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

Andriy Shkilko*
Wilfrid Laurier University

Konstantin Sokolov
Wilfrid Laurier University

First version: March 2016
This version: November 2016

Abstract: 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, we study a series of exogenous weather-related episodes that temporarily remove 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, resulting in liquidity improvements.

JEL classification: G10; G14

*Correspondence: ashkilko@wlu.ca (A. Shkilko).

We thank Jim Angel, Matt Barron, Robert Battalio, Michael Brolley, Jean-Edouard Colliard, Sean Foley, Thierry Foucault, Austin Gerig, Michael Goldstein, Björn Hagströmer, Terrence Hendershott, Sophie Moinas (Central Bank Workshop discussant), Andreas Park, Fabricio Perez, Ryan Riordan, René Wells, Bart Yueshen, Yelin Zhang, Haoxiang Zhu, and the audiences at the 12th Annual Central Bank Workshop on the Microstructure of Financial Markets, Northern Finance Association annual meeting, the University of South Australia, University of Sydney and Wilfrid Laurier University for insightful comments. Stuart Hinson from the NOAA National Climatic Data Center and Jireh Ray from the CME have provided detailed guidance on the use of data.

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

Abstract: 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, we study a series of exogenous weather-related episodes that temporarily remove 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, resulting in liquidity improvements.

1. Introduction

In recent years, information processing and order transmission speeds in financial markets have increased significantly and are measured in fractions of a second. Being faster than others has important advantages. First, the fastest trader is the first to change his limit orders in response to new information, thus avoiding being picked off by others. Second, he himself may choose to pick off slower traders. Given the importance of being the fastest, trading firms generously invest in new technology in a race to bring information transmission speeds closer to the speed of light.

This speed race creates a marketplace where some firms are faster than others. In its concept release on equity market structure, the SEC (2010) notes that differences in speeds between market participants may hurt liquidity. Harris (2013) echoes this concern and points out that if liquidity providers are even marginally slower than the fastest traders, they are at risk of being adversely selected. Recognizing this risk, liquidity providers will quote wider spreads. Several theory models support this notion, suggesting that when speed differentials between traders exist, adverse selection may increase and liquidity may become more expensive (e.g., Hoffmann, 2014; Biais, Foucault and Moinas, 2015; Budish, Cramton and Shim, 2015; Foucault, Hombert and Roşu, 2016).

In this study, we use a previously unexplored dependence between precipitation (i.e., rain and snow) and information transmission speed to build a multi-year time series of intraday speed differentials. Specifically, we use the fact that precipitation disrupts microwave networks used by select traders to transmit information between Chicago and New York. During such disruptions, these traders must fall back on the slower fiber-optic cable, losing the speed advantage. We show that when this happens, adverse selection, trading costs and volatility decline. As such, rain clouds come with a silver lining.

The first microwave network that linked the markets in 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 transmissions was limited to a small group of trading firms, because the Federal Communications Commission, citing airwave congestion, restricted the number of network licenses. 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 information transmission speeds and others do not. Our results linking precipitation episodes to lower adverse selection and trading costs come from this period.

In winter of 2012-2013, one of the microwave technology providers in the Chicago-New York corridor introduced a new business model that democratized microwave transmissions. Instead of selling bandwidth that traders could use to outpace others, the provider began to use its microwave network to transmit the latest price updates between Chicago and New York and sell them to anyone on a subscription basis. As a result, the advantages previously enjoyed by select firms that had access to microwave networks were diminished. We find that once information transmission is democratized in this manner, adverse selection, trading costs and volatility decline. In summary, precipitation events serve as short-term speed equalizers, while the democratization event has similar effects in the long run.

O'Hara (2015) and Brogaard, Hendershott and Riordan (2016) show that fast informed traders often use limit orders. Yao and Ye (2015) find that modern traders use speed to establish time priority in limit order queues. In the meantime, theory models often assume that fast traders choose marketable orders to pick off slower traders (e.g., Biais, Foucault and Moinas, 2015; Budish, Cramton and Shim, 2015; Foucault, Hombert and Roşu, 2016; Foucault, Kozhan and Tham, 2016). Our results reconcile these notions. First, consistent with the first group of studies, we confirm that quotes usually contribute more to price discovery than trades. Second, consistent with the second group, trade price impacts decline significantly when fast traders lose their speed advantage, whether due to precipitation or democratization. Taken together, these findings suggest that even though limit orders usually lead price discovery, there remains ample room for marketable orders to bring information into prices. We note that the use of marketable orders should be particularly widespread in assets with tight spreads, in which the traders' ability to post aggressively priced limit orders is diminished. Consistent with this argument, when the microwave networks are disrupted, the largest adverse selection reductions occur in assets with tight spreads.

Our results suggest that liquidity providers are not always on the forefront of the latest technology. As such, this study provides a complementary perspective to that of Brogaard, Hagströmer, Nordén and Riordan (2015), who show that in the Swedish market speed advantages of colocation are mainly sought by liquidity suppliers and as such benefit liquidity. Focusing on the U.S. market, we find that even though faster traders may occasionally choose to supply liquidity, the net effect of speed differentials on market quality is unfavourable. This 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. In our setting, traders who transmit information discovered in the futures market to the underlying equities (effectively arbitraging prices in Chicago and New York) mainly generate marketable orders that adversely select liquidity providers.

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 existing literature is as follows. First, we shed new light on predictions of theory models that examine trader speed differentials. Second, we provide new insights into order choices of the fastest traders. Third, we offer evidence complementary to existing empirical research that finds that market makers are the ones most interested in accessing new trading technology. Finally, we describe a new panel approach to measuring the speed of information transmission that, to our knowledge, has not been previously examined in the literature.

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

2. Institutional background and related literature

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 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 (hereafter, MWNs) advertise round-trip information transmission speeds that are about 30% faster than their fiber-optic competitors.

Although faster than cable, MWNs have a disadvantage – they are relatively easily disrupted. Among the most important MWN disruptors are rain droplets and snowflakes, especially when rainfall/snowfall is substantial. During weather disruptions, traders who use MW links lose their speed advantage and must either stop trading or transition from microwave to fiber transmissions. This transition is automatic and does not require human involvement. As such, precipitation serves as a natural exogenous equalizer of information transmission speed.

2.2. Information transmission speed and liquidity

We use the susceptibility of MWNs to precipitation disruptions to examine the effects of differential trader speeds on liquidity. The speed-related effects have been extensively modeled in recent literature, with Menkveld (2016) providing a comprehensive review. For instance, Biais, Foucault and Moinas (2015), Budish, Cramton and Shim (2015), Foucault, Hombert and Roşu (2016) and Foucault, Kozhan and Tham (2016) model a market where some traders receive and act on new information faster than market makers. These traders generate adverse selection that in turn may force market makers to seek higher compensation for providing liquidity, thereby increasing liquidity costs for all market participants.

Menkveld and Zoican (2015) suggest that negative liquidity effects may arise even when liquidity demanders and liquidity providers are equally fast. In their setting, there are three types of market participants: fast informed liquidity takers, fast informed liquidity providers, and slower uninformed traders. Allowing the already fast traders to become even faster harms liquidity because the probability of encountering a slow trader declines, causing market makers to widen their quotes.

Hoffmann (2014) and Jovanovic and Menkveld (2015) show that when some market makers become fast they can avoid being adversely selected and therefore may 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 if one assumes declining marginal gains from trade. Du and Zhu (2015) and Roşu (2015) suggest that when some traders are faster than others, volatility may increase.

2.3. Information flows between the futures and equity markets

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 Chicago and New York 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 to some readers because arbitrageurs are often viewed as liquidity providers who enhance market efficiency. Specifically, 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 therefore 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. Data and sample

Our analysis is based on millisecond DTAQ data. The sample period spans four years, from January 2011 through December 2014. The first two years (2011-2012) are characterised by the proliferation of 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 very similar. For illustration, Figure 1 reports tower locations of three select networks connecting Chicago to the New York data centers. The data on tower locations is 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 1 only maps three microwave networks; FCC data show that all networks follow similar paths.

[Figure 1]

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 2). In the robustness section, we examine station samples of different sizes.

[Figure 2]

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 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.

[Figure 3]

Statistics reported in Panel A of Table 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.

[Table 1]

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; Hasbrouck, 2003; Roll, Schwartz and Subrahmanyam, 2007). Given that the most active futures contracts track baskets of securities (e.g., among the most active E-minis are those tracking the S&P 500 and the NASDAQ-100 indexes), our focus in the equity market is on the ETFs. As long as price discovery via futures contracts 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. The results show that price discovery from index futures to the underlying equity ETFs is anything but trivial.

We use millisecond DTAQ data for two asset samples. The first (small) sample consists of five ETFs: SPY (SPDR S&P500), XLF (Financial Select Sector SPDR), TLT (iShares 20+ Year Treasury Bond), SDS (ProShares UltraShort S&P500), and GLD (SPDR Gold Shares). These assets are among the most active in the New York equity markets and are closely related to the active futures contracts that trade in Chicago. SPY in particular is linked to the E-mini – the most actively traded index futures. The upside of using this sample is that Laughlin, Aguirre and Grundfest (2014) show that price discovery in its five constituents strongly depends on Chicago-New York information transfers.[1] The downside is the small size of the cross section.

To address the problem posed by the cross section size, we examine an additional (large) sample that includes 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 1).

Many ETFs in our sample track the same baskets of securities as the CME futures contracts. For example, the QQQ ETF and the CME's E-mini NASDAQ 100 futures track the same index.

[1] Laughlin, Aguirre and Grundfest (2014) also use VXX (iPath S&P500 VIX ETF). We use this ETF to proxy for intraday levels of the VIX index, so we do not include it in the sample.

The remaining ETFs do not have perfect futures counterparts, but track baskets similar to those of major CME contracts. As an 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.[2] As such, we expect these ETFs to react to information discovered in Chicago as long as the information is relevant to some of the constituents in the underlying basket.

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. Table 2 reports descriptive statistics for the two samples: the small sample of 5 ETFs and the large sample of 100 ETFs. Because precipitation data are in 15-minute intervals, we aggregate the millisecond DTAQ data accordingly.

An average ETF in the small sample has 21,603 NBBO updates every 15 minutes, equivalent to 24 updates per second. In addition, this ETF trades 3,228 times every 15 minutes, for a total volume of 1,308,086 shares (Panel A). Because the ETFs used by Laughlin, Aguirre and Grundfest (2014) are among the most active, quoting and trading activity in the sample of 100 ETFs (Panel B) is expectedly less intensive. Specifically, an average ETF in this sample has 5,305 NBBO updates every 15 minutes, equivalent to about 6 updates per second, and trades 500 times every 15 minutes, for a total volume of 190,522 shares.

[2] The CME delisted E-mini Russell 1000 futures contract in 2007 when Russell Investments sold licensing rights to the Intercontinental Exchange. The CME relisted the contract in 2015.

[Table 2]

3.4. Picking-off risk

Recent literature suggests that fast informed traders increasingly choose to trade via limit orders. For instance, Brogaard, Hendershott and Riordan (2015) show that limit orders play a significant role in price discovery in the Canadian market. Using a U.S. dataset, O'Hara (2015) also argues that fast informed traders often prefer limit to marketable orders. She however suggests 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, she may choose to pick off the outstanding ask quote despite having to pay the spread. As such, picking-off risk should be higher in assets with binding tick sizes.

Our samples, and especially the smaller sample of very active ETFs, are quite liquid and therefore are likely to be constrained by the minimum tick size. Panel A of Table 2 shows that the NBBOs in the small sample average 1.1 cents, with more than half of NBBOs at exactly 1 cent. Tick size is often binding in the large sample as well, with at least 25% of the NBBOs at 1 cent (Panel B). Given these constraints, trade-related price discovery and the associated picking-off risk should be of considerable importance.

To further examine this assertion, we compute two metrics. First, we estimate a share of price discovery attributable to trades. Second, we compute the price impacts of trades. The share of price discovery metric follows Hasbrouck's (1991 a,b) and decomposes the efficient price variance into the trade-related and trade-unrelated components. The details on this calculation are in the Appendix. The results in Table 2 show that the trade-related component amounts to 29.1% in the small sample and 29.6% in the large sample. As such, new information is incorporated into prices through trades rather often, and therefore concerns with the picking-off risks 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 certain time after the trade and the midquote at the time of the trade: $PRIMP_t = 2q_t(mid_{t+\gamma} - mid_t)$, 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 4 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, a significant share of information is incorporated into the midquotes within a second after the trade and the incorporation is, expectedly, faster in more frequently traded (small sample) ETFs. This said, information incorporation continues beyond the first second in both samples and takes up to 60 seconds in the less active (large) ETFs. As such, although it may be tempting to think that full quote adjustments in modern markets happen in sub-second periods, the data suggest that this is not the case. To account for this characteristic, we focus on 15-second intervals, with robustness checks examining intervals between 1 and 60 seconds.

[Figure 4]

It may not be immediately obvious that there should be enough adverse selection in ETFs to warrant this result. After all, ETFs are baskets of many securities, and as such the idiosyncratic risk associated with these securities is relatively low. This said, as long as sufficient amounts of macro information are present, the price impacts in ETFs should be quite sizeable. In fact, the price impacts reported in Figure 4 are comparable to those obtained for individual stocks. Specifically, in our two samples price impacts are 30-40% of the effective spread. In a study that examines a recent sample of large U.S. equities, Chakrabarty, Jain, Shkilko and Sokolov (2015) find that the price impact is 35% 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.

3.5. Trading costs and liquidity provider revenues

Table 2 also reports liquidity costs and liquidity provider revenues proxied by, respectively, 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: $ESP_t = 2q_t(p_t - mid_t)$. In turn, RSP_t is computed as twice the difference between the effective spread and the price impact. We volume-weight effective and realized spreads. Although the median effective spread is equal to the NBBO spread in the small sample (Panel A of Table 2), the average effective spread is almost twice as large as the NBBO. This statistic suggests that trades in the small sample occasionally occur outside of the best quotes, perhaps due to the use of ISO orders as described by Chakravarty, Jain, Upson and Wood (2012). These orders are permitted, in some circumstances, to take liquidity located beyond the best quotes (Rule 611 of Reg NMS).

4. Empirical findings

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 authors model traders, who are faster than others, and therefore their limit orders are less exposed to picking-off risk. Others model traders who use their speed advantage to pick off the limit orders of others, resulting in greater adverse selection. In this section, we aim to better understand which of these assumptions prevails in our setting.

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 the regression models control for asset fixed effects. Additionally, all time-series variables are standardized by year, and the standard errors are double-clustered along the asset and time dimensions.

Table 3 shows that price impacts decline during network disruptions. For example in the large sample, significant amounts of precipitation captured by $PRECIP2$ are associated with a 0.047 standard deviations, or 7.05%, decline in price impacts (Panel B of Table 3). As such, it appears that the fast traders prefer marketable orders on average.

[Table 3]

MWN disruptions slow down information transmission by mere milliseconds. Is it surprising that we observe changes in price impacts over 15-second periods? We argue that it is not. Price impacts are based on midquotes and hence proxy for the speed with which limit orders adjust to new information. Not all limit order traders are fast, and Figure 4 shows that full adjustments normally happen over multi-second periods. More importantly however, our focus is on how (not how fast) information is incorporated into prices. A smaller 15-second price impact therefore means that when fast traders lose their speed advantage, more information is incorporated into prices through quotes than through trades.

Per our earlier suggestion, fast traders may be forced to use marketable orders 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 C 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. These results are in line with Carrion (2013) and Hendershott and Riordan (2013), who show that algorithmic traders tend to supply liquidity when spreads are wide. In our setting however, liquidity taking by the fastest traders appears to dominate even when the tick size is not binding.

4.2. Trading costs and liquidity provider revenues

Some theory models assume that traders use speed advantages to post better priced limit orders. A loss of the speed advantage may force these traders to price limit orders less aggressively to compensate for greater picking-off risk, and trading costs may increase. Other models assume that traders use speed advantages to pick off the limit orders of others. A loss of the speed advantage may reduce adverse selection and therefore attract better priced limit orders. Consequently, trading costs may decline. We note that the abovementioned models are equilibrium models. It is not immediately clear how quickly an equilibrium may arise, and if it may arise at all given that the precipitation episodes are relatively short. Easley and O'Hara (1992) describe price adjustment as a process of learning by market participants. In their model, traders learn by observing the time between trades, with spreads decreasing gradually as this time increases. Whether trading costs have enough time to adjust to lower adverse selection during microwave network disruptions is an empirical question that we examine next.

In Table 3, we report eq. 1 regression results for effective and realized spreads. The data show that effective spreads decline during MWN disruptions, especially in the large sample where the tick size is less binding. The $PRECIP2$ coefficient suggests that effective spreads decline by 0.043 standard deviations (Panel B). Expectedly, this result is more pronounced for the least constrained assets (Panel C). 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, it appears that the length of an average precipitation episode is sufficient for the spreads 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.

When it comes to realized spreads, they too decline, by 0.021 standard deviations, during $PRECIP2$ events. As such, network disruptions appear to not only reduce liquidity costs, but also reduce liquidity provider revenues. There are several possible explanations for this result. One explanation involves latent liquidity that stays on the sidelines when the speed differentials are present. Chakrabarty, Jain, Shkilko and Sokolov (2015) use limit order book data to show that favorable conditions incentivise liquidity providers to reposition existing liquidity from the deeper layers of the book to the inside. An alternative explanation may have to do with MWN traders switching from taking liquidity to supplying new limit orders. Since our data do not contain trader account information or limit order book information, we must rely on future research to reconcile these explanations.

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, reducing the arbitrageurs' trading interest. The regression results in Table 4 are consistent with this notion. For instance, in the large sample, the number of trades declines by 0.072 standard deviations during $PRECIP1$ events. Trading volume also declines; by 0.042 standard deviations.

[Table 4]

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 model by Roşu (2015) suggests that as fast traders pick off market makers' quotes, volatility may increase. Du and Zhu (2015) 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 price during an interval scaled by the average price, declines by 0.118 (0.109 standard deviations) in the large (small) sample during $PRECIP2$ events.

As we point our earlier, fast traders have more maneuvering room in assets with wider spreads. In such assets, new information may be incorporated into prices through both marketable orders and better priced limit orders. Meanwhile, in assets whose spreads are constrained by the minimum tick size, fast traders are not always able to improve the NBBO and occasionally have to rely on marketable orders. Naturally, these considerations should affect changes in trading activity during network downtimes. Panel C shows that the number of trades and trading volume decline in the most constrained assets, yet remain the same in the least constrained assets. These results corroborate two of our earlier conjectures: (i) that fast traders use fewer marketable orders in the least constrained stocks and (ii) that trading volume may increase in response to lower trading costs.

Overall, the results suggest that even though lower spreads may attract additional trading interest, trading volume generated by this interest is not larger than the lost arbitrage volume. One possibility is that the disruptions are not long enough or 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 though the high cost periods are occasionally interrupted by the lower 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.

4.4. The futures market during microwave network disruptions

Our main analysis focuses on the effects of speed differentials on the equity markets in New York. In this section, we ask if the differentials also affect the futures market in Chicago. On the one hand, fast traders may carry information not only from futures to equities but also in the opposite direction. If this is the case, speed differentials may negatively affect CME liquidity. On the other hand, prior research shows that futures provide the lion's share of price discovery in index instruments. If so, it would be wasteful to use the limited microwave bandwidth to transmit information from the ETFs to futures. As such, speed differentials may not have much of an effect on the CME liquidity.

To examine futures liquidity, we obtain 2012 intraday CME data for four e-mini contracts: S&P 500, S&P MidCap 400, Nasdaq 100 and Financial Sector Select. Intraday data for these four contracts are sold by the CME as a bundle. First, we ask how the futures and equity markets react to each other's trades. Using S&P 500 contract as an example, Figure 5 reports a significant reaction of the ETF to trades in the futures, but not vice versa. The ETF reaction begins about five milliseconds after a CME trade, the time required for a signal from Chicago to reach New York. As such, it appears that the futures market leads equities.

[Figure 5]

To confirm this graphic result, we compute information shares as in Hasbrouck (1995) and Westerlund, Reese and Narayan (2014) for all four futures-ETF pairs. The details of the methodologies are in the Appendix. These methodologies corroborate the graphic result; in all sample contracts, price discovery occurs virtually entirely on the CME (not tabulated). Finally, we examine CME liquidity during precipitation episodes. Consistent with limited price discovery from ETFs to futures, the data show that precipitation does not affect adverse selection and trading costs on the CME (Table 5).

[Table 5]

4.5. Democratization of MWN access

In late December 2012 – early January 2013, one of the microwave technology providers disrupted the business model used by the MWN firms. Instead of selling bandwidth on its network to select traders, it 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 at least equally as fast as (if not faster than) the existing MWNs. The offer was soon replicated by other providers, and the market for microwave transmissions was democratized. Put differently, existing microwave users lost their speed advantage.

Given our findings so far, democratization of access to higher information transmission speeds may lead to two outcomes. First, the relation between precipitation and market quality observed in the 2011-2012 period may diminish. Second, the democratization should 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 6, we report the coefficients of the $PRECIP2$ variable obtained from estimating eq. 1 during the post-democratization period. The results confirm the abovementioned expectations. Precipitation episodes no longer have an effect on price impacts, effective spreads, realized spreads, volatility and trading activity. The change is observed for both samples (small and large) and for the most and least constrained subsamples. As such, democratization was a significant market disruptor.

[Table 6]

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

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

where $DEPVAR_{it}$ is one of the following variables (price impacts, effective spreads, realized spreads, the number of trades, traded volume, or volatility) in asset i on day t, $POST$ 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. 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. The regression results for the large sample indicate that price impacts, effective spreads, realized spreads and volatility decline post-democratization (Table 7). As such, the event study findings are consistent with the panel findings discussed earlier.

[Table 7]

One notable difference between the event study and the panel findings comes from the trade-related variables: number of trades and volume. In the large sample, both of these variables increase post-democratization. As such, the results corroborate our earlier suggestion that lower trading costs may generate new trading interest over periods of time that are longer than the weather-related MWN disruptions. Finally, the results for the most and least constrained samples generally confirm our earlier findings. Notably, the loss of arbitrage opportunities in the most constrained ETFs (including the small sample ETFs) appears to be equivalent to the gain from new trading interest, resulting in the absence of changes in trading activity.

It may not be immediately clear what drove the technology provider to disrupt the status quo in the market for microwave transmissions. Cespa and Foucault (2014) argue that it may be in data providers' best interests to restrict dissemination of pricing data to select traders. Their model shows that information that is accessible to many is less valuable to the few, who may be willing to pay a premium for the exclusive use of this information. Our investigation into the technology firm's motives suggests that the firm believed that increasing the 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 maintain a sufficiently large customer base. It remains unclear whether this belief was correct given that many other connectivity providers launched similar offerings in the months after the democratization, possibly driving down the price of the service. This said, the firm continues to offer the service to this day and has expanded it to several continents.

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 reflects 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 following earnings announcements. As such, the reduction in adverse selection that we document in the previous section 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 MWN disruptions.

To examine this possibility, we recalculate the $PRECIP2$ variable to capture time periods when the networks are disrupted, yet the moods of traders in Chicago and New York are not altered. 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 large ETF sample in 2011-2012 and report the results in the *mood control* specification in Panel A of Table 8. The effects are consistent with those reported in the earlier tables. The results for the small sample and the two subsamples are similar. As such, trader moods do not seem to be the source of our findings.

[Table 8]

Our sample of weather stations is selected to capture the area closely surrounding the MWN paths. As with any such selection procedure, it is important to show that the results are not driven by the choice of the specific set of stations. The *mood control* specification takes the first step in this direction by restricting the sample to Ohio stations. In two additional Table 8 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 then narrow in the afternoon (Figure 6). Notably, intraday precipitation too follows a reverse J-pattern, with precipitation amounts being lower in the morning hours. Since the results in previous section point to a negative relation between precipitation and spreads, we must make sure that the findings are not due to these intraday patterns.

[Figure 6]

We examine this possibility in two additional specifications in Panel A of Table 8. 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.

Specifications that focus on the afternoon period bring an extra benefit as they take away the possibility that the results are driven by an occasional morning fog. Like rain and snow, fog droplets disrupt microwave transmissions, yet they are suspended in the air and may not register accurately with the weather stations. Along the MWN paths, fog is mainly observed at night and in the early morning hours before the markets open, as such it is not a significant concern for our main analysis. Still, it is encouraging that the results remain strong during the afternoon periods when fog is normally absent.

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 events, such as atmospheric pressure changes or dew accumulations, extend 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 8 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). Corroborating this reasoning, the results for $MPRECIP$ are weaker than those reported earlier for $PRECIP$, yet the results for $MPRECIP1$ and $MPRECIP2$ are generally 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 9 we report eq. 1 regression results for the equally-weighted variables. The results reported earlier hold for both samples (small and large) and both the most and least tick-constrained sub-samples.

[Table 9]

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 9 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 10, we estimate eq. 1 for two additional variables – 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 trading costs are low, effective spreads may not be fully indicative of changes in available liquidity. Table 10 shows that quoted spreads decline when the networks are down across all sample groups aside from the small sample. 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 10]

Table 10 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 our previous discussions. In the most constrained ETFs, there is not always room to improve the spread, but there is room to improve quoted depth. Notably, the depth does not increase in the least constrained ETFs; it appears that in these assets latent liquidity invoked by the network disruptions mainly improves the spreads.

6. Conclusions

This study examines the effects of speed differentials on liquidity. During our sample period, microwave networks stretched from Chicago to 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.

Finally, the results shed light on latent liquidity. We show that when speed differentials among traders decline due to precipitation, the emergence of latent liquidity narrows spreads more than one would expect based only on the decline in adverse selection. We also find that in assets where spread reductions are not possible due to the binding tick size, latent liquidity improves quoted depths.

Our results are confirmed in an event-study setting. In winter of 2012-2013, one of the technology providers 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 one-time event had positive consequences for market quality similar to precipitation-related network disruptions. This result further supports the claim that the technological race that leads to a market with speed differentials may be suboptimal for market quality.

The technological race continues to drive spending in the trading industry. A recent example is a new data transmission tower proposed by the telecommunications company Vigilant

Global to connect the U.K. and European markets. The tower will be among the tallest structures in the U.K. and will rival the height of the Eiffel Tower. It 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.

References

Baron, Matthew, Jonathan Brogaard, Björn Hagströmer and Andrei Kirilenko, 2016, Risk and return in high-frequency trading, Working paper.

Biais, Bruno, Thierry Foucault, and Sophie Moinas, 2015, Equilibrium fast trading, Journal of Financial Economics 116, 292-313.

Boehmer, Ekkehart, Kingsley Fong, and Julie Wu, 2015, International evidence on algorithmic trading, Working paper.

Bongaerts, Dion, Lingtian Kong, and Mark Van Achter, 2016, Trading speed competition: Can the arms race go too far? Working paper.

Brogaard, Jonathan, Björn Hagströmer, Lars Nordén, and Ryan Riordan, 2015, Trading fast and slow: Colocation and liquidity, Review of Financial Studies 28, 3407-3443.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2014, High-frequency trading and price discovery, Review of Financial Studies 27, 2267-2306.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2015, Price discovery without trading: Evidence from limit orders, Working paper.

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

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

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

Chakrabarty, Bidisha, Pankaj Jain, Andriy Shkilko, and Konstantin Sokolov, 2015, Speed of market access and market quality: Evidence from the SEC naked access ban, Working paper.

Chakravarty, Sugato, Pankaj Jain, James Upson and Robert Wood, 2012, Clean sweep: Informed trading through Intermarket Sweep Orders, Journal of Financial and Quantitative Analysis 47, 415-435.

Conrad, Jennifer, Sunil Wahal, and Jin Xiang, 2015, High-frequency quoting, trading, and the efficiency of prices, Journal of Financial Economics 116, 271-291.

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

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

Chan, Kalok, 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, Kee, and Chairat Chuwonganant, 2014, Uncertainty, market structure, and liquidity, Journal of Financial Economics 113, 476-499.

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

Du, Songzi, and Haoxiang Zhu, 2015, Welfare and optimal trading frequency in dynamic double auctions, Working paper.

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

Foucault, Thierry, Johan Hombert, and Ioanid Roşu, 2016, News trading and speed, Journal of Finance 71, 335-382.

Foucault, Thierry, Roman Kozhan, and Wing Wah Tham, 2016, Toxic arbitrage, Review of Financial Studies, forthcoming.

Goetzmann, William, Dasol Kim, Alok Kumar, and Qin Wang, 2015, Weather-induced mood, institutional investors, and stock returns, Working paper.

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

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

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

Harris, Larry, 2013, What to do about high-frequency trading, Financial Analysts Journal 69 (2), 6-9.

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

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

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

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

Hasbrouck, Joel, and Gideon Saar, 2013, Low-latency trading, Journal of Financial Markets 16, 646-679.

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

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

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

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

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

Jovanovic, Boyan, and Albert Menkveld, 2015, Middlemen in limit order markets, Working paper.

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

Laughlin, Gregory, Anthony Aguirre, and Joseph Grundfest, 2014, Information transmission between financial markets in Chicago and New York, Financial Review 49, 238-312.

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

Malinova, Katya, Andreas Park, and Ryan Riordan, 2014, Do retail traders suffer from high frequency traders, Working paper.

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

Menkveld, Albert, and Marius Zoican, 2015, Need for speed? Exchange latency and liquidity, Working paper, VU University Amsterdam.

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

Pesaran, Hashem, 2006, Estimation and inference in large heterogeneous panels with a multifactor error structure, Econometrica 74, 967-1012.

Roll, Richard, Eduardo Schwartz, and Avanidhar Subrahmanyam, 2007, Liquidity and the law of one price: The case of the futures-cash basis, Journal of Finance 62, 2201-2234.

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

Securities and Exchange Commission, 2010, Concept release on equity market structure, Release No. 34-61358.

Westerlund, Joakim, Simon Reese, and Paresh Narayan, 2014, A factor analytical approach to price discovery, Working paper.

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

Table 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: 100 ETFs	
Equities	
US index	50
International index	22
Interest rate products	20
Metals	4
Real estate	1
Other	3

Table 2. Market activity statistics

The table contains summary statistics for two samples: the sample of 5 ETFs (Panel A) and 100 ETFs (Panel B). Statistics are derived from the millisecond DTAQ data and aggregated into 15-minute intervals to match the 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 defined as the difference between the lowest offer quote and the highest bid quote across all markets. In Panel B, we separate 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: 5 ETFs					
# NBBO updates	21,603	17,538	11,012	14,083	27,224
# trades	3,228	3,797	1,146	1,426	3,838
volume, sh.	1,308,086	1,361,369	347,474	815,054	1,809,586
price, $	80.92	62.92	28.58	70.17	144.03
trade size, sh.	511	394	233	342	674
volatility, %	0.336	0.221	0.190	0.272	0.360
NBBO, $	0.011	0.005	0.010	0.010	0.011
trade price disc., %	0.291	0.156	0.185	0.246	0.341
price impact, $	0.008	0.007	0.006	0.009	0.012
effective spread, $	0.020	0.058	0.010	0.011	0.013
realized spread, $	0.011	0.059	0.000	0.003	0.006
Panel B: 100 ETFs					
# 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
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. 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 5 ETFs in the small sample, Panel B examines 100 ETFs in the large sample, and Panel C separately 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.

	$PIMP$			ESP			RSP		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Small sample									
$PRECIP$	-.012***			-.003			-.008*		
	(.004)			(.004)			(.004)		
$PRECIP1$		-.035***			-.011			-.024**	
		(.011)			(.011)			(.011)	
$PRECIP2$			-.051***			-.023*			-.028**
			(.014)			(.014)			(.012)
VIX	.110***	.111***	.111***	.052***	.052***	.052***	-.007	-.007	-.007
	(.004)	(.004)	(.004)	(.004)	(.004)	(.004)	(.018)	(.018)	(.018)
Panel B: Large 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 C: 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 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 5 ETFs in the small sample, Panel B examines 100 ETFs in the large sample, Panel C separately 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: Small sample									
$PRECIP$	-.021			-.035***			-.024***		
	(.013)			(.014)			(.008)		
$PRECIP1$		-.106***			-.138***			-.089***	
		(.039)			(.040)			(.027)	
$PRECIP2$			-.112**			-.146***			-.109***
			(.045)			(.046)			(.032)
VIX	.143***	.145***	.143***	.106***	.108***	.107***	.144***	.145***	.144***
	(.053)	(.053)	(.053)	(.033)	(.033)	(.033)	(.036)	(.036)	(.036)
Panel B: Large 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 C: 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 5. Microwave connectivity and trading costs on the CME

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 the futures contract 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, 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 2012 and cover four e-mini futures contracts: S&P 500, S&P MidCap 400, Nasdaq 100 and Financial Sector Select. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% levels.

	$PIMP$			ESP			RSP		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$PRECIP$	-.026			-.043			-.001		
	(.017)			(.034)			(.002)		
$PRECIP1$		-.057			-.011			-.005	
		(.037)			(.081)			(.006)	
$PRECIP2$			-.039			-.115			-.016
			(.044)			(.083)			(.012)
VIX	.030***	.029***	.030***	.017***	.018***	.016***	.022	.022	.023
	(.006)	(.006)	(.006)	(.007)	(.008)	(.007)	(.023)	(.023)	(.023)

Table 6. 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 four groups of assets: (i) 5 ETFs in the small sample, (ii) 100 ETFs in the large sample, and (iii/iv) terciles of ETFs 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
small sample	.006	.011	.008	.024	.008	-.002
	(.019)	(.011)	(.013)	(.035)	(.033)	(.028)
large 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 7. Network 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:

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

where $DEPVAR_{it}$ is one of the following variables (price impacts, effective spreads, realized spreads, number of trades, traded volume, or volatility) in asset i on day t; $POST$ 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 (in parentheses) are double-clustered along the asset and time dimensions. We examine four groups of assets: (i) 5 ETFs in the small sample, (ii) 100 ETFs in the large sample, and (iii/iv) terciles of ETFs 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
Panel A: Small sample						
small sample	-.475***	-.337***	-.148*	.173	.028	-.318
	(.176)	(.097)	(.089)	(.223)	(.266)	(.216)
large sample	-.295***	-.342***	-.262***	.159**	.186***	-.253**
	(.071)	(.058)	(.057)	(.065)	(.062)	(.113)
most constr.	-.370***	-.238***	-.102	.019	.106	-.325***
	(.122)	(.068)	(.081)	(.110)	(.100)	(.130)
least constr.	-.319***	-.352***	-.320***	.158**	.167***	-.173*
	(.075)	(.108)	(.105)	(.067)	(.066)	(.103)

Table 8. 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 sample of 100 ETFs. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% levels.

	$PIMP$	ESP	RSP	trades	volume	volatility
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 9. 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 four groups of assets: (i) 5 ETFs in the small sample, (ii) 100 ETFs in the large sample, and (iii/iv) terciles of ETFs 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
small sample	-.071***	-.120**	-.093***	-.110***	.021	-.025
	(.015)	(.050)	(.019)	(.017)	(.047)	(.028)
large 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 10. 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 four groups of assets: (i) 5 ETFs in the small sample, (ii) 100 ETFs in the large sample, and (iii/iv) terciles of ETFs for which the tick size is the most (least) binding. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% levels.

	QSP	$DEPTH$
small sample	-.009	-.004
	(.039)	(.023)
large sample	-.065***	.014
	(.020)	(.029)
most constr.	-.026**	.118***
	(.012)	(.045)
least constr.	-.105***	-.013
	(.034)	(.034)

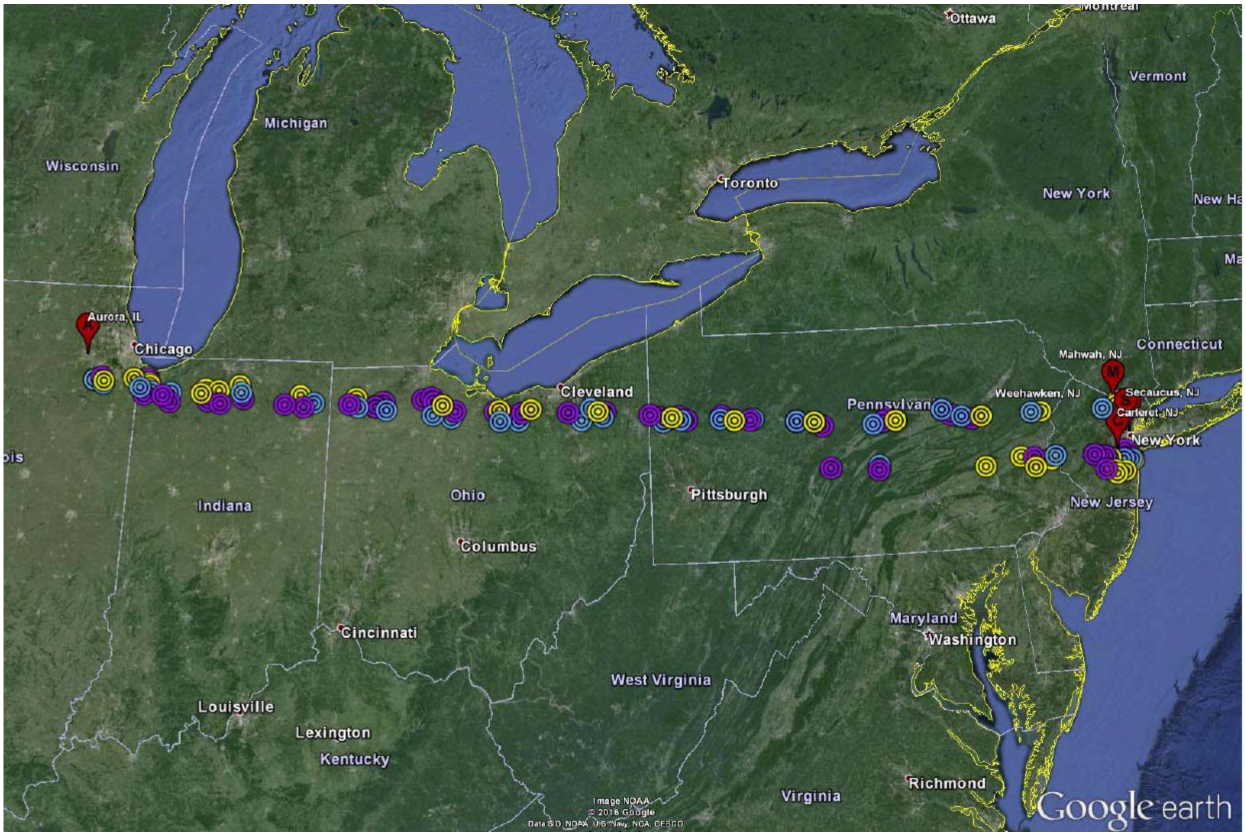

Figure 1. Microwave network paths

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).

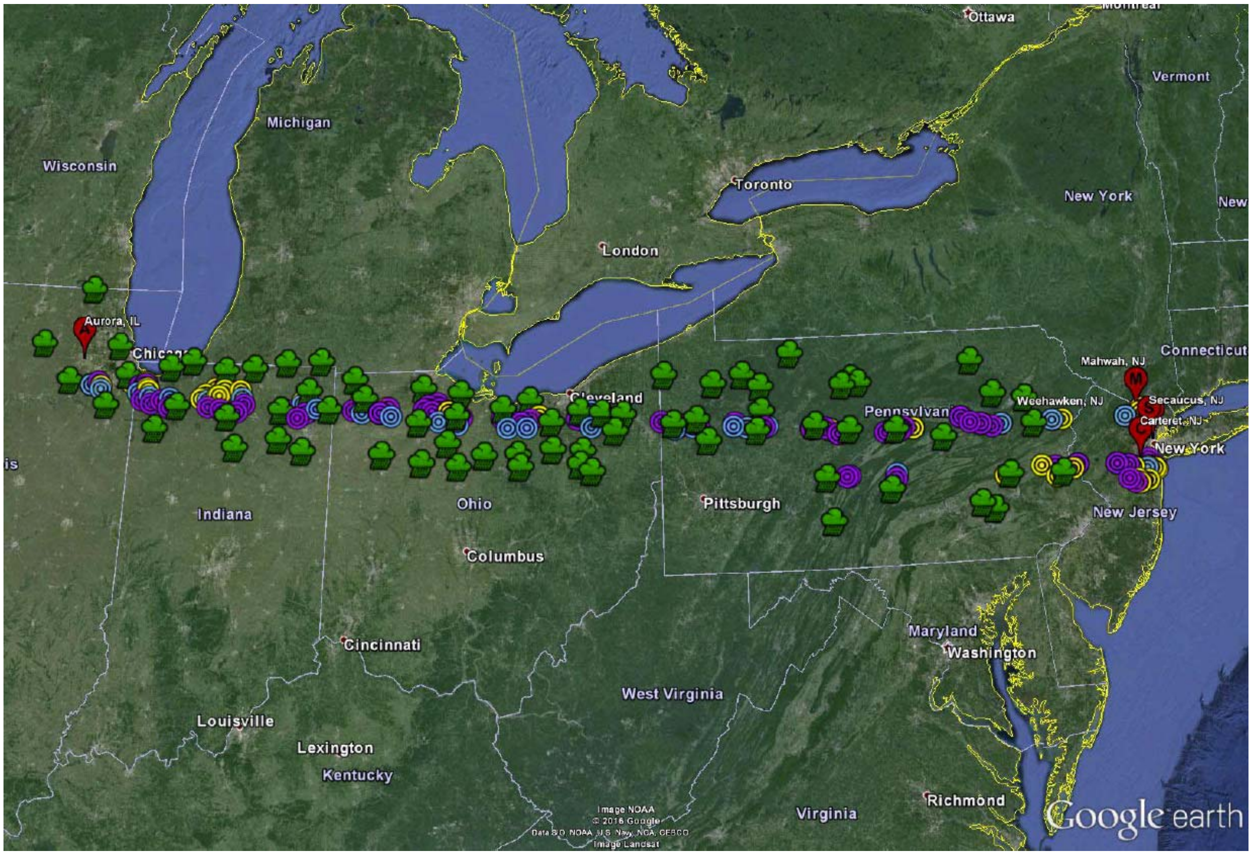

Figure 2. Locations of microwave networks and weather stations
The figure maps the weather stations (green icons) located near the microwave network paths. Station data are obtained from the National Oceanic and Atmospheric Administration. The red markers indicate locations of the CME's data center in 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. 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.

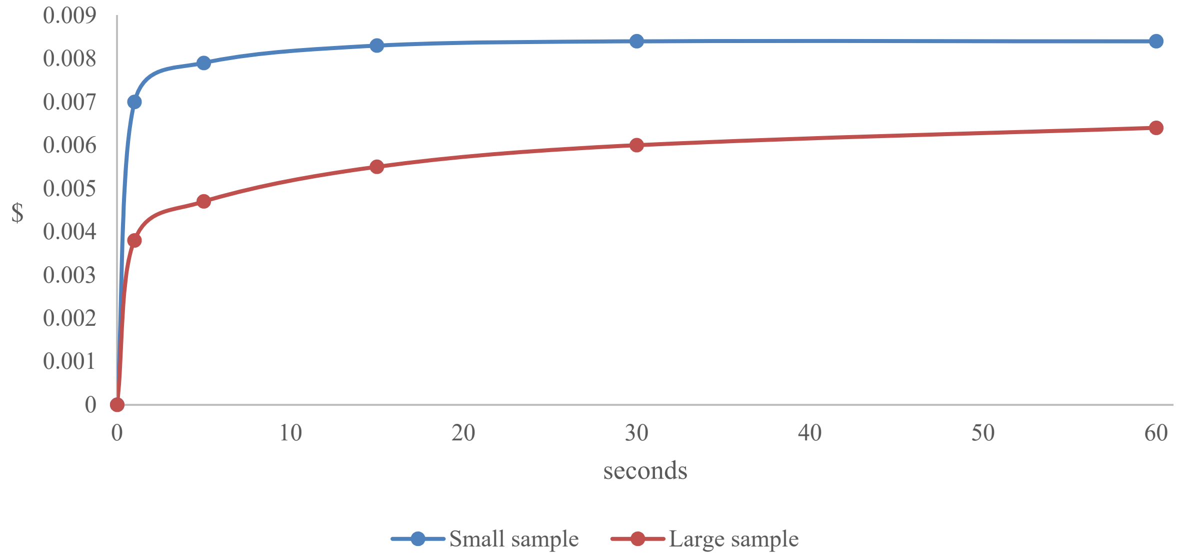

Figure 4. Price impacts

The figure reports price impacts computed as the signed scaled difference between a midquote at a certain time after the trade and the midquote at the time of the trade: $PRIMP_t = q_t(mid_{t+\gamma} - mid_t)/mid_t$, 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 $\gamma \in \{1s, 5s, 15s, 30s, 60s\}$.

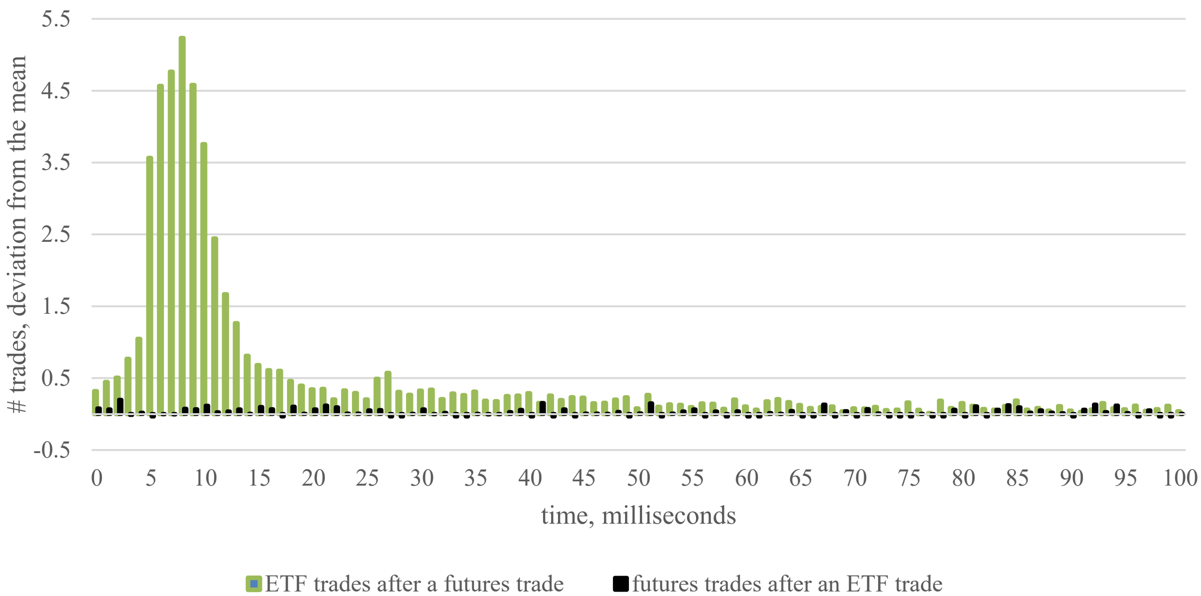

Figure 5. Interaction between futures and equities (S&P 500 index)
The figure contains the results from two tests. The first test counts the average number of ETF trades following a futures trade at t_0 (green bars), and the second test counts the average number of futures trades following an ETF trade at t_0 (black bars). In the first (second) test, we focus on the standalone t_0 futures (ETF) trades – those not preceded by another futures (ETF) trade in the previous 100 milliseconds. For illustration purposes, we demean the trade counts.

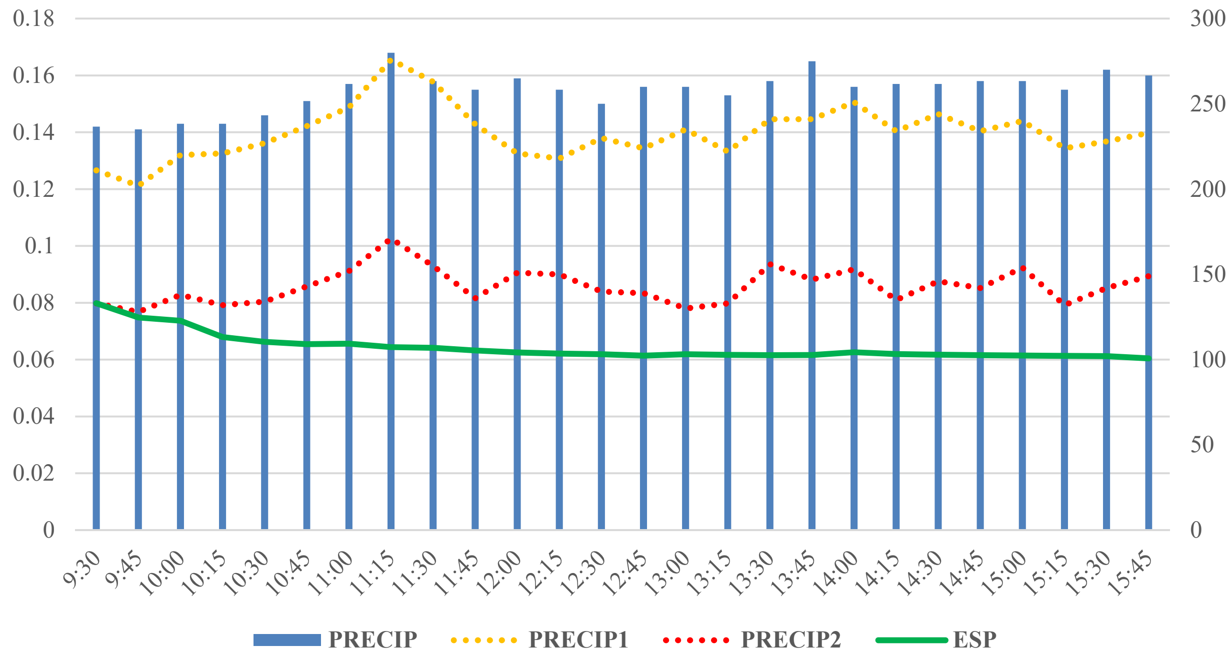

Figure 6. Intraday patterns

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" by A. Shkilko and K. Sokolov

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:

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

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:

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

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.

To confirm IS results, we use the panel information share (PIS) methodology developed by Westerlund, Reese and Narayan (2014). The key difference between IS and PIS is that PIS relies on the common factor model instead of the VECM. Specifically, for each of the two markets we estimate the unobserved factors and loadings of the following model:

$$\Delta p_{i,t} = \lambda_i f_t + u_{i,t,}$$

where f_t is the common price innovation component, and λ_i is the market-specific loading. Following Pesaran (2006), we estimate factors $\hat{f}_t = \frac{\sum_{i=1}^N \Delta p_{i,t}}{N}$, loadings $\hat{\lambda}_i = \frac{\sum_{t=2}^T \Delta p_{i,t} \hat{f}_t}{\sum_{t=2}^T \hat{f}_t^2}$ and obtain the cumulative residuals $\widehat{U}_{i,t} = \sum_{n=2}^t \hat{u}_{i,n}$. Then panel information share is computed as:

$$\widehat{PIS}_i = \frac{\hat{\lambda}_i^2 \hat{\sigma}_\eta^2 \hat{\sigma}_{U,i}^{-2}}{\sum_{n=1}^N \hat{\lambda}_i^2 \hat{\sigma}_\eta^2 \hat{\sigma}_{U,n}^{-2}},$$

where $\hat{\sigma}_\eta^2 = \frac{\sum_{t=2}^T \hat{f}_t^2}{T}$ and $\hat{\sigma}_{U,i}^2 = \frac{\sum_{t=2}^T \widehat{U}_{i,t}}{T}$.