equal to one if the 10-second interval t in stock i is identified as an EPM and is equal to zero otherwise, Ret_{it} is the absolute return, Vol_{it} is the traded share volume, Spr_{it} is the percentage quoted spread, and $\boldsymbol{Lags_{kit-\sigma}}$ is a vector of lags for the dependent and each of the independent variables, with $\sigma \in \{1, 2, \ldots, 10\}$. The variables in the vector are indexed with a subscript k. All variables are standardized at the stock level.

Because the coefficients on the 1_{EPM} dummy are related to returns, they should be interpreted jointly with those on the Ret variable. For example in column 1 of Table 2.7, the estimated coefficient on the Ret variable confirms that HFTs usually demand liquidity in the direction of return. In the meantime, the 1_{EPM} dummy shows that HFTs reduce liquidity demand during EPMs, with the incremental effect of -0.798 standard deviations. Having established the basic result, we next turn to HFT activity during the previously identified EPM types. Column 2 shows that during both transitory and permanent EPMs the normally positive HFT-return relation is significantly reduced. In column 3, we find the same result for the standalone and co-EPMs, yet the decline is much greater for the standalone EPMs. Similar results emerge in column 4 that accounts for EPM magnitude; the normally positive relation between HFT behavior and returns is reduced, more so during the largest EPMs. Overall, even in cases when they demand liquidity during the EPM episodes (the co-EPM case), HFTs demand considerably less than they normally would.

2.3.8. HFT-return relation within the 10-second intervals

The 10-second event windows are quite long given the speed of modern trading and may conceal nefarious aspects of HFT behavior. Yang and Zhu (2015) propose and van Kervel and Menkveld (2015) show that HFTs are able to recognize trading patterns after a period of time and switch from supplying liquidity to demanding it. Although van Kervel and Menkveld (2015) focus on time horizons that are much longer than ours, even one second is a long enough time for HFT algorithms to re-evaluate a trading strategy. It is therefore possible that HFTs supply liquidity at the beginning of EPMs yet exacerbate their tail ends.

To examine this possibility, in Figure 2.7 we plot second by second cumulative returns, HFT, and nHFT activity centered on the largest one-second return during an average EPM. The figure shows that prices continue to move in the direction of the largest return for several seconds afterwards. If HFT algorithms had been designed to quickly switch from liquidity supply to demand after observing large price changes, they would have had sufficient time to do so. The figure contains no evidence of HFT^{NET} switching to positive values. If anything, it remains slightly negative.[14]

Although the time aggregation period in Figure 2.7 is finer than that used in the remainder of the study, it is still long relative to the usual timing of HFT interactions. It is therefore possible that HFT^{NET} is positive at the very beginning of some EPMs, perhaps for a few micro- or milliseconds. Because pinpointing the exact time when an EPM begins is next to impossible, we are unable to examine this issue. This said, even if short-lived HFT liquidity demand exists at early EPM stages, the economic effect of such demand is economically small and does not register in the data.

2.3.9. Profitability of liquidity provision during EPMs

The data show that HFTs usually provide liquidity to nHFTs during both transitory and permanent EPMs. Since HFTs choose to do so, liquidity provision should be profitable. How are these profits derived? During positive permanent EPMs as described in Figure 2.4, if a trader limits liquidity provision to the size of his existing long inventory, he will have bought low and sold high. If however he provides liquidity indiscriminately, in the amount larger than the existing long inventory, he may accumulate a money-losing short position. The same logic, but in reverse, applies to negative permanent EPMs.

[14] As in Figure 2.3, non-permanent return components are evident, suggesting lagged adjustment in prices. In addition to generating momentum after t=0, thus adjustment may generate smoothing prior to t=0.

During transitory price movements, when the price first moves up and then down (Figure 2.4), a skilled trader may profit by initially selling high to the impatient buyers and then buying low when the price reverses. The literature shows that providing liquidity during such reversals is profitable (Hendershott and Seasholes, 2007; Nagel, 2012; So and Wang, 2014). This strategy does not require pre-existing inventory as profits are derived from the inventory accumulated during the EPM. In summary, it is possible that HFTs profit from both permanent and transitory EPMs. Next, we examine the data for evidence of such profits.

Specifically, we estimate HFT trading revenues on EPM days and compare them to the days without EPMs. We follow the approach used by Sofianos (1995), Menkveld (2013), and Brogaard, Hendershott and Riordan (2014) and assume that for each sample stock and each day HFTs start and end the day with zero inventory, and that all inventory accumulated by the end of the day is sold at the closing midpoint. We compute HFT revenue for each stock and each day as:

$$\pi HFT = -\sum_{n=1}^{N} HFT_n \times I_n \times P_n + invHFT_N \times P_N, \tag{4}$$

where HFT_n is the number of shares traded by HFTs during the n^{th} transaction, I is the indicator equal to 1 for buy trades and -1 for sell trades, P_n is the trade price, $invHFT_N$ is the inventory accumulated by HFTs before the end of the day, and P_N is the end of day midquote. Following Brogaard, Hendershott and Riordan (2014), we adjust transaction prices by the taker fee of \$0.00295 and the maker rebate of \$0.0028, although the results are robust to other levels of maker-taker fees and to omitting the fees. The first term of Equation 4 represents cash flows to HFTs throughout the day, and the second term assigns a value to the end-of-day inventory.

To assess the impact of EPMs on daily HFT revenues, we estimate the following panel regression for each stock i on day t:

$$\pi HFT_{it} = \alpha + \beta nEPM_{it} + \varepsilon_{it}, \tag{5}$$

where $nEPM$ captures the number of EPMs. The results are simple, and we report them here rather than in a separate table. The α estimate suggests than HFTs capture $3,834 in revenue per stock on an average day; whereas the β estimate indicates that the revenue becomes $219 greater with each EPM. As such, HFT activities during EPMs are potentially profitable. In addition to the general case in Equation 5, we compute the β estimates for all EPM breakdowns (i.e., permanent, transitory, standalone, co-EPMs, and for four magnitude quartiles). Due to the noisiness of profit calculations, the β estimates for the breakdowns are statistically insignificant. The revenue results should therefore be interpreted with caution. A conservative interpretation would suggest that there is no evidence of HFT losses on average due to EPMs and some evidence of profits.

2.3.10. HFT activity and future EPMs

Other research has suggested that HFTs trigger EPMs. Golub, Keane and Poon (2013) report that mini-crashes in individual stocks have increased in recent years and suggest a link between these crashes and HFT. Leal, Napoletano, Roventini and Fagiolo (2014) model a market in which HFTs play a fundamental role in generating flash crashes. To shed light on this issue, we use probit regressions to model the probability of an EPM as a function of lagged values of HFT^{NET}, return, volume and spread:

$$Prob\ (EPM = 1)_{it} = \alpha + \beta_1 HFT_{it-1}^{NET} + \beta_2 Ret_{it-1} + \beta_3 Vol_{it-1} + \beta_4 Spr_{it-1} + \varepsilon_{it}, \tag{6}$$

where all variables are as previously defined and are lagged by one interval.

The results are in Table 2.8 and show no evidence of HFT being associated with a higher probability of future EPMs. On the contrary, HFT is associated with a lower EPM probability. For instance in column 1, the marginal effect of the HFT^{NET} variable implies that the probability of an EPM decreases by 0.6% of the unconditional probability with every standard deviation increase in the pre-EPM HFT^{NET}.

2.4. Robustness

2.4.1. Alternative EPM identification techniques

Earlier, we discuss two alternative methods of EPM identification. The first method identifies EPMs as the 99.9^{th} percentile of raw returns, and the second method uses the Lee and Mykland (2012) methodology. In Table 2.9, we report a brief summary of results arising from these two methodologies. The results are qualitatively similar to those obtained from the main sample.

2.4.2. Alternative return distributions

Figure 2.1 points to a significant intraday pattern in the number of EPMs, which is consistent with the well-known phenomenon whereby returns are large in the early morning and then level off. It therefore may be useful to check if conditioning the EPM definition on return distributions that are allowed to vary intraday affects our conclusions. Put differently, an early morning return may look extreme with respect to the afternoon returns, but unremarkable with respect to a distribution of price changes in the first half-hour of trading.

To examine this issue, we split the sample into seven intervals: 9:35-10:00, 10:00-11:00, 11:00-12:00, 12:00-13:00, 13:00-14:00, 14:00-15:00 and 15:00-15:55. We then define EPMs as the

99.9th percentile of Equation 1 residuals in each interval. This approach produces a more even distribution of EPMs throughout the day than that in Figure 2.1. We then examine HFT behavior for these newly defined EPMs. The results are in Table 2.10 and support those obtained for the original sample; HFTs provide liquidity during an average EPM. We obtain similar results when we use two instead of seven intervals (i.e., 9:35-10:00 and 10:00-15:55).

2.5. Conclusion

We provide novel evidence on the stability of liquidity supply by high frequency traders (HFTs), a dominant subset of liquidity providers in modern markets. HFTs are endogenous liquidity providers (ELPs) and do not have the obligation to supply liquidity during stressful times. We show that HFTs are net suppliers of liquidity to non-HFTs (nHFTs) during extreme price movements (EPMs). HFTs supply liquidity even during the most extreme EPMs and the EPMs that result in permanent price changes.

However, HFT liquidity supply is sensitive to multiple EPMs, as HFTs on average switch to demanding liquidity when multiple stocks simultaneously undergo EPMs and when EPMs persist throughout the day. The switch is due to the liquidity supplying strategies being more risk averse than the liquidity demanding strategies. During episodes of multiple simultaneous EPMs, position risk accumulated by HFTs is significantly higher than normal, likely leading to the reduction in their activity, particularly on the supply side. We find some evidence of HFTs' earning positive revenues on days with EPMs. Despite this, the results show that HFTs do not appear to cause EPMs.

While beyond the scope of this paper, more research can help generalize or qualify the findings. For instance, it will be important to know whether the practice of HFT evolved in a way whereby what was true in the late 2000s is no longer the case. Also, it is important to understand how changes in market structure, such as the introduction of limit-up limit-down trading rules or the arrival of a new venue that provides protection to liquidity providers impacted ELPs.

2.6. References

Aitken, M., Cumming, D. and Zhan F., 2015, High frequency trading and end-of-day price dislocation, *Journal of Banking and Finance* 59, 330-349.

Allen, L., Bali, T. and Tang, Y., 2012, Does systemic risk in the financial sector predict future economic downturns? *Review of Financial Studies* 25, 3000-3036.

Amihud, Y. and Mendelson, H., 1980. Dealership market: Market-making with inventory, *Journal of Financial Economics* 8, 31-53.

Anand, A. and Venkataraman, K., 2016, Market conditions, fragility, and the economics of market making, *Journal of Financial Economics* 121, 327-349.

Andersen, T. G., Bollerslev, T., Diebold, F. and Ebens, H., 2001, The distribution of realized stock return volatility, *Journal of Financial Economics* 61, 43–76.

Angel, J., Harris, L. and Spatt, C., 2011, Equity trading in the 21st century, *Quarterly Journal of Finance* 1, 1-53.

Bessembinder, H., Carrion, A., Tuttle, L. and Venkataraman, K., 2016, Liquidity, resiliency and market quality around predictable trades: Theory and evidence, *Journal of Financial Economics* 121, 142-166.

Bessembinder, H., Hao, J. and Lemmon, M., 2011, Why designate market makers? Affirmative obligations and market quality, Working paper.

Bongaerts, D. and Van Achter, M., 2015, High-frequency trading and market stability, Working paper.

Brogaard, J. A., Hendershott, T. and Riordan, R., 2014, High-frequency trading and price discovery, *Review of Financial Studies* 27, 2267-2306.

Brunnermeier, M and Pedersen, L., 2005, Predatory trading, *Journal of Finance* 60, 1825-1863.

Carrion, A., 2013, Very fast money: High-frequency trading on the NASDAQ, *Journal of Financial Markets* 16, 680-711.

Cespa, J. and Vives, X., 2015, The welfare impact of high frequency trading, Working paper.

CFTC-SEC, 2011, Recommendations regarding regulatory responses to the market events of May 6, 2010.

Chaboud, A., Chiquoine, B., Hjalmarsson, E. and Vega, C., 2014, Rise of machines: Algorithmic trading in the foreign exchange market, *Journal of Finance* 2045-2084.

Chan, K., Christie, W. and Schultz, P., 1995, Market structure and the intraday pattern of bid-ask spreads for NASDAQ securities, *Journal of Business* 68, 35-60.

Clark-Joseph, A., Ye, M. and Zi, C., 2016, Designated market makers still matter: Evidence from two natural experiments, *Journal of Financial Economics*, forthcoming.

Comerton-Forde, C., Hendershott, T., Jones, C., Moulton, P. and Seasholes, M., 2010, Time variation in liquidity: The role of market-maker inventories and revenues, *Journal of Finance* 65, 295-331.

Conrad, J., Wahal, S. and Xiang, J., 2015, High-frequency quoting, trading, and the efficiency of prices, *Journal of Financial Economics* 116, 271-291.

Egginton, J., 2014, The declining role of NASDAQ market makers, *Financial Review* 49, 461-480.

Golub, A., Keane, J. and Poon, S., 2013, High-frequency trading and mini flash crashes, Working paper.

Grossman, S. and Miller, M., 1988, Liquidity and market structure, *Journal of Finance* 43, 617-633.

Harris, L., 2013, What to do about high-frequency trading, *Financial Analysts Journal* 69, 6-9.

Hasbrouck, J. and Saar, G., 2013, Low-latency trading, *Journal of Financial Markets* 16, 646-679.

Hendershott, T. and Seasholes, M., 2007, Market maker inventories and stock prices, *American Economic Review* 97, 210-214.

Holden, C. and Jacobsen, S., 2014, Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions, *Journal of Finance* 69, 1747-1785.

Jones, C., 2013, What do we know about high-frequency trading? Working paper.

Kirilenko, A., Kyle, A., Samadi, M. and Tuzun, T., 2016, The flash crash: The impact of high-frequency trading on an electronic market, *Journal of Finance*, forthcoming.

Korajczyk, R. and Murphy, D., 2015, High frequency market making to large institutional trades, Working paper.

Leal, S., Napoletano, M., Roventini, A. and Fagiolo G., 2014, Rock around the clock: An agent-based model of low- and high-frequency trading, Working paper.

Lee, S. and Mykland, P., 2012, Jumps in equilibrium prices and market microstructure noise, *Journal of Econometrics* 168, 396–406.

Malinova, K., Park, A. and Riordan, R., 2014, Do retail traders suffer from high-frequency traders? Working paper.

Menkveld, A., 2013, High-frequency trading and the new-market makers, *Journal of Financial Markets* 16, 712-740.

Nagel, S., 2012, Evaporating liquidity, *Review of Financial Studies* 25, 2005-2039.

O’Hara, M., Yao, C. and Ye, M., 2014, What’s not there: Odd-lots and market data, *Journal of Finance* 69, 2199-2236.

Raman, V., Robe, M. and Yadav, P., 2014, Electronic market makers, trader anonymity and market fragility, Working paper.

So, E. and Wang, S., 2014, News-driven return reversals: Liquidity provision ahead of earnings announcements, *Journal of Financial Economics* 114, 20-35.

Sofianos, G., 1995, Specialist gross trading revenues at the New York Stock Exchange, Working paper.

Menkveld, A. and van Kervel, V., 2015, High-frequency trading around large institutional orders, Working paper.

Yang, L. and Zhu, H., 2015, Back-running: Seeking and hiding fundamental information in order flows, Working paper.

Table 2.1. Summary statistics

The table reports summary statistics for the sample of extreme price movements (EPMs) (Panel A) and for the full sample of 10-second intervals (Panel B). $Absolute\ Return$ is the absolute value of the 10-second midpoint return. $Total\ (HFT)\ Trades$ is the number of (HFT) trades during the interval. $Dollar\ Volume$ and $Share\ Volume$ are the total dollar and share volume traded during the interval. $Quoted\ Spread$ and $Relative\ Spread$ are quoted and relative quoted NBBO spreads, respectively in dollars and percentage points. All statistics are averaged over the 10-second sampling intervals.

Panel A: Extreme price movements

	Mean	**Median**	**Std. Dev.**
Absolute Return, %	0.478	0.436	0.188
Total Trades	72.19	42.00	88.33
Total HFT Trades	57.29	32.00	72.89
Dollar Volume	462,950	166,929	998,832
Share Volume	15,361	5,300	31,778
Quoted Spread, $	0.044	0.015	0.138
Relative Spread, %	0.076	0.063	0.154
N	45,200		

Panel B: Full sample

Absolute Return, %	0.028	0.009	0.048
Total Trades	18.1	11.0	18.7
Total HFT Trades	15.8	10.0	15.4
Dollar Volume	76,076	11,701	230,661
Share Volume	1,987	292	6,045
Quoted Spread, $	0.026	0.010	0.057
Relative Spread, %	0.046	0.041	0.032
N	45.2 M		

Table 2.2. Liquidity supply and demand around EPMs

The table reports directional trading volume around extreme price movements (EPMs). Time interval t is the 10-second EPM interval. In addition, we report the results for the two time intervals preceding the EPM and two subsequent time intervals. HFT^{D} ($nHFT^{D}$) is the difference in liquidity-demanding HFT (nHFT) volume in the direction of the EPM and liquidity-demanding volume against the direction of the EPM. HFT^{S} ($nHFT^{S}$) is the difference in liquidity-providing volume against the direction of the EPM and liquidity-providing volume in the direction of the EPM. HFT^{NET} ($nHFT^{NET}$) is the difference between HFT^{D} and HFT^{S} ($nHFT^{D}$ and $nHFT^{S}$). *p*-values are in parentheses. *** and ** indicate statistical significance at the 1% and 5% levels.

	t-20	**t-10**	**t**	**t+10**	**t+20**
HFT^{NET}	-20.2	-73.9***	-242.7***	-112.8***	-33.7
	(0.32)	(0.00)	(0.00)	(0.00)	(0.11)
HFT^{D}	-76.5***	-143.5***	2296.3***	-273.6***	-63.1***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
HFT^{S}	56.3***	69.6***	-2539.0***	160.8***	29.4
	(0.00)	(0.00)	(0.00)	(0.00)	(0.12)
$nHFT^{NET}$	20.2	73.9***	242.7***	112.8***	33.7
	(0.32)	(0.00)	(0.00)	(0.00)	(0.11)
$nHFT^{D}$	-64.6*	-63.6	5369.0***	613.6***	296.0***
	(0.05)	(0.13)	(0.00)	(0.00)	(0.00)
$nHFT^{S}$	84.8**	137.5***	-5126.4***	-500.8***	-262.3***
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)

Table 2.3. Transitory and permanent EPMs

The table reports summary statistics for transitory and permanent EPMs. Transitory EPMs revert by more than 2/3 of the EPM return in the following 30 minutes. Permanent EPMs do not revert by more than 1/3 in the same interval. Because we exclude EPMs that revert by the amount between 1/3 and 2/3, the total number of EPMs in this table is 85.8% of that reported in Panel A of Table 1. Panel B reports HFT^{NET} around the two EPM types.

Panel A: Summary statistics

	transitory		permanent	
	mean	**std. dev.**	**mean**	**std. dev.**
Absolute Return, %	0.481	0.188	0.481	0.187
Total Trades	70.90	87.79	69.78	85.64
Total HFT Trades	55.97	71.35	55.54	71.90
Dollar Volume	456,326	1,022,813	434,572	947,261
Share Volume	14,576	29,516	14,470	29,250
Quoted Spread, $	0.047	0.147	0.046	0.140
Relative Spr., %	0.079	0.144	0.080	0.157
N	17,915		20,848	

Panel B: HFT^{NET}

	t-20	**t-10**	**t**	**t+10**	**t+20**
transitory	-72.0**	-144.2***	-363.0***	-121.6***	-92.5***
permanent	35.5	-3.0	-303.5***	-110.8***	12.0

Table 2.4. EPM magnitude quartiles

Panel A divides EPMs into quartiles by return magnitude, from smallest to largest. Panel B contains HFT^{NET} statistics.

Panel A: Summary statistics

	Q1 (small)		Q2	
	mean	**std. dev.**	**mean**	**std. dev.**
Absolute Return, %	0.385	0.094	0.415	0.102
Total Trades	59.96	67.79	63.77	72.67
Total HFT Trades	48.29	57.05	51.09	60.33
Dollar Volume	365,702	764,091	390,139	819,044
Share Volume	12,191	24,783	12,947	25,700
Quoted Spread, $	0.040	0.114	0.041	0.118
Relative Spr., %	0.071	0.090	0.073	0.095
N	11,280		11,320	
	Q3		**Q4 (large)**	
Absolute Return, %	0.465	0.116	0.645	0.261
Total Trades	70.69	81.24	94.28	118.37
Total HFT Trades	56.08	67.08	73.67	97.28
Dollar Volume	455,307	977,999	640,282	1,316,028
Share Volume	14,806	29,864	21,488	42,633
Quoted Spread, $	0.044	0.119	0.051	0.188
Relative Spr., %	0.077	0.107	0.082	0.258
N	11,280		11,320	

Panel B: HFT^{NET}

	t-20	**t-10**	**t**	**t+10**	**t+20**
Q1	-4.9	-106.2**	-115.7**	-102.2**	-6.9
Q2	-8.5	-90.3**	-60.3	-108.4***	-7.1
Q3	-70.9**	-65.5	-248.7***	-89.8**	-44.7
Q4	3.4	-33.6	-545.5***	-150.6***	-76.0

Table 2.5. Standalone and co-EPMs

Panel A divides EPMs into standalone and co-EPMs, with the latter group capturing EPMs that occur simultaneously in several stocks. Panel B contains HFT^{NET} statistics.

Panel A: Summary statistics

	standalone		co-EPMs	
	mean	**std. dev.**	**mean**	**std. dev.**
Absolute Return, %	0.487	0.198	0.471	0.181
Total Trades	88.34	106.18	60.04	69.64
Total HFT Trades	68.13	87.24	49.15	58.58
Dollar Volume	611,337	1,237,578	351,352	753,084
Share Volume	21,109	40,462	11,038	22,238
Quoted Spread, $	0.049	0.124	0.040	0.148
Relative Spr., %	0.084	0.147	0.069	0.159
# stocks			3.5	2.78
N	19,402		25,798	

Panel B: HFT^{NET}

	t-20	**t-10**	**t**	**t+10**	**t+20**
standalone	25.5	-31.4	-1295.7***	-101.2**	-37.1
co-EPMs	-54.5***	-105.8***	549.3***	-121.5***	-31.1

Table 2.6. Standalone and Co-EPMs, EPM sequences

The table reports HFT^{NET} for standalone and co-EPMs (Panel A) and for EPM sequences (Panel B). EPM sequences are strings of same-directional EPMs during the trading day, with column (4) identifying the position of a particular EPM in the sequence. p-values are in parentheses. Asterisks *** and ** indicate statistical significance at the 1% and 5% levels.

Panel A: Standalone and co-EPMs			**Panel B: EPM sequences**		
	HFT^{NET}	**# obs.**		**HFT^{NET}**	**# obs.**
(1)	(2)	(3)	(4)	(5)	(6)
1	-1,295***	19,402	1st	-717***	10,221
2	-55	7,326	2nd	-526***	5,679
3	389***	4,353	3rd	-419***	3,931
4	542***	2,980	4th	-315**	2,982
5	288***	2,210	5th	-209	2,379
6	698***	1,602	6th	-229	2,001
7	630***	1,274	7th	316	1,710
8	973***	888	8th	-21	1,483
9	1,411***	891	9th	91	1,303
10	796***	690	10th	69	1,145
11+	1,684***	3,584	11th+	177***	12,366

Table 2.7. Net HFT activity and EPMs

The table reports estimated coefficients from the following regression:

$$HFT^{NET}{}_{it} = \alpha_i + \beta_1 1_{EPM_{it}} + \beta_2 Ret_{it} + \beta_3 Vol_{it} + \beta_4 Spr + \boldsymbol{Lags_{kit-\sigma}\gamma_{k\sigma}} + \varepsilon_{it},$$

where HFT^{NET} is the difference between HFT^{D} and HFT^{S}; the dummy 1_{EPM} is equal to one if a 10-second interval t is identified to contain an EPM and is equal to zero otherwise; $1_{EPM\text{-}TRANSITORY}$ and $1_{EPM\text{-}PERMANENT}$ are dummies that capture the two EPM types; $1_{EPM\text{-}STANDALONE}$ captures the standalone EPMs; $1_{CO\text{-}EPM}$ captures EPMs that occur simultaneously in two or more sample stocks; $1_{EPM\text{-}Q1}$ through $1_{EPM\text{-}Q4}$ identify four EPM quartiles by magnitude, from the smallest to the largest; Ret is the absolute return; Vol is the total trading volume; Spr is the percentage quoted spread; and $\boldsymbol{Lags_{kit-\sigma}}$ is a vector of σ lags of the dependent variable and each of the independent variables, with $\sigma \in \{1, 2, \dots, 10\}$ and the variables indexed with a subscript k. All non-dummy variables are standardized on the stock level. Regressions are estimated with stock fixed effects. p-Values associated with the double-clustered standard errors are in parentheses. *** and ** denote statistical significance at the 1% and 5% levels.

	(1)	(2)	(3)	(4)
1_{EPM}	-0.798***			
	(0.00)			
$1_{EPM\text{-}TRANSITORY}$		-0.783***		
		(0.00)		
$1_{EPM\text{-}PERMANENT}$		-0.816***		
		(0.00)		
$1_{EPM\text{-}STANDALONE}$			-1.437***	
			(0.00)	
$1_{CO\text{-}EPM}$			-0.305***	
			(0.00)	
$1_{EPM\text{-}Q1}$				-0.487***
				(0.00)
$1_{EPM\text{-}Q2}$				-0.561***
				(0.00)
$1_{EPM\text{-}Q3}$				-0.799***
				(0.00)
$1_{EPM\text{-}Q4}$				-1.397***
				(0.00)
Ret	0.080***	0.080***	0.080***	0.081***
	(0.00)	(0.00)	(0.00)	(0.00)
Vol	0.065***	0.065***	0.066***	0.065***
	(0.00)	(0.00)	(0.00)	(0.00)
Spr	-0.009***	-0.009***	-0.009***	-0.009***
	(0.00)	(0.00)	(0.00)	(0.00)
Adj. R^2	0.01	0.01	0.01	0.01

Table 2.8. EPM determinants

The table reports the coefficients and the marginal effects from a probit model of EPM occurrence:

$$Prob\ (EPM = 1)_{it} = \alpha + \beta_1 HFT_{it-1}^{NET} + \beta_2 Ret_{it-1} + \beta_3 Vol_{it-1} + \beta_4 Spr_{it-1} + \varepsilon_{it},$$

where the dependent variable is equal to one if an interval t contains an extreme price movement and zero otherwise. All independent variables are lagged by one interval. HFT^{NET} is the share volume traded in the direction of the price movement minus the share volume traded against the direction of the price movement for all HFT trades, Ret is the absolute return, Vol is total traded volume, Spr is the percentage quoted spread. All variables are standardized on the stock level. The marginal effects are scaled by a factor of 1,000. p-Values are in parentheses. *** and ** indicate statistical significance at the 1% and 5% levels.

	All	**Standalone**	**Co-EPMs**	**Permanent**	**Transitory**
	(1)	(2)	(3)	(4)	(5)
Intercept	-3.237***	-3.446***	-3.382***	-3.430***	-3.469***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
HFT^{NET}_{t-1}	-0.002***	-0.002**	-0.004***	-0.001	-0.005***
Marginal Effect	-0.006	-0.003	-0.008	-0.002	-0.006
	(0.00)	(0.03)	(0.00)	(0.23)	(0.00)
Controls	Yes	Yes	Yes	Yes	Yes
Pseudo-R^2	0.15	0.12	0.14	0.13	0.12

Table 2.9. EPMs defined using alternative methodologies

Panel A reports HFT^{NET} for the sample of EPMs defined using the 99.9th percentile of raw returns. Panel B reports HFT^{NET} for the sample defined using the Lee and Mykland (2012) methodology.

Panel A: 99th percentile of raw returns

	HFT^{NET}
all	-299.3***
transitory	-457.6***
permanent	-323.2***
Q1	-110.8*
Q2	-145.5***
Q3	-293.7***
Q4	-655.5***
standalone	-1296.9***
co-EPMs	446.4***

Panel B: Lee-Mykland (2012)

	HFT^{NET}
all	-892.6***
transitory	-973.2***
permanent	-1063.9***
Q1	-648.8***
Q2	-748.6***
Q3	-859.2***
Q4	-1312.5***
standalone	-1811.9***
co-EPMs	850.7***

Table 2.10. EPMs defined intraday

The table examines HFT behavior around EPMs defined for different time-of-the-day distributional cutoffs. First, we split the day into the following seven intervals: 9:35-10:00, 10:00-11:00, 11:00-12:00, 12:00-13:00, 13:00-14:00, 14:00-15:00 and 15:00-15:55. For each interval, we select returns above the 99.9th percentile. This definition allows EPMs to be evenly distributed within an average day. For the newly defined EPMs, we report the average HFT^{NET} statistics as in the previous tables. We also report the statistics for the 9:35-10:00 and 10:00-15:55 intervals. Asterisks *** and ** indicate statistical significance at the 1% and 5% levels.

	t-20	**t-10**	**t**	**t+10**	**t+20**
Seven intervals	9.9	47.3**	-267.3***	-130.6***	-61.2***
9:35-10:00	-20.1	114.4	-607.4***	-236.3**	-98.3
10:00-15:55	12.1	42.6**	-243.4***	-123.0***	-59.0***

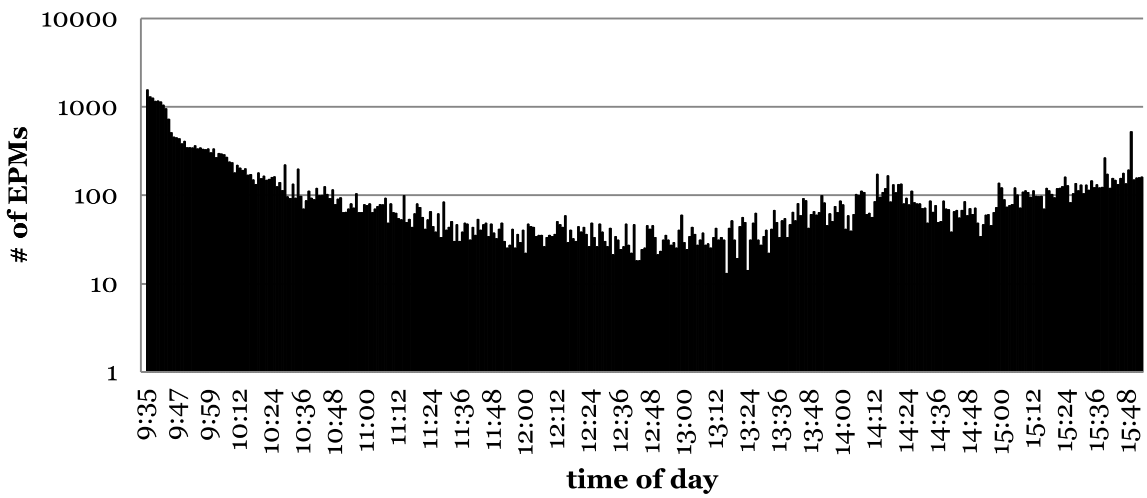


Figure 2.1. Intraday distribution of EPMs

The figure contains a minute-by-minute intraday distribution of EPMs. The scale of the vertical axis is logarithmic.

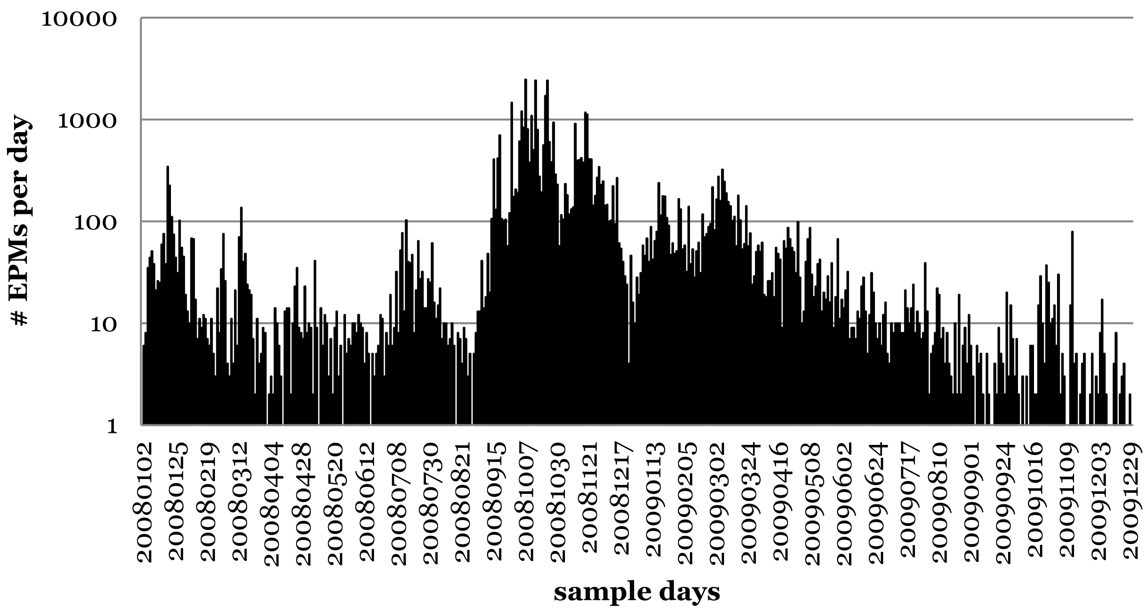


Figure 2.2. Daily distribution of EPMs

The figure contains the daily distribution of sample EPMs identified during the 2008-2009 period. The scale of the vertical axis is logarithmic.

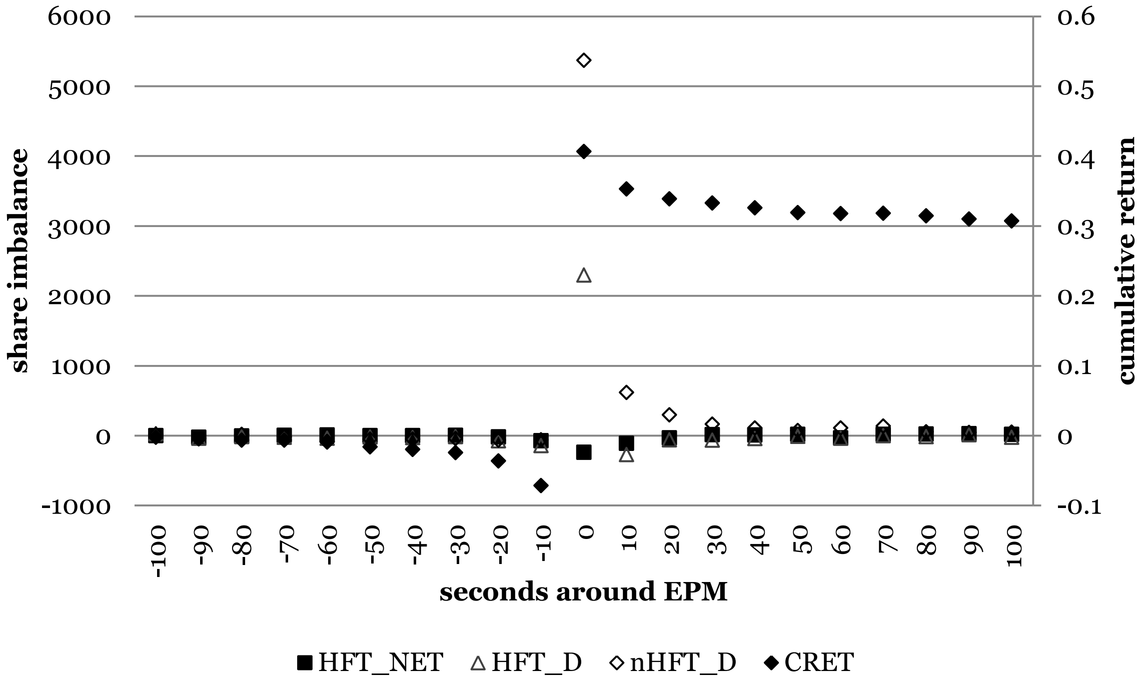


Figure 2.3. HFT and nHFT activity around EPMs

The figure displays the average return path and trading activity around the sample EPMs. HFT^{D} ($nHFT^{D}$) is liquidity demanded by HFTs (nHFTs) in the direction of the EPM (in # shares) minus liquidity demanded against the direction of the EPM. HFT^{NET} is the net effect of HFT liquidity demand and supply. CRET is the cumulative return. The figure includes both positive and negative EPMs, and for exposition purposes we invert the statistics for the latter.

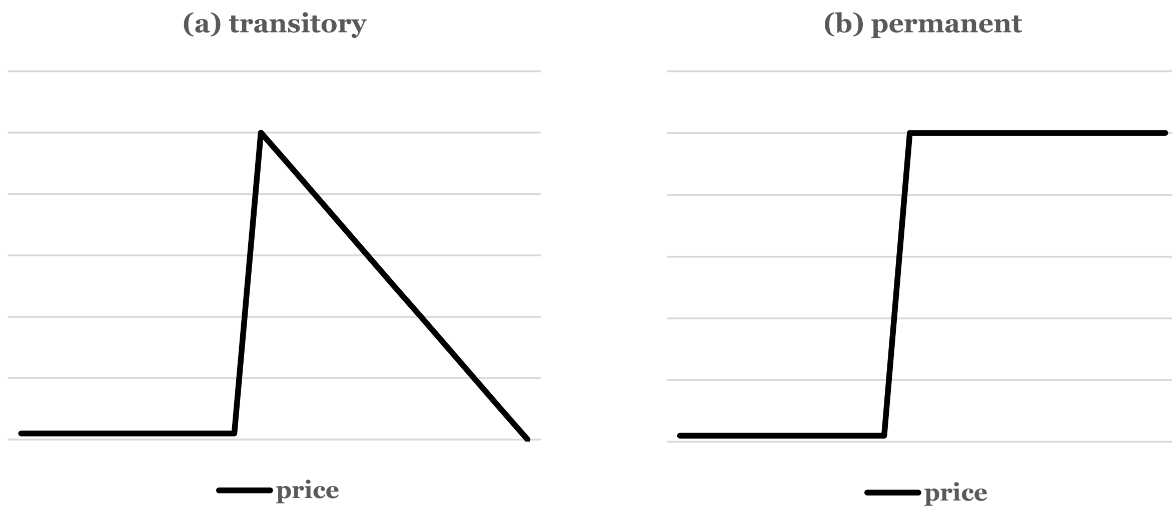


Figure 2.4. EPM types, an illustration

The figure describes two EPM types according to the associated price patterns: (a) a transitory EPM that reverses after a period of time and (b) a permanent EPM that does not reverse.

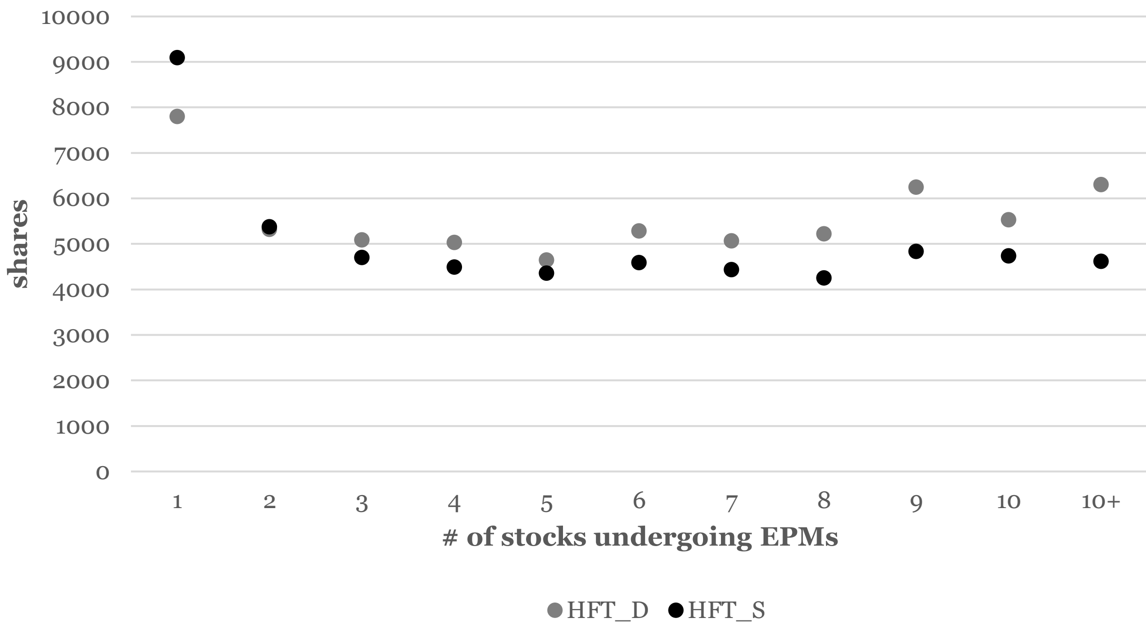


Figure 2.5. HFT^D and HFT^S during co-EPMs

The figure reports per stock HFT^D and HFT^S during the standalone EPMs and co-EPMs.

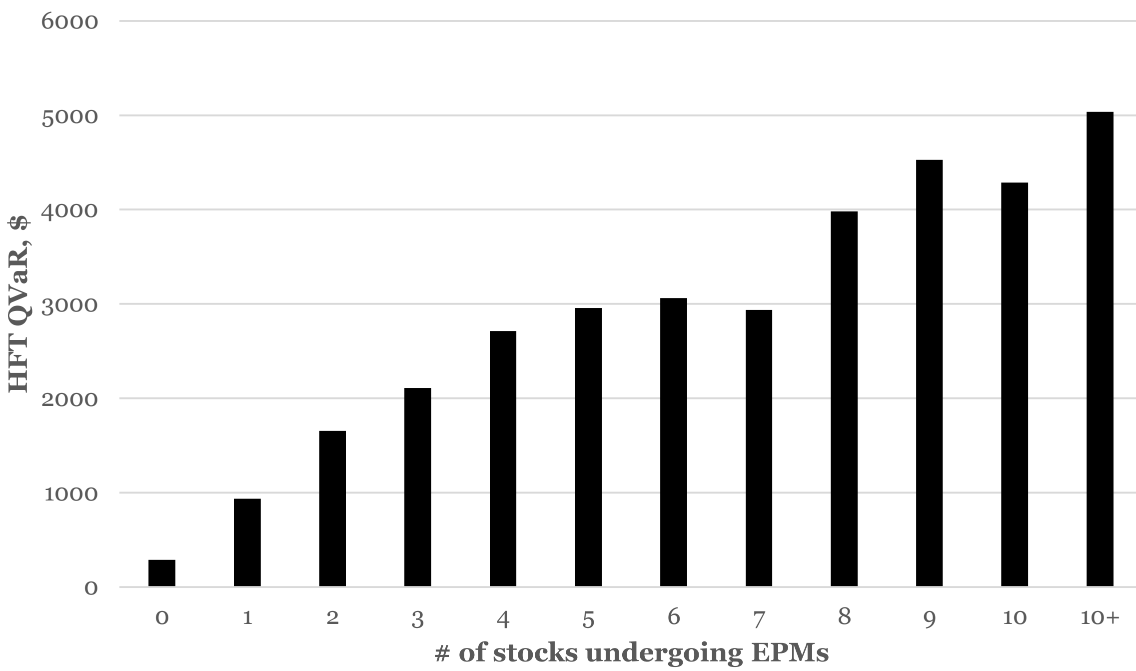


Figure 2.6. QVaR

The figure reports the quasi-value at risk (QVaR) accumulated by HFTs during 0, 1, 2, … and 10+ simultaneous EPMs. We begin by estimating the 99th percentile of the 10-second absolute return of the portfolio of sample stocks. Then, we estimate average returns for each stock i, Ret_i^{tail}, during the instances of portfolio tail returns. The contribution of individual stocks to the portfolio tail returns varies slowly. With this in mind, we use the previous day's composition of portfolio tail returns as a proxy for the expected composition on day d. We then compute intraday QVaR as follows:

$$QVaR_t = -\min\left(\sum_{i}^{N} Ret_{id-1}^{tail} \times DINV_{it}, \sum_{i}^{N} -Ret_{id-1}^{tail} \times DINV_{it}\right),$$

where $DINV_{it}$ is the dollar inventory in stock i accumulated by HFTs during the interval t valued at the last midquote of the interval.

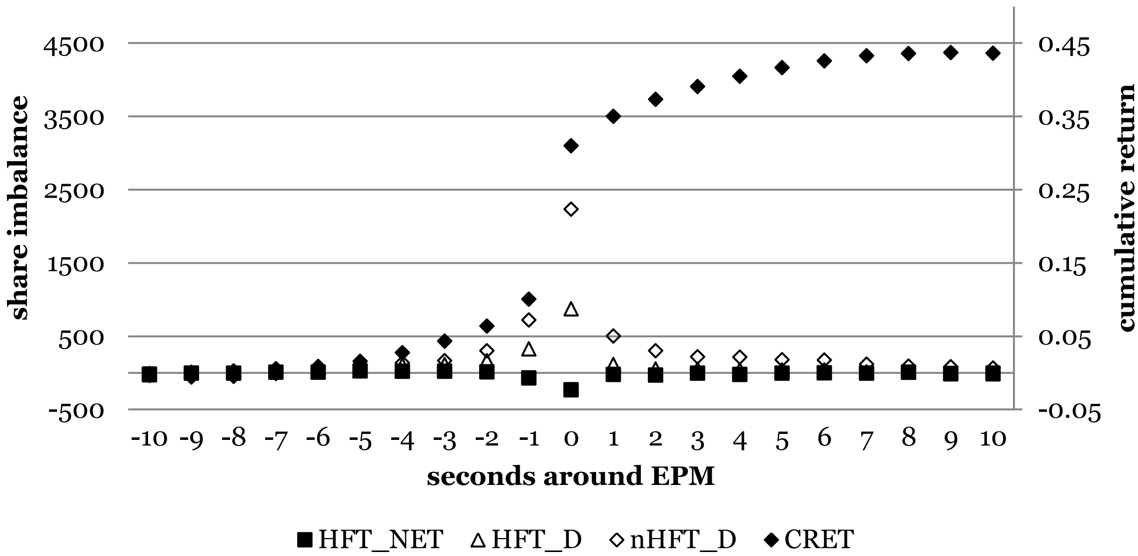


Figure 2.7. HFT and nHFT activity during EPMs, a second by second view

The figure displays the average second by second price path and trading activity during [-10; +10]-second windows centered on the largest one-second EPM return. HFT^{D} ($nHFT^{D}$) is liquidity demanded by HFTs (nHFTs) in the direction of the EPM (in # shares) minus liquidity demanded against the direction of the EPM. HFT^{NET} is the net effect of HFT liquidity demand and supply. CRET is the cumulative return. The figure includes both positive and negative EPMs, and for exposition purposes we invert the statistics for the latter.

Chapter 3. EVERY CLOUD HAS A SILVER LINING: FAST TRADING, MICROWAVE CONNECTIVITY AND TRADING COSTS

3.1. Introduction

Competition on relative speed is a defining characteristic of modern markets, where trading firms spend generously to gain sub-second speed advantages over their rivals. Speed-improving technology is expensive and sometimes only available to a select few, leading to speed differentials. A rich theory literature suggests two possible effects of such differentials on liquidity.[15] On the one hand, being faster may allow liquidity providers to avoid adverse selection and to manage inventory more efficiently. As a result, liquidity may improve. Alternatively, the differentials may allow some traders to pick off stale limit orders, impairing liquidity. To shed new light on these possibilities, we examine a multi-year time series of exogenous shocks to speed differentials. The results show that when the differentials exist liquidity is impaired.

We examine information transmission between financial markets in Chicago and New York, where signals are sent via two channels: a fiber-optic cable and several microwave networks. Microwave networks are about 30% faster than cable, and have two important characteristics. First, in 2011-2012 (the first two years of our four-year sample period) they are only accessible by a select group of traders. Second, precipitation (i.e., rain and snow) disrupts them. The first characteristic creates a two-tiered market, where some traders are faster than others. The second characteristic intermittently eliminates the speed advantage of the fastest tier. We show that when the microwave networks are functional, the speed advantage is used primarily to pick off stale limit orders in the course of latency arbitrage. When precipitation eliminates the speed advantage adverse selection and trading costs decline by more than 7%, while volatility declines by about 6%.

[15] See Hoffmann (2014), Biais, Foucault and Moinas (2015), Foucault, Hombert and Roşu (2016), Foucault, Kozhan and Tham (2016), Menkveld and Zoican (2016), Aït-Sahalia and Sağlam (2017).

The liquidity improvements have two explanations. The first is mechanical: when the microwave users lose speed advantage they stop picking off standing limit orders, and the remaining liquidity takers receive better executions. The second explanation allows for improvements in liquidity supply, whereby limit orders are priced more aggressively when adverse selection is low. The data support both explanations. First, both trade price impacts and trading volume decline, and second, realized spreads decline and limit order aggressiveness increases during microwave disruptions.

On balance, the results suggest that microwave users prefer to take liquidity rather than supply it. Many theory models recognize this preference, yet some empirical studies find that fast traders often trade via limit orders.[16] There are two possible explanations for liquidity taking in our setting. First, the execution probability of limit orders is relatively low, especially for the short-lived latency arbitrage opportunities. Second, in many liquid stocks spreads are narrow and order queues are long, further reducing execution chances. Corroborating the latter explanation, the largest reductions in adverse selection during microwave disruptions occur in assets with narrow spreads.

In addition to posting new orders, speed advantages should allow traders to timely cancel stale limit orders. The data however suggest that an average limit order trader does not possess such advantages, likely because the microwave connections are available only to a select few in 2011-2012. Notably, the status quo changes in winter of 2012-2013, when a technology provider McKay Brothers democratizes microwave transmissions. Instead of selling microwave bandwidth that traders use to outpace others, the firm begins to use its network to transmit the latest price updates and sell them to anyone on a subscription basis. As a result, the speed advantages previously enjoyed by select firms are diminished. We find that once information transmission is democratized

[16] See for instance, O'Hara (2015), Yao and Ye (2015), Brogaard, Hendershott and Riordan (2016), Chordia, Green and Kottimukkalur (2016).

in this manner precipitation stops having an effect on trading costs, suggesting an elimination of the speed differential between an average liquidity taker and an average supplier. Furthermore, democratization leads to a one-time reduction in adverse selection and trading costs.

Our results point to a negative relation between speed differentials and liquidity. As such, they provide a complementary perspective to that of Brogaard, Hagströmer, Nordén and Riordan (2015), who show that in the Swedish market speed differentials resulting from colocation are mainly sought by market makers and therefore benefit liquidity. The authors suggest that although colocation is most attractive to market makers, technology that increases information transmission speeds between markets may be sought by other traders, such as latency arbitragers, leading to negative liquidity effects. Our study also corroborates the findings of Baron, Brogaard, Hagströmer and Kirilenko (2016) and Foucault, Kozhan and Tham (2016), who suggest that modern arbitragers often use marketable orders, thus increasing order flow toxicity and impairing liquidity.

Although the financial economics literature has previously explored the effects of weather on trader behavior, these effects have been mainly ascribed to investor mood. Although we examine a different weather-induced regularity, a technological one, it is important that we address the possibility that our results come from slower information processing attributed to weather-induced moods of traders in Chicago and New York (deHaan, Madsen and Piotroski, 2015). To do so, we show that our results are robust to focusing exclusively on precipitation in Ohio, a state that hosts all microwave network paths yet has a relatively low concentration of financial firms. We also confirm the robustness of the results to various sample selection procedures and to alternative precipitation variables.

Our contribution to the literature is as follows. First, we shed new light on the predictions of theory models that examine speed differentials and provide empirical evidence on the models' insights into (i) order choices of the fastest traders and (ii) the liquidity suppliers' response to lower adverse selection risk. Second, we offer evidence complementary to existing empirical research that examines the relation between speed differentials and liquidity. Some market participants claim that faster markets are unconditionally better; our results suggest that the benefits are conditional on how speed advancements are used. Finally, we describe a new approach to measuring exogenous variation in relative speed in modern markets that, to our knowledge, has not been previously examined.[17]

The remainder of the paper is as follows. Section 2 discusses the history and physics of information transmission, the state of the trading speed literature, and latency arbitrage between the futures and equity markets. Section 3 describes the data and sample. Section 4 discusses the main empirical tests. Section 5 reports robustness tests. Section 6 concludes.

3.2. Institutional background and related literature

3.2.1. History and physics of information transmission between Chicago and New York

In the world of ultra-fast trading, the physics of signal transmission plays an important role. The most common way to transmit information over long distances is via a fiber-optic cable. The first such cable between Chicago and New York was laid in the mid-1980s; however, its path was not optimal for ultra-fast communications. The cable was placed along the existing rail lines, making multiple detours from a straight line, going south to Pittsburgh and thereby exceeding the straight-line distance between Chicago and New York by about 300 miles. Realizing potential latency reduction from a more linear setup, a technology company Spread Networks laid another

[17] Koudijs (2015, 2016) uses adverse weather events to study information transmission between London and Amsterdam in the 18th century.

cable in 2010. The new cable had significantly fewer detours, went through the Appalachian Mountains and shaved valuable milliseconds off the signal transmission time.

Although fiber is a very fast transmission medium, it is not the fastest. Because microwaves travel faster through air than photons do through fiber, a network of microwave towers placed in a straight line can shave additional milliseconds off the signal transmission time. At the time of this study, microwave networks advertise round-trip information transmission speeds that are about 30% faster than their fiber-optic competitors. Specifically, during the sample period, microwave networks transmit information from Chicago to New York in about 4.5 ms, whereas fiber transmission takes about 6.5 ms.

Although faster than cable, microwaves have a disadvantage – they are relatively easily disrupted. Among the disruptors are rain droplets and snowflakes, especially when rainfall/snowfall is substantial. During such disruptions, traders who use microwave links lose their speed advantage and must either stop trading or transition from microwave to fiber transmissions. Industry insiders suggest that mainly the former happens; certain strategies are switched off when firms realize that their speed advantage is temporarily lost. The switch is automatic and does not require human involvement.

The first microwave network that linked Chicago and New York was operational at the end of 2010, with several additional networks built in 2011 and 2012. During this period, access to microwave transmission speeds was limited to a small group of trading firms, because the Federal Communications Commission restricted the number of network licenses citing airwave congestion. As such, the 2011-2012 period provides us with a unique opportunity to examine a two-tiered marketplace where some traders have access to the fastest speeds and others do not. Our results linking precipitation episodes to lower adverse selection and trading costs come from this period.

3.2.2. Information transmission speed and market quality

The speed-related effects have been extensively modeled in recent literature; Menkveld (2016) provides a comprehensive review. Bernales (2014), Biais, Foucault and Moinas (2015), Foucault, Hombert and Roşu (2016) and Foucault, Kozhan and Tham (2016) model a market where speed differentials result in fast traders' generating adverse selection for slower limit order traders. Limit order traders in turn seek higher compensation for providing liquidity, thereby increasing liquidity costs for all market participants. Budish, Cramton and Shim (2015) and Menkveld and Zoican (2016) show that, even in absence of speed differentials between the fast traders, sequential order processing and increases in exchange engine speeds lead to adverse selection of liquidity providers, increasing transaction costs.

Hoffmann (2014) and Jovanovic and Menkveld (2015) show that when some market makers become fast they avoid being adversely selected and increase liquidity supply. In Hoffmann (2014) however, slower market makers become more exposed to adverse selection and widen their quotes. Depending on the relative size and competitiveness of the two groups, speeding up of select market makers may have both positive and negative consequences. Bongaerts, Kong and Van Achter (2016) show that both liquidity takers and liquidity makers will engage in speed competition. Roşu (2015) and Du and Zhu (2016) suggest that when some traders are faster than others, volatility may increase.

3.2.3. Information flow between futures and equities

We focus on information transmission between Chicago and New York. In the U.S., most futures contracts trade on the Chicago Mercantile Exchange (CME), particularly in its data center in Aurora, IL. Meanwhile, equities mainly trade at data centers that are located in New Jersey, close to

New York City. During our sample period, the NYSE data center is in Mahwah, NJ; Nasdaq data center is in Carteret, NJ; BATS is in Weehawken, NJ; and Direct Edge is in Secaucus, NJ. To continue with academic tradition, throughout the paper we refer to the two locales as Chicago and New York.

Information transmission between the two market centers is driven by fast arbitrageurs. Our data show that when microwave technology allows these arbitrageurs to speed up, both price impacts and trading costs increase. This result may appear counterintuitive because arbitrageurs are often viewed as liquidity providers who enhance market efficiency. Several theory models suggest that arbitrageurs may respond to supply and demand shocks faster and more effectively than traditional market makers thereby improving liquidity (Holden, 1995; Gromb and Vayanos, 2002, 2010). Guided by the insights of Grossman and Stiglitz (1980), these models assume that arbitrageurs are passive and provide liquidity when it is required by noise traders.

Recent theory relaxes this assumption and allows arbitrageurs to demand liquidity when it is profitable. Foucault, Kozhan and Tham (2016) model a market in which arbitrageurs are faster than market makers. When arbitrageurs trade to enforce the law of one price, they often expose market makers to adverse selection risk. As in Copeland and Galai (1983), market makers require compensation for the risk of being adversely selected, and liquidity becomes more expensive. Foucault, Kozhan and Tham (2016) conclude that although arbitrage makes prices more efficient, it may hurt liquidity. This conclusion echoes the result in Roll, Schwartz and Subrahmanyam (2007), who find that arbitrage opportunities Granger-cause illiquidity.

3.3. Data and sample

Our main analysis is based on the millisecond DTAQ data. The sample period spans four years, from January 2011 through December 2014. The first two years (2011-2012) are characterised by limited access to microwave technology. The latter period (2013-2014) captures the time after the technology was democratized.

To achieve the fastest speeds, microwave networks follow paths that are as straight as possible and therefore rather similar. For illustration, Figure 3.1 reports tower locations of three select networks connecting Chicago to the New York data centers. The data on tower locations are obtained from the Federal Communications Commission (https://www.fcc.gov). Going east from the CME data center, the networks pass through Illinois, Indiana, Ohio, western Pennsylvania and then split in eastern Pennsylvania, with the southern branches going to Nasdaq's data center in Carteret and the northern branches going to the NYSE in Mahwah. To avoid clutter, Figure 3.1 maps three microwave networks; FCC data show that all networks follow similar paths.

3.3.1. Precipitation data

We obtain precipitation data from the National Oceanic and Atmospheric Administration (http://www.noaa.gov). The data contain precipitation statistics collected by weather stations across the U.S., in 15-minute intervals. The data also contain precise station locations. The stations report in local time, so for stations in Illinois and northwestern Indiana located in the Central time zone we add one hour to report times to match DTAQ time stamps. A standard piece of equipment at every station is a precipitation tank equipped with an automatic gauge that measures accumulated precipitation. We focus on data collected by 83 stations located along the Chicago-New York corridor (Figure 3.2). In the robustness section, we examine station samples of different sizes.

We note that although it may only rain over Indiana or Ohio, the entire microwave network will be disrupted. A relatively narrow weather front like the one in Figure 3.3 will result in weather stations located within the front reporting high levels of precipitation. In the meantime, stations located outside the front will report no precipitation. To capture relatively narrow bands of intense precipitation, our main independent variable $PRECIP$ is computed as the sum of precipitation amounts reported by all stations. We examine alternative specifications in the robustness section.

Statistics reported in Panel A of Table 3.1 indicate that an average 15-minute sampling interval sees 0.155 mm of precipitation. The distribution is rather skewed, with a median of 0.07, indicating that periods of low precipitation are occasionally interrupted by significant rain or snow. We note that microwave networks are only disrupted when precipitation is substantial. We therefore focus on high levels of precipitation and compute two additional metrics, $PRECIP1$ and $PRECIP2$, that capture intervals when precipitation is 0.5 and 1 standard deviations above the mean. The two groups contain, respectively, 17% and 10.5% of all intervals, and $PRECIP1$ and $PRECIP2$ events last on average 54 and 49 minutes. As such, significant precipitation is observed rather frequently but ends quickly, forming a time series with sufficient variability.

3.3.2 Asset samples

The importance of information flows between the futures markets in Chicago and the equity markets in New York is well recognized in the literature. Some studies find that futures markets lead price discovery (Kawaller, Koch, and Koch, 1987; Chan, 1992). Others suggest that information may flow both ways (Chan, Chan and Karolyi, 1991; Roll, Schwartz and

Subrahmanyam, 2007). Hasbrouck (2003) shows that the direction of information flow depends on the futures trading activity; for the most active contracts, futures dominate price discovery. Given that the most active futures contracts track baskets of securities, our focus in the equity market is on the ETFs. As long as price discovery via futures is non-trivial, the speed of information transmission between Chicago and New York should matter for trading costs in ETFs. In a later section, we examine the direction of price discovery between the two markets in more detail.

We use millisecond DTAQ data for a sample of 100 most actively traded ETFs. Among these, 50 ETFs track U.S. equity indexes; 22 – international indexes; 20 – corporate or treasury interest rate indexes; 4 – metals (i.e., gold and silver); 1 – a real estate portfolio; and 3 – other assets (Panel B of Table 3.1). Many ETFs in our sample track the same baskets of securities as the CME futures contracts (e.g., the QQQ ETF and the E-mini Nasdaq-100 futures). Others track baskets similar to those of major CME contracts. For example, the iShares Russell 1000 ETF does not have a corresponding CME futures contract; however, a portion of price discovery in this ETF comes from futures on other indexes such as the S&P 500.[18]

3.3.3. DTAQ data and summary statistics

Following Holden and Jacobsen (2014), we combine the DTAQ NBBO and Quote files to obtain the complete NBBO record and merge the resulting dataset with the Trade file. We sign trades using the Lee and Ready (1991) algorithm and exclude the first and the last five minutes of each trading day to avoid the influence of the opening and closing procedures. Panel A of Table 3.2 reports market activity statistics. Precipitation data are in 15-minute intervals, and we aggregate the statistics accordingly. An average ETF has 5,305 NBBO updates every 15 minutes, equivalent to

[18] The CME delisted E-mini Russell 1000 futures contract in 2007 and relisted it in 2015.

about 6 updates per second. In addition, this ETF trades 500 times every 15 minutes, for a total volume of 190,522 shares.

3.3.4. Precipitation and information transmission speed

Our empirical tests are based on the premise that precipitation disrupts microwave networks and slows down the fastest traders. In Figure 3.4, we illustrate such disruptions. The figure reports the number of equity trades that follow a futures trade when precipitation is zero or very low ($PRECIP2$<0, orange bars), and when it is high ($PRECIP2$>1, blue bars). The number of trades is standardized for each asset to allow for cross-sectional comparability. To reduce serial correlation effects, we focus on the standalone futures trades, those not preceded by another futures or equity trade in the previous 100 milliseconds. Futures trades come from a 2012 CME dataset that contains four E-mini contracts: S&P 500, S&P MidCap 400, Nasdaq 100 and Financial Sector Select. Intraday data for these contracts are sold by the CME as a bundle. ETF trades are from DTAQ.

Figure 3.4 shows that when precipitation is low, trading activity in the equity market picks up 5 ms after a futures trade. Meanwhile, when precipitation is heavy, equity trading begins 7 ms after a futures trade. During the sample period, microwave networks transmit information from Chicago to New York in about 4.5 ms, whereas fiber transmission takes about 6.5 ms. As such, the results corroborate the notion that precipitation slows traders by the 2-ms difference between the microwave and fiber speeds.

Until mid-2015, U.S. exchanges are not required to synchronize their clocks (Bartlett and McCrary, 2016). During our sample period, the CME clock lags DTAQ by about one millisecond, and we adjust for this lag. Without the adjustment, it appears that equity trading during periods of zero precipitation picks up 4 ms after a futures trade, which is not possible as it would require microwave speed to be equal to or faster than the speed of light. Importantly, the adjustment does not affect the evidence of the 2-ms speed advantage of the fastest traders. This is because the adjustment affects both $PRECIP2<0$ and $PRECIP2>1$ equally, and therefore the difference between them remains the same.

3.3.5. Picking-off risk

Recent literature suggests that fast informed traders often trade via limit orders. Brogaard, Hendershott and Riordan (2016) show that limit orders submitted by fast traders play a significant role in price discovery. O'Hara (2015) also suggests that fast informed traders often prefer limit to marketable orders. Both studies however point out that most traders do not resort to one order type exclusively, but rather use them interchangeably depending on the circumstances.

One of such circumstances is the constraint introduced by the minimum tick size. A binding tick size provides a strong incentive for fast traders to use marketable orders. Assume that a fast trader learns that an asset is underpriced. She wants to buy, but if the tick size is binding she cannot raise the outstanding bid without locking or crossing the market. Given these considerations, and if her signal is sufficiently strong, she may choose to consume liquidity (pick off the outstanding ask quote) despite having to pay the spread. As such, picking-off risk may be higher in assets with binding tick sizes.

The very active ETFs in our sample are quite liquid and therefore are likely to be constrained by the minimum tick size. Panel B of Table 3.2 shows that the average NBBO is 1.9 cents, with a median of 1.2 cents. Given these constraints, trade-related price discovery and the associated picking-off risk may be important. In subsequent tests, we subdivide assets into two categories:

most and least constrained. To do so, we divide the assets into terciles according to the average NBBO. An average (median) NBBO in the first tercile is 1.0 (1.0) cent, whereas it is 3.6 (1.8) cents in the third tercile. We define the first tercile as the most tick constrained and the third tercile as the least constrained.

To further examine the issue of picking-off risk, we compute two metrics. First, we estimate a share of price discovery attributable to trades. Second, we compute trade price impacts. The former metric follows Hasbrouck's (1991 a,b) and decomposes the efficient price variance into the trade-related and trade-unrelated components. The details of this calculation are in the Appendix. Panel B of Table 3.2 shows that the trade-related component is 29.6%. As such, new information is incorporated into prices through trades rather frequently, and therefore concerns with the picking-off risk are warranted.

Our second proxy for the picking-off risk is the conventional price impact metric, computed on a round-trip basis as twice the signed difference between the midquote at a time after the trade and the midquote at the time of the trade: $PRIMP_t = 2q_t\left(mid_{t+\gamma} - mid_t\right)$, where q_t is the Lee and Ready (1991) trade direction indicator, mid_t is the midquote computed as $(NBBO\ Ask_t + NBBO\ Bid_t)/2$, and γ indicates the time elapsed since the trade. Recent research uses γs of just a few seconds. For instance, O'Hara (2015) suggests that 5- to 15-second intervals may be the most useful, whereas Conrad, Wahal and Xiang (2015) use price impacts up to 20 seconds.

To check if intervals of these lengths are practical in our setting, Figure 3.5 traces price impacts for 60 seconds after a trade. The results clarify our understanding of price dynamics on two levels. First, the data show that price impacts are greater than zero, corroborating the earlier assertion that non-trivial amounts of information are incorporated into prices through trades.

Second, although a significant proportion of information is incorporated into quotes within a second after the trade, incorporation continues at a slower pace up to 60 seconds. In the remainder of the tests, we focus on 15-second intervals, with robustness checks examining intervals between 1 and 60 seconds.

It may not be immediately obvious that there is enough adverse selection in ETFs to warrant non-zero price impacts. We suggest that as long as sufficient amounts of macro information are present, price impacts in ETFs may be quite sizeable. In a study that examines a recent sample of large U.S. equities, Chakrabarty, Jain, Shkilko and Sokolov (2016) report price impacts that are 35% of the effective spread. In Table 3.2, the ETF price impacts are 31% of the effective spread. As such, adverse selection is a non-trivial component of ETF trading costs and is comparable to the levels found in equities.

3.3.6. Trading costs and liquidity provider revenues

Table 3.2 also reports liquidity costs and liquidity provider revenues proxied by effective spreads, ESP_t, and realized spreads, RSP_t. ESP is computed as twice the signed difference between the prevailing midquote and the trade price, p_t: $ESP_t = 2q_t(p_t - mid_t)$. RSP_t is computed as the difference between the effective spread and the price impact. We volume-weight effective and realized spreads. The average (median) effective spread is 1.9 (1.0) cents and the average (median) realized spread is 1.3 (0.7) cents.

3.4. Empirical findings

3.4.1. Connectivity disruptions and picking-off risk

When the microwave networks are fully functional, their users have a speed advantage. Theory models make several assumptions as to how this advantage may be used. Some models assume that the fastest traders can better manage adverse selection risk. Others suggest that speed advantages are used to generate such risk by picking off slower traders. In this section, we aim to better understand which of these assumptions prevails.

If speed advantages allow fast traders to pick off outstanding limit orders, connectivity disruptions should result in lower price impacts. Alternatively, if fast connections are used to incorporate the latest information into quotes, the disruptions may be accompanied by larger price impacts. Certainly, it is possible that both explanations have merit, and our data allow us to gauge which of them prevails. We focus on the 2011-2012 period when the microwave networks allowed for speed differentials among traders. The post-democratization period (2013-2014) is examined in a later section. Chung and Chuwonganant (2014) and Malinova, Park and Riordan (2014) argue that VIX is a first-order determinant of trading activity and liquidity, and we use their insight in a regression setup as follows:

$$DEPVAR_{it} = \alpha_0 + \beta_1 PRECIP_t + \beta_2 VIX_t + \varepsilon_{it}, \tag{1}$$

where $DEPVAR$ is the price impact; $PRECIP$ is total precipitation in the Chicago-New York corridor; and VIX is the intraday volatility index proxied by the iPath S&P500 VIX ST Futures ETF that tracks VIX. As discussed earlier, we also use $PRECIP1$ and $PRECIP2$ to identify the most significant precipitation events. All asset-specific variables are standardized (by demeaning and scaling by the standard deviation for each stock), so regression models control for asset fixed effects. Additionally, the standard errors are double-clustered along the asset and time dimensions.

Table 3.3 shows that price impacts decline during network disruptions. Significant amounts of precipitation captured by $PRECIP2$ are associated with a 0.047 standard deviations, or 7.1%, decline in price impacts (Panel A). It therefore appears that microwave users prefer marketable orders to limit orders.

Per our earlier suggestion, marketable orders may be the only choice when the tick size is binding. If this is so, microwave network disruptions will have a larger effect on price impacts in the most constrained assets. The results in Panel B are consistent with this expectation. Price impacts in the most constrained ETFs decline by 0.051 standard deviations, whereas they decline by only 0.039 standard deviations in the least constrained ETFs. It therefore appears that fast traders use more limit orders when the tick size allows. This said, even in the least constrained ETFs liquidity taking is preferred by microwave users as evident from the decline in price impacts.

3.4.2. Trading costs and liquidity provider revenues

To the extent that liquidity providers use speed advantages to avoid being adversely selected, microwave connectivity disruptions should cause them to widen spreads. If however speed advantages are used mainly to pick off standing orders, precipitation should result in reduced adverse selection, and spreads may narrow. The tests discussed earlier provide support to the picking-off story, so we expect trading costs to decline during precipitation episodes. We however note that it is not clear how quickly liquidity providers adjust to lower adverse selection, and if they adjust at all given that precipitation episodes are relatively short. Easley and O'Hara (1992) describe price adjustment as gradual learning. In their model, market makers do not immediately know if informed traders are active, but learn over time. Whether such learning happens in our setting is an empirical question.

In Table 3.3, we report eq. 1 coefficient estimates for effective and realized spreads. Effective spreads decline by 0.043 standard deviations, or 7.2%, during heavy precipitation episodes ($PRECIP2$ in Panel A). Expectedly, this result is more pronounced for the least constrained assets (Panel B), where the spreads have room to decline. The results are consistent with predictions of the models that emphasize the picking-off risk and are also informative about the speed of adjustment to changing levels of adverse selection. Specifically, the length of an average precipitation episode appears sufficient for liquidity providers to adjust. It is however unclear if this adjustment reflects an equilibrium. In a later section, we report results based on an exogenous shock that resulted in the long-term reduction in speed differentials. This shock further improves our understanding of the equilibrium effects.

We note that there is an alternative, mechanical, explanation to the decline in effective spreads. Aggressively priced limit orders may be added to the book equally often when the microwaves are up and when they are down. Since such orders are consumed less when the arbitrage strategies are switched off, they are more readily available to the remaining liquidity takers, who obtain better prices. In a later section, we examine trades and limit order book data to show that both explanations have merit.

Realized spreads also decline, by 0.021 standard deviations, or 5.3%, during $PRECIP2$ events. As such, network disruptions not only reduce liquidity costs, but also reduce liquidity provider revenues. Similarly to the effective spread result, there are two possible explanations. On the one hand, liquidity supply may become more competitive when picking-off risk is reduced. Chakrabarty, Jain, Shkilko and Sokolov (2016) show that when the adverse selection risk declines, liquidity providers reposition orders from the deeper layers of the book to the inside. On the other hand, the effect may be mechanical. When microwaves allow latency arbitrageurs to pick off inside quotes, the spread increases by at least one tick. Given the coarseness of this change, the resulting realized spreads may be unduly large.

3.4.3. Trading activity and volatility

The literature often assumes that lower trading costs attract additional trading interest and therefore result in higher trading volume. In our setting, this assumption will not necessarily hold. This is because aside from lower costs, network disruptions lead to a reduction in the number of picking-off opportunities and consequently the volume generated by latency arbitrage. The regression results in Panel A of Table 3.4 are consistent with this notion. The number of trades declines by 0.072 standard deviations during $PRECIP2$ events. Trading volume also declines; by 0.042 standard deviations.

The finance literature has not yet come to a consensus on the relation between electronic trading and volatility. While some studies report that the relation is negative (Hasbrouck and Saar, 2013; Brogaard, Hendershott and Riordan, 2014), others find it to be positive (Boehmer, Fong and Wu, 2015). Closest to our setting, a theory model by Roşu (2015) suggests that as fast traders pick off market makers' quotes, volatility may increase. Du and Zhu (2016) also show that when some traders are faster than others, liquidity shocks result in greater volatility. Our results are consistent with these insights; volatility, which we define as the difference between the high and low prices during an interval scaled by the average price, declines by 0.118 standard deviations, or 5.8%, during $PRECIP2$ events.

In assets with wider spreads new information may be incorporated into prices through both marketable and aggressive limit orders. Meanwhile, in assets whose spreads are constrained by the minimum tick size, fast traders must rely on marketable orders. Naturally, these considerations should affect changes in trading activity during network disruptions. Panel B shows that the number of trades and trading volume decline in the most constrained assets, yet remain unchanged in the least constrained assets. These results corroborate two of our earlier conjectures: (i) fast traders use fewer marketable orders in the least constrained ETFs, and (ii) trading volume may increase in response to lower trading costs compensating for some of the volume lost when the arbitrage strategies are switched off.

Overall, the results suggest that even though lower spreads may attract additional trading interest, trading volume generated by this interest is smaller or equal to the lost arbitrage volume. One possibility is that the disruptions are not long enough or not sufficiently predictable for additional trading interest to emerge. A trading strategy that is highly sensitive to transaction costs may not be viable in a high cost environment, even if high cost periods are occasionally interrupted by low cost periods. This said, an extended period of lower spreads may make the strategy viable, thus generating new trading interest. In a later section, we examine this possibility by studying an event that resulted in a long-lasting loss of speed advantage by the network users.

3.4.4. Limit order aggressiveness

Earlier, we suggested that the reduction in trading costs may have two explanations: (i) the mechanical one, whereby standing orders are not picked off as frequently during microwave disruptions and (ii) the one based on the emergence of latent liquidity. The results in the previous section corroborate the first explanation. In this section, we use ITCH limit order book data provided by Nasdaq to examine the second explanation.

Specifically, we ask if the proportion of aggressively priced limit orders increases during microwave disruptions. We compute two metrics: the number of orders that (i) match the prevailing NBBO and (ii) the number of orders that match or improve the NBBO. We scale both metrics by the number of total order submissions. Eq. 1 coefficient estimates reported in Table 3.5 confirm that the number of aggressively priced limit orders increases during precipitation episodes. For orders that match the NBBO, the increase is seen in the full sample as well as the most and the least constrained subsamples (Panel A). For orders that match or improve the NBBO, the coefficients are insignificant for the most constrained ETFs. This result is not surprising given that improving the NBBO in such ETFs is often impossible given the constraints of the tick size.

3.4.5. Information asymmetry in the futures market

Do speed differentials also affect the picking-off risk in futures contracts? On the one hand, fast traders may carry information both from futures to equities and in the opposite direction. On the other hand, prior research suggests that futures provide the lion's share of price discovery in index instruments. If so, using limited microwave bandwidth to transmit information from ETFs to futures may be wasteful. If this is the case, speed differentials may not have much of an effect on the CME.

To examine this issue, we use the 2012 CME data described in Section 3.4 and compute information shares as in Hasbrouck (1995) for the four futures-ETF pairs. The details of the methodology are in the Appendix. The results are consistent with earlier studies, in that price discovery occurs mainly on the CME; the CME information shares are in the [0.64; 0.82] range. Second, in Panel A of Table 3.6 we examine the price impacts in futures during precipitation episodes and find no relation.

Statistical power is a concern in the analysis above given the small size of the cross-section. To address this concern, in Panel B we replicate the results for the sample of four corresponding ETFs. Similarly to the main sample, precipitation is associated with reductions in price impacts.

3.4.6. Democratization of MWN access

In late December 2012 – early January 2013, a microwave technology provider McKay Brothers, disrupted the business model used by the microwave firms. Instead of selling bandwidth on its network to select traders, McKay Brothers began selling information transmitted by the network to everyone who was willing to pay a nominal fee. Subscribers to this service obtained access to an affordable and non-exclusive channel of information transmission that was among the fastest in the industry. The offer was soon replicated by other providers, and the market for microwave transmissions was democratized. Put differently, microwave users lost the speed advantage they enjoyed in 2011-2012.

Democratization of access to microwave transmission speeds may lead to two outcomes. First, the relation between precipitation and market quality observed in 2011-2012 may diminish because access to superior speeds is no longer limited to a small group of traders. Second, democratization may result in market quality changes similar to those observed during precipitation events. In this sense, precipitation episodes in 2011-2012 may be viewed as periods of short-term democratization, whereas the 2012-2013 event may be viewed as long-term democratization.

In Table 3.7, we report the coefficients of the $PRECIP2$ variable obtained from estimating eq. 1 during the post-democratization period. The results confirm expectations. Precipitation episodes no longer have an effect on price impacts, effective spreads, realized spreads, volatility and trading activity. The change is observed for the full sample and for the most and least constrained subsamples. As such, democratization appears to be a significant market disruptor.

Given the significance of democratization, the liquidity effects associated with the loss of the speed advantage may reappear around the event. To examine if this is the case, we estimate an event study regression model that compares market quality and activity variables in a three-month pre-event window (September – November 2012) and a three-month post-event window (February – April 2013). We exclude December 2012 and January 2013 to allow for a transition period, however the results are similar when these months are included. The regression is set up as follows:

$$DEPVAR_{it} = \alpha_0 + \gamma_t + \beta_1 POST_t + \beta_2 VIX_t + \varepsilon_{it}, \quad (2)$$

where $DEPVAR_{it}$ is one of the following variables (price impacts, effective spreads, realized spreads, the number of trades, traded volume, volatility and stock price) in asset i on day t, γ_t denotes a time trend, $POST$ is a dummy variable that equals to one in February-April 2013, and VIX is the volatility index. All variables are standardized. The model controls for asset fixed effects, and the standard errors (in parentheses) are double-clustered along the asset and time dimensions.

The main variable of interest in eq. 2 is $POST$ as it captures the difference between the pre- and post-democratization periods. Consistent with expectations, the results for the full sample in Table 3.8 indicate that price impacts, effective spreads, realized spreads and volatility decline post-democratization. We must however note an important caveat. Democratization was a single event that affected all assets at the same time. As such, we are unable to eliminate the possibility that the results are driven by a confounding event(s) unrelated to democratization. Although we are unaware of any such events, the event study results should be interpreted with due caution.

A notable difference between the event study and the panel findings discussed earlier comes from the trade-related variables: the number of trades and volume. Recall that in the full sample these variables decline during precipitation events (Table 3.4), likely because latency arbitrage is diminished. In Table 3.8, these variables do not change post-democratization in the full sample, corroborating our earlier suggestion that lower trading costs may encourage new trading interest over long periods of time. The new interest offsets the loss of arbitrage volume. The results pointing to new trading interest are observed even in the sample of the most constrained ETFs where trade and volume losses during precipitation-related disruptions are the largest. Although Table 3.8 shows a declining number of trades in the most constrained assets, the change in volume is statistically insignificant, suggesting that new trading interest arises even in these assets.

The market structure literature often suggests that lower trading costs translate into higher stock prices (e.g., Amihud and Mendelson, 1986; Easley, Hendershott and Ramadorai, 2014). If so, liquidity improvements caused by democratization may result in price appreciations. Table 3.8 indicates that such appreciations indeed occur.

Given the private benefits from exclusive network access that accrued to the select few trading firms in 2011-2012, democratization is a curious case of the market fixing itself. It is however not immediately clear why McKay Brothers chose to disrupt the status quo. Cespa and Foucault (2014) note that it is in a data provider's best interest to restrict dissemination of pricing data only to select traders. Their model shows that if information is accessible to many, it is less valuable to the few who may be willing to pay a premium for the exclusive use. It is our understanding that McKay Brothers was driven by the following two motives. First, the firm believed that even if others were to replicate its offering, there would be enough pricing information for everyone to transmit given the large numbers of traded instruments and the low bandwidth of microwave links. Second, the firm argued that growing its customer base was more profitable than providing restricted access to a small group of clients. As long as the firm maintained its latency advantage, it expected to always retain its customers. Many other connectivity providers launched similar offerings in the months after the democratization, possibly driving down the price of the service. This said, McKay Brothers continues to offer the service to this day and has expanded it to several continents.

3.5. Robustness

A rich literature examines the effects of weather on the behavior of market participants and finds that poor weather is associated with investor pessimism, which is reflected in stock returns (Hirshleifer and Shumway, 2003). The pessimism affects even the sophisticated investors (Goetzmann, Kim, Kumar and Wang, 2015). Furthermore, deHaan, Madsen and Piotroski (2015) show that pessimistic moods induced by poor weather often delay equilibrium price adjustments. As such, the reduction in adverse selection during precipitation episodes may be attributed (at least in part) to slower price discovery caused by the poor weather in Chicago and/or New York rather than to the microwave disruptions.

To examine this possibility, we recalculate the $PRECIP2$ variable to capture periods when the networks are disrupted, yet the moods of traders in Chicago and New York are not affected. Specifically, we compute $PRECIP2$ that satisfies the following two conditions: (i) only weather stations in Ohio indicate high levels of precipitation, and (ii) weather stations in the western and eastern parts of the Chicago-New York corridor indicate near-zero precipitation. We then re-estimate eq. 1 for the 2011-2012 sample and report the results in the *mood control* specification in Panel A of Table 3.9. The effects are consistent with those reported in the earlier tables. As such, trader moods do not seem to be the source of our findings.

Our sample of weather stations is selected to capture the area closely surrounding the microwave paths. As with any such selection procedure, it is important to show that the results are not driven by the specific set of stations. The *mood control* specification takes the first step in this direction by restricting the sample to the Ohio stations. In two additional Table 3.9 specifications, we show that using information from an expanded area surrounding the MWN paths leads to similar conclusions, while precipitation in the placebo area over Colorado, Utah and Wyoming (far removed from the Chicago-New York corridor) has no effect on the variables of interest.

Information asymmetry, trading costs and trading activity vary throughout the day. For instance, effective spreads follow an intraday J-pattern, with wider spreads in the morning that become narrower in the afternoon (Figure 3.6). Notably, intraday precipitation too follows a reverse J-pattern, with precipitation amounts being lower in the morning hours. Since the results in the previous section point to a negative relation between precipitation and spreads, we need to establish that the findings are not due to these intraday patterns.

We examine this possibility in two additional specifications in Panel A of Table 3.9. First, we focus on the afternoon period, when spreads and precipitation are relatively flat. Our results hold for every variable of interest. Second, the results continue to hold when we add intraday fixed effects to eq. 1. As such, the relations between precipitation and spreads observed in the earlier sections are independent of intraday patterns.

Recall that the $PRECIP$ variable estimates total precipitation in the Chicago-New York corridor. This variable is well-suited to capture periods of high precipitation over small areas, but may occasionally acquire high values if relatively minor precipitation extends over the entire corridor. This possibility is the reason for our focus on $PRECIP2$ that captures very high precipitation totals not likely to be achieved through anything other than significant precipitation.

To provide another alternative to $PRECIP$, in Panel B of Table 3.9 we report the results using the average precipitation per station, $MPRECIP$, and its variations, $MPRECIP1$ and $MPRECIP2$, that capture periods when average precipitation is 0.5 and 1 standard deviations above the mean. We note that although these variables mitigate the abovementioned concern, they potentially reduce our ability to detect relatively narrow bands of strong precipitation, especially those accompanied by near-zero precipitation in the rest of the corridor (Figure 3.3). Corroborating this reasoning, the results for $MPRECIP$ are weaker than those reported earlier for $PRECIP$, yet the results for $MPRECIP1$ and $MPRECIP2$ are equally as strong as those for their counterparts computed using total precipitation.

In the main analysis, we compute the effective spreads and their components on a volume-weighted basis. As such, large trades have a stronger effect on the estimates than small trades. To shed more light on the effects of network disruptions on small trades, in Table 3.10 we report eq. 1 regression results for the equally-weighted variables (specifications $EW\_$). The results reported earlier hold.

The results for the volume-weighted effective spreads and their components reported in earlier tables use raw dollar metrics. Naturally, raw spreads may vary in the price of the asset. Although our regressions account for the overall price levels by using asset fixed effects, intraday price changes remain unaccounted for. The $VWP\_$ specifications in Table 3.10 address this issue using effective spreads, price impacts and realized spreads scaled by the midquote at the time of trade. The results corroborate those reported in the earlier tables.

In the previous sections, we discuss the effects of network disruptions on effective spreads. We also show that the effects differ between the assets most and least constrained by the minimum tick size. In Table 3.11, we estimate eq. 1 for two additional variables – the quoted NBBO spread and quoted depth. Whereas effective spreads capture the realized trading costs, the quoted spreads summarize liquidity that is available at all times. As long as investors choose to trade when costs are low, effective spreads may not be fully indicative of changes in available liquidity. Table 3.11 shows that quoted spreads decline when the networks are down across all sample groups. The coefficients repeat the patterns reported for effective spreads, with quoted spreads declining more for the least constrained ETFs, in which more price improvement is possible.

Table 3.11 also reports the results for quoted NBBO depth, which increases during network disruptions, but only for the most constrained ETFs. This finding is consistent with the earlier discussion. In the most constrained ETFs, there is not always room to improve the spread, so liquidity improvements are reflected in quoted depth. There is certainly an alternative mechanical explanation in that less depth is consumed by arbitrage strategies during precipitation episodes, therefore allowing for larger depth averages. Notably, depth does not increase in the least constrained ETFs.

3.6. Conclusions

This study examines the effects of speed differentials on liquidity. During our sample period, microwave networks located between Chicago and New York allow for the fastest information transmission and are only available to select trading firms. When it rains or snows in the area between the two cities, the networks are disrupted because rain droplets and snowflakes block the microwave paths. With the networks temporarily down, information transmission falls back onto the fiber-optic cable – a more reliable, yet slower transmission medium – effectively eliminating the speed advantages of the fastest traders. We show that when this happens, adverse selection and trading costs decline. This result is consistent with predictions of theory models that show that speed differentials among traders may be associated with lower liquidity.

Our results also shed new light on traders' order choices. Recent research suggests that informed fast traders may prefer to trade via limit orders. Our results confirm that this is the case, yet this preference varies in the cross-section. Specifically, in assets with binding tick sizes, trading on short-lived information through limit orders is difficult due to long queues. In such assets, traders prefer marketable orders.

Our results are confirmed in an event-study setting. In winter of 2012-2013, a technology firm, McKay Brothers, democratized microwave transmissions by introducing a new business model. Instead of selling bandwidth on its network, the firm began selling information on both sides of the Chicago-New York corridor. This event had positive liquidity consequences similar to precipitation-related network disruptions.

The technological race continues to drive spending in the trading industry. Recent examples include new data transmission towers to connect the U.K. and European markets. The towers will be among the tallest structures in the U.K. and will rival the Eiffel Tower. They will provide trading firms with a completely unobstructed optical and radio line of sight, never previously offered in Europe, increasing signal transmission speed. In the meantime, traders in the U.S. have been switching from microwave transmissions to more reliable, yet costly, laser links. Our findings shed light on the possible consequences of these developments.

3.7. References

Aït-Sahalia, Y. and Sağlam, M., 2017, High frequency market making: Optimal quoting, Working paper.

Amihud, Y. and Mendelson, H., 1986, Asset pricing with liquidity risk, Journal of Financial Economics 77, 375-410.

Baron, M., Brogaard, J., Hagströmer, B. and Kirilenko, A., 2016, Risk and return in high-frequency trading, Working paper.

Bartlett, R. and McCrary, J., 2016, How rigged are stock markets? Evidence from microsecond timestamps, Working paper.

Bernales, A., 2014, Algorithmic and high frequency trading in dynamic limit order markets, Working paper.

Biais, B., Foucault, T. and Moinas, S., 2015, Equilibrium fast trading, Journal of Financial Economics 116, 292-313.

Boehmer, E., Fong, K. and Wu, J., 2015, International evidence on algorithmic trading, Working paper.

Bongaerts, D., Kong, L. and Van Achter, M., 2016, Trading speed competition: Can the arms race go too far? Working paper.

Brogaard, J., Hagströmer, B., Nordén, L. and Riordan, R., 2015, Trading fast and slow: Colocation and liquidity, *Review of Financial Studies* 28, 3407-3443.

Brogaard, J., Hendershott, T. and Riordan, R., 2014, High-frequency trading and price discovery, *Review of Financial Studies* 27, 2267-2306.

Brogaard, J., Hendershott, T. and Riordan, R., 2016, Price discovery without trading: Evidence from limit orders, Working paper.

Budish, E., Cramton, P. and Shim, J., 2015, The high-frequency trading arms race: Frequent batch auctions as a market design response, *Quarterly Journal of Economics* 130, 1547-1621.

Carrion, A., 2013, Very fast money: High-frequency trading on the NASDAQ, *Journal of Financial Markets* 16, 680-711.

Cespa, G. and Foucault, T., 2014, Sale of price information by exchanges: Does it promote price discovery? *Management Science* 60, 148-165.

Chakrabarty, B., Jain, P., Shkilko, A. and Sokolov, K., 2016, Speed of market access and market quality: Evidence from the SEC naked access ban, Working paper.

Chordia, T., Green, C. and Kottimukkalur, B., 2016, Do high frequency traders need to be regulated? Evidence from trading on macroeconomic announcements, Working paper.

Conrad, J., Wahal, S. and Xiang, J., 2015, High-frequency quoting, trading, and the efficiency of prices, *Journal of Financial Economics* 116, 271-291.

Copeland, T. and Galai, D., 1983, Information effects on the bid-ask spread, *Journal of Finance* 38, 1457-1469.

Chan, K., Chan, K.C. and Karolyi, A., 1991, Intraday volatility in the stock index and stock index futures markets, *Review of Financial Studies* 4, 657-684.

Chan, K., 1992, A further analysis of the lead-lag relationship between the cash market and stock index futures market, *Review of Financial Studies* 5, 123-151.

Chung, K. and Chuwonganant, C., 2014, Uncertainty, market structure, and liquidity, *Journal of Financial Economics* 113, 476-499.

deHaan, E., Madsen, J. and Piotroski, J., 2015, Do weather-induced moods affect the processing of earnings news? Working paper.

Du, S., and Zhu, H., 2016, What is the optimal trading frequency in financial markets? *Review of Economic Studies*, forthcoming.

Easley, D., Hendershott, T. and Ramadorai, T., 2014, Leveling the trading field, *Journal of Financial Markets* 17, 65-93.

Easley, D. and O'Hara, M., 1992, Time and process of security price adjustment, *Journal of Finance* 47, 577-605.

Foucault, T., Hombert, J. and Roşu, I., 2016, News trading and speed, *Journal of Finance* 71, 335-382.

Foucault, T., Kozhan, R. and Tham, W., 2016, Toxic arbitrage, *Review of Financial Studies*, forthcoming.

Goetzmann, W., Kim, D., Kumar, A. and Wang, Q., 2015, Weather-induced mood, institutional investors, and stock returns, Working paper.

Gromb, D. and Vayanos, D., 2002, Equilibrium and welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics* 66, 361-407.

Gromb, D. and Vayanos, D., 2010, Limits of arbitrage: The state of the theory, *Annual Review of Financial Economics* 2, 251-275.

Grossman, S. and Stiglitz, J., 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393-408.

Hagströmer, B. and Nordén, L., 2013, The diversity of high-frequency traders, *Journal of Financial Markets* 16, 741-770.

Hasbrouck, J., 1991a, Measuring the information content of stock trades, *Journal of Finance* 46, 179-207.

Hasbrouck, J., 1991b, The summary informativeness of stock trades: An economic analysis, *Review of Financial Studies* 4, 571-595.

Hasbrouck, J., 1995, One security, many markets: Determining the contributions to price discovery, *Journal of Finance* 50, 1175-1199.

Hasbrouck, J., 2003, Intraday price formation in the U.S. equity index markets, *Journal of Finance* 58, 2375-2399.

Hasbrouck, J. and Saar, G., 2013, Low-latency trading, *Journal of Financial Markets* 16, 646-679.

Hendershott, T. and Riordan, R., 2013, Algorithmic trading and the market for liquidity, *Journal of Financial and Quantitative Analysis* 48, 1001-1024.

Hirshleifer, D. and Shumway, T., 2003, Good day sunshine: Stock returns and the weather, *Journal of Finance* 58, 1009-1032.

Hoffmann, P., 2014, A dynamic limit order market with fast and slow traders, *Journal of Financial Economics* 113, 156-169.

Holden, C., 1995, Index arbitrage as cross-sectional market making, *Journal of Futures Markets* 15, 423-455.

Holden, C. and Jacobsen, S., 2014, Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions, *Journal of Finance* 69, 1747-1785.

Jovanovic, B. and Menkveld, A., 2015, Middlemen in limit order markets, Working paper.

Kawaller, I., Koch, P. and Koch, T., 1987, The temporal price relationship between S&P 500 futures and the S&P 500 index, *Journal of Finance* 5, 1309-1329.

Koudijs, P., 2015, Those who know most: Insider trading in 18th century Amsterdan, *Journal of Political Economy* 123, 1356-1409.

Koudijs, P., 2016, The boats that did not sail: Asset price volatility in a natural experiment, *Journal of Finance* 71, 1185-1226.

Lee, C. and Ready, M., 1991, Inferring trade direction from intraday data, Journal of Finance 46, 733–746.

Malinova, K., Park, A. and Riordan, R., 2014, Do retail traders suffer from high frequency traders, Working paper.

Menkveld, A., 2016, The economics of high-frequency trading: Taking stock, *Annual Review of Financial Economics*, forthcoming.

Menkveld, A. and Zoican, M., 2016, Need for speed? Exchange latency and liquidity, *Review of Financial Studies*, forthcoming.

O'Hara, M., 2015, High frequency market microstructure, *Journal of Financial Economics* 116, 257-270.

Roll, R., Schwartz, E. and Subrahmanyam, A., 2007, Liquidity and the law of one price: The case of the futures-cash basis, *Journal of Finance* 62, 2201-2234.

Roşu, I., 2015, Fast and slow informed trading, Working paper.

Yao, C. and Ye, M., 2015, Why trading speed matters: A tale of queue rationing under price controls, Working paper.

Table 3.1. Descriptive statistics

The table reports descriptive statistics for precipitation and for the sample of 100 ETFs. In Panel A, $PRECIP$ is the variable that captures total precipitation recorded by the weather stations along the Chicago-New York corridor. Along with precipitation statistics (in mm per a 15-minute sampling interval), we report the percent share of intervals with $PRECIP$ greater than 0.5 standard deviations ($PRECIP1$) and with $PRECIP$ greater than 1 standard deviation ($PRECIP2$). Finally, we report the length of an average period with consecutive $PRECIP1$ and $PRECIP2$ as well as the percent share of days with episodes of $PRECIP1$ or $PRECIP2$. Panel B classifies 100 sample ETFs into categories according to the underlying asset basket.

Panel A: Precipitation	
$PRECIP$, mm/interval	
mean	0.155
median	0.070
std. dev.	0.218
% intervals with $PRECIP1$	17.0
% intervals with $PRECIP2$	10.5
length $PRECIP1$, min	54.2
length $PRECIP2$, min	49.1

Panel B: ETF sample	
Equities	
US index	50
International index	22
Interest rate products	20
Metals	4
Real estate	1
Other	3

Table 3.2. Market activity statistics

The table contains summary statistics for the sample of 100 ETFs. Statistics are derived from the millisecond DTAQ data and aggregated into 15-minute intervals to match precipitation data. Volatility is defined as the difference between the high and low price in a 15-minute interval scaled by the average price. Trade price discovery is the percentage of efficient price variance that may be attributed to trades (Hasbrouck, 1991). The National Best Bid and Offer (NBBO) is defined as the difference between the lowest offer quote and the highest bid quote across all markets. We divide the assets into terciles by their average NBBO. Assets with the smallest (largest) NBBOs are considered the most (least) tick-constrained. Price impact is defined as twice the signed difference between the NBBO midquote 15 seconds after the trade and the midquote at the time of the trade. Effective spread is twice the signed difference between the trade price and the corresponding midquote. Realized spread is the difference between the effective spread and the corresponding price impact.

	mean	std. dev.	25%	median	75%
Panel A: Activity statistics					
# NBBO updates	5,305	8,169	608	2,470	6,953
# trades	500	1,233	39	113	432
volume, sh.	190,522	485,076	13,438	32,171	120,444
price, $	71.69	36.67	42.08	69.61	92.06
trade size, sh.	448	852	246	311	425
volatility, %	0.154	0.076	0.111	0.158	0.195

Panel B: Trading cost statistics

NBBO, $	0.019	0.024	0.010	0.012	0.019
most constrained	0.010	0.003	0.010	0.010	0.010
least constrained	0.036	0.037	0.018	0.028	0.042
trade price disc., %	0.296	0.159	0.194	0.261	0.350
price impact, $	0.006	0.009	0.000	0.004	0.009
effective spread, $	0.019	0.032	0.010	0.011	0.018
realized spread, $	0.013	0.033	0.002	0.007	0.014

Table 3.3. Microwave connectivity and trading costs

The table contains coefficient estimates from the following panel regression:

$$DEPVAR_{it} = \alpha_0 + \beta_1 PRECIP_t + \beta_2 VIX_t + \varepsilon_{it},$$

where $DEPVAR_{it}$ is one of the following variables: the price impact, $PIMP$, the effective spread, ESP, or the realized spread, RSP, in asset i; $PRECIP$ is total precipitation in the Chicago-New York corridor; and VIX is the volatility index. We also use $PRECIP1$ and $PRECIP2$ to identify the most significant precipitation events. All variables are standardized (by demeaning and scaling by the standard deviation for each stock) and as such the regression models control for asset fixed effects, and the standard errors (in parentheses) are double-clustered along the asset and time dimensions. The data are from the 2011-2012 period. Panel A examines the full sample of 100 ETFs, and Panel B examines the assets for which the minimum tick size is the most (least) binding. For this test, we separate the assets into terciles by their average NBBO on the previous day. Assets with the smallest (largest) NBBOs are considered the most (least) tick-constrained. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% levels.

	$PIMP$			ESP			RSP		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Full sample									
$PRECIP$	-.010***			-.010***			-.005**		
	(.004)			(.003)			(.002)		
$PRECIP1$		-.035***			-.041***			-.024***	
		(.012)			(.010)			(.007)	
$PRECIP2$			-.047***			-.043***			-.021***
			(.013)			(.011)			(.008)
VIX	.035***	.035***	.035***	.057***	.058***	.057***	.036***	.036***	.036***
	(.009)	(.009)	(.009)	(.008)	(.008)	(.008)	(.006)	(.006)	(.006)

Panel B: Effects of $PRECIP2$ for assets that are the most (least) constrained by the minimum tick size (large sample)

	$PIMP$		ESP		RSP	
	most	least	most	least	most	least
$PRECIP2$	-.051***	-.039***	-.023***	-.079***	-.006	-.058***
	(.017)	(.010)	(.008)	(.020)	(.007)	(.017)

Table 3.4. Microwave connectivity, trading activity and volatility

The table contains coefficient estimates from the following panel regression:

$$DEPVAR_{it} = \alpha_0 + \beta_1 PRECIP_t + \beta_2 VIX_t + \varepsilon_{it},$$

where $DEPVAR_{it}$ is one of the following three variables (the number of trades, traded volume, or volatility) in asset i during a 15-minute interval t; $PRECIP$ is total precipitation in the Chicago-New York corridor; and VIX is the volatility index. We also use $PRECIP1$ and $PRECIP2$ to identify the most significant precipitation events. All variables are standardized and as such the regression models control for asset fixed effects, and the standard errors (in parentheses) are double-clustered along the asset and time dimensions. The data are from the 2011-2012 period. Panel A examines the full sample of 100 ETFs, Panel B examines the assets for which the minimum tick size is the least (most) binding. For this test, we separate the assets into terciles by their average NBBO on the previous day. Assets with the smallest (largest) NBBOs are considered the most (least) tick-constrained. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% levels.

	trades			volume			volatility		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Full sample									
$PRECIP$	-.010***			-.012***			-.025**		
	(.006)			(.004)			(.010)		
$PRECIP1$		-.070***			-.044***			-.103***	
		(.020)			(.013)			(.032)	
$PRECIP2$			-.072***			-.042***			-.118***
			(.023)			(.015)			(.036)
VIX	.079***	.079***	.079***	.049***	.050***	.049***	.185***	.186***	.185***
	(.015)	(.015)	(.015)	(.009)	(.009)	(.009)	(.024)	(.024)	(.024)

Panel B: Effects of $PRECIP2$ for assets that are the most (least) constrained by the minimum tick size (large sample)

	trades		volume		volatility	
	most	least	most	least	most	least
$PRECIP2$	-.111***	-.015	-.064***	-.010	-.119***	-.109***
	(.034)	(.021)	(.025)	(.013)	(.038)	(.032)

Table 3.5. Limit order aggressiveness: ITCH sample

The table contains coefficient estimates β_1 from the following panel regression estimated using ITCH limit order book data:

$$DEPVAR_{it} = \alpha_0 + \beta_1 PRECIP_t + \beta_2 VIX_t + \varepsilon_{it},$$

where $DEPVAR_{it}$ is one of the following two variables: (i) the number of limit orders that match the NBBO in ETF i during period t (Panel A), and (ii) the number of orders that match or improve the NBBO (Panel B). Both variables are scaled by the number of order submissions. $PRECIP$ is total precipitation in the Chicago-New York corridor; and VIX is the volatility index. We also use $PRECIP1$ and $PRECIP2$ to identify the most significant precipitation events. All variables are standardized and as such the regression models control for asset fixed effects, and the standard errors (in parentheses) are double-clustered along the asset and time dimensions. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% levels. Note: results reported here are derived from 2012; we are in the process of adding the results from 2011.

	Panel A: NBBO match			Panel B: NBBO match or improve		
	full sample	most constr.	least constr.	full sample	most constr.	least constr.
	(1)	(2)	(3)	(4)	(5)	(6)
$PRECIP$	.017***	.025***	.038***	.006	.008	.041***
	(.004)	(.004)	(.006)	(.004)	(.005)	(.006)
$PRECIP1$	.054***	.080***	.095***	.028***	.006	.107***
	(.007)	(.008)	(.011)	(.006)	(.010)	(.010)
$PRECIP2$	.040***	.068***	.113***	.063***	.010	.126***
	(.012)	(.013)	(.013)	(.012)	(.015)	(.013)

Table 3.6. Microwave connectivity and information asymmetries in futures and equities: small sample

The table contains coefficient estimates from the following panel regression:

$$PIMP_{it} = \alpha_0 + \beta_1 PRECIP_t + \beta_2 VIX_t + \varepsilon_{it},$$

where $PIMP_{it}$ is the price impact; $PRECIP$ is total precipitation in the Chicago-New York corridor; and VIX is the volatility index. We also use $PRECIP1$ and $PRECIP2$ to identify the most significant precipitation events. All variables are standardized, and as such the regression models control for asset fixed effects, and the standard errors (in parentheses) are time-clustered. The futures sample (Panel A) is from 2012 and includes four e-mini contracts: S&P 500, S&P MidCap 400, Nasdaq 100 and Financial Sector Select. The equities sample (Panel B) is from the same time period and includes the four corresponding ETFs. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% levels.

	Panel A: futures			Panel B: equities		
	(1)	(2)	(3)	(4)	(5)	(6)
$PRECIP$	-.004			-.037***		
	(.009)			(.006)		
$PRECIP1$		-.003			-.078***	
		(.022)			(.018)	
$PRECIP2$			-.014			-.079***
			(.031)			(.022)

Table 3.7. Post-democratization period

The table reports the β_1 coefficient estimates from the following panel regression:

$$DEPVAR_{it} = \alpha_0 + \beta_1 PRECIP2_t + \beta_2 VIX_t + \varepsilon_{it},$$

where $DEPVAR_{it}$ is one of the following four variables (price impacts, effective spreads, realized spreads, number of trades, traded volume, or volatility) in asset i during a 15-minute interval t; $PRECIP2$ is a dummy variable that captures periods when total precipitation in the Chicago-New York corridor exceeds one standard deviation; and VIX is the volatility index. All variables are standardized and as such the regression models control for asset fixed effects, and the standard errors (in parentheses) are double-clustered along the asset and time dimensions. The data are from the 2013-2014 period. We examine three groups of assets: (i) 100 ETFs in the full sample, and (ii/iii) the ETF terciles for which the tick size is the most (least) binding. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% levels.

	$PIMP$	ESP	RSP	trades	volume	volatility
	(1)	(2)	(3)	(4)	(5)	(6)
full sample	.007	.001	-.003	.027	.007	.016
	(.013)	(.012)	(.009)	(.018)	(.012)	(.033)
most constr.	-.016	.003	.010	.024	.009	.001
	(.015)	(.008)	(.007)	(.023)	(.018)	(.031)
least constr.	.014	.002	-.006	.031	.005	.023
	(.012)	(.021)	(.016)	(.021)	(.009)	(.032)

Table 3.8. Democratization: Event study

The event window spans the months of September 2012 to April 2013. In this window, the months of September, October and November capture the period prior to the democratization, and the months of February, March and April capture the post-democratization period. We report the coefficient estimates β_1 from the following panel regression:

$$DEPVAR_{it} = \alpha_0 + \gamma_t + \beta_1 POST_t + \beta_2 VIX_t + \varepsilon_{it},$$

where $DEPVAR_{it}$ is one of the following variables (price impacts, effective spreads, realized spreads, number of trades, traded volume, volatility or stock price) in asset i on day t; γ_t denotes a time trend, $POST$ is a dummy variable that equals to one in February-April 2013; and VIX is the volatility index. All variables are standardized and as such the regression models control for asset fixed effects, and the standard errors are double-clustered along the asset and time dimensions. We examine three groups of assets: (i) 100 ETFs in the full sample, and (ii/iii) the ETF terciles for which the tick size is the most (least) binding. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% levels.

	$PIMP$	ESP	RSP	trades	volume	volatility	price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
full sample	-.484***	-.711***	-.546***	-.071	.045	-.836***	.319***
	(.122)	(.127)	(.125)	(.129)	(.125)	(.244)	(.083)
most constr.	-.590***	-.454**	-.190	-.475***	-.174	-1.09***	.496***
	(.185)	(.179)	(.184)	(.171)	(.177)	(.254)	(.135)
least constr.	-.448***	-.965***	-.905***	.084	.095	-.542**	.284**
	(.100)	(.181)	(.178)	(.120)	(.123)	(.222)	(.129)

Table 3.9. Robustness: alternative sampling and regression setup

The table reports the β_1 coefficient estimates from the following panel regression:

$$DEPVAR_{it} = \alpha_0 + \beta_1 PRECIPx_t + \beta_2 VIX_t + \varepsilon_{it},$$

where $DEPVAR_{it}$ is one of the following variables (price impacts, effective spreads, realized spreads, number of trades, traded volume, or volatility) in asset i during a 15-minute interval t; $PRECIP2$ is a dummy variable that captures periods when total precipitation in the Chicago-New York corridor exceeds one standard deviation (Panel A); and VIX is the volatility index. We examine several specifications of the model. The *mood control* specification restricts precipitation episodes to those occurring in Ohio when precipitation is near-zero in the eastern and western parts of the Chicago-New York corridor. The *expanded area* specification uses additional weather stations, forming a wider area around the corridor. The *placebo area* specification uses data from the weather stations located in Colorado, Utah and Wyoming, away from the corridor. The *afternoon only* specification uses data between noon and the market close. The *intraday FE* specification adds intraday fixed effects. Finally, Panel B replaces total precipitation across all stations with the average precipitation per station $MPRECIP$, and its two variations, $MPRECIP1$ and $MPRECIP2$, which are dummies that capture episodes when the average precipitation is more than 0.5 and 1 standard deviation removed from the mean. All variables are standardized and as such the regression models control for asset fixed effects, and the standard errors (in parentheses) are double-clustered along the asset and time dimensions. The data are from the 2011-2012 period, and we examine the full sample of 100 ETFs. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% levels.

	$PIMP$	ESP	RSP	trades	volume	volatility
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: $PRECIPx = PRECIP2$						
mood control	-.060***	-.061***	-.026***	-.094***	-.053***	-.166***
	(.013)	(.012)	(.009)	(.024)	(.016)	(.035)
expanded area	-.034***	-.040***	-.020**	-.055**	-.032**	-.087**
	(.013)	(.012)	(.008)	(.023)	(.015)	(.039)
placebo area	.006	-.012	-.001	.015	.003	-.036
	(.016)	(.025)	(.019)	(.024)	(.016)	(.038)
afternoon only	-.061***	-.063***	-.028***	-.080***	-.048***	-.147***
	(.015)	(.014)	(.010)	(.026)	(.018)	(.040)
intraday FE	-.054***	-.060***	-.028***	-.067***	-.043***	-.141***
	(.012)	(.012)	(.008)	(.021)	(.014)	(.035)
Panel B: $PRECIPx \in \{MPRECIP, MPRECIP1, MPRECIP2\}$						
$MPRECIP$	-.007*	-.006*	-.003	-.014**	-.009**	-.013
	(.004)	(.004)	(.002)	(.006)	(.004)	(.011)
$MPRECIP1$	-.024**	-.027**	-.013*	-.051***	-.030**	-.057*
	(.012)	(.011)	(.007)	(.020)	(.013)	(.034)
$MPRECIP2$	-.043***	-.039***	-.014*	-.067***	-.037***	-.096***
	(.012)	(.011)	(.008)	(.021)	(.014)	(.035)

Table 3.10. Robustness: alternative variables of interest

The table reports the β_1 coefficient estimates from the following panel regression:

$$DEPVAR_{it} = \alpha_0 + \beta_1 PRECIP2_t + \beta_2 VIX_t + \varepsilon_{it},$$

where $DEPVAR_{it}$ is one of the following three variables: price impacts ($PRIMP$), effective spreads (ESP) and realized spreads (RSP). Each variable is computed as equally-weighted ($EW\_$) or volume-weighted scaled by the corresponding midquote ($VWP\_$); $PRECIP2$ is a dummy variable that captures periods when total precipitation in the Chicago-New York corridor exceeds one standard deviation; and VIX is the volatility index. All variables are standardized and as such the regression models control for asset fixed effects, and the standard errors (in parentheses) are double-clustered along the asset and time dimensions. The data are from the 2011-2012 period. We examine three groups of assets: (i) 100 ETFs in the full sample, and (ii/iii) the ETF terciles for which the tick size is the most (least) binding. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% levels.

	EW_PIMP	VWP_PIMP	EW_ESP	VWP_ESP	EW_RSP	VWP_RSP
	(1)	(2)	(3)	(4)	(5)	(6)
full sample	-.064***	-.059***	-.089***	-.067***	-.008	-.032***
	(.018)	(.014)	(.021)	(.015)	(.012)	(.010)
most constr.	-.079***	-.069***	-.086***	-.035***	.030*	-.005
	(.026)	(.020)	(.020)	(.010)	(.018)	(.008)
least constr.	-.046***	-.047***	-.105***	-.109***	-.067***	-.077***
	(.014)	(.011)	(.030)	(.027)	(.026)	(.022)

Table 3.11. Quoted spread and inside depth

The table reports the β_1 coefficient estimates from the following panel regression:

$$DEPVAR_{it} = \alpha_0 + \beta_1 PRECIP2_t + \beta_2 VIX_t + \varepsilon_{it},$$

where $DEPVAR_{it}$ is one of the following four variables (NBBO spread or NBBO inside depth) in asset i during a 15-minute interval t; $PRECIP2$ is a dummy variable that captures periods when total precipitation in the Chicago-New York corridor exceeds one standard deviation; and VIX is the volatility index. All variables are standardized and as such the regression models control for asset fixed effects, and the standard errors (in parentheses) are double-clustered along the asset and time dimensions. The data are from the 2011-2012 period. We examine three groups of assets: (i) 100 ETFs in the full sample, and (ii/iii) the ETF terciles for which the tick size is the most (least) binding. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% levels.

	QSP	$DEPTH$
	(1)	(2)
full sample	-.065***	.014
	(.020)	(.029)
most constr.	-.026**	.118***
	(.012)	(.045)
least constr.	-.105***	-.013
	(.034)	(.034)

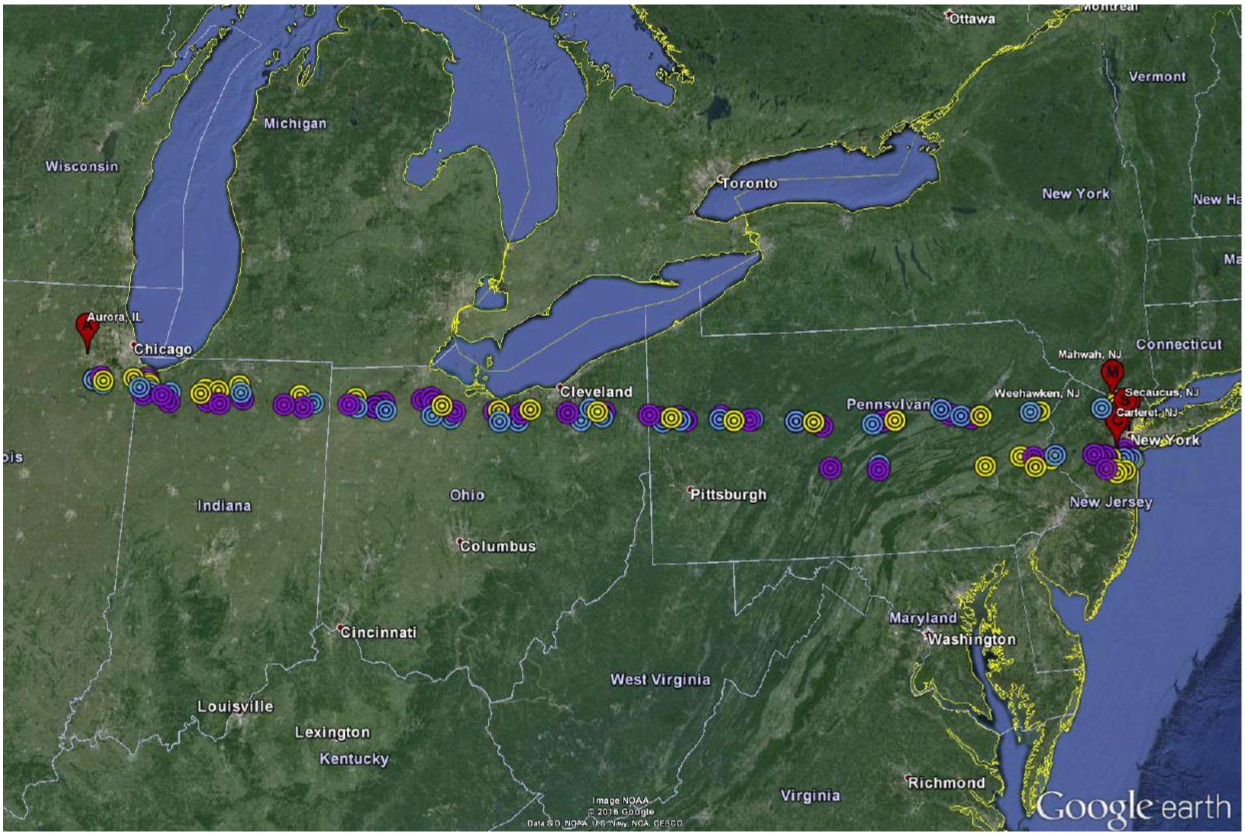


Figure 3.1. Microwave network paths

The figure maps tower locations of three microwave networks (blue, yellow and purple icons) obtained from the Federal Communications Commission. There are more than three microwave networks between Chicago and New York during our sample period; however, we plot only three to avoid clutter. The remaining networks follow very similar paths. The red markers indicate locations of the CME's data center in Aurora, IL (marker A); the NYSE data center in Mahwah, NJ (marker M); Nasdaq data center in Carteret, NJ (marker C); BATS data center in Weehawken, NJ (marker W); and Direct Edge data center in Secaucus, NJ (marker S).

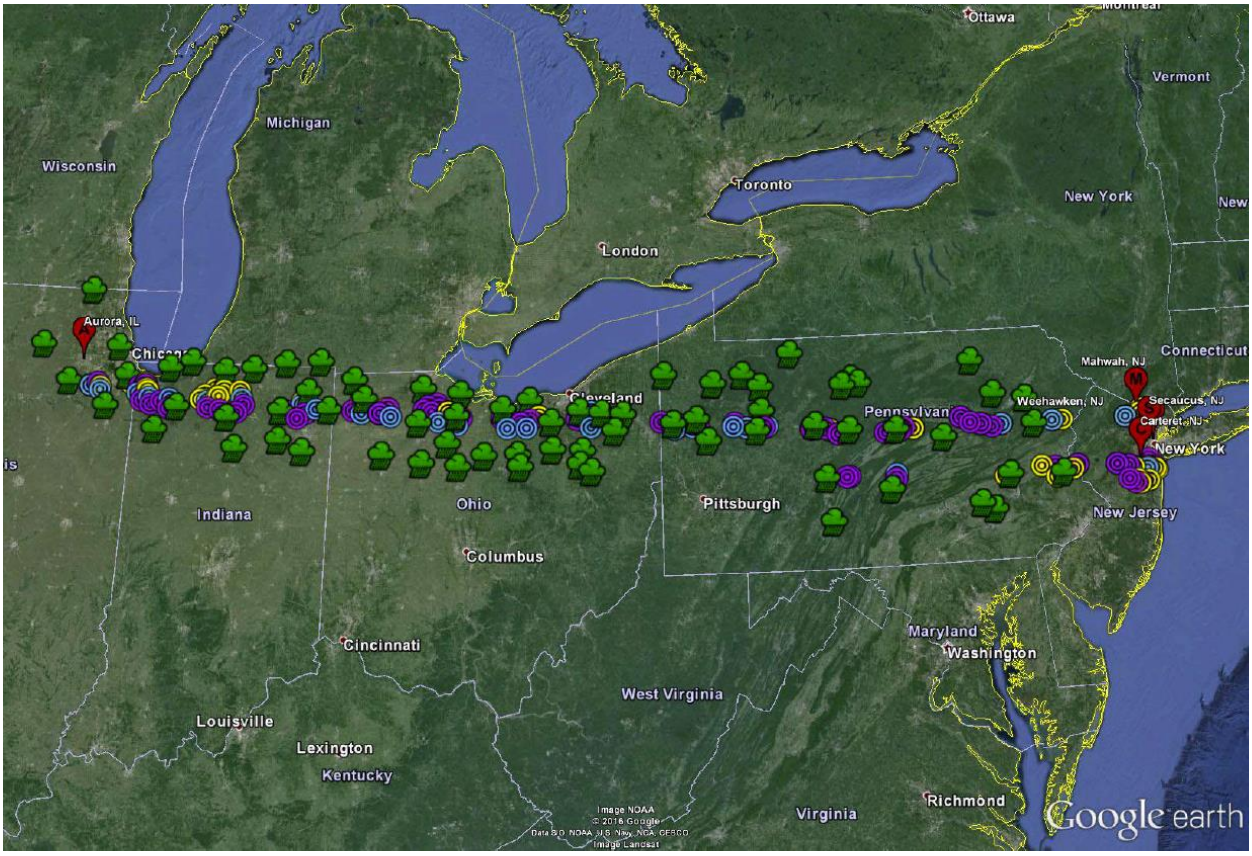


Figure 3.2. Locations of microwave networks and weather stations

The figure maps the weather stations (green icons) located near the microwave network paths. Station data are obtained from the National Oceanic and Atmospheric Administration. The red markers indicate locations of the CME's data center in Aurora, IL (marker A); the NYSE data center in Mahwah, NJ (marker M); Nasdaq data center in Carteret, NJ (marker C); BATS data center in Weehawken, NJ (marker W); and Direct Edge data center in Secaucus, NJ (marker S).

Figure 3.3. A typical weather front

As a weather front moves over the microwave paths, it disrupts data transmission forcing trading firms to fall back on the fiber-optic cable.

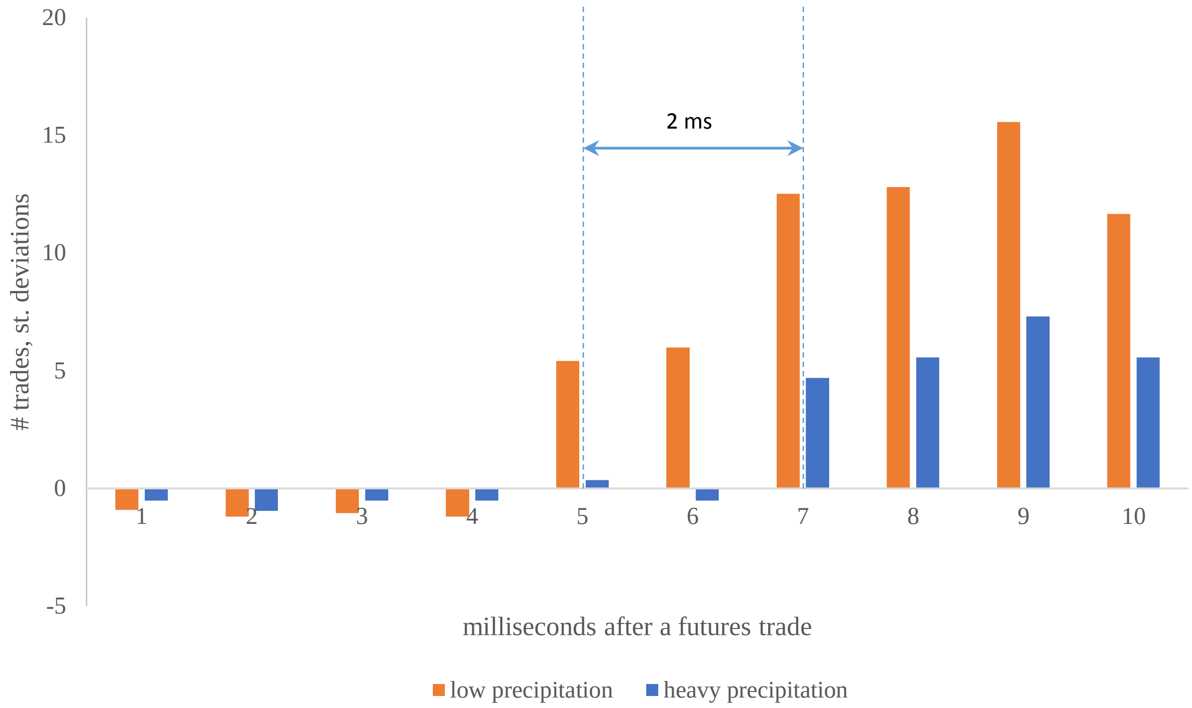


Figure 3.4. Equity trades after a futures trade during low and heavy precipitation episodes

The figure reports a timeline of equity trades that follow a futures trade. Orange bars represent periods of zero or very low precipitation ($PRECIP2$<0), and blue bars represent periods of heavy precipitation ($PRECIP2$>1) when the microwave networks are disrupted. The number of trades is standardized on an asset by asset basis to allow for cross-sectional comparability. We focus on the standalone futures trades (t = 0), those not preceded by another futures or equity trade in the previous 100 milliseconds. Note: light covers the distance from Chicago to New York in 4 milliseconds (ms), the microwave signal covers this distance in about 4.5 ms, and the same signal takes 6.5 ms to cover the distance through fiber. During our sample period, the CME clock lags DTAQ by about one millisecond, and we adjust for this lag. The 2-ms advantage of the fastest traders is evident even without the adjustment.

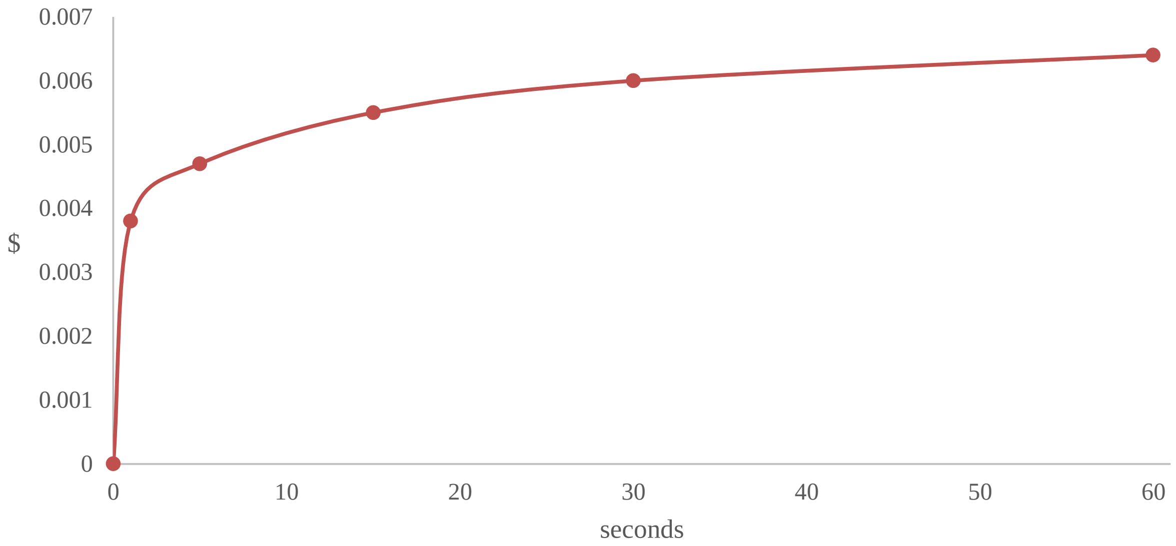


Figure 3.5. Price impacts

The figure reports price impacts computed as the signed difference between a midquote at a certain time after the trade and the midquote at the time of the trade: $PRIMP_t = 2q_t\left(mid_{t+\gamma} - mid_t\right)$, where q_t is the trade direction indicator, mid_t is the midquote computed as $(NBBO\ Ask_t + NBBO\ Bid_t)/2$, and γ indicates the time elapsed since the trade, with $\gamma \in \{1s, 5s, 15s, 30s, 60s\}$.

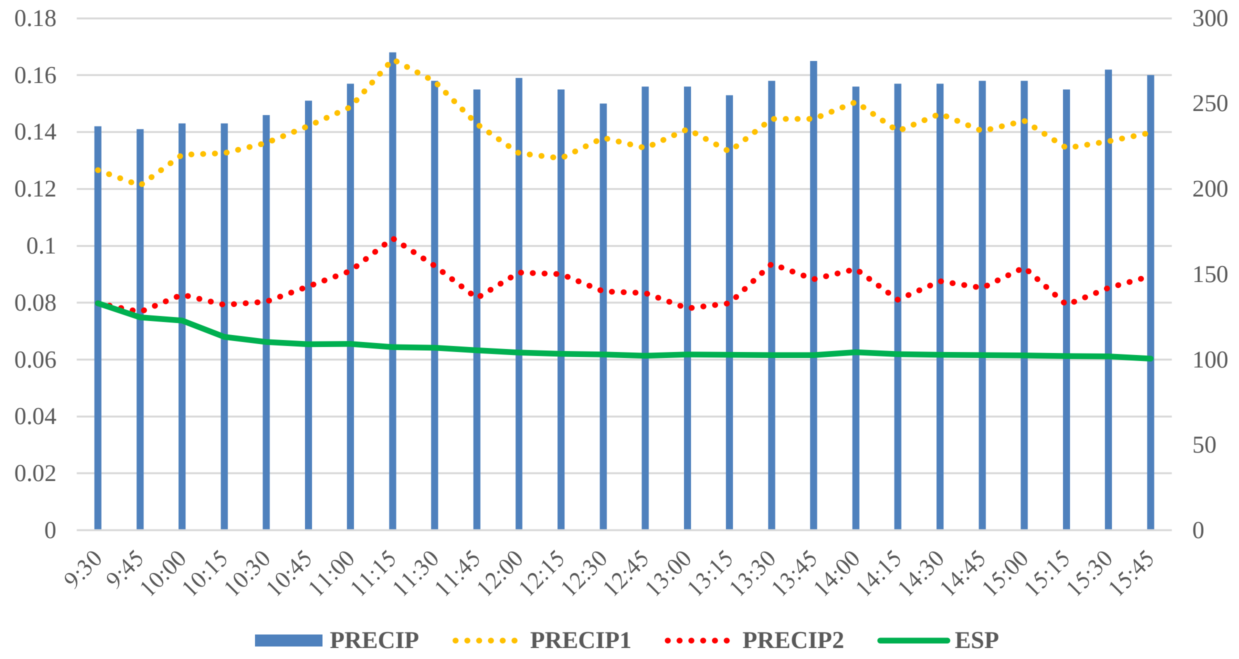


Figure 3.6. Intraday patterns

The figure reports intraday patterns for $PRECIP$ (in mm average per intraday period, left axis), $PRECIP1$ and $PRECIP2$ (both in number of occasions per intraday period, right axis), and ESP (scaled by 10000 for display purposes, right axis).

Appendix to Every cloud has a silver lining: Fast trading, microwave connectivity and trading costs

A1. Price discovery via trades and quotes

To examine price discovery via trades and quotes, we use the methodology described in Hasbrouck (1991 a,b) to decompose the efficient price variance into the trade-related and trade-unrelated components. We begin with an assumption that the observed midquotes p_t follow a random walk with two components:

$$p_t = m_t + s_t,$$

where m_t is the efficient price (the expectation of price conditioned on all available information at time t), and s_t is a deviation of the price from the efficient price. We then estimate the VAR with ten lags as follows:

$$r_t = a_1 r_{t-1} + a_2 r_{t-2} + \cdots + b_0 q_t + \ b_1 q_{t-1} + b_2 q_{t-2} + \cdots \ + v_{r,t}$$

$$q_t = c_1 r_{t-1} + c_2 r_{t-2} + \cdots + d_1 q_{t-1} + d_2 q_{t-2} + \cdots \ + v_{q,t},$$

where r_t is the difference in log-midquotes, and q_t is a vector of three trade-related variables, including a trade direction indicator, signed volume and signed square root of volume. The VAR is then converted into the VMA model:

$$r_t = a_1^* v_{r,t-1} + a_2^* v_{r,t-2} + \cdots + b_0^* v_{q,t} + b_1^* v_{q,t-1} + b_2^* v_{q,t-2} + \cdots$$

$$q_t = c_1^* v_{r,t-1} + c_2^* v_{r,t-2} + \cdots + d_1^* v_{q,t-1} + d_2^* v_{q,t-2} + \cdots \ ,$$

and the total variance of the random walk component is given by:

$$\sigma_w^2 = (1 + \textstyle\sum_{i=1}^{\infty} a_i^*)^2 \sigma_r^2 + (\sum_{i=0}^{\infty} b_i^*)\Omega(\sum_{i=0}^{\infty} b_i^{*\prime}),$$

where the first term corresponds to the trade-unrelated component of the efficient price innovation, and the second term corresponds to the trade-related component of this innovation. The model is estimated in event time, with ts indexing every new midquote.

A2. Information share estimation

To compute information shares using the methodology in Hasbrouck (1995), we first estimate the following vector error correction model (VECM) for each futures-ETF pair:

$$\Delta p_{f,t} = \alpha_1 \Delta p_{f,t-1} + \cdots + \alpha_k \Delta p_{f,t-k} + \beta_0 \Delta p_{e,t-1} + \cdots + \beta_k \Delta p_{e,t-k} + g_1 (p_{f,t-1} - p_{e,t-1} - \mu) + u_{f,t}$$
$$\Delta p_{e,t} = \gamma_1 \Delta p_{f,t-1} + \cdots + \gamma_k \Delta p_{f,t-k} + \delta_0 \Delta p_{e,t-1} + \cdots + \delta_k \Delta p_{e,t-k} + g_2 (p_{f,t-1} - p_{e,t-1} - \mu) + u_{e,t},$$

where $\Delta p_{f,t}$ ($\Delta p_{e,t}$) is the difference between the current and lagged prices of the futures (ETF), and μ is the mean difference between the price of the futures and the ETF.

In the second step, we obtain the VMA representation of the above model:

$$\Delta p_{f,t} = a_0 u_{f,t} + \cdots + a_k u_{f,t-k} + b_0 u_{e,t} + \cdots + b_k u_{e,t-k}$$
$$\Delta p_{e,t} = c_0 u_{f,t} + \cdots + c_k u_{f,t-k} + d_0 u_{e,t} + \cdots + d_k u_{e,t-k}$$

and add the coefficients $A = \sum_0^k a_i$ and $B = \sum_0^k b_i$. Next we obtain the covariance matrix of the residuals:

$$\Omega = \begin{bmatrix} \sigma_f^2 & \sigma_{ef} \\ \sigma_{ef} & \sigma_e^2 \end{bmatrix},$$

and finally, the information share (IS) of the futures market is calculated as:

$$IS_f = \frac{A^2 \sigma_f^2}{\sigma_W^2}, \text{ where } \sigma_W^2 = \begin{bmatrix} A \\ B \end{bmatrix}' \begin{bmatrix} \sigma_f^2 & \sigma_{ef} \\ \sigma_{ef} & \sigma_e^2 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}.$$

Since some price innovations happen in both markets within the same millisecond, $\sigma_{ef} \neq 0$. To address this, we follow Hasbrouck (1995) and orthogonalize Ω. Orthogonalization maximises (minimises) the variance of the futures market and gives the upper (lower) bound of the true variance.