The Sensitivity of Trading to the Cost of Information

Alex Frino[a], Ognjen Kovacevic[b*], Vito Mollica[b], Robert I. Webb[c]

[a] School of Accounting, Economics and Finance, Faculty of Business, University of Wollongong, Wollongong, NSW, Australia

[b] Macquarie Graduate School of Management, Department of Applied Finance, Macquarie University, Sydney, NSW, Australia

[c]McIntire School of Commerce, University of Virginia, Charlottesville, Virginia, U.S.

Abstract

This study examines the impact of a modification to data feed pricing schedule on the price discovery of competing venues. To this end, we use three exogenous events arising from the staggered increase of price quotation fee on Chicago Mercantile Exchange to test the theoretical predictions of Cespa & Foucault (2014) who note that price discovery is "…determined by the fee charged by exchanges for price information". After controlling for known determinants, we observe a decrease in the efficiency of price discovery following increases in the acquisition costs of exchange's data feeds, in line with the theory.

We would like to thank Eugenio Piazza, Tom Steffen, and the participants at the research seminars at the Capital Markets Cooperative Research Centre in Sydney, Australia for their helpful suggestions. We thank the Derivative Markets Research Centre for funding of the research. This Research is funded by the ASX under Corporations regulation 7.5.88 through the use of excess funds from the Sydney Futures Exchange Fidelity Fund. Any errors or omissions are the responsibility of the authors alone.

*****Corresponding author, Macquarie Graduate School of Management, 99 Talavera Rd, North Ryde NSW 2113, Australia. Email: ognjen.kovacevic@hdr.mq.edu.au

1. INTRODUCTION

Exchanges own the trade and quote information generated on their platforms in the process of trading. In the U.S, these property rights were established in early 20^{th} century and give exchanges the right to set the selling price for their data feeds. In competitive markets, and in the era of nanosecond latency markets in particular, the value of new information fades quickly highlighting the necessity of having direct access to exchange's real-time price information in addition to the processing power providing for swift analysis and order execution. The current financial environment is characterized in literature as having an *"ever increasing pace of both information gathering and the actions prompted by this information."* (Hasbrouck & Saar, 2013), and obtaining direct feeds to exchanges' trading is regarded as beneficial in both futures and equities trading as it gives *"… high speed traders an ability to see the market with more clarity…"* (O'Hara, 2015).

The cost of subscription to exchange's direct data feeds can be substantial, while exchanges earn a fifth of their global $30bn revenue from data selling.[1] The apparent lack of price competitiveness in the data market raises concern, as noted in the U.S. Department of the Treasury Report (2017): *"…the market for proprietary data feeds is not fully competitive … , data from one exchange's feed cannot substitute for data from another exchange's feed".* In addition, recent increases of data feed prices resulted in an uproar of trading firms petitioning the SEC and requesting stricter control over data fees. Securities and Exchange Comission (2018a; 2018b) responded with two decisions blocking the proposed changes to data fees. The Commission, however, cited lack of supporting information as a reason for the rejection of changes to the fee structure and distanced themselves from qualifying the data fees as *"unreasonable"* or *"unfair"*.

Repercussions of *"unfair"* profit maximizing behavior of the exchanges in data fee setting are more than profit redistribution from traders to exchanges. The charges for price quotations create differential access to real-time data across market participants and can exert significant effects on market efficiency and the behavior of speculative prices. For instance, Easley, O'Hara, & Yang (2016) argue: "*selling price data increases the cost of capital and volatility, worsens market efficiency and liquidity, and discourages the production of fundamental information relative to a world in which all traders freely observe prices*." Cespa & Foucault (2014) note that

[1] https://www.ft.com/content/d8c2743e-549f-11e9-91f9-b6515a54c5b1

as a result of exchanges charging a high market data fee to restrict access to its price quotations *"…price discovery is not as efficient as it would be with free price information."*

This paper uses three natural experiments to test the theoretical predictions of Cespa & Foucault (2014). Specifically, we examine the impact of the staggered modification of the data feed pricing schedule by the CME Group (CME), giving rise to three separate events, on liquidity and the efficiency of price discovery. We test the conclusion of Cespa & Foucault (2014) and Easley et al. (2016) that increases in data-fee price adversely affect liquidity and price discovery in light of these events by analyzing the behavior of traditional liquidity metrics – spreads and depth – and popular price discovery metrics – Information Share, Component Share, and Information Leadership Share. Our analysis focuses on the behavior of WTI futures contract traded on the CME relative to the same contract on Intercontinental Exchange (ICE). Futures markets, not adopting a fee-rebate structure in place in equity markets, provide a well-ordered setting to analyze the market impact of changing data feed fees, in isolation from noise stemming from liquidity provision/taking incentives.

Contrary to theoretical predictions of Cespa & Foucault (2014) and Easley et al. (2016), our results do not identify conclusive adverse changes in liquidity following the data feed fee increase. Our findings, however, corroborate the Cespa & Foucault (2014) predictions in terms of efficiency of price discovery. We find that in the period following the increases in the data feed fee on January 1, 2015 and 2016, price discovery on CME decreases, as measured by all three price discovery metrics used in the literature and after controlling for known determinants. As predicted by the Cespa & Foucault (2014) model, the increase in the data fees leads to the reduction in the number of traders purchasing price information, as noted in the 2016 CME Annual Report, reducing the efficiency of the price discovery.[2]

We document an improvement in price discovery on CME following March 1, 2014. We explain this phenomenon as rational behavior by professional traders: following the announcement that after March 1, 2014 all new traders are required to pay the full price for data, a host of new traders joins the market in order to enjoy free data access until January 1, 2015 and access at a reduced fee until January 1, 2016. We test this interpretation by using a known proxy for

[2] Chicago Mercantile Exchange Annual Report (2016) notes that the increase in data revenues "… was partially offset by some rationalization as customer firms transitioned into full-priced offerings, as well as reductions in overall screen counts at member banks."

algorithmic trading and find evidence of an increased presence of fast traders in the period leading up to the event. Such a result confirms the reasoning behind the Cespa & Foucault (2014) model – with a wider distribution of price quotations, the efficiency of price discovery increases.

This paper is organized as follows. Section 2 provides an overview of literature related to market data ownership, pricing, and the impact of fees on market efficiency. Section 3 describes the event and data. Section 4 describes the research method and the empirical model. Section 5 reports the results of the analysis. Section 6 provides a conclusion.

2. LITERATURE OVERVIEW

2.1. Market data as property

Exchanges ownership of the price quotation data is established in U.S. Supreme Court decisions. Mulherin, Netter, and Overdahl (1991) and Webb (2003) review the origin of the property rights the futures exchanges have over their price quotations in the USA. Notable early U.S. Supreme Court cases which establish the price quotations are property of the exchanges that produced them are the *Board of Trade v. Christie Grain & Stock Co., 198 U.S. 236* (1905) and *Hunt v. New York Cotton Exchange : 205 U.S. 322* (1907).

The former involves a claim made by the CBOT against an alleged bucket shop[3] that obtained its quotes without CBOT's authorization[4]. The defense argued that CBOT was itself operating a bucket shop which invalidated its property rights to the quotes. Ruling in favor of the CBOT, the court decision states that even in case the CBOT operated an illegal bucket shop *"... it does not follow that it should not be protected in this suit"*. Moreover, the majority opinion of the Court, written by Justice Holmes, underlines that *"...the plaintiff's collection of quotations is entitled to the protection of the law. It stands like a trade secret. The plaintiff has the right to keep the work which it has done, or paid for doing, to itself."*

[3]A bucket shop is "an establishment, nominally for the transaction of a stock exchange business, or business of similar character, but really for the registration of bets, or wagers, usually for small amounts, on the rise or fall of the prices of stocks, grain, oil, etc., there being no transfer or delivery of the stock or commodities nominally dealt in". (*Gatewood v. North Carolina, 27 S.Ct 167, 168,* 1906)

[4] At the time of the proceeding, data was distributed by telegraph companies that purchased the rights to distribution of quote data from the exchanges, subject to restrictions on its redistribution.

The notion of property rights over price quotations was further recognized in the *Hunt v. New York Cotton Exchange : 205 U.S. 322* (1907) decision. Ruling in favor of the New York Cotton Exchange, the decision states that "*…the exchange has the right to keep them to itself or have them distributed under conditions established by it"*. As noted by Webb (2003), unlike the previous case where the decision hinged on the implied breach of contract by a subscriber to the exchange market data, in this case *"…the exchange's ownership of its real-time price quotations is central in the Court's decision"*.

The decision in *Board of Trade v. Christie Grain & Stock Co., 198 U.S. 236* (1905) reflects on the nature of the property rights, and, in effect, envisions these rights to be *transitory*. As noted in the majority opinion: "*Time is of the essence in matters like this, and it fairly may be said that, if the contracts with the plaintiff are kept, the information will not become public property until the plaintiff has gained its reward. A priority of a few minutes probably is enough".* While the decision in the Hunt v. New York Cotton Exchange was grounded in CBOT v. Christie Grain & Stock Co., it does not comment on the duration of the property rights, and as noted by Webb (2003) *"…The idea that price quotations information becomes part of the public domain after a transitory period during which exchanges can exploit its value seems to have been lost in subsequent decisions"*. Arguing that in the U.S. exchange property right over its price quotations is not time limited, while such limits exist *"…on the rights to most forms of intellectual property"*, Webb (2003) suggests that exchange's ownership over the ticker data should be limited to only a few minutes.

2.2. Market Data Fee Issues

Expressions of concern over the monopoly enjoyed by the exchanges and the freedom they have in setting the data fee structure are reported in media outlets[5], petitions to the SEC[6], and Treasury reports to the President of the U.S

[5] Articles on this topic are printed by every major financial media outlet. See Bloomberg, the Financial Times, Reuters, and Business Insider (https://goo.gl/uHvHMf, https://goo.gl/BbNNy9, https://goo.gl/fsCzBH, https://goo.gl/cSQdGh, https://goo.gl/LchCML)

[6] While the number of these petitions is very high *("During the pendency of the action that resulted in the Data Fee Decision, over 400 other challenges to market data and market access fees were filed with the Commission.",* Clayton, 2018), only one of them resulted in the SEC taking action, at the time of writing of this paper.

Rulemaking petition concerning market data fees (2017)[7] argues that regardless of the cost of the data, traders are forced to acquire market data from the exchanges if they are to obey the legal requirements[8] and remain competitive in the market place.[9] This leads to a *de facto* monopoly, and a significant and accelerating price increase. They ask for higher disclosure requirements which would force the exchanges to disclose details on the revenues and costs of market data and for a review of the data fee structure to ensure the fees are *"fair and reasonable"*. In addition, the petition calls for disallowing of immediate enforceability of fee changes.[10] *Petition for Rulemaking Regarding Market Data Fees* (2018) makes similar requests calling for higher disclosure on exchanges' market data operating costs, a more detailed filings with respect to market data fee structures providing for justification for any changes, and public notice and comment period before the fee change approval.

U.S. Department of the Treasury (2017) acknowledges the regular price increases in recent years and also recognizes the lack of competitiveness in the market for property data feeds. The report states that *"For use in making routing and trading decisions for active or institutional size order flow, data from one exchange's feed cannot substitute for data from another exchange's feed"*. In addition, they report that many brokers and dealers feel obliged to purchase the proprietary data in order to meet the best execution obligations[11] and conclude recommending stricter approval of fee changes by the

[7] Bloomberg, Citigroup Citadel Securities, Investors Exchange (IEX), Morgan Stanley, UBS, and the Vanguard group are among the 24 undersigned companies from HFT, banking, exchange, and data provision sectors

[8] "SEC's Display Rule requires vendors and broker-dealers to display consolidated data from all the market centers that trade a stock. Vendors and broker-dealers, therefore, must purchase the consolidated data feeds from securities information processors ("SIPs"), which are owned and operated by the exchanges themselves.

[9] Exchanges sell proprietary data feeds, in addition to SIP data feeds. The Petition notes that even though proprietary data is more expensive than the SIP data feed *"…they also include much more comprehensive data, including "depth-of-book" and "imbalance" data, and generally are subject to less latency. In today's high-speed electronic markets, many broker-dealers, market makers, hedge funds, data distributors, and a wide array of other market participants have concluded that they must purchase proprietary data feeds from exchanges in order to remain commercially competitive".*

[10] An amendment of the Securities Exchange Act of 1934 made by the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 enables changes of SIP data (feed distributed by the securities information processors: "SIPs") fee to be effective immediately upon filing, as opposed to the previous rule that envisioned a period allowing for public notice and comment

[11] In order not to be in breach of the best execution obligations, broker-dealers use the best available data to execute at the best available price. In addition, the customers themselves may demand that firms use proprietary data feeds to identify the best prices (U.S. Department of the Treasury, 2017).

Commission, and enabling competing consolidators to provide alternatives to the proprietary data (depth of the book and last sale) which should provide faster consolidation, distribution, and lower prices.

A limited number of researchers provides support for the monopolistic market for price quotation data. Simonov (1999), presents a model in which a competitive market for information is not better than a monopolistic one, neither in terms of the quality of signal being traded, nor in terms of financial payoff to traders, since the information providers capture all of the surplus, as in the case of a monopoly. Mulherin et al. (1991) note that it is misleading to label an exchange's exercise of its property rights as a monopoly. They posit that the price discovery mechanism is established by the exchanges and is a product of their work – allowing the exercise of those property rights enables exchanges to reap the gains from technological innovation and foster their growth. Finally, Angel & McCabe (2018) delve into the ethics of the data-selling issue, prompted by the NY Attorney General's statement referring to the sale of faster access to financial data as *"Insider Trading 2.0"*. Referring to a case of sale of the Consumer Sentiment Data by University of Michigan, they posit that the early purchasers of data essentially subsidize the free, albeit delayed, release of the data to general public[12], highlighting a positive role of the price discrimination inherent in the data sale. Furthermore, they argue that since the access to information was available to everyone willing to pay for it, there are no grounds for labelling these practices as unfair. Finally, they underline that since information production is a costly activity, it must be rewarded, in line with the theory of Grossman & Stiglitz (1980).

Two recent SEC decisions (Securities and Exchange Comission, 2018a; 2018b), referring to both consolidated and proprietary data, provide evidence that the importance of data fees and their impact on the market has not been overlooked by the Commission. In SEC (2018a) the Commission rejects the proposed change in the consolidated data pricing schedule allowing for an increase in the Enterprise Cap made by participating exchanges (Participants).[13,14] The basis for this decision is a lack

[12] A comparison that can be extended to the current exchange policies of releasing the data with a 10-20-minute delay.

[13] The Decision identifies the Participating exchanges as: Cboe BYX Exchange; Inc.; Cboe BZX Exchange, Inc.; Cboe EDGA Exchange, Inc.; Cboe EDGX Exchange, Inc.; Chicago Board Options Exchange, Incorporated; Chicago Stock Exchange, Inc.; Financial Industry Regulatory Authority, Inc.; Investors Exchange LLC; Nasdaq BX, Inc.; Nasdaq ISE, LLC; Nasdaq PHLX LLC; The Nasdaq Stock Market LLC; New York Stock Exchange LLC; NYSE Arca, Inc.; NYSE American LLC; NYSE National, Inc.

[14] The increase in the Enterprise Cap was coupled with a decrease in the per-quote-packet charges. Participants argued that the change was designed to maintain status quo, following a consolidation in the industry that allowed combined entities to benefit from substantial fee decreases under the previous Enterprise Cap regime.

of information and justification that would enable the Commission to determine whether the proposed change is consistent with the Exchange Act and would result in "reasonable" and "not unreasonably discriminatory" fees, that would *"...not impose an undue or inappropriate burden on competition ... "*.[15]

In SEC (2018b), the Commission takes action as a response to a petition filed against fees charged by NYSE Arca and Nasdaq for their proprietary depth-of-book data. Exchanges were required to demonstrate that the new fees are "reasonable and not unreasonably discriminatory". Similar to SEC (2018a) decision, the Commission finds that the exchanges fail to establish i) *that their need to attract order flow constrains their pricing of the depth-of-book products* ii) *that availability of alternatives constrains their pricing of the depth-of-book products* iii) *a basis other than competitive forces to demonstrate that the fees at issue are fair and reasonable* (SEC, 2018b) leading to setting aside of the fees covered in the case. It is important to note that while the SEC bases their decision on lack of evidence supporting the proposal of fee increase, they did not find the fees unfair or unreasonable. Further, the Commission sent 400 other challenges to market data and access fees (filed during the pendency of SEC, 2018b) to the Exchanges requesting further comment, signaling a new era of policy in which, while exchanges have the freedom of setting their own fees, market participants are free to challenge them, and possibly have them reversed.

2.3. Data Fees and Price Discovery

Cespa & Foucault (2014) study how the pricing of exchange's price quotation feeds affects the price discovery process, and trade profit distributions under different scenarios. They propose a model envisioning three types of traders: insider and outsider speculators, and liquidity traders, where the difference between the former two groups is based on whether or not they purchase the access to the exchanges "real-time" price ticker or trade using a "lagged ticker".[16] They predict that the efficiency of price discovery increases with the proportion of insiders and decreases with latency of the data feed released to the speculators not purchasing the real time feed. This stems

[15] It is interesting to note that, even though two comment letters responding to the price increase notice of filing argued there is inadequate support for changes in the proposal, the Participants declined to provide further clarifications in their response. This is also noted in the discussion section of the release that cites the lack of information both in the proposal and the participants' response as the rationale of the decision.

[16] Real-time price ticker contains the updated price history while the lagged ticker lacks information on orders from a certain number of previous periods

from their observation that given higher fraction of speculators trading on real time information, not only will their impact on prices be stronger via their trades, but also indirectly, by sending stronger signals to the speculators trading based on a lagged ticker. The latency of the lagged data feed, while not affecting insiders, will adversely affect price discovery by reducing the informativeness of the dataset the outsiders base their trades on. Finally, Cespa & Foucault (2014) note that the exchanges, in order to maximize their profits, can influence the number of traders purchasing the direct access to its market data by changing the real-time data fee. Depending on the sensitivity of liquidity traders to trading losses, exchange's pricing strategies differ. Cespa & Foucault (2014) conclude that in most of the situations, the profit maximizing behavior of exchanges negatively affects the price discovery, while the impact of an increase in data fees on price discovery is always adverse, since it decreases the number of informed traders and slows the process of information integration into prices.

Extending the reasoning in Cespa & Foucault (2014), the model of Easley et al. (2016), considers the data sold by the exchanges and private data produced by the traders themselves as complements and not substitutes.[17] They underline that if exchanges charge profit maximizing prices for their data, some traders will be restricted from obtaining the real-time quotes leading to a reduction in market efficiency and liquidity, and an increase in volatility and cost of capital levels. In addition, unwilling to purchase exchange data due to higher prices, traders reduce their own fundamental data production, further exacerbating the adverse impact of market data fees on price discovery.

3. EVENT AND DATA

<INSERT GRAPH 1>

Three events analyzed in this paper result from a staggered modification of the CME data feed pricing schedule; Graph 1 reports the timeline of these changes. Starting March 1, 2014, the CME put forward a new marked data pricing schedule, changing a fee for access to its real-time price data of $85 to professional and $3 to non-professional traders per month, per user, per

[17] Bergemann, Bonatti, & Smolin (2018) provide a broader model focusing primarily on data selling outside of the scope of exchange data. Their data buyer already possesses private information, the quality thereof determines her willingness to pay for supplemental data.

platform, per exchange.[18] The fee is charged to all new subscribers; those operating under the data fee waiver before March 1, 2014 are "grandfathered" into the program and continue to benefit from the 100% discount in subscriber fees until January 1, 2015 when the discount is lowered to 50%. Starting January 1, 2016 all existing fee waivers are cancelled, and all traders pay the full data fee.

This gives rise to three separate events, two of which: January 1, 2015 and 2016 are de facto price increases implemented via a reduction and removal of the fee waiver. Based on the predictions of Cespa & Foucault (2014) and Easley et al. (2016) models, these two events are expected to adversely impact liquidity and price discovery on the CME.

Conversely, the first event, leaves all traders under the waiver prior to March 1, 2014 unaffected and impacts only new traders subscribing to market data feed afterwards. It could be argued, therefore, that there should be little to no impact on liquidity and price discovery around the first event. However, as the cutoff date was announced in advance, it could be the case that traders, behaving rationally, opted to subscribe for data prior to March 1, in order to be grandfathered into the program and enjoy a 100% discount on data fees until January 1, 2015 and a 50% discount until January 1, 2016 causing an increase in the number of traders purchasing the data feed in the period leading up to the event. In such a scenario, Cespa & Foucault (2014) and Easley et al. (2016) theory predicts an overall increase in liquidity and efficiency of price discovery due to a higher number of traders trading based on the finer quality information. The behavior of liquidity and price discovery around event 1 is therefore ambiguous and hinges on the decision of traders to subscribe to the data feed in the period leading up to March 1, 2014.

In order to assess the impact of data fee increases on efficiency of price discovery on the CME, we analyze the behavior of its most traded commodity contract, the WTI oil futures. WTI futures are, in addition to the CME traded on the ICE, with nearly identical contract specifications, warranting cointegrated price series. Futures contracts on CME are traded through Globex platform Sunday - Friday from 6:00 p.m. to 5:00 p.m. next day ET. WTI oil contracts on ICE trade from 8:00 p.m. to 6:00 p.m. next day ET. Both exchanges trade the WTI contract at 1000 barrels per contract, with price quotations expressed in US Dollars and cents, and minimum price fluctuation of $0.01 per

[18] Monthly fees for the non-professional traders are capped at $15.

barrel. The only difference between the contract stems from the delivery method: WTI oil futures are physically deliverable on CME, while they are cash settled on ICE.

Our sample periods contain 125 trading days around the three events – March 1, 2014; January 1, 2015 and January 1, 2016 – corresponding to approximately 6 months of trading pre and post data price increases. Data are sourced from Thompson Reuters Tick History (TRTH) trade and quote database and refer to the chain of nearest to maturity contracts. In addition, we obtain TRTH data listing the daily open interest value for the near WTI contract for each of the exchanges.

Federal holidays and Sundays are removed from our sample. We remove 1 hour before and after the overlapping periods of trading, to remove potential beginning/end of day trading patterns. This leaves us with five 19-hour trading days a week with a gap in the middle of the day from 4:00 p.m. to 9:00 p.m. ET (1 hour before the close of CME trading and 1 hour after the start of the ICE trading, including the non-overlapping 3 hours are deleted).

4. METHOD

4.1. Model Specification

4.1.1. Liquidity

We use difference-in-difference to estimate the impact of increase in data fees on the market liquidity as follows:

$$Liq_t = \beta_1 + \beta_2 Exchange_t + \beta_3 Event_t + \beta_4 DID_t + \beta_5 Volatility_t + \beta_6 OPINT_t \qquad (1)$$

where Liq_t can take values of bid-ask spread, effective, realized, and adverse spread, and best depth; $Exchange_t$ is the dummy variable that takes value of 1 for CME and 0 for ICE; $Event_t$ takes the value of 1 after the event (March 1, 2014; January 1, 2015; January 1, 2016) and 0 in the period before the event; and DID_t is the interaction between the two dummies explaining the changes in liquidity on CME resulting from the data fee increase. We control for the known liquidity determinants (Demsetz, 1968; Harris, 1994; Mcinish and Wood, 1992; Chordia, Roll, and Subrahmanyam, 2000): $Volatility_t$ is the daily mid-quote volatility; and $OPINT_t$ is the daily open interest (in '00000). Regressions are computed on a sample extending 250 trading days, 125 before and after the event, corresponding to six months of trading pre and post the event.

4.1.2. Price Discovery

We estimate the impact of increase in data fees on the efficiency of price discovery as follows:

$$PD_t = \beta_1 + \beta_2 Event_t + \beta_3 RelBAS_t + \beta_4 RelOPINT_t \quad (2)$$

where PD_t is logit transformation of either of the three price discovery metrics analyzed (IS, CS, or ILS); $Event_t$ is the dummy variable that takes value of 1 in the period after the event (March 1, 2014; January 1, 2015; January 1, 2016) and 0 in the period before the event; $RelBAS_t$ is the average daily bid-ask spread (measured in ticks) on CME relative to the one on ICE; and $RelOPINT_t$ is the fraction of daily open interest on CME relative to daily open interest on ICE. We follow Frijns, Gilbert, & Tourani-Rad (2015) in applying the logit transformation of the price discovery metric and the selection of the controls. Regressions are computed on a sample of 250 observations, 125 before and after the event.

4.2. Variable Measurement

4.2.1. Spread Measures and Best Depth

Daily average bid-ask spread (measured in ticks) is calculated as follows:

$$tick\ spread_t = \frac{\sum_{i=1}^{n} \frac{ask_{i,t} - bid_{i,t}}{minimum\ tick\ size}}{N_t} \quad (3)$$

where $ask_{i,t}$, $bid_{i,t}$, represent best quotes at time *i* during the day, $minimum\ tick$ is \$0.01, and N_t is the total number of best quotes during the day.

Daily average effective, realized, and adverse spread measures (all measured in ticks) are calculated based on Glosten (1987) as follows:

$$effective\ spread_t = \frac{\sum_{i=1}^{n} \frac{direction_{i,t}(p_{i,t} - midquote_{i,t})}{minimum\ tick\ size}}{N_t} \quad (4)$$

$$realized\ spread_t = \frac{\sum_{i=1}^{n} \frac{direction_{i,t}(p_{i,t} - midquote_{i+x,t})}{minimum\ tick\ size}}{N_i} \quad (5)$$

$$adverse\ spread_t = \frac{\sum_{i=1}^{n} \frac{direction_{i,t}(midquote_{i+x,t} - midquote_{i,t})}{minimum\ tick\ size}}{N_t} \quad (6)$$

where $minimum\ tick$ and N_t are as above, $p_{i,t}$ represents the price of the trade taking place at time *i* during the day, and $midquote_{i,t}$ is calculated as:

$$midquote_{i,t} = \frac{ask_{i,t} + bid_{i,t}}{2} \tag{7}$$

$midquote_{i+x,t}$ is the prevailing midquote x seconds after the trade took place and $direction_{i,t}$ identifies the initiator of the trade based on the Lee & Ready (1991) algorithm.[19] It is equal to -1 if the trade is seller initiated and 1 if it is buyer initiated. We estimate measures of realized and adverse spread at 1, 5, 30, and 60-second horizons.

Best depth is the sum of contracts available to trade at best bid and ask and is averaged across the day.

4.2.2. Price Discovery

We quantify price discovery with three measures commonly used in literature: Information Share (Hasbrouck, 1995), Component Share (Gonzalo & Granger, 1995), and Information Leadership Share (Yan & Zivot, 2010; Putniņš, 2013). Each of the three metrics relies on the existence of a cointegrating relationship between two price series. All three metrics allow for short term deviation of price series one from the another, while convergence to the intrinsic relationship connecting them (the value of the underlying asset) is assumed in the long run. Both IS, and CS are derived from the parameters of Vector Error Correction Model (VECM):

$$\begin{aligned} \Delta p_{1,t} &= \alpha_1\left(p_{1,t-1} - p_{2,t-1}\right) + \sum_{i=1}^{200} \gamma_i \Delta p_{1,t-i} + \sum_{j=1}^{200} \delta_i \Delta p_{2,t-j} + \varepsilon_{1,t} \\ \Delta p_{2,t} &= \alpha_2\left(p_{1,t-1} - p_{2,t-1}\right) + \sum_{k=1}^{200} \varphi_i \Delta p_{1,t-k} + \sum_{m=1}^{200} \phi_i \Delta p_{2,t-m} + \varepsilon_{2,t} \end{aligned} \tag{8}$$

where as p_1 and p_2 the natural log of the prevailing mid-quote is used in order to remove the bid-ask bounce. As shown by Putniņš (2013), and following Baillie, Geoffrey Booth, Tse, & Zabotina (2002), CS is calculated from the normalized orthogonal to the vector error correction coefficients:

$$CS_1 = \gamma_1 = \frac{\alpha_2}{\alpha_2 - \alpha_1}\ , \qquad CS_2 = \gamma_2 = \frac{\alpha_1}{\alpha_1 - \alpha_2} \tag{9}$$

[19] Lee and Ready (1991) algorithm identifies any trade that took place at a price below the prevailing midquote as seller initiated, and as buyer initiated if the price is above the midquote

While it is possible for CS to end up negative in certain cases, this lacks economic sense. As noted by Harris, McInish, & Wood (2002) market which always impounds the efficient price first should have a CS of 1. We therefore truncate to 1 every value of CS greater than 1, and to 0 every negative value.

Information Share is found using the Cholesky factorization (M) of covariance matrix of reduced VECM errors (Ω), $\Omega = MM'$:

$$M = \begin{pmatrix} m_{11} & 0 \\ m_{21} & m_{22} \end{pmatrix} \quad (10)$$

where the IS is found as:

$$IS_1 = \frac{(\gamma_1 m_{11} + \gamma_2 m_{12})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2}, \qquad IS_2 = \frac{(\gamma_2 m_{22})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2} \quad (11)$$

Since the order of the price series in the VECM affects the results of the Cholesky decomposition, the followed procedure is the one used by Baillie et al. (2002): IS is found as a simple average of the two values estimated for each possible ordering.

Yan & Zivot (2010) demonstrate in detail how to interpret these two measures and their shortcomings. They note that both IS and CS are biased towards the less noisy market, while only IS provides information on relative informativeness of a market. The two measures, therefore, accurately describe the first mover split between two markets only if those two markets have similar levels of trade noise present. In addition, Yan & Zivot (2010), propose a new metric combining the two existing measures with a purpose of cancelling out the lower-noise-bias:

$$IL_1 = \left| \frac{IS_1}{IS_2} \frac{CS_2}{CS_1} \right|, \qquad IL_2 = \left| \frac{IS_2}{IS_1} \frac{CS_1}{CS_2} \right| \quad (12)$$

Putniņš (2013) scales these two variables to 100%:

$$ILS_1 = \frac{IL_1}{IL_1 + IL_2}, \qquad ILS_2 = \frac{IL_2}{IL_1 + IL_2} \quad (13)$$

and provides evidence through extensive simulations that the new metric accurately identifies the market where the information gets impounded into prices regardless of the informativeness/noise characteristics of the two price series.

5. RESULTS

5.1. Descriptive Statistics

Table 1 reports the summary statistics of liquidity and trading metrics for a sample period extending three years of trading in the front contract, July 1, 2014 to Jun 30, 2016, sampled 19 hours a day, with a 5-hour gap from 4:00 p.m. to 9:00 p.m. ET. We observe that the overall liquidity and trading activity is much higher on the CME than on ICE: spreads on ICE are almost double the ones on CME, while the number of contracts available for trade at best quotes is half the number. In addition, the trading activity on CME is almost 4 times higher than on ICE, as depicted in average daily open interest values.

<INSERT TABLE 1>

5.2. Liquidity Regression Results

Tables 2 through 6 report the results of the diff-in-diff analysis of the impact of the staggered modification of the data pricing schedule on liquidit. Coefficient estimates for bid-ask spread, best depth, and effective spread are presented in panels A, B, and C of Table 2, respectively. Tables 3 through 6 present the decomposition of the spread on realized and adverse component; we estimate the realized and adverse components of the spread 1,5,30, and 60 seconds following the trade. The regression results prove to be inconclusive, contrary to the predictions of Cespa & Foucault (2014) and Easley et al. (2016). The increase in the price of data feed does not significantly impact the liquidity in the market around Events 1 and 2. Final removal of the data fee waiver causes an increase in the effective and realized spread, indicating a higher fee for liquidity provision stemming, in part, from the lower price impact, as seen in the decrease of adverse spread following Event 3, consistent with a decrease in activity of speculative traders. This effect is, however, coupled with a statistically significant increase in available depth on the CME, therefore having an ambiguous effect on overall liquidity.

<INSERT TABLE 2>

<INSERT TABLE 3>

<INSERT TABLE 4>

<INSERT TABLE 5>

<INSERT TABLE 6>

5.3. Price Discovery Regression Results

Table 7 reports coefficient estimates and p-values for Equation 1. Panels A, B, and C report results for CS, IS, and ILS, respectively. Each of the price discovery measures is expressed in terms of the proportion of price discovery on CME.

Table 7 documents that, after controlling for relative bid-ask spread and relative open interest, the increase in the price of data adversely affects the efficiency of price discovery. These results imply that following the staggered removal of the data fee waiver on the CME (Events 2 and 3), a fraction of the traders decides to stop purchasing the data, which lowers the informational efficiency of the CME relative to ICE. This is in line with Cespa & Foucault (2014) theory, and shows that profit maximizing behavior of the exchanges hurts the quality of the market. In addition, the decrease in the number of traders purchasing access to the data feed is confirmed in the Chicago Mercantile Exchange Annual Report (2016) noting that the increase in data revenues *"... was partially offset by some rationalization as customer firms transitioned into full-priced offerings, as well as reductions in overall screen counts at member banks."*

<INSERT TABLE 7>

The increase in price discovery following Event 1 is explained by strategic behavior of rational traders. After learning that starting March 1, 2014 the fee will be charged to all traders previously not having subscription to the data feed and being part of the fee waiver program, traders sign up for the service in the period leading up to the event, thus increasing the presence of informed agents. Our interpretation is confirmed by the behavior of Algorithmic Trading (AT) Volume, a proxy for AT activity used in Hendershott, Jones, & Menkveld (2011).[20] AT proxy demonstrates an increased presence of fast traders in the period leading up to March 1, 2014, as

[20] AT proxy is calculated as follows:

$$AT_t = \frac{messages_t}{trades_t}$$

where $messages_t$ is the frequency of best quote updates and $trades_t$ is the number of trades on day *t*. In addition, the same metric is used as a proxy for HFT in Malceniece, Malcenieks, & Putniņš (2019). Regardless of the slightly different interpretations of this metric in Hendershott et al. (2011) and Malceniece et al. (2019), in our context it fits the purpose of identifying the activity of "fast" traders that by definition require access to real-time price information.

seen on Graph 2: AT activity increases on CME, despite of the declining trend on ICE. In addition, Table 8 reports results of a diff-in-diff analysis of AT behavior in the period leading up to Event 1. Statistically significant DID coefficient corroborates our interpretation: algorithmic traders joined the exchange prior to the March 1, 2014 cutoff date in order to be able to enjoy the waiver until the end of 2014. This result further corroborates Cespa & Foucault (2014) assertion that the efficiency of price discovery is a function of competition between traders having access to real time quote information.

<INSERT GRAPH 2>

<INSERT TABLE 8>

6. CONCLUSION

Our analysis examines how changes in the price of data fee affect liquidity and efficiency of price discovery. We empirically verify the theory of Cespa & Foucault (2014) that posits that increases in data fees draw some of the traders away from obtaining the price information, which adversely affects the speed and efficiency of incorporation of new information in price series. We do not, however, find supporting evidence of an adverse effect of data fee increase on market liquidity.

Our findings provide corroborating empirical evidence to the Cespa & Foucault (2014) conclusions on the relevance of data fees as price discovery determinants and indicate that the regulators should pay close attention to the issue of data fees since they not only redistribute income from the traders to the exchanges, but also affect the quality of the market and price discovery as one of market's most important functions.

REFERENCES

Angel, J. J., & McCabe, D. M. (2018). Insider Trading 2.0? The Ethics of Information Sales. *Journal of Business Ethics*, *147*(4), 747–760. https://doi.org/10.1007/s10551-016-3391-4

Baillie, R. T., Geoffrey Booth, G., Tse, Y., & Zabotina, T. (2002). Price discovery and common factor models. *Journal of Financial Markets*, *5*(3), 309–321. https://doi.org/10.1016/S1386-4181(02)00027-7

Bergemann, D., Bonatti, A., & Smolin, A. (2018). The Design and Price of Information. *American Economic Review*, *108*(1), 1–48. https://doi.org/10.1257/aer.20161079

Board of Trade v. Christie Grain & Stock Co., 198 U.S. 236. (1905). Retrieved from http://cdn.loc.gov/service/ll/usrep/usrep198/usrep198236/usrep198236.pdf

Cespa, G., & Foucault, T. (2014). Sale of Price Information by Exchanges: Does It Promote Price Discovery? *Management Science*, *60*(1), 148–165.

Chicago Mercantile Exchange Annual Report. (2016). Retrieved from http://www.annualreports.com/HostedData/AnnualReportArchive/c/NASDAQ_CME_2016.pdf

Chordia, T., Roll, R., & Subrahmanyam, A. (2000). Co-Movements in Bid-Ask Spreads and Market Depth. *Financial Analysts Journal*, *56*(1), 23–27. https://doi.org/10.2469/faj.v56.n5.2386

Clayton, J. (2018). Statement on Market Data Fees and Market Structure. Retrieved February 7, 2019, from https://www.sec.gov/news/public-statement/statement-chairman-clayton-2018-10-16

Demsetz, H. (1968). The Cost of Transacting. *The Quarterly Journal of Economics*, *82*(1), 33–53. https://doi.org/10.2307/1882244

Easley, D., O'Hara, M., & Yang, L. (2016). Differential Access to Price Information in Financial Markets. *Journal of Financial and Quantitative Analysis*, *51*(4), 1071–1110. https://doi.org/10.1017/S0022109016000491

Frijns, B., Gilbert, A., & Tourani-Rad, A. (2015). The determinants of price discovery: Evidence from US-Canadian cross-listed shares. *Journal of Banking and Finance*, *59*, 457–468.

Gatewood v. North Carolina, 27 S.Ct 167, 168. (1906).

Glosten, L. R. (1987). Components of the Bid-Ask Spread and the Statistical Properties of Transaction Prices. *The Journal of Finance*, *42*(5), 1293–1307.

Gonzalo, J., & Granger, C. (1995). Estimation of Common Long-Memory Components in Cointegrated Systems. *Journal of Business & Economic Statistics*, *13*(1), 27–35. https://doi.org/10.2307/1392518

Grossman, S. J., & Stiglitz, J. E. (1980). On the Impossibility of Informationally Efficient Markets. *The American Economic Review*, *70*(3), 393–408. Retrieved from https://www.jstor.org/stable/1805228

Harris, F. H. deB., McInish, T. H., & Wood, R. A. (2002). Security price adjustment across exchanges: an investigation of common factor components for Dow stocks. *Journal of Financial Markets*, *5*(3), 277–308. https://doi.org/10.1016/S1386-4181(01)00017-9

Harris, L. E. (1994). Minimum Price Variations, Discrete Bid-Ask Spreads, and Quotation Sizes. *The Review of Financial Studies*, *7*(1), 149–178. https://doi.org/10.1093/rfs/7.1.149

Hasbrouck, J. (1995). One Security, Many Markets: Determining the Contributions to Price Discovery. *The Journal of Finance*, *50*(4), 1175–1199. https://doi.org/10.2307/2329348

Hasbrouck, J., & Saar, G. (2013). Low-latency trading. *Journal of Financial Markets*, *16*, 646–679. https://doi.org/10.1016/j.finmar.2013.05.003

Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does Algorithmic Trading Improve Liquidity? *The Journal of Finance, 66*(1), 1–33.

Hunt v. New York Cotton Exchange, 205 U.S. 322. (1907). Retrieved from https://supreme.justia.com/cases/federal/us/205/322/

Lee, C. M. C., & Ready, M. J. (1991). Inferring Trade Direction from Intraday Data. *The Journal of Finance*, *46*(2), 733–746. https://doi.org/10.2307/2328845

Malceniece, L., Malcenieks, K., & Putniņš, T. J. (2019). High Frequency Trading and Co-Movement in Financial Markets. *Journal of Financial Economics (Forthcoming)*. https://doi.org/10.2139/ssrn.2827148

Mcinish, T. H., & Wood, R. A. (1992). An Analysis of Intraday Patterns in Bid/Ask Spreads for NYSE Stocks. *The Journal of Finance*, *47*(2), 753–764. https://doi.org/10.1111/j.1540-6261.1992.tb04408.x

Mulherin, J. H., Netter, J. M., & Overdahl, J. A. (1991). Prices Are Property: The Organization of Financial Exchanges from a Transaction Cost Perspective. *Journal of Law & Economics*, *34*, 591–644. https://doi.org/https://doi.org/10.1086/467237

O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics*, *116*(2), 257–270. https://doi.org/10.1016/J.JFINECO.2015.01.003

Petition for Rulemaking Regarding Market Data Fees and Request for Guidance on Market Data Licensing Practices; Investor Access to Market Data. (2018). Retrieved from https://www.sec.gov/rules/petitions/2018/petn4-728.pdf

Putniņš, T. J. (2013). What do price discovery metrics really measure? *Journal of Empirical Finance*, *23*, 68–83. https://doi.org/10.1016/j.jempfin.2013.05.004

Rulemaking petition concerning market data fees. (2017). Retrieved from https://www.sec.gov/rules/petitions/2017/petn4-716.pdf

Securities and Exchange Comission. (2018a). *Release No. 34-83148*. Washington, D.C. Retrieved from https://www.sec.gov/rules/sro/nms/2018/34-83148.pdf

Securities and Exchange Comission. (2018b). *Release No. 84432*. Washington, D.C. Retrieved from https://www.sec.gov/litigation/opinions/2018/34-84432.pdf

Simonov, A. (1999). *Competition In Markets For Information*. Fontainebleau, France. Retrieved from http://www.andreisimonov.com/pdf/comp.pdf

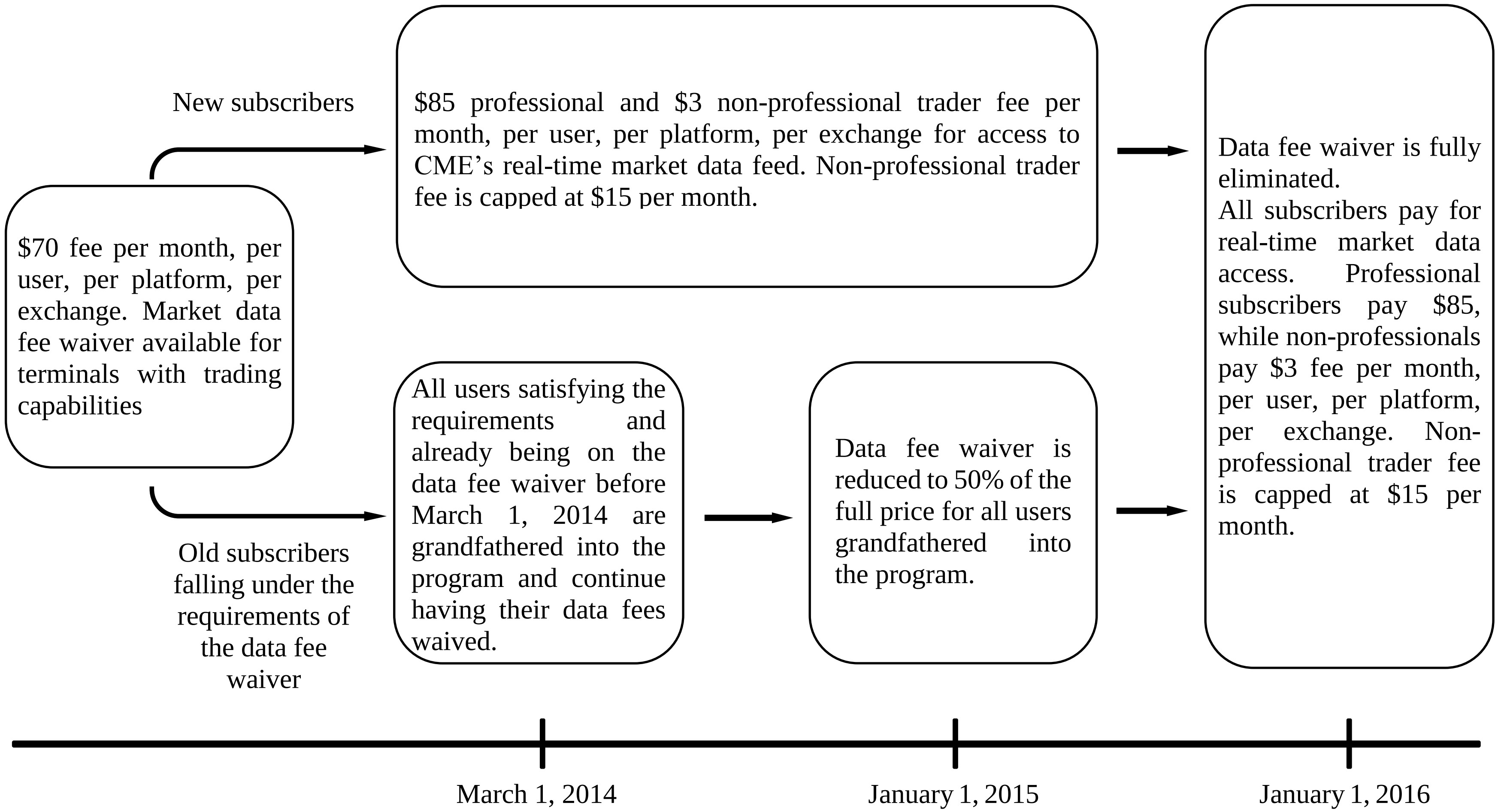
U.S. Department of the Treasury. (2017). *A Financial System That Creates Economic Opportunities*. Retrieved from https://www.treasury.gov/press-center/press-releases/documents/a-financial-system-capital-markets-final-final.pdf

Webb, R. I. (2003). Transitory real-time property rights and exchange intellectual property. *Journal of Futures Markets*, *23*(9), 891–913. https://doi.org/10.1002/fut.10097

Yan, B., & Zivot, E. (2010). A structural analysis of price discovery measures. *Journal of Financial Markets*, *13*(1), 1–19. https://doi.org/10.1016/J.FINMAR.2009.09.003

Graph 1

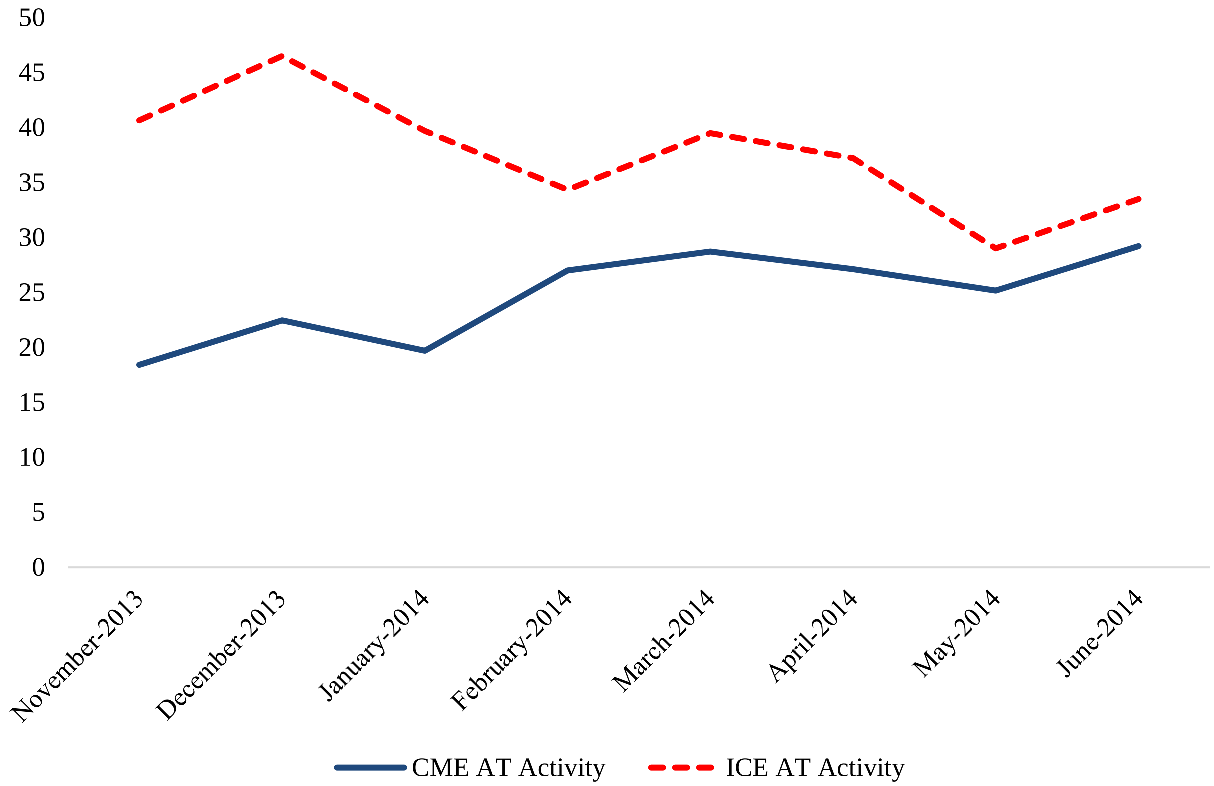
Timeline of CME's data feed fee changes



Note: This Graph reports the timeline of data feed fee changes in the period extending March 1, 2014 to January 1, 2016.

Graph 2

AT Activity leading up to March 1, 2014



Note: This Graph reports the monthly averages of the AT proxy on CME and ICE across the sample spanning November 1, 2013 – June 30, 2014. AT proxy is calculated as follows:

$$AT_t = \frac{messages_t}{trades_t}$$

where $messages_t$ is the frequency of best quote updates and $trades_t$ is the number of trades on day *t*.

TABLE 1

Summary statistics

	Mean	Median	Std Dev	Q1	Q3
Panel A: CME Summary Statistics					
Bid-Ask Spread	1.3058	1.2189	0.3534	1.1778	1.2839
Best Depth	29.6616	26.7434	11.8745	21.2327	34.1422
Open Interest (in '00000)	2.7265	2.8346	1.5495	1.5643	3.7208
Volatility	0.2201	0.1317	0.2636	0.0759	0.2639
Panel B: ICE Summary Statistics					
Bid-Ask Spread	2.3032	2.2456	0.3462	2.0893	2.4080
Best Depth	15.5687	13.8245	5.1901	11.9376	18.2392
Open Interest (in '00000)	0.6617	0.6601	0.2588	0.4869	0.8137
Volatility	0.2275	0.1365	0.2721	0.0759	0.2840

Note: This Table reports summary statistics for trading in WTI futures on CME and ICE reported in Panels A and B, respectively. The sample period extends three years of trading, July 1, 2014 to Jun 30, 2016, sampled 19 hours a day, with a 5-hour gap from 4:00 p.m. to 9:00 p.m. ET. Bid-ask spread is the daily average relative spread in bps; best depth is the sum of contracts available to trade at best bid and ask, averaged across the day; open interest is the total number of futures contracts outstanding at the end of the day; and volatility is the daily mid-quote volatility.

TABLE 2

Quoted Spread, Best Depth and Effective Spread

	Event 1		Event 2		Event 3	
	Coefficient Estimate	*p-value*	*Coefficient Estimate*	*p-value*	*Coefficient Estimate*	*p-value*
Panel A: Bid-Ask Spread						
Intercept	2.4670	0.0000	2.3332	0.0000	2.2998	0.0000
Event	-0.1852	0.0000	0.1288	0.0076	-0.1061	0.0010
Exchange	-0.6744	0.0000	-0.6981	0.0000	-0.7419	0.0000
Interaction	0.1114	0.0446	-0.0253	0.7145	0.0460	0.3132
Volatility	0.2765	0.0056	0.2083	0.0001	0.0313	0.4271
Open Interest (in '00000)	-0.2047	0.0000	-0.1643	0.0000	-0.0935	0.0000
R-squared	0.7293		0.6737		0.7993	
Panel B: Best Depth						
Intercept	15.1076	0.0000	19.2533	0.0000	11.7755	0.0000
Event	7.5939	0.0000	-7.0384	0.0000	5.8649	0.0000
Exchange	4.9267	0.0000	5.5249	0.0000	16.9225	0.0000
Interaction	0.3065	0.6707	5.3914	0.0000	10.7106	0.0000
Volatility	-2.2702	0.0796	-4.9680	0.0000	-4.7587	0.0000
Open Interest (in '00000)	1.2638	0.0000	0.6530	0.0145	1.5501	0.0000
R-squared	0.6436		0.5098		0.8127	
Panel C: Effective Spread						
Intercept	0.0119	0.0000	0.0115	0.0000	0.0101	0.0000
Event	-0.0009	0.0250	0.0003	0.2967	-0.0019	0.0000
Exchange	-0.0028	0.0000	-0.0037	0.0000	-0.0037	0.0000
Interaction	0.0000	0.9513	0.0000	0.9024	0.0014	0.0000
Volatility	0.0016	0.1370	0.0021	0.0000	0.0010	0.0000
Open Interest (in '00000)	-0.0006	0.0036	-0.0004	0.0001	-0.0001	0.0136
R-squared	0.2662		0.5973		0.7421	

Note: This table reports results of the difference-in-difference analysis based on equation:

$$Liq_t = \beta_1 + \beta_2 Exchange_t + \beta_3 Event_t + \beta_4 DID_t + \beta_5 Volatility_t + \beta_6 OPINT_t$$

where Liq_t is either bid-ask spread, best depth, or effective spread, presented Panels A, B, and C, respectively; $Exchange_t$ is the dummy variable that takes value of 1 for CME and 0 for ICE; $Event_t$ takes the value of 1 after the event (March 1, 2014; January 1, 2015; January 1, 2016) and 0 in the period before the event; DID_t is the interaction between the Event and Exchange dummies; $Volatility_t$ is the daily mid-quote volatility; and $OPINT_t$ is the daily open interest (in ‘00000). Regressions are computed on a sample of 250 trading days, 125 before and after the event, corresponding to six months pre and post the event.

TABLE 3

Realized and Adverse 1-second Spread

	Event 1		**Event 2**		**Event 3**	
	Coefficient Estimate	*p-value*	*Coefficient Estimate*	*p-value*	*Coefficient Estimate*	*p-value*
Panel A: Realized Spread 1sec						
Intercept	0.0074	0.0000	0.0068	0.0000	0.0066	0.0000
Event	-0.0004	0.3453	0.0005	0.0143	-0.0026	0.0000
Exchange	-0.0044	0.0000	-0.0057	0.0000	-0.0056	0.0000
Interaction	-0.0002	0.7274	-0.0003	0.2906	0.0028	0.0000
Volatility	-0.0004	0.7298	0.0004	0.1399	0.0004	0.0440
Open Interest (in '00000)	-0.0004	0.0301	0.0000	0.9794	0.0001	0.0077
R-squared	0.3783		0.7442		0.7338	
Panel B: Adverse Spread 1sec						
Intercept	0.0046	0.0000	0.0047	0.0000	0.0033	0.0000
Event	-0.0005	0.0000	-0.0003	0.1112	0.0007	0.0000
Exchange	0.0016	0.0000	0.0021	0.0000	0.0011	0.0000
Interaction	0.0002	0.2938	0.0004	0.1482	-0.0013	0.0000
Volatility	0.0019	0.0000	0.0018	0.0000	0.0006	0.0000
Open Interest (in '00000)	-0.0001	0.0162	-0.0004	0.0000	-0.0001	0.0003
R-squared	0.4582		0.3379		0.1846	

Note: This table reports results of the difference-in-difference analysis based on equation:

$$Liq_t = \beta_1 + \beta_2 Exchange_t + \beta_3 Event_t + \beta_4 DID_t + \beta_5 Volatility_t + \beta_6 OPINT_t$$

where Liq_t is either realized or adverse 1-second spread, presented Panels A and B, respectively; $Exchange_t$ is the dummy variable that takes value of 1 for CME and 0 for ICE; $Event_t$ takes the value of 1 after the event (March 1, 2014; January 1, 2015; January 1, 2016) and 0 in the period before the event; DID_t is the interaction between the Event and Exchange dummies; $Volatility_t$ is the daily mid-quote volatility; and $OPINT_t$ is the daily open interest (in '00000). Regressions are computed on a sample of 250 trading days, 125 before and after the event, corresponding to six months pre and post the event.

TABLE 4

Realized and Adverse 5-second Spread

	Event 1		**Event 2**		**Event 3**	
	Coefficient Estimate	*p-value*	*Coefficient Estimate*	*p-value*	*Coefficient Estimate*	*p-value*
Panel A: Realized Spread 5 sec						
Intercept	0.0073	0.0000	0.0068	0.0000	0.0065	0.0000
Event	-0.0005	0.2656	0.0005	0.0353	-0.0025	0.0000
Exchange	-0.0045	0.0000	-0.0058	0.0000	-0.0056	0.0000
Interaction	-0.0002	0.7042	-0.0003	0.3331	0.0027	0.0000
Volatility	-0.0005	0.6634	0.0003	0.2200	0.0004	0.0516
Open Interest (in '00000)	-0.0004	0.0598	0.0000	0.5992	0.0001	0.0016
R-squared	0.3750		0.7357		0.7317	
Panel B: Adverse Spread 5 sec						
Intercept	0.0046	0.0000	0.0048	0.0000	0.0034	0.0000
Event	-0.0005	0.0005	-0.0002	0.2389	0.0006	0.0000
Exchange	0.0017	0.0000	0.0022	0.0000	0.0011	0.0000
Interaction	0.0002	0.2788	0.0004	0.1925	-0.0012	0.0000
Volatility	0.0021	0.0000	0.0018	0.0000	0.0005	0.0001
Open Interest (in '00000)	-0.0002	0.0021	-0.0004	0.0000	-0.0001	0.0000
R-squared	0.4235		0.3290		0.1708	

Note: This table reports results of the difference-in-difference analysis based on equation:

$$Liq_t = \beta_1 + \beta_2 Exchange_t + \beta_3 Event_t + \beta_4 DID_t + \beta_5 Volatility_t + \beta_6 OPINT_t$$

where Liq_t is either realized or adverse 5-second spread, presented Panels A and B, respectively; $Exchange_t$ is the dummy variable that takes value of 1 for CME and 0 for ICE; $Event_t$ takes the value of 1 after the event (March 1, 2014; January 1, 2015; January 1, 2016) and 0 in the period before the event; DID_t is the interaction between the Event and Exchange dummies; $Volatility_t$ is the daily mid-quote volatility; and $OPINT_t$ is the daily open interest (in ‘00000). Regressions are computed on a sample of 250 trading days, 125 before and after the event, corresponding to six months pre and post the event.

TABLE 5

Realized and Adverse 30-second Spread

	Event 1		**Event 2**		**Event 3**	
	Coefficient Estimate	*p-value*	*Coefficient Estimate*	*p-value*	*Coefficient Estimate*	*p-value*
Panel A: Realized Spread 30 sec						
Intercept	0.0072	0.0000	0.0067	0.0000	0.0065	0.0000
Event	-0.0006	0.1501	0.0002	0.4534	-0.0024	0.0000
Exchange	-0.0047	0.0000	-0.0060	0.0000	-0.0058	0.0000
Interaction	0.0000	0.9755	-0.0001	0.6948	0.0026	0.0000
Volatility	-0.0006	0.6027	0.0003	0.3599	0.0004	0.1374
Open Interest (in '00000)	-0.0002	0.2297	0.0002	0.1098	0.0002	0.0010
R-squared	0.3478		0.6636		0.6583	
Panel B: Adverse Spread 30 sec						
Intercept	0.0047	0.0000	0.0048	0.0000	0.0036	0.0000
Event	-0.0003	0.0783	0.0001	0.8106	0.0005	0.0013
Exchange	0.0019	0.0000	0.0024	0.0000	0.0020	0.0000
Interaction	0.0000	0.9438	0.0002	0.5666	-0.0012	0.0000
Volatility	0.0022	0.0000	0.0019	0.0000	0.0006	0.0038
Open Interest (in '00000)	-0.0003	0.0001	-0.0005	0.0000	-0.0003	0.0000
R-squared	0.2832		0.2511		0.1850	

Note: This table reports results of the difference-in-difference analysis based on equation:

$$Liq_t = \beta_1 + \beta_2 Exchange_t + \beta_3 Event_t + \beta_4 DID_t + \beta_5 Volatility_t + \beta_6 OPINT_t$$

where Liq_t is either realized or adverse 30-second spread, presented Panels A and B, respectively; $Exchange_t$ is the dummy variable that takes value of 1 for CME and 0 for ICE; $Event_t$ takes the value of 1 after the event (March 1, 2014; January 1, 2015; January 1, 2016) and 0 in the period before the event; DID_t is the interaction between the Event and Exchange dummies; $Volatility_t$ is the daily mid-quote volatility; and $OPINT_t$ is the daily open interest (in '00000). Regressions are computed on a sample of 250 trading days, 125 before and after the event, corresponding to six months pre and post the event.

TABLE 6

Realized and Adverse 60-second Spread

	Event 1		Event 2		Event 3	
	Coefficient Estimate	*p-value*	*Coefficient Estimate*	*p-value*	*Coefficient Estimate*	*p-value*
Panel A: Realized Spread 60 sec						
Intercept	0.0074	0.0000	0.0066	0.0000	0.0065	0.0000
Event	-0.0009	0.0593	0.0003	0.4031	-0.0023	0.0000
Exchange	-0.0050	0.0000	-0.0059	0.0000	-0.0058	0.0000
Interaction	0.0002	0.7963	-0.0002	0.6953	0.0024	0.0000
Volatility	-0.0019	0.0989	0.0001	0.7650	0.0001	0.6202
Open Interest (in '00000)	-0.0001	0.6869	0.0002	0.1413	0.0002	0.0009
R-squared	0.3295		0.5670		0.5980	
Panel B: Adverse Spread 60 sec						
Intercept	0.0045	0.0000	0.0049	0.0000	0.0035	0.0000
Event	-0.0001	0.7762	0.0000	0.9709	0.0004	0.0274
Exchange	0.0022	0.0000	0.0022	0.0000	0.0021	0.0000
Interaction	-0.0002	0.4939	0.0002	0.5783	-0.0010	0.0004
Volatility	0.0035	0.0000	0.0020	0.0000	0.0008	0.0012
Open Interest (in '00000)	-0.0005	0.0000	-0.0005	0.0000	-0.0003	0.0000
R-squared	0.2399		0.1895		0.1431	

Note: This table reports results of the difference-in-difference analysis based on equation:

$$Liq_t = \beta_1 + \beta_2 Exchange_t + \beta_3 Event_t + \beta_4 DID_t + \beta_5 Volatility_t + \beta_6 OPINT_t$$

where Liq_t is either realized or adverse 60-second spread, presented Panels A and B, respectively; $Exchange_t$ is the dummy variable that takes value of 1 for CME and 0 for ICE; $Event_t$ takes the value of 1 after the event (March 1, 2014; January 1, 2015; January 1, 2016) and 0 in the period before the event; DID_t is the interaction between the Event and Exchange dummies; $Volatility_t$ is the daily mid-quote volatility; and $OPINT_t$ is the daily open interest (in ‘00000). Regressions are computed on a sample of 250 trading days, 125 before and after the event, corresponding to six months pre and post the event.

TABLE 7

Regression Results

Panel A: Component Share (CS)

	Event 1		**Event 2**		**Event 3**	
	Coefficient Estimate	*p-value*	*Coefficient Estimate*	*p-value*	*Coefficient Estimate*	*p-value*
Intercept	3.7873	0.0000	6.9273	0.0000	6.3017	0.0000
Event	0.7528	0.0086	-0.4072	0.1443	-0.8446	0.0000
Relative BAS	-4.6348	0.0000	-5.1608	0.0000	-7.4525	0.0000
Relative Open Interest	0.3901	0.0049	-0.0109	0.9076	-0.1025	0.0593
R squared	0.2143		0.1692		0.4161	

Panel B: Information Share (IS)

	Event 1		**Event 2**		**Event 3**	
	Coefficient Estimate	*p-value*	*Coefficient Estimate*	*p-value*	*Coefficient Estimate*	*p-value*
Intercept	3.6954	0.0000	7.1559	0.0000	4.7072	0.0000
Event	1.1067	0.0000	-0.8885	0.0029	-1.1401	0.0000
Relative BAS	-3.2255	0.0001	-4.7084	0.0000	-4.2547	0.0000
Relative Open Interest	0.7681	0.0000	0.4431	0.0000	0.1970	0.0002
R squared	0.3240		0.2529		0.3714	

Panel C: Information Leadership Share (ILS)

	Event 1		**Event 2**		**Event 3**	
	Coefficient Estimate	*p-value*	*Coefficient Estimate*	*p-value*	*Coefficient Estimate*	*p-value*
Intercept	-0.1837	0.8744	0.4573	0.6953	-3.1889	0.0048
Event	0.7079	0.0575	-0.9626	0.0431	-0.5910	0.0909
Relative BAS	2.8185	0.0209	0.9047	0.5017	6.3956	0.0000
Relative Open Interest	0.7559	0.0000	0.9080	0.0000	0.5991	0.0000
R squared	0.0817		0.1241		0.1269	

Note: This table reports results of the following equation:

$$PD_t = \beta_1 + \beta_2 D_t + \beta_3 RelBAS_t + \beta_4 RelOPINT_t$$

PD_t is logit transformation of either of the three price discovery metrics of interest (CS, IS, or ILS in Panels A,B, and C respectively); D_t is the dummy variable that takes value of 1 after the event and 0 before the event; $RelBAS_t$ is the average daily percentage bid-ask spread on CME relative to the one on ICE; and $RelOPINT_t$ is the fraction of daily open interest on CME relative to daily open interest on ICE. Event 1 is the March 1, 2014 CME removal of the data fee waiver for all new traders; Event 2 is the January 1, 2015 CME increase of data fees for traders on the waiver of 50% of the professional data fee ($85); Event 3 is the January 1, 2016 removal of data fee waiver for all traders. All regressions are based on 250 trading days: 125 pre and post event date.

TABLE 8

AT Activity leading up to Event 1

	Coefficient Estimate	*p-value*
Intercept	61.3234	0.0000
Event	-16.5351	0.0040
Exchange	-5.6081	0.3925
Diff-in-diff	23.8313	0.0040
Volatility	-20.1880	0.1287
Open Interest (in ‘00000)	-14.7010	0.0000
R-squared	0.3999	

Note: This table reports results of the difference-in-difference analysis based on equation:

$$AT_t = \beta_1 + \beta_2 Exchange_t + \beta_3 Event_t + \beta_4 DID_t + \beta_5 Volatility_t + \beta_6 OPINT_t$$

where AT_t is the Algorithmic Trading proxy of Hendershott et al. (2011); $Exchange_t$ is the dummy variable that takes value of 1 for CME and 0 for ICE; $Event_t$ takes the value of 1 in the period of 25 days prior to event, and 0 in the period 26-50 days before the event; DID_t is the interaction between the Event and Exchange dummies; $Volatility_t$ is the daily mid-quote volatility; and $OPINT_t$ is the daily open interest (in ‘00000). Regression is computed on a sample of 50 observations.