Competition and Exchange Data Fees

Jonathan Brogaard, James Brugler and Dominik Rösch *

November 12, 2020

Abstract

Exchanges are monopolist suppliers of their own order book data. We examine three events where exchanges begin charging a fee for order book data for the first time and test whether or not these fees affect their market share in a difference-in-differences setting. We find that the introduction of fees leads to a fall of market share of around 5-7.5 percent. Examining volume executed by intermarket sweep orders, transaction cost decompositions and information shares indicates that order routing decisions of informed traders are relatively more sensitive to order book data fees than other trader categories.

JEL Codes: G12, G14

Keywords: Data Fees, Exchange Competition

* Brogaard is at the David Eccles School of Business, University of Utah, Email: brogaardj@eccles.utah.edu; Phone: +1 (801) 581-7463. Brugler is at the University of Melbourne, Email: james.brugler@unimelb.edu.au; Phone: +61 3 8344 1164. Rösch is at the State University of New York at Buffalo, Email: drosch@buffalo.edu.

1 Introduction

A major source of revenue for exchanges is the sale of their data. For example, in recent years Nasdaq sells over 90 data products contributing to almost 20% of their total revenue.[1] Exchanges sell faster access to their core data—such as best bid and ask or last trade prices—and proprietary non-core data—such as bid and ask quotes worse than the best ones.

Recently, a spat has erupted between the SEC (and many investors) and the exchanges over whether data fees are too high. The data that exchanges sell is to some degree a public good (Alan and Schwartz, 2013) and plays an important role in the price discovery process. Investors argue that exchanges are exclusive processors of essential information and therefore high data fees are a way to extract economic rents from a captured audience.[2] Exchanges argue that competition amongst the exchanges means investors choose to not subscribe to their data feeds and focus on trading on other venues. In competitive markets rents should diminish and any exchange charging too high data fees would be competed out of business. The concerns therefore boil down to the question whether the equity market exchange structure is competitive (Spatt, 2020).

This article tests whether investors change their behavior due to fee changes, indicating that exchanges face competition, and how the changes impact market quality.

One simple means of examining the claims is to study how market share reacts to data fee changes. An increase in fees corresponding to a decrease in market share is evidence

[1] SEC (2016, p.11). The exchange data environment and fee structure is complex. For example, see `https://www.nasdaqtrader.com/Trader.aspx?id=Global_Market_Data` or `https://www.nyse.com/market-data` for a list of the products offered by NASDAQ and NYSE, respectively. In addition to the data itself other services, such as colocation, are also available from many of the exchanges.

[2] The Managed Funds Association and The Alternative Investment Management Association filled a petition regarding the SEC ruling on August 22, 2018 that sums up the concern: "The Associations are concerned that exchanges as exclusive processors are charging unreasonable fees for market data products, and as a consequence, restricting trade and harming competition." (MFA and AIMA, 2018).

that exchanges are subject to market forces to some degree. We study three discrete data fee events in the U.S. Equity markets that occurred on ARCA, Direct Edge, and BATS. We find that when a trading venue increases its data fees, it loses market share. Next, we investigate the question why market-share drops. Traders decide where to post limit-orders and where to route intermarket sweep orders (ISOs). But traders have little control over where market-orders are routed, because of Reg NMS market-orders have to be executed at the exchange posting the national best bid or offer (NBBO). Therefore, market-share can drop because traders avoid posting limit orders, resulting in the exchange quoting prices less often at NBBO and therefore executing less market-orders. Market-share can also drop because traders route less ISOs to the treated exchange.

The results are consistent with investors being able to pick and choose their data needs and being able to vote with their wallets. Besides trading volume, the fee introducing exchange decreases the amount of time it has quotes at the NBBO. Finally, we also observe a decline in ISOs on the fee introducing exchange.

Of course, it is challenging to establish causality between a drop in market-share and the introduction of the fee. For example, exchanges often have higher market-share in specific segments of the market and their market share is therefore likely subject to fads. The dot-com boom in 2001 disproportionately affected technology companies listed at the Nasdaq. In general, unobserved heterogeneity can lead to an omitted-variable bias and preclude causal inference. We pursue a difference-in-difference methodology where the treated stock is the stock trading on the exchange introducing the fee and the control group is the same stock traded on all other exchanges. Instead of controlling for potentially endogenous control variables, which would preclude causal inference (Lewbel, 2019, p.853), we address omitted variables using many fixed effects (following Gormley and Matsa, 2014).

Our results suggest that trading venues must optimize a trade off between increasing data fees and losing investors subscriptions. The results show that some investors choose not to subscribe (i.e. some investors are sensitive to price at the extensive margin).

We next turn to understanding who, at the margin, is changing their trading behavior. While we cannot directly observe the behavior of different types of investors, we can infer whose behavior changed based on changes in market quality. We more fully explore the theoretical mechanisms that would link different types of traders to different market quality outcomes. We focus on three standard categories of market participants: market makers, informed traders, and liquidity traders.

Theory suggests that with less information from a given venue, an investor faces greater execution risk when sending an order to that venue (because they have less information with which to assess execution probability) and also face greater information asymmetry against the traders who do subscribe. Therefore, it is more likely for an investor who does not subscribe to a venues' data feed to route their orders elsewhere. Different categories of traders' orders will have different characteristics. By examining the average characteristics of trades on venues introducing fees compared with other venues, we can infer which category of traders are more sensitive to market data fees.

While informed and uniformed traders use passive limit orders to enter trades, the use of ISOs tends to be by informed traders Chakravarty et al. (2012). The earlier finding that ISO trades decline even more than non-ISO trades is the first indicator that informed investors are the ones disproportionally changing their behavior.

For further evidence, we use the same setup as we do when examining the change in volume and find that the introduction of a data fee results in price impact falling and realized spread increasing. These changes are most consistent with informed traders decreasing their

trading on the venue increasing a data fee. Note that we can only make statements about the relative change among the market participants. It could be, for example, that the introduction of a data fee impacts the trading activity of both market makers and informed traders, but that the informed traders are impacted to a greater degree.

The final piece of evidence comes from calculating the information share coming from the fee-introducing venue versus the other venues. The information share estimates from which exchange price innovations materialize. The results indicate that the venue with the new fee experiences a decline in its information share, consistent with informed trading choosing not to use its venue after the data fee is put into place.

To further enrich how data feeds are used and how data fees alter that usage we examine impact on large versus small stocks. The trading volume, time at national best bid and offer, and non-ISO volume all decline substantially more for small than large stocks. In other words, trading volume at the treated exchange for large and for small stocks seems to drop for different reasons. Large stocks are affected because traders do not route ISOs and small stocks are affected because traders send less limit orders to the treated exchange. Less limit orders results in less time the exchange quotes prices at NBBO and therefore attracts less market-orders than have to be routed to the exchange with best prices (Reg NMS). Consistent with this we find that the information share of large stocks drops more than for small stocks.

Our paper adds to the literature on exchange data fees. A broader literature examines other dimensions of exchange competition. For example, Ramos and von Thadden (2008) focuses on how transaction costs impacts exchange competition and Pagnotta and Philippon (2018) studies the role of speed among exchanges. Hendershott and Jones (2005) study the removal of market data for three ETFs trading on Island ECN. Like us, they document a large

and significant drop volume traded for these ETFs on Island ECN after the venue ceases to display order book data to any traders for these securities. An important difference between this event and our study is that, in their case, order book information is unavailable at any price to any trader. In our study, market data is available for a price, and so both access to market data and subsequent order routing decisions are strategic decisions. Hendershott and Jones (2005) find that liquidity traders are less likely to route their orders to Island ECN after market data is removed, while we find that informed traders are more sensitive to the existence of fees.

Our focus on data fees specifically most closely relates to Jones (2018) and Hendershott, Rysman, and Schwabe (2020). Jones (2018) argues that though market data is valuable to market participants, market data fees are a small cost to the industry overall and the market for data is characterized by "robust competition". Hendershott et al. (2020) shows that the NYSE gains market share after the introduction of new market data products, the NYSE Integrated Feed, demonstrating that this information is valuable to traders. They further demonstrate the existence of a positive externality for market data: traders who do not subscribe also increase their trading on the NYSE after the Integrated Feed begins. Our results are consistent with these findings. We contribute to this literature by first documenting that the introduction of fees for market data leads to lower market share, and by identifying informed traders as the most affected trader categories after fees are introduced.

2 Three Fee Introductions

Fees for order book information are complex and frequently changing. Due to the complexity of fees, different market participants face different total costs for market data and it is difficult

to summarize the average cost of market data in a simple way. We study three events where exchanges that introduce fees for the first time. These events are advantageous because we can clearly identify instances where the cost of market data unambiguously increase for all market participants.

The first event is the introduction of market data fees by ARCA on 1 January 2009. Prior to this date, order book information from ARCA was provided free of charge. In 2006, the exchange filed with the SEC to establish fees for its market data (SEC, 2008). The filing was subsequently approved in December 2008 and fees were introduced in 2009. These initial fees included a $750 per month access fee, a monthly fee of $15 per device for professional subscribers for ETFs and equities data as governed by the CTA Plan and $15 per device and month for data covering other equities securities governed by the Nasdaq UTP plan. For non-professional subscribers, the monthly fees per device were $5 for both categories.

The second event is the introduction of fees for depth of order book data for trading on the Direct Edge EDGX exchange. This real-time data contains information on all displayed orders, executions and order modifications and was initially provided free-of-charge to all Member firms. On 1 May 2012, Direct Edge introduced monthly fees of $500 for internal use of the feed and $2,500 for users who wish to distribute the feed externally (CBOE, 2012).

The final event is the introduction of market data fees by BATS exchange on 1 July 2013. BATS fee structure initially comprised of a total charge of $1,500 per month for depth of book data for internal use only, covering both the BZX and BYZ exchange (Reuters, 2013). For top of book and last sale data, these fees total $1,000 across both sub-exchanges. For firms that wish to distribute the data externally, the total monthly costs would be $7,500 for depth of book data and $5,000 for top of book and last sale data.

3 Hypothesis Development

Traders faced with fees for order book information must decide whether or not the expected benefit of having this information outweighs the explicit costs of purchasing it. It may be that order book information from venues introducing fees is so important that practically all market participants are forced to subscribe to the data after fees are introduced. If so, we argue there should be no change in market conditions (market shares, trading costs across venues etc) after fees are introduced compared with prior to the introduction of fees. Nothing has changed in terms of the amount or distribution of information throughout the trading economy. Order submission strategies that were optimal in the absence of fees would remain optimal when fees are introduced. In this case, the venue introducing fees suffers no loss of market share and fees are just a mechanism to transfer surplus from the trading process from traders to the exchange.

If instead, at least some traders are sensitive to the introduction of fees at the extensive margin (i.e. they choose not to subscribe to order book data), then we argue that they should become less likely to submit orders to that venue and the venue introducing fees will lose market share. This leads to the first testable hypothesis regarding data fees and trading outcomes: at least some traders are sensitive to the cost of order book information and choose to route their orders elsewhere after data fees are introduced, leading to lower market share.

We then consider the incentives of different trader types to subscribe to fees and how they may rationally respond when fees are introduced. The presence of Reg NMS means that market orders will be routed to the exchange with the best prices. Therefore access to order book information will most likely affect the decision by traders to submit limit orders or to submit intermarket sweep orders (ISOs), where traders with direct market access have

the discretion to exclude particular exchanges. Market shares could fall because exchanges attract fewer limit orders after fees (and so spend less time at the NBBO and have less market orders routed to it) or because they attract fewer ISOs.

The expected benefit of having order book information can differ for different trader types. For example, market makers providing liquidity without this information face an important source of information asymmetry compared with other market makers who do have this information. These market makers will face severe adverse selection problems from not being able to update their quotes after changes in the state of the order book. Though order book information would appear to be a key requirement for any market maker wishing to provide liquidity on a particular venue, it does not follow that all market makers would necessarily subscribe to the order book information for a given venue. Instead, smaller market makers who make low levels of profit on the venue may rationally not subscribe and cease to provide liquidity on that venue. If costs of entry in liquidity provision are sufficiently low that the marginal liquidity provider makes close to zero expected profit, then the marginal provider will rationally withdraw from that particular venue when fees are introduced.

For informed traders who seek to earn positive expected profits by trading on private information, the decision to subscribe to data fees has some similar features to the decision for market makers. The expected benefit of subscribing for informed traders can come from enhanced information about the state of the market-wide order book, which can be informative about the value of the asset (e.g. if limit orders contain information about fundamental value as in Brogaard, Hendershott, and Riordan 2019). The benefit of subscribing can also come from the ability to condition limit orders and ISO routing decisions based on the state of the order book and design trading strategies that minimize costs accordingly.

The fee-introducing venues that we study have sizeable market shares (around 10% on

average), but it is not obvious that informed traders need to condition on the limit orders placed on these venues to better estimate the fundamental value of the security. Clearly trading costs are a very important consideration for informed traders, as these costs act as a tax on the expected return from information acquisition. If having order book information can help minimize these costs, then informed traders may choose to pay the fees. The fees themselves however also act as a (lump sum) tax on expected profits, so informed traders may well choose to not pay this tax and route orders to competitor venues after fees are introduced. A break-even condition when information acquisition is costly would imply that the marginal informed trader will not be able to pay the fee and still earn non-negative profits.

In many standard market microstructure models such as Kyle (1985), liquidity traders are assumed to either demand or supply shares in a way that is inelastic to price. In other words, these traders submit a market order regardless of price or trading costs. While this is clearly a highly stylized model of uninformed trader behavior, it is useful to consider how fees for order book information may affect such a trader. Presumably, such a liquidity trader with inelastic demand or supply would not choose to pay a lump sum tax in the form of order book data fees when they choose not to condition their orders on current information. They may still route orders some fraction of their orders to the venue or may exclude the venue from their routing decisions. A liquidity trader could instead rationally choose to subscribe if they believe doing so will allow them to trade cheaply on the venue and this leads to lower overall transaction costs compared with not having this information. The rational liquidity trader would have to analyze the costs and benefits of the order book data in a similar way to the informed trader and decide accordingly.

4 Data and Summary Statistics

Data for our analysis is obtained from the NYSE Trade and Quote (TAQ) database. The TAQ data contain trades and quotes for all securities listed on the NYSE, the American Stock Exchange (AMEX), the Nasdaq National Market System and Small Cap issues. Using the TAQ data, we calculate a set of liquidity and trading activity variables at the stock-day level separately for the exchange introducing the fee and for all other trading venues in three month windows either side of the fee introduction date. The sample includes all stocks in the S&P 500, S&P 400 and S&P 600 indices (collectively, the S&P 1500 index). Ticker identifiers for S&P 1500 members over the relevant sample periods are obtained from Thomson Reuters Tick History, maintained by Refinitiv.

The set of variables we analyse are the dollar volume traded, the dollar volume traded via intermarket sweep orders (ISOs), the time at NBBO, the information share of the treated exchange, the dollar average trade size, the average percent effective spread, percent realized spread, and percent price impact across trades. For all variables we estimate one on the trading venue introducing the market fee (the "treated" venue) and one across all other trading venues.

Log dollar volume is defined as the log of the total dollar value traded across the N_{ijt} trades in stock i on venue j and date t:

$$Log\,Dollar\,Volume_{ijt} = \ln \sum_{n=1}^{N_{ijt}} vol_{ijt}^{n} \tag{1}$$

Similarly, we compute the dollar volume traded via ISOs, but in this case we only sum over all ISO trades, i.e., trades with a condition containing "F" in the TAQ database.

To estimate the time each venue quotes prices at the NBBO, we estimate the NBBO

following Holden and Jacobsen (2014) and compare the NBBO each second to the quotes of the exchange. If the bid and the ask are equal to the national best bid and ask prices, this variable takes the value of one. We then sum up the time at NBBO over the whole continuous trading session (and drop stock-days with a continuous trading session of less than 6.5 hours). Note, that when using the time at NBBO we only have one observations per stock-day for the treated exchange. The time at NBBO for the other exchanges is automatically given by one minus the seconds the treated exchange quotes at NBBO prices.

The effective spread, realized spread and price impact for trade n in stock i on venue j and date t are respectively defined as:

$$ES^n_{ijt} = 100 \times \frac{D^n_{ijt}\left(P^n_{ijt} - Mid^n_{ijt}\right)}{Mid^n_{ijt}} \tag{2}$$

$$RS^n_{ijt} = 100 \times \frac{D^n_{ijt}\left(P^n_{ijt} - Mid^{n+5}_{ijt}\right)}{Mid^n_{ijt}} \tag{3}$$

$$PI^n_{ijt} = 100 \times \frac{D^n_{ijt}\left(Mid^{n+5}_{ijt} - Mid^n_{ijt}\right)}{Mid^n_{ijt}} \tag{4}$$

where D^n_{ijt} is an indicator variable equal to one if the n^{th} trade in stock i on venue j and date t is a customer buy and negative one if the trade is a customer sell, P^n_{ijt} is the trade price, M^n_{ijt} is the prevailing midquote at the time of the trade and M^{n+5}_{ijt} is the prevailing midquote five minutes after the trade time. Customer trade directions are assigned according to the Lee and Ready (1991) algorithm. For each trade-level variable defined in Equations (2), (3) and (4), our stock-day-venue value is the average across all trades in that stock-day-venue.[3] All variables are Winsorised at the 1% level to limit the effect of outliers on our results.

We also estimate the information share (IS) of the treated exchange from midpoint prices

[3]We gratefully use SAS code provided by Craig Holden and Stacey Jacobsen to calculate each of our variables, as described in Holden and Jacobsen (2014), with slight modifications to separate estimation into different venues.

prevailing trades. We estimate the IS of the treated exchange as the contribution of that exchange to the total variance of the common (random-walk) component, following Hasbrouck (1995). As in Hasbrouck (1995) we derive a lower and an upper band of the IS, because when prices at different venues move at the same time the IS cannot be uniquely attributed to any venue. In the regression we use the average of the lower and upper bound and also estimate a modified IS following Lien and Shrestha (2009).

To illustrate how we define our variables, ARCA first introduced fees for order book information on January 1, 2009. For every stock in the S&P 1500 index and every trading date in a three month window either side of this date, we calculate each activity or liquidity variable for all trades and quotes on ARCA and also for all trades and quotes on all other trading venues excluding ARCA. We then use the stock-day averages of trading activity and liquidity variables for trades and quotes on ARCA as our treated group and stock-day averages using all trades and quotes from other venues as the control group. We do this for each of our three events corresponding to fee introduction on ARCA, Direct Edge and BATS respectively. Our final sample consists of 1,774 stocks to trades in S&P 1500 stocks on ARCA, BATS and Direct Edge and the control venues, over three six month periods around each fee introduction event (368 trading days in total).[4] Table 1 contains summary statistics for each variable across the entire pooled sample and averages split by before and after fee introduction.

Table 1 about here

Average volume traded across all stock-day-venues rises slightly from around $32.3m before fee introduction to $34.9m in the post-fee introduction period. Quoted spreads trend

[4]The final number of stock venues is slightly larger than 1,500 because index members change over time.

slightly lower from around 36 bps in the pre-event period to 32.3bps in the post-event period, while dollar depth is largely unchanged. Average trade costs (effective spreads) are very similar in the pre- and post-event periods, though there are small decreases in average realized spreads and increases in price impact. These summary statistics primarily capture general trends in trading activity across the three sample periods.

Table 2 presents summary statistics split by venue type: exchanges introducing the fee (treated) vs. all other venues (control). Panel A contains averages across stock-days for trading on the treated exchanges while Panel B contains averages across stock-days on the controls venues. Trading activity on the treated venue represents approximately 12% of total volume traded across all stock-venue-days in our sample with $7.95m worth of stock traded on the treated exchange on average compared with $59.5m on other venues. There is also less liquidity at the quotes for the treated exchange compared with across all other venues (as measured by total depth available at the best bid and offer and the quoted spread at the best bid and offer). The average quoted spread across all stock-day-venues is 34 bps, however the corresponding levels are 48bps for the treated exchange and 20 bps for other venues. The average depth at the best bid and ask is $17,600 across all stock-venue-days, but is $8,890 and $24,160 for the treated exchange and the other venues respectively. The control group aggregates total liquidity across multiple venues (including the primary exchange for that stock and all other trading venues). It is not surprising that liquidity as measured at the quotes appears to be worse on the treated exchange than in the control group but due to aggregation across venues for the control group, this does not indicate that the treated exchange has substantially poorer liquidity at the quotes than any other specific venue.

Table 2 about here

There are generally small differences in average effective spreads, realized spreads or price impact between the treated and control exchanges. The average effective spread is close to 10.5bps for both the treated and control, which on average can decomposed into 7bps due to price impact and 3-4bps that accrues to the market maker as realized spread. Effective spreads trend slightly higher for trades on both the treated and the control exchanges, though the rise is larger for control exchanges. Realized spreads increase and price impact falls on both venue categories, but the effects are more pronounced for the treated exchange relative to the control venues. Prior to fee introduction, average realized spreads are approximately 0.6bps lower on the treated exchanges while after the fee introduction, they are 0.3bps smaller. Price impact is on average 0.15bps higher for trades on the treated exchanges prior to fee introduction but are 0.2bps lower after.

Figure 1 plots the average trading activity in the pre- and post-event periods. The y-axis depicts the average of log dollar value traded by venue type while the x-axis depicts the number of days before and after the fee-introduction date. The "treated" line is the daily simple average of log dollar volume on the treated exchanges across all stock-days in our sample. The "control" line is the equivalent average across the same stocks, but traded on all other venues.

In-line with Tables 1 and 2, average volumes traded are significantly lower on the treated exchanges in both the pre-event and post-event periods, and also appear to be trending slightly lower in both categories over the pre-event sample period, reflecting a general trend in total trading activity that is unlikely to be related to the introduction of fees.

Figure 1 about here

5 Do Market Participants Respond to Data Fees?

We first use a difference-in-differences approach to causally identify the effect of fees for order book information on trading volumes. Our strategy treats the introduction of fees as natural experiments that exogenously raise the price for order book information on one exchange but not others. For each of our three events, we construct a sample of stock-day observations of trading volume and costs for the venue introducing the fee and a sample of stock-day observations for all other venues where the stock is traded. We assign stock-days on the venue introducing fees as the treated group and stock-days on all other venues as the control group. We then pool the sample together across all three events and estimate the effect of fee introduction using difference-in-difference estimators stock-day fixed effects and a treatment dummy:

$$y_{ijt} = \alpha_{jt} + \beta_1 Treat_j + \beta_2 Treat_j \times Post_t + \varepsilon_{ijt} \tag{5}$$

where y_{ijt} is the trading activity or market quality for stock i on venue j (either the treated or control exchange) and date t, $Treat_j$ is a dummy for trades taking place on treated exchanges, $post_t$ is a dummy taking the value 1 if the date is after the fee introduction for venue j. The parameter β_2 identifies the causal effect of fee introduction in the absence of confounding trends or selection to treatment status based on unit-specific unobservables. In our case units are determined at the stock-venue level and so selection into the treatment or control group cannot plausibly depend on unobservable effects.

We also estimate a model with two-way stock-venue and date fixed effects. This model

is given by:

$$y_{ijt} = \alpha_{ij} + \gamma_t + \beta Treat_j \times Post_t + \varepsilon_{ijt} \tag{6}$$

where α_{ij} is a stock-venue fixed effect, γ_t is a daily fixed effect and D_j and $post_t$ are defined as per Equation (5). The key parameter of interest is β which again captures the causal effect of the treatment (fee introduction) on the trading activity and market quality for trading on the treated exchange relative to the control venues.

Our two econometric specifications in Equations (5) and (6) are similar insofar as the both use the introduction of market data fees on an exchange as treatments that affect trading on that exchange while trading on other venues act as the control group. The key difference between the specification relates to what kind of variation is controlled for by the fixed effects.

In Equation (5), each stock-day pair has its own intercept term that applies to observations across all venues for that stock and day. This specification is robust to any unobserved effects at the stock-day level (i.e. stock-day shocks that affect trading across all venues). Average differences across the treated and control groups are then captured by including the treatment dummy itself ($Treat_j$), which imposes the constraint that relative differences within stocks across venues is constant across stocks.

In Equation (6), our fixed effects allow for separate intercept terms for trading in each stock on each venue (ARCA, Direct Edge, BATS vs. all control venues). This is potentially important because average levels of trading or market quality display large differences across venues within a stock (e.g. consider the log volume of AAPL traded on ARCA vs. all other venues) and across stocks within a venue. Our time effects capture market-wide differ-

ences that affect all stock-venue pairs equally. This specification corresponds to a standard difference-in-differences regression with unit and time fixed effects, where each stock-day-venue combination is a separate observation. Note that the $_j$ dummy is not included in Equation (6) since it is co-linear with the stock-venue fixed effects.

In both models, we pool our data across all three events and estimate a single treatment effect corresponding to their average effect of fee introduction on trading activity and market quality. Pooling across the three events helps to limit the reliance on parallel trends insofar as similar confounding trends would need to be present across all three events to contaminate our results in a systematic way. The evidence presented in Figure 1 suggests that trading volumes closely track each other over both the pre-event and post-event period, supporting the parallel trends identifying assumption. We estimate all standard errors by clustering at the stock level.

5.1 Effect of fees on market shares and trade size

Estimates from our differences-in-differences regressions for log dollar value traded are presented in Columns (1) and (2) of Table 3. These regressions test whether venues that introduce fees for order book information suffer a fall in market share relative to other venues. Our empirical approach is designed to causally identify the effect of fees on market share and execution quality. As we argue in Section 3 a change in market share reflects that at least some traders do not subscribe to the fees and prefer to route some of their orders to other exchanges, relative to the pre-fee environment. Our log volume regressions are therefore tests of the hypothesis that some traders are sensitive to the cost of order book information and exchanges face downward sloping demand curves for order book information. Column (1) refers the treatment effect estimated using stock-day fixed effects while Column (2) refers to

the model estimated using stock-venue and date fixed effects.

Table 3 about here

The point estimates for the effect of fee introduction on log volume in Table 3 are highly consistent across specifications — venues lose 5-7.5% of their existing volume on average after introducing fees, relative to venues that do not introduce fees.[5] The effect is highly statistically significant. In terms of dollar value, a 5-7.5% fall corresponds to around $400,000 to $600,000 less daily volume for the average stock on the treated venues in our sample, or $78,500 to $120,000 for the median stock. The coefficient estimates in Table 3 demonstrate that at least some market participants are sensitive to the introduction of fees and prefer to route their orders to other venues after fees are introduced.

Columns (3) and (4) perform the same tests but where the dependent variable is the log of average trade size. These regressions tell us whether or not the trades that take place on the treated exchange tend to be larger or smaller after order book fees are introduced. The point estimates for both treatment effects are economically small and suggest that trade sizes fell by about 0.3% after fees are introduced. The effect is only statistically significant under the stock-day fixed effect specification. Together these provide weak evidence that trade sizes change.

5.2 Effect of fees on intermarket sweep orders and time at NBBO

In columns (5), (6), and (7) of Table 3 we estimate potential reasons for the causal effect of the treatment (fee introduction) on trading activity. Columns (5) and (6) estimate the differences-in-differences regressions explaining log dollar value similar as for columns (1)

[5]Since the dependent variable is in logs, the treatment effect parameter captures the difference-in-differences in logs and so captures percentage changes in volumes.

and (2) but we now also distinguish between ISO and non-ISO trades using an indicator variable. In other words, instead of explaining the dollar value of the treated versus the control exchange, given by y_{ijt} in Eqs. (5) and (6), we explain $y_{ijt,ISO}$ with ISO equal to zero or one if $y_{ijt,ISO}$ refers to, respectively, the non-ISO or ISO dollar value of stock i in exchange j on day t.

The interaction "$Treat \times Post$" indicates a drop in the dollar value of non-ISO trades of around 4%. The triple interaction term "$ISO \times Treat \times Post$" captures the incremental effect of the treatment on ISO trades. Both specifications indicate a further drop in the dollar value of ISO trades of around 5%, which is statistical significant with t-statistics below -6. In other words, regressions in columns (5) and (6) indicate that fees mainly affect the market-share of ISO trades.

Column (7) of Table 3 estimates the effect of the treatment on the time the treated exchange quotes prices at the NBBO. Because we only have one observation per stock-day we cannot include day fixed effects. Results indicate that the treatment causes a drop in the time an exchange is at NBBO of around 6% (with a t-stat of -11.21).

6 Which Market Participants Respond to Data Fees?

Given that venues face a downward sloping demand curve for order book information, we next ask whether some groups of traders are more sensitive to fees than others. As per Section 3, we focus on three categories of traders that are present in many canonical market microstructure models: informed traders, liquidity traders and market makers. Evidence in Table 3 suggests that fewer ISO trades execute on the treated exchange, which is consistent with informed traders not subscribing, but also that treated exchanges spend less time at the NBBO after fees. Informed traders may choose to not use the treated exchanges because

liquidity is on average worse after fees are introduced.

6.1 Evidence from transaction cost decompositions

To further distinguish between these different categories, we examine how the average informativeness of customer trades (price impact), dealer revenue from market making (realized spreads) and total trading costs (effective spreads) change around fee introduction. We now try to map how three variables (average trading costs, the informativeness of trade arrivals and market maker profits from liquidity provision) are likely to change if one group of traders has a higher propensity to not to subscribe.

First, for informed traders, whose trades contribute to the price discovery process, we argue that if they are less likely to subscribe to order book fees and reroute their orders (especially ISOs) elsewhere, the primary effect will be a fall in average price impact per trade for venues introducing fees. Market makers subsequently face less adverse selection when providing liquidity and earn higher revenue per trade from market making, all else being equal. Average effective spreads may decrease or remain unchanged depending on the degree of competition between market makers.

Analogous but opposite arguments apply to liquidity traders, whose trades by definition carry no informational content. If this category is less likely to subscribe to order book fees and therefore route their orders elsewhere, then average price impact per trade is expected to increase for venues introducing fees. Market makers are more likely to trade against informed traders and therefore face more adverse selection when providing liquidity than prior to fees. Average effective spreads must increase to compensate for the higher costs of providing liquidity, all else being equal. Revenue per trade either falls or is unchanged depending on the degree of competition between market makers.

Lastly, if market makers are relatively more sensitive to the cost of order book information, then then the degree of competition in the provision of liquidity is expected to fall as these market makers can no longer compete on a level playing field with those that do subscribe. Average market maker revenue per trade (i.e. for the market makers that do subscribe) and average trading costs increase after fees are introduced. Price impact may increase or decrease depending on whether informed or liquidity traders are more sensitive to trading costs. If informed traders are more (less) sensitive to trading costs, then price impact is expected to fall (rise).

We summarize these three mechanisms and their predictions for trading costs, market making profits and trade price impact in Table 4.

Table 4 about here

Figure 2 depicts the average effective spreads by venue type in the pre- and post-event periods. Before fees are introduced, there is no notable gap between the two series. After fees are introduced, we observe a small but persistent gap between the average trading cost for a trade on treated exchanges and the control exchanges. Like log volume in Figure 1, the two series track each other very closely over the pre- and post-event periods.

Figure 2 about here

Plots of the effective spread decomposition are contained in Figure 3. Panel A contains market maker realized spreads while Panel B contains price impact. Realized spreads are persistently lower for trades on the treated exchanges prior to fee introduction and this gap falls after fees are introduced. Price impacts are slightly higher for trades on the treated exchanges before the fee introduction and slightly lower after. For all four variables, the time-

series of stock-day averages across venues are highly correlated and appear to be following common trends.

Figure 3 about here

Parameter estimates for the effect of fees on average price impact, effective spreads and realized spread are presented in Table 5. Columns (1) and (2) present results for average price impact with stock-day fixed effects and stock-venue and day fixed effects respectively. Columns (3) and (4) present these results for effective spreads. Columns (5) and (6) present the results for realized spreads.

Table 5 about here

Due to the complex interactions between informed trading, trading costs and venue selection, no single variable in isolation can perfectly distinguish between one trader category over the other two. We choose to first examine average price impact because detecting a significant positive or negative effect of fees on this variable can most clearly rule out either the informed or liquidity trader categories. In both econometric specifications, the effect of fees on average price impact is negative and highly significant. The size of the parameter estimate is approximately -0.4 bps across both specifications, which corresponds to approximately 6% of the average trade price impact across all stock-day-venues in our sample. The size of the t-statistic is well in excess of the 1% threshold in all four specifications.

Columns (1) and (2) of Table 5 demonstrates that trades on treated exchange become less informed on average after the introduction of data fees, relative to trades on other venues with no change in data fees. This evidence rules out that liquidity traders routing decisions are more sensitive to order book fees than other trader categories. Liquidity traders either

are more willing to pay for order book information than market maker or informed traders, or are less reliant on order book information in deciding where to route their orders and so continue to use the treated exchange that introduce fees despite not observing the state of the order book.

Columns (1) and (2) of Table 5 cannot directly distinguish between market makers or informed traders being more sensitive to data fees. Arguably, the simplest explanation for decreased price impact is that informed traders route their orders to other venues after fees are introduced. In this case, market makers face less adverse selection costs when providing liquidity. Depending on the degree of competition in liquidity provision, market makers will pass this saving onto customers via lower effective spreads and / or earn higher revenue per trade due to lower their lower costs.

However, decreased price impact is also consistent with an alternative explanation based on the behavior of market makers. If some marker makers choose not to subscribe to order book information, they will suffer from information asymmetry with market makers who do subscribe. Standard Winner's Curse arguments can induce these market makers to compete less aggressively when providing liquidity on the venue introducing fees which leads to higher trading costs and dealer revenue per trade. Price impact could plausibly increase in response to higher trading costs if informed traders are more sensitive to these costs than liquidity traders.

Whether informed traders or market makers are relatively more sensitive to order book fees can be distinguished by the effect of data fees on effective spreads, which measure trading costs, and realized spreads, which measure average dealer revenue per trade. If informed traders are more likely to route their orders elsewhere after fees, then either effective spreads should decrease, realized spreads should increase or both. If market makers are choosing

not to subscribe, thus lowering the degree of competition between market makers, effective spreads and realized spreads should increase. If informed traders are also more sensitive to trading costs than uninformed traders, then price impact could fall, consistent with Columns (1) and (2) Table 5.

Coefficient estimates for effective spreads — Columns (3) and (4) show a very small reduction in average trading costs on the venues introducing fees for order book information relative to other venues. The economic magnitude of the effect is around 0.05 bps per trade, compared with a pre-fee average of around 11bps. The statistical significance of the effect is mixed.

For realized spreads, we detect a strongly significant positive effect of approximately 0.3 bps per trade. The size of the effect is highly consistent and the magnitude of the t-statistics exceed the 1% thresholds across specifications.

In Columns (3) - (4) of Table 5, the reduction in the costs of market making due to reduced average price impact almost entirely accrue to the market makers in the form of higher realized spreads, with only a very small fraction of the savings (if any) being passed on to customers. These patterns combined with the fall in total volume and average price impact are consistent with the trading decisions of informed traders being more sensitive to data fees than liquidity traders or market makers. Informed traders appear to be less likely to pay the fixed costs of subscribing to data fees and then choose to route their other venues after fees are introduced. Market makers capture higher profits on a per-trade basis, though their total revenue from market making on the venue introducing the fee is also affected by lower volumes.

Crucially, Table 5 is inconsistent with the key effect of fees being less market maker competition. In this scenario, a fall in price impact requires first for average trading costs

to rise and for this rise to disproportionately affect order routing decisions of informed traders. Table 5 indicates that effective spreads are either lower or unchanged after fees are introduced. Even if routing decisions of informed traders are more sensitive to trading costs, we find no evidence that these increased on the venue introducing fees relative to the pre-fee environment.

6.2 Evidence from Hasbrouck information shares

Previous results indicate a drop in informed trading at the treated exchange. A more direct test is to estimate the contribution of the treated exchange to the information share. Table 6 presents results for information share measured following Hasbrouck (1995) and Lien and Shrestha (2009). Similar as when using the time at NBBO, information share aggregates to one when measured over all exchanges and therefore does not allow us to estimate panel regressions with day fixed effects. When estimating panel regressions with stock fixed effects, both specifications indicate a drop in the information share of the treated exchange of 3% to 4% with t-stats of -6.12 and -5.93, respectively.

6.3 Evidence from sample splits

Our analysis so far uses all trades from every stock in the S&P1500 indices over our sample period. Since the unit of observation is at either the stock-day or stock-day-venue level, trades in small stocks (that tend to be fewer in number) get similar weight to the trades in large stocks. In this section, we split our analysis into two samples based on membership in the large-cap S&P500 index vs. the mid-cap and small-cap S&P400 and S&P600 indices.

The purpose of these regressions are two fold. First, we are interested to see if our results are capturing effects primarily in small or large stocks. If the effect is concentrated in small

stocks, it may complicate our ability to make inference about market quality more broadly, given that these stocks constitute a relatively small fraction of total trading. Second, this analysis can help us to understand for what kinds of stocks order book data is most relevant for, large stocks or small stocks.

The hypothesized outcome is unclear. While the fundamental economics are the same for all stocks, we think that the additional condition of size may result in further differential effects. There are at least three possible scenarios, all regarding the choice of some informed traders to decline subscribing to the data fees. Scenario 1 is that the subset of informed traders that stop receiving the data were active across the two sets of stocks and found the data to be equally valuable regardless of the stock size. Under this scenario we expect to find minimal differences between the two groups.

Scenario 2 is that the subset of informed traders that stop receiving the data were disproportionately active in small stocks, or that informed traders found the data to be disproportionately informative for small stocks. When these investors stop receiving the data they stop making informed trades more in small stocks than in large stocks. Scenario 2 therefore suggests that small stocks should be more impacted than large stocks.

Scenario 3 the the inverse of Scenario 2. Instead of the informed traders who stop receiving the data being mainly in small stocks, they are mainly in large stocks, or the data feed is more informative for large stocks. These traders stop making informed trades more in large stocks than in small stocks. This final scenario suggests that large stocks should be more impacted than small stocks.

Table 7 contains results for log volume, trade size, transaction cost decompositions, ISO volume and time at NBBO split into these two samples. Panel A contains results for S&P500 stocks (large stocks). Panel B contains results for non-S&P500 stocks (small stocks). All

other details are as per Tables 3 and 5.

Table 7 about here

The results are reported in Table 7. There are differences across the types of stock. In general, the results for the small stocks are larger than those for small stocks. For example, the volume decline reported in Column (1) is -3.49% for large stocks, but is -7.09% for small stocks. In Column (5) the ISO volume drops -6.9% for large stocks, but declines -10.8% for small stocks.

For the transaction costs decompositions, we also find a difference. The fall in price impact is larger for small stocks than large stocks, as expected if more informed traders are being impacted in small stocks. The effective spread effect is not consistently statistically significant, and switches signs between the large and small stock sample. This is still consistent with the interpretation of a decline in informed trading, as Table 4 makes clear we do not have a predicted direction for effective spread. Finally, dealers appear to capture a greater amount of revenue per trade for large stocks. If informed traders in small stocks were dis proportionally impacted, we would expect the opposite.

In Table 8 we repeat the S&P 500 v. non-S&P 500 analysis for the information share analysis.

Table 8 about here

While both large and small stocks experience a fall in information share on the venue with the data fee introduction, the decline for small stocks is less than the decline for large stocks (around 1% compared to around 8% for large stocks), which is inconsistent with the notion that small stocks are dis proportionally impacted. This result is consistent with results in

Table 7 that the decline in trading activity for small stocks is more attributable to less orders routed to the treated exchange because its prices are less often at the NBBO (8% versus 4.5% for large stocks), while the decline in trading activity for large stocks is more attributable to traders using less ISO (ISOs for small stocks drop by around twice non-ISOs, ISOs for large stocks drop by almost four times non-ISOs).

Overall the evidence suggests small stocks experience a larger decline in trading activity than do large stocks around the introduction of data fees and different reasons for the decline.

7 Conclusion

This paper explores the role of data fees on exchange competition. While exchange data has properties of a public good, the exchanges control it and can, with the approval of the Securities and Exchange Commission, set a price on it. We show that when a trading venue introduces a fee on its data, it experiences a decline in trading activity. This shows that the demand for such data is not perfectly inelastic and so exchanges experience at least some trade off between higher volume and higher data fees.

By examining different measures of liquidity we can infer the type of trader that is predominantly impacted by the data fee introduction. We find that the introduction of a data fee results in price impact falling and realized spread increasing, with negative but small changes to effective spreads. These changes are most consistent with informed traders decreasing their trading on the venue increasing a data fee. Whether this change materially impacts the amount of information acquired, or simply the distribution of where the informed traders choose to trade on the information, is an open question.

References

Alan, N. S. and R. A. Schwartz (2013). Price discovery: The economic function of a stock exchange. *Journal of Portfolio Management 40*(1), 124–132.

Brogaard, J., T. Hendershott, and R. Riordan (2019). Price discovery without trading: Evidence from limit orders. *The Journal of Finance 74*(4), 1621–1658.

CBOE (2012). Notice of Filing and Immediate Effectiveness of Proposed Rule Change Relating to Fees Associated with Receipt of the EDGX Book Feed. `https://cdn.cboe.com/resources/regulation/rule_filings/approved/2012/SR-EDGX-2012-14-Revised.pdf`. Accessed: 2020-09-22.

Chakravarty, S., P. Jain, J. Upson, and R. Wood (2012). Clean sweep: Informed trading through intermarket sweep orders. *Journal of Financial and Quantitative Analysis*, 415–435.

Gormley, T. A. and D. A. Matsa (2014). Common errors: How to (and not to) control for unobserved heterogeneity. *Review of Financial Studies 27*, 617–661.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *Journal of Finance 50*(4), 1175–1199.

Hendershott, T. and C. M. Jones (2005). Island goes dark: Transparency, fragmentation, and regulation. *The Review of Financial Studies 18*(3), 743–793.

Hendershott, T., M. Rysman, and R. Schwabe (2020). Stock exchanges as platforms for data and trading. *Working Paper*.

Holden, C. W. and S. Jacobsen (2014). Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. *The Journal of Finance 69*(4), 1747–1785.

Jones, C. M. (2018). Understanding the market for u.s. equity market data. *Working Paper*.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, 1315–1335.

Lee, C. M. and M. J. Ready (1991). Inferring trade direction from intraday data. *The Journal of Finance 46*(2), 733–746.

Lewbel, A. (2019). The identification zoo: Meanings of identification in econometrics. *Journal of Economic Literature 57*, 835–903.

Lien, D. and K. Shrestha (2009). A new information share measure. *Journal of Futures Markets 29*(4), 377–395.

MFA and AIMA (2018). Petition for rulemaking regarding market data fees and request for guidance on market data licensing practices.

Pagnotta, E. S. and T. Philippon (2018). Competing on speed. *Econometrica 86*(3), 1067–1115.

Ramos, S. B. and E.-L. von Thadden (2008). Stock exchange competition in a simple model of capital market equilibrium. *Journal of Financial Markets 11*(3), 284–307.

Reuters (2013). BATS exchanges to start charging for market data. `https://www.reuters.com/article/batsglobalmarkets-marketdata-fees-idUSL2N0D52AC20130418`. Accessed: 2020-09-22.

SEC (2008). Self-Regulatory Organizations; NYSE Arca, Inc.; Order Setting Aside Action by Delegated Authority and Approving Proposed Rule Change Relating to NYSE Arca Data. `https://www.sec.gov/rules/sro/nysearca/2008/34-59039.pdf`. Accessed: 2020-09-22.

SEC (2016). INITIAL DECISION RELEASE NO. 1015 ADMINISTRATIVE PROCEEDING FILE NO. 3-15350. `https://www.sec.gov/alj/aljdec/2016/id1015bpm.pdf`. Accessed: 2020-11-11.

Spatt, C. S. (2020). Is equity market exchange structure anti-competitive? *working paper, Tepper School of Business*.

Figure 1: Market Shares Around Fee Introduction

This figure plots the average log dollar volume traded across the stocks in our sample in three month windows around fee introduction, split by venues introducing fees and other venues. The y-axis reports the average log dollar value traded across all S&P1500 stocks by day, pooled across the three fee introduction events. The x-axis reports the number of trading days before and after the introduction of fee. Log dollar volume by day traded on the three exchanges introducing a fee is depicted in the "treated" line. Log dollar volume by day traded on all other venues not introducing fees is depicted in the "control" line.

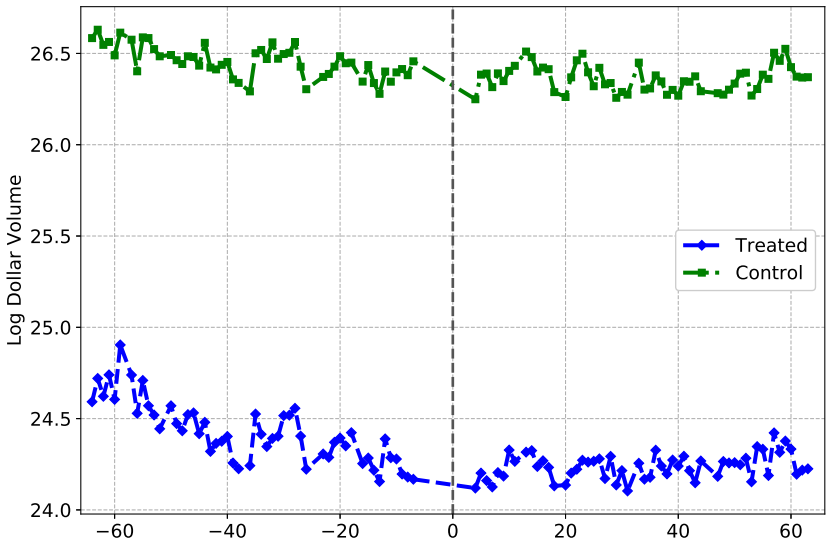


Figure 2: Effective Spreads Around Fee Introduction

This figure plots the average effective spreads across the stocks in our sample in three month windows around fee introduction, split by venues introducing fees and other venues. The y-axis reports the average effective spread for trades in all S&P1500 stocks by day, pooled across the three fee introduction events. The x-axis reports the number of trading days before and after the introduction of fee. Effective spreads on the three exchanges introducing a fee is depicted in the "treated" line. Effective spreads on all other venues not introducing fees is depicted in the "control" line.

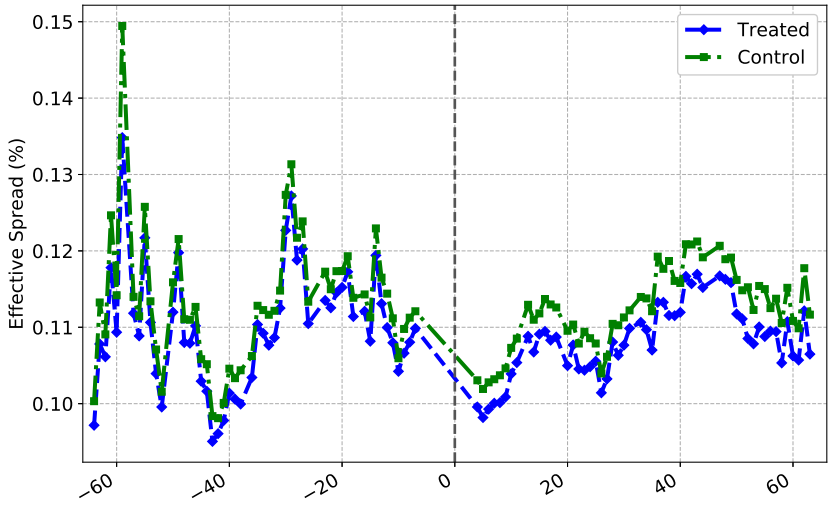
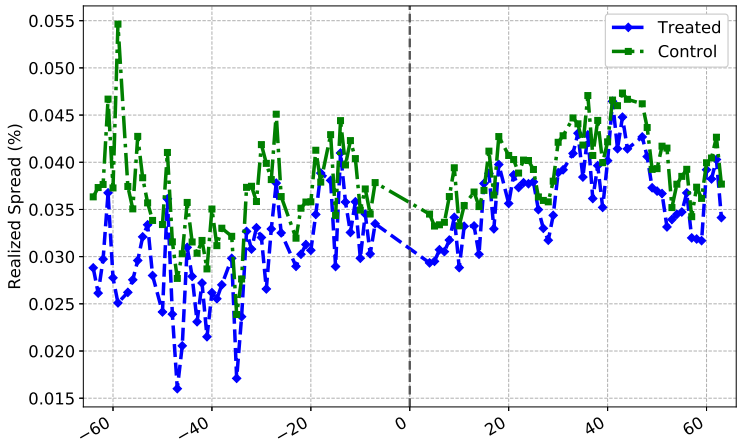


Figure 3: Spread Decomposition Around Fee Introduction

This figure plots the average realized spreads (Panel a) and price impact (Panel b) across the stocks in our sample in three month windows around fee introduction, split by venues introducing fees and other venues. The y-axis reports the average realized spread and price impact for trades in all S&P1500 stocks by day, pooled across the three fee introduction events. The x-axis reports the number of trading days before and after the introduction of fee. Realized spreads and price impact on the three exchanges introducing a fee are depicted in the "treated" line. Realized spreads and price impact on all other venues not introducing fees are depicted in the "control" line.

(a) Realized Spreads



(b) Price Impact

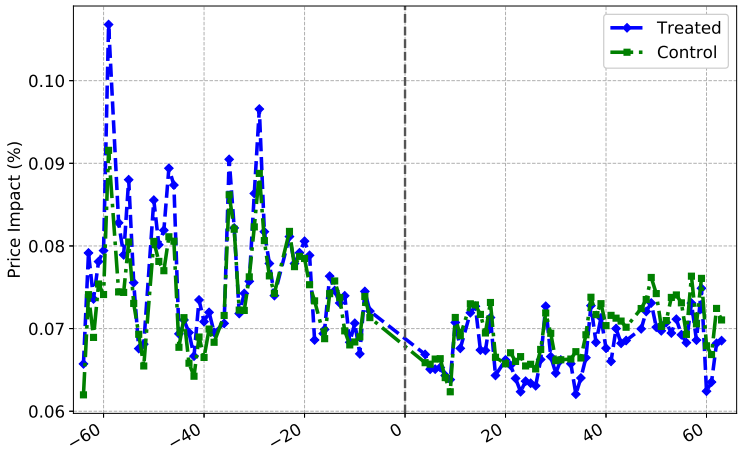


Table 1: Stock-Day Trading Summary Statistics

This table reports means, standard deviations, minimums, maximums and 25th, 50th and 75th quantiles for trading activity and liquidity variables for our sample, split into trading on treated exchange and all other venues. The sample includes all stock-venue-day combinations for S&P 1500 members over three month windows around the introduction of fees on ARCA, BATS and Direct Edge, where venues are divided into the treated exchanges introducing the fee and all other trading venues. Dollar volume traded is the total volume traded by stock-day-venue measured in millions of dollars. The quoted spread is the time-weighted average of the difference between the best bid and ask, divided by the midquote (expressed in %). Dollar depth is the time-weighted average of the dollar volume available at the best bid and ask measured in thousands of dollars. Effective spread, realized spread and price impact are defined as per Section 4. All variables are winsorized at the 1% level.

	Mean	SD	Min	25%	50%	75%	Max	Mean Pre-Fee	Mean Post-Fee
Dollar Value Traded (mil)	33.598	86.251	0.0157	1.1380	5.4985	24.177	735.21	32.283	34.941
Log Dollar Value Traded	15.435	2.1789	9.6608	13.944	15.520	17.000	20.415	15.367	15.505
Quoted Spread (%)	0.3413	0.4415	0.0203	0.0874	0.1877	0.4069	2.9518	0.3595	0.3226
Dollar Depth (000)	17.853	31.631	0.6738	5.1847	9.1645	16.917	281.42	17.877	17.828
Log Dollar Depth	9.1857	1.0051	6.5129	8.5535	9.1231	9.7361	12.547	9.1886	9.1828
Effective Spread (%)	0.1061	0.1286	0.0133	0.0380	0.0673	0.1227	1.0433	0.1064	0.1058
Realized Spread (%)	0.0348	0.0823	-0.2349	0.0051	0.0204	0.0480	0.6283	0.0370	0.0325
Price Impact (%)	0.0681	0.1040	-0.1048	0.0148	0.0371	0.0834	0.7552	0.0660	0.0701
Trade size ($)	5,234.9	4,640.5	305.55	2,341.7	4,060.6	6,662.6	36818	5,116.1	5,356.2
Log Trade Size	8.2549	0.8114	5.7221	7.7587	8.3091	8.8043	10.513	8.2187	8.2918

Table 2: Stock-Day Trading Summary Statistics — Split by Treated Exchange and Control

This table reports means, standard deviations, minimums, maximums and 25th, 50th and 75th quantiles for trading activity and liquidity variables for our sample. The sample includes all stock-venue-day combinations for S&P 1500 members over three month windows around the introduction of fees on ARCA, Direct Edge and BATS, where venues are divided into those introducing the fee (treated) and all other trading venues (control). Dollar volume traded is the total volume traded by stock-day-venue measured in millions of dollars, reported separately for all trades and for only intermarket sweep orders (ISOs). The quoted spread is the time-weighted average of the difference between the best bid and ask, divided by the midquote (expressed in %). Dollar depth is the time-weighted average of the dollar volume available at the best bid and ask measured in thousands of dollars. Effective spread, realized spread, price impact, seconds at NBBO, information shares (Hasbrouck, 1995), and modified information shares (Lien and Shrestha, 2009) are defined as per Section 4. All variables are winsorized at the 1% level.

	Mean	SD	Min	25%	50%	75%	Max	Mean Pre-Fee	Mean Post-Fee
Panel A: Trading on Treated Exchanges									
Dollar Value Traded (mil)	7.9487	27.302	0.0157	0.4030	1.5718	6.0632	735.21	8.5862	7.3247
Log Dollar Value Traded	14.239	1.9045	9.6608	12.906	14.267	15.617	20.415	14.323	14.157
Quoted Spread (%)	0.4826	0.5374	0.0203	0.1314	0.3046	0.6070	2.9518	0.4462	0.5182
Dollar Depth ('000s)	11.603	19.260	0.6738	4.2116	7.1874	12.186	281.42	11.450	11.754
Log Dollar Depth	8.8947	0.8893	6.5129	8.3456	8.8801	9.4081	12.547	8.8832	8.9061
Effective Spread (%)	0.1042	0.1291	0.0133	0.0362	0.0643	0.1199	1.0433	0.1040	0.1044
Realized Spread (%)	0.0326	0.0912	-0.2349	-0.0017	0.0191	0.0515	0.6283	0.0296	0.0355
Price Impact (%)	0.0679	0.1147	-0.1048	0.0088	0.0339	0.0868	0.7552	0.0709	0.0649
Trade size ($)	4,494.4	4,058.8	305.55	2,036.9	3,522.9	5,705.5	36818	4,613.6	4,377.8
Log Trade Size	8.1071	0.8040	5.7221	7.6192	8.1670	8.6492	10.513	8.1446	8.0704
Seconds at NBBO	6,605	6,616	1.000	1,231	4,068	10,216	23,400	6,515	6,693
Information Share	0.267	0.186	0.000	0.099	0.241	0.429	0.9997	0.274	0.259
Modified Information Share	0.259	0.198	0.000	0.087	0.220	0.408	0.9999	0.265	0.254
Panel B: Trading on Control Venues									
Dollar Value Traded (mil)	59.482	113.40	0.0157	5.0929	17.533	60.595	735.21	61.532	57.475
Log Dollar Value Traded	16.642	1.7252	9.6608	15.443	16.679	17.919	20.415	16.698	16.588
Quoted Spread (%)	0.1986	0.2443	0.0203	0.0660	0.1283	0.2425	2.9518	0.1980	0.1993
Dollar Depth ('000s)	24.160	39.454	0.6738	6.6877	11.956	23.491	281.42	24.264	24.057
Log Dollar Depth	9.4793	1.0296	6.5129	8.8080	9.3890	10.064	12.547	9.4850	9.4737
Effective Spread (%)	0.1080	0.1280	0.0133	0.0400	0.0703	0.1253	1.0433	0.1076	0.1085
Realized Spread (%)	0.0370	0.0721	-0.2349	0.0090	0.0212	0.0450	0.6283	0.0355	0.0385
Price Impact (%)	0.0683	0.0920	-0.1048	0.0190	0.0395	0.0807	0.7552	0.0694	0.0671
Trade size ($)	5,982.1	5,052.8	305.55	2,724.4	4,697.7	7,662.1	36818	6,105.4	5,861.3
Log Trade Size	8.4040	0.7912	5.7221	7.9100	8.4548	8.9441	10.513	8.4404	8.3684
Seconds at NBBO	16,812	6,615	1.000	13,210	19,352	22,182	23,400	16,898	16,727
Information Share	0.733	0.186	0.000	0.571	0.760	0.901	1.000	0.726	0.741
Modified Information Share	0.741	0.198	0.000	0.592	0.780	0.913	1.000	0.735	0.746

Table 3: Difference-in-Differences Regressions: Market shares and trade size

This table reports coefficients (t-statistics) for the effect of fee introduction on log dollar volume traded, log of trade size, volume executed via intermarket sweep orders (ISOs) and otherwise, and seconds spent at NBBO. Columns (1) through (6) present difference-in-differences regressions of log dollar volume and log dollar trade size around fee introduction. The treated sample includes daily trading in all S&P1500 stocks on the venue introducing order book fees in three month windows around fee introduction. The control sample includes the daily trading in the same stocks on all other venues not introducing fees. In columns (1) through (4) each stock-day-venue combination contributes one observation to the regression. In columns (5) and (6) we further distinguish volume executed via ISOs and non-ISOs. In column (7) we only have one observation per stock-day, the number of seconds the treated exchange is at the NBBO. Observations across each of the three fee introduction events are pooled into a single regression. Standard errors are clustered at the stock level.

	Log Volume		Log Trade Size		Log Volume		Log Seconds at NBBO
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treat × Post	-0.0586***	-0.0730***	-0.0034***	-0.0038	-0.040***	-0.038***	
	(-19.6)	(-10.0)	(-3.22)	(-1.08)	(-12.20)	(-5.21)	
ISO × Treat × Post					-0.051***	-0.050***	
					(-9.53)	(-6.88)	
Post							-0.068***
							(-11.21)
Treat	-2.3646***		-0.2944***		-2.516***		
	(-311)		(-153)		(-384)		
ISO					-0.390***	-0.459***	
					(-52.91)	(-67.57)	
ISO × Post					0.011***	0.014***	
					(2.76)	(2.91)	
ISO × Treat					0.378***	0.499***	
					(38.00)	(52.66)	
N_{obs}	1,030,173	1,030,173	1,030,173	1,030,173	1,776,261	1,776,447	468,133
R^2	0.97	0.93	0.98	0.91	0.95	0.89	0.75
Stock × Treat FE	-	X	-	X	-	X	X
Day FE	-	X	-	X	-	X	-
Stock × Day FE	X	-	X	-	X	-	-

Table 4: Predicted transaction costs decompositions vs. trader categories

This table contains predictions for changes to price impact, effective spreads and realized spreads when different trader categories are assumed to be relatively more sensitive to the introduction of order book fees than the other categories. The three trader categories are market makers, informed traders and liquidity traders.

Category	Price Impact	Effective Spreads	Realized Spreads
Market Makers	$\uparrow / \downarrow$	$\uparrow$	$\uparrow$
Informed Traders	$\downarrow$	$\uparrow / \downarrow$	$\uparrow$
Liquidity Traders	$\uparrow$	$\uparrow$	$\uparrow / \downarrow$

Table 5: Difference-in-Differences Regressions: Transaction Costs Decomposition

This table reports coefficients (t-statistics) for difference-in-differences regressions of price impact, effective spreads and realized spreads around fee introduction. The treated sample includes trades in all S&P1500 stocks on the venue introducing order book fees in three month windows around fee introduction. The control sample includes the trades in the same stocks on all other venues not introducing fees. Each stock-day-venue combination contributes one observation to the regression. Observations across each of the three fee introduction events are pooled into a single regression. Standard errors are clustered at the stock level.

	Price Impact		Effective Spread		Realized Spread	
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	0.0016***		-0.0036***		-0.0061***	
	(5.75)		(-24.4)		(-19.7)	
Treat × Post	-0.0038***	-0.0039***	-0.0005***	-0.0006	0.0030***	0.0030***
	(-13.3)	(-6.25)	(-4.27)	(-0.68)	(10.8)	(4.97)
N_{stocks}	1,774	1,774	1,774	1,774	1,774	1,774
T	368	368	368	368	368	368
N_{obs}	1,030,173	1,030,173	1,030,173	1,030,173	1,030,173	1,030,173
R^2	0.85	0.53	0.99	0.81	0.76	0.29
Stock × Treat FE	-	X	-	X	-	X
Day FE	-	X	-	X	-	X
Stock × Day FE	X	-	X	-	X	-

Table 6: Information shares

This table reports changes in Hasbrouck (1995) and Modified (Lien and Shrestha, 2009) Information Shares after fee introduction. The sample includes all S&P1500 stocks on the venue introducing order book fees in three month windows around fee introduction. *Post* is an indicator variable equal to one after the fee introduction and zero before. Each stock-day contributes one observation to the regression. Observations across each of the three fee introduction events are pooled into a single regression. Standard errors are clustered at the stock level.

	Information Shares	Modified Information Shares
	(1)	(2)
Post	-0.031***	-0.036***
	(-6.12)	(-5.93)
N_{obs}	464,690	464,690
R^2	0.55	0.44
Stock FE	X	X

Table 7: Difference-in-Differences Regressions — split by S&P500 membership

This table reports coefficients (t-statistics) for the effect of fee introduction split by S&P500 membership. All details are as per Tables 3 and 5 other than that we split our sample into two based on membership of the S&P500 index. Results for stocks that are members of the S&P500 index are presented in Panel A. Results for stocks in the S&P400 and S&P600 indices are presented in Panel B.

	Log Volume		Log Trade Size		Log Volume		Log Seconds at NBBO	Price Impact		Effective Spread		Realized Spread	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel A: S&P 500 Stocks													
Treat × Post	-0.0349***	-0.0491***	0.0007	0.0001	-0.018***	-0.017*		-0.0034***	-0.0034***	0.0008***	0.0007	0.0038***	0.0038***
	(-7.37)	(-5.13)	(0.40)	(0.03)	(-3.53)	(-1.78)		(-16.2)	(-5.26)	(10.8)	(0.92)	(17.8)	(12.3)
ISO × Treat × Post					-0.051***	-0.051***							
					(-6.78)	(-6.16)							
Post							-0.046***						
							(-5.35)						
Treat	-2.2825***		-0.3115***		-2.461***			-0.0001		-0.0018***		-0.0029***	
	(-183)		(-94.9)		(-251)			(-0.30)		(-15.2)		(-9.61)	
ISO					-0.320***	-0.365***							
					(-34.37)	(-41.08)							
ISO × Post					0.001	-0.000							
					(0.13)	(-0.01)							
ISO × Treat					0.416***	0.486***							
					(29.30)	(36.36)							
Panel B: Non-S&P 500 Stocks													
Treat × Post	-0.0709***	-0.0854***	-0.0050***	-0.0055	-0.052***	-0.049***		-0.0043***	-0.0045***	-0.0012***	-0.0014	0.0028***	0.0028***
	(-18.8)	(-8.91)	(-3.76)	(-1.21)	(-12.11)	(-5.09)		(-10.0)	(-5.06)	(-6.66)	(-0.99)	(6.91)	(2.85)
ISO × Treat × Post					-0.056***	-0.055***							
					(-7.49)	(-5.13)							
Post							-0.080***						
							(-9.86)						
Treat	-2.4094***		-0.2854***		-2.544***			0.0026***		-0.0045***		-0.0078***	
	(-260)		(-122)		(-301.95)			(6.25)		(-20.9)		(-17.5)	
ISO					-0.436***	-0.523***							
					(-44.06)	(-60.80)							
ISO × Post					0.021***	0.017**							
					(3.78)	(2.38)							
ISO × Treat					0.338***	0.510***							
					(26.96)	(43.77)							
Stock × Treat FE	-	X	-	X	-	X	X	-	X	-	X	-	X
Day FE	-	X	-	X	-	X	-	-	X	-	X	-	X
Stock × Day FE	X	-	X	-	X	-	-	X	-	X	-	X	-

Table 8: Information shares — split by S&P 500 membership

This table reports changes in Hasbrouck (1995) and Modified (Lien and Shrestha, 2009) Information Shares after fee introduction split by S&P500 membership. All details are as per Tables 6 other than that we split our sample into two based on membership of the S&P500 index. Results for stocks that are members of the S&P500 index are presented in Panel A. Results for stocks in the S&P400 and S&P600 indices are presented in Panel B.

	Information Share	Modified Information Shares
	(1)	(2)
Panel A: S&P 500 Stocks		
Post	-0.076***	-0.075***
	(-8.17)	(-6.94)
Stock FE	X	X
Panel B: Non-S&P 500 Stocks		
Post	-0.010*	-0.017**
	(-1.74)	(-2.42)
Stock FE	X	X