Securities Market Design: Exchange Access Fee, Tick Size, and Dynamic Price Limit

Author:
Lin, Yiping

Publication Date:
2018

DOI:
https://doi.org/10.26190/unsworks/3353

Securities Market Design: Exchange Access Fee, Tick Size, and Dynamic Price Limit

Yiping LIN

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

School of Banking and Finance

Australian School of Business

University of New South Wales

January, 2018

THE UNIVERSITY OF NEW SOUTH WALES
Thesis/Dissertation Sheet

Surname or Family name: **Lin**

First name: **Yiping** Other name/s:

Abbreviation for degree as given in the University calendar: **PhD**

School: **School of Banking and Finance** Faculty: **UNSW Business School**

Title: **Securities Market Design: Exchange Access Fee, Tick Size, and Dynamic Price Limit**

Abstract

The impact of market structure designs on market quality is of interest to academics, practitioners and regulators. Using three exogenous events and with the benefit of proprietary data, this thesis builds on theoretical models and examines the impact of three important market structure designs - exchange access fee, tick size, and dynamic price limit - on market quality.

First, I provide a theoretical model of exchange fee structures to show that in a competitive environment, traders optimise the degree of information in their trades to fully exploit fees/rebates. I examine the 'natural' experiment of a unilateral Nasdaq exchange fee reduction, and find evidence consistent with the theoretical model which indicates that the expected 'washing-out' of the fee changes is offset by the flight of highly-informed market orders to the remaining highest rebate venues. Far from fees washing-out, there is a redistribution of informed traders across venues.

Second, I extend the theoretical model to examine the impact of the 2016 U.S. tick size pilot. I show that the information content of trades increases more in markets that subsidise liquidity providers. Moreover, markets subsidising liquidity consumers and off-exchange trades are the major beneficiary of the sizeable rise in the tick size that acts as an additional transaction tax paid especially by liquidity traders and a corresponding subsidy to limit orders. This sizeable increase in transaction costs means that those uninformed traders that remain in the lit market during the pilot are far more price sensitive, encouraging them to flee venues subsidising liquidity providers in favour of cheaper venues subsidising liquidity consumers.

Third, I analyse the impact of the dynamic price limit rule - Limit Up Limit Down (LULD) and High Frequency Trading (HFT) behaviour around price limit on markets with different fee structures. I find LULD interferes with trading activity but curbs short term volatility without delaying the price discovery process. The magnet effect exists and the impact is stronger when approaching the upper price limit. Also, HFT trading activity decreases on market subsidising liquidity providers after the LULD halt which is driven by liquidity taking.

ORIGINALITY STATEMENT

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed

Date

COPYRIGHT STATEMENT

'I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.
I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only).
I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.'

Signed

Date

AUTHENTICITY STATEMENT

'I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.'

Signed

Date

To my family

Acknowledgements

My PhD study has been a fantastic, life changing experience and I am very grateful to the extensive network of individuals who have supported me throughout the journey. Their guidance, advice, support and encouragement are greatly acknowledged.

First, I am indebted to my chair academic supervisor Professor Peter L. Swan for his consistently strong support and guidance of my PhD research. It has been a great pleasure to work and discuss with him and learn from his extensive expertise. I would like to express my deepest gratitude for his enthusiasm, encouragement and numerous discussions with me on my research. I also want to thank Dr Frank Hatheway for his constructive advices, continuously strong support, and data provision during my PhD journey. The other special thanks must also go to Professor Jerry Parwada for his generous support and guidance.

Second, I would like to express my great gratitude to the Capital Markets CRC (CMCRC) and my industry sponsor, Nasdaq, Inc. for providing me with opportunities and scholarship. I would especially like to thank Professor Mike Aitken and Dr Shan Ji from CMCRC for their ongoing strong support and encouragement throughout every stage of my PhD life. My sincere thanks also must go to Dr Jeffrey Smith and Dr Claude Courbois from Nasdaq for sharing with me their vast knowledge of the U.S. financial markets and helpful comments throughout this research.

Third, I am also obliged to many academics for their mentoring and assistance with various chapters in this dissertation. In particular, a hearty thank-you must go to Professors Rick Harris, Ronald Masulis, and Dr Vito Mollica for their helpful feedback in improving my research. Thanks also goes to Professors Dave Michayluk and George Wang for their encouragement and support from the early stages of my PhD.

Finally, I am truly indebted and thankful to my parents, in-laws and friends. I am especially grateful to Edell for her wholehearted support and for making this dissertation possible. Thank you for being my best friend, my partner in love, my constant encouragement, and my inspiration.

Abstract

The impact of market structure designs on market quality is of interest to academics, practitioners and regulators. Using three exogenous events and with the benefit of proprietary data, this thesis builds on theoretical models and examines the casual impact of three important market structure designs - exchange access fee, tick size, and dynamic price limit - on market quality.

First, I provide a theoretical model of the exchange fee structures to show that in a competitive environment, traders optimize the degree of information in their trades to fully exploit fees/rebates. I examine the 'natural' experiment of a unilateral Nasdaq exchange fee reduction, and find evidence consistent with the theoretical model which indicates that the expected 'washing-out' of the fee changes is offset by the flight of highly-informed market orders to the remaining highest rebate venues. Far from fees washing-out, there is a redistribution of informed traders across venues.

Second, I extend the theoretical model to examine the impact of the 2016 U.S. tick size pilot. I show that the information content of trades increases more in markets that subsidise liquidity providers. Moreover, markets subsidising liquidity consumers and off-exchange trades are the major beneficiaries of the sizeable rise in the tick size that acts as an additional transaction tax paid especially by liquidity traders and a corresponding subsidy to limit orders. This sizeable increase in transaction costs means that those uninformed traders that remain in the lit market during the pilot are far more price sensitive, encouraging them to flee venues subsidising liquidity providers in favour of cheaper venues subsidising liquidity consumers.

Third, I analyse the impact of the intraday dynamic price limit rule – Limit Up Limit Down (LULD) and High Frequency Trading (HFT) behaviour around price limits on markets with different fee structures. Using difference-in-differences and propensity score matching, I find LULD interferes with trading activity but curbs short-term volatility without delaying the price discovery process. The magnet effect exists and the impact is stronger when approaching the upper price limit. Also, HFT trading activity decreases on market subsidising liquidity providers after the LULD halt which is driven by liquidity taking.

Table of Contents

List of Tables

List of Figures

Abbreviations

ATS-Alternative Trading System

CBOE-Chicago Board Options Exchange

CMCRC-Capital Markets Cooperative Research Centre

CPS-Cents Per 100 Shares Traded

DDD-Difference-in-Difference-in-Differences

DID-Difference-in-Differences

ECN-Electronic Communication Network

ETF-Exchange Traded Fund

FINRA-Financial Industry Regulatory Authority

HFT-High Frequency Trading

IPO-Initial Public Offering

LOB-Limit Order Book

LULD-Limit Up Limit Down

MQD-Market Quality Dashboard

Nasdaq-National Association of Securities Dealers Automated Quotations

NBB-National Best Bid

NBBO-National Best Bid and Offer

NBO-National Best Offer

NMS-National Market System

NYSE-New York Stock Exchange

SEC-Securities and Exchange Commission

SIP-Securities Information Processor

TRTH-Thomson Reuters Tick History

Chapter 1 Introduction

1.1 Motivation

The impact of the following three important market structure designs on market quality is of interest to academics, industries and regulators: exchange access fee, tick size, and price limit (Angel, Harris and Spatt, 2013). To analyse those impacts on market quality, exogenous events of the Nasdaq access fee pilot, the Securities and Exchange Commission (SEC) mandated nickel tick size pilot and the Limit Up Limit Down (LULD) rule, and proprietary[1] and public data are used.

A. Exchange Access Fee

Three are three major exchange access fee structures operated in global equity markets: traditional *taker-taker*, which charge both liquidity suppliers and consumers when there is a trade; *maker-taker*, which reward liquidity suppliers and charge liquidity consumers; and *taker-maker ('inverted'),* which reward liquidity consumers and charge liquidity suppliers. The maker-taker and the inverted fee structure originates in U.S. financial markets. One of the most important rules in U.S. equity market is the Regulation National Market System (Reg NMS) which the three main provisions are Rule 610 ("Access Rule"), Rule 611 ("Order Protection Rule"), and Rule 612 ("Sub-Penny Rule"). An access fee is a fee paid to an exchange or electronic communication network (ECN) by the trader or broker who uses a

[1]I thank Nasdaq, Inc. and CMCRC MQD for providing data and the views herein are not intended to represent the views of Nasdaq, Inc. or CMCRC, its employees, or directors. Any errors or omissions are the responsibility of the author alone.

marketable order to initiate a transaction with a resting order. The access rule prohibits market centres from charging more than 0.3 cents per share in access fees. The order protection rule requires participants in the NMS, which consists of exchanges, dealers, and alternative trading systems, to honour the publicly displayed prices posted at other centres by either matching those prices or routing orders to the other centres, as long as they are electronically accessible. The sub-penny rule prohibits disseminating sub-penny quotes for stock price above $1 (SEC, 2005).

Outstanding question whether the composition of maker-taker fee matters while holding the net fee constant remains. Angel, Harris, and Spatt (2011) propose that fee-rebate schemes wash-out, specifically, makers tend to offer better prices, which on average reduces the bid-ask spreads, to earn the liquidity rebate. In a competitive trading environment, the access fee offsets the narrower average quoted spreads so that takers are no better or worse off on average. Likewise, makers are also no better or worse off on average as the liquidity rebates offsets the narrower quoted spreads. The true economic spread (cum fee bid-ask spread) should be adjusted to the access fee, which is the quoted spread plus twice the access fee for simultaneously buying and selling using marketable orders. Thus, only the difference between rebates and fees (net fee) matters as traders simply adjust their quoted prices so that the net prices paid or received on average are the same. Then maker-taker pricing model appears to only reduce quoted bid-ask spreads and thereby obfuscating true economic spread which makes it more challenging for traders to recognize their true trading costs.

Colliard and Foucault (2012) and Malinova and Park (2015) underscore theoretically and empirically that only net fee change matters because the maker-taker fee breakdown is irrelevant. Colliard and Foucault (2012) predicts that a decrease in the net trading fee should lower cum fee bid-ask spreads, however, its impact on raw quoted spreads (i.e., not adjusted for taker fee) depends on whether this decrease is caused by reducing the maker rebate or taker fee. The raw quoted spread falls in the former case and increase in the latter. Testing these predictions is difficult because maker-taker fee breakdown often changes simultaneously in reality with the net exchange fee which makes identification of the effects of the components change and the net fee change challenging. Malinova and Park (2015) solved the data problem and test the predication in a monopoly trading environment. They find that posted quotes adjust after the change in maker-taker fee composition, but the true transaction costs for liquidity demanders remain unaffected. Also, traders use aggressive orders more frequently when bid-ask spreads decrease, and adverse selection costs decline.

In contrast, Foucault, Kadan, and Kandel (2013) claim that trading volume may increase or decrease depending on the model parameters even without change in the net fee, as a fixed tick size prevents prices from neutralizing the effect of the maker rebate. Also, maker-taker venues benefit from a rise in the tick size from 1 to 12 cents because a finite minimum tick size prevents prices from fully neutralizing the effect of the make rebate. In addition, Chao, Yao, and Ye (2017a) show that traders can choose prices that perfectly neutralize any fee division in the absence of a minimum tick. However, in the presence of a minimum tick size constraint, fee structures could be used to constrain or segment the market if liquidity maker preferences are sufficiently heterogeneous which prevents perfect neutralization.

In addition, over the past decade, exchange-listed securities have traded more volume in off-exchange venues. There is a public assertion that the shift in trading away from lit markets is caused by high exchange access fees? To test this, Nasdaq initiated the unilateral exchange access fee pilot, which lowered both maker rebates and taker fees by identical amounts for selected stocks while holding the net fee unchanged. This also provides ideal environment and data to examine the remaining question about the impact of the component of maker-taker fee change.

Moreover, policy questions are foremost in this maker-taker fee debate. The SEC has questioned whether "rebates [are] unfair to long-term investors because they necessarily will be paid primarily to [high-frequency] proprietary firms engaging in passive market making strategies. Or do they generally benefit long-term investors by promoting narrower spreads and more immediately accessible liquidity?" (SEC, 2010). In addition, the SEC (2010) summarizes that one high frequency trading (HFT) strategy is passive market making to earn the liquidity rebate. However, there seems to be no existing literature providing direct evidence.

B. Tick Size

When there is an exogenous shock in the minimum tick size, do exchange access fees matter? The conventional view promulgated by Angel, Harris, and Spatt (2011, 2015) and Colliard and Foucault (2012) is that these tax-subsidy arrangements simply wash-out in competitive markets, without affecting trade prices or quotes. Taking it further, Foucault, Kadan, and Kandel (2013) present a model of trader monitoring in which tax-subsidy arrangements wash-out in a zero-minimum tick regime but become more efficacious for maker-taker venues and traders the higher the tick size. Likewise, Chao, Yao, and Ye (2017a) prove that fee structures wash-out with a zero-minimum tick or continuous pricing in the absence of informed traders, to show that a monopoly exchange can survive competition between exchange operators with different fee structures venues provided price changes are discrete. Using the 2015 Nasdaq fee experiment, this thesis shows a maker-taker fee and rebate set at the (capped) maximum level is most beneficial for market efficiency and the market share of a venue.

Using SEC mandated nickel ticket size pilot which began on October, 2016 as an exogenous shock, this thesis models and empirically tests how minimum tick size change impact venues with different exchange access fee. I show that, far from making existing fee structures work better, the minimum tick size itself operates as an excessively-severe tax on liquidity. In addition, this thesis examines the impact of the absolute minimum tick increase on market overall using difference-in-differences methodology.

C. Price limit

Circuit breakers aim to restrict extreme daily price movement, and it is a commonly used mechanism among stock exchanges that continuously monitor the market and trigger a trading halt when the price of a security or an index goes beyond a predetermined level. The market circuit breaker system was introduced after the market crash in October 1987. Trading halts (market or security level) and price limits (order rejection or volatility interruption) are two main types of circuit breakers (Abad and Pascual, 2013). After the May 2010 'Flash Crash', the SEC implemented the Single Stock Circuit Breakers (SSCBs), then replaced by the Limit Up Limit Down (LULD) on April 8, 2013. The importance of circuit breakers is highlighted again in today's financial markets, such as immense fluctuations of equity prices in August 2015, the four-day China stock markets crash in January 2016, and the Brexit referendum in June 2016 caused significant market activity.

Using LULD as an exogenous shock and propensity score matching, this thesis also aims to analyse the impact of this new dynamic price limit rule, and the role of HFT around price limits in maker-taker and taker-maker markets. To evaluate the impact of LULD comprehensively, the following five hypotheses are tested: trading interference, volatility spillover, delayed price discovery, magnet effects, and the role of HFT around the price limit on the maker-taker vs. taker-maker market. Consistent with the existing literature, the former three hypotheses are tested via comparing trading activity patterns, volatility levels, and price continuation and reversal patterns. For the magnet effect, I test both speed and magnitude aspects as well as on exchange with different access fee structures. To further examine the impact of price band setting on magnet effect, a subset of sample stocks switching above and below the $3 threshold is constructed and examined. In addition, how HFT trading behaviour changes around price limit are tested. Overall, LULD interferes with trading activity, but curbs short term volatility without delaying the price discovery process. The magnet effect exists when the trade price approaches the price limit. In addition, HFT trading activity dropped after the LULD trading halt, which was driven by the decrease in liquidity taking in the maker-taker market, while no changes in the inverted market.

1.2 Structure

The remainder of this dissertation is structured as follows. Chapter 2 reviews the existing market microstructure literature on exchange access fee, tick size, price limits. Chapter 3 provides a theoretical model of maker-taker fee structures and provide empirical evidence to show that why exchange access fee matter. Chapter 4 examines the impact of U.S. tick size pilot on markets with different exchange access fee structures. Chapter 5 analyses the impact of the intraday dynamic price limit rule - LULD on U.S. equity markets and HFT trading behaviour around price limits. Chapter 6 concludes with a summary of the main findings and discussion of key policy implications.

Chapter 2 Literature Review

This chapter reviews the theoretical and empirical literature on three critical equity market designs: exchange access fee, tick size and price limit. Section 2.1 summarizes the theoretical and empirical literature of exchange access fee. Section 2.2 focuses on the impact of tick size change overall, as well as its impact across different lit market access fee structures and off-exchange venues. Section 2.3 examines the impact of price limits, specifically, trading interference, volatility spillover, delayed price discovery, and the magnet effect hypothesis. Section 2.4 concludes with a discussion of our main contributions to the literature.

2.1 Exchange Access Fee

In the fragmented equity trading environment, trading venues commonly use exchange access fees to compete for order flow; therefore, understanding the impact of exchange access fee structures has become increasingly important in the new competitive environment (Malinova and Park, 2015). Around global equity markets, there are three main types of exchange access fees: taker-taker, maker-taker and taker-maker (inverted). The International Organization of Securities Commissions (IOSCO) defines maker-taker fees as "a pricing model whereby the maker of liquidity, or passive [limit] order, is paid a rebate and the taker of liquidity, or aggressive [market] order, is charged a fee." Maker-taker fees for protected quotes in the equity markets are bound by Rule 610 of Regulation NMS[2], which

[2] If the price of a protected quotation is less than $1, the fee cannot exceed 0.3% of the quotation price. See SEC Rule 610.

caps fees at 30 cents per 100 shares traded. Maker rebates aim to both improve liquidity, by rewarding its provision, and increase trading volume.

The maker-taker payment model originated with electronic trading venues in the late 1990s (Harris, 2013). Cardella, Hao, and Kalcheva (2015) report that the Island electronic communication network (ECN) decided in 1997 to pay a maker rebate to brokers who added liquidity and to charge a taker fee for brokers who took liquidity so as to compete with Nasdaq[3]. Soon, Island market share of Nasdaq trades increased significantly from approximately 3 percent in 1997 to almost 13 percent in 1999[4]. Since then, other U.S. alternative trading systems (ATSs) and lit markets followed this fee model especially when Attain ECN charged non-subscribers an access fee of 150 cents per 100 shares traded (CPS) in 1998 (Harris, 2013). Up to date, three U.S. exchanges adopted the inverted fee model, offering rebates for taking liquidity and conversely charging a higher fee for adding liquidity. On August 19, 2016, IEX became the latest exchange in the U.S. trading landscape and created a special taker-taker fee model.

Theoretical

Exchange access fee comprises a sizeable proportion of overall trading costs, given the typical bid-ask spread in a liquid stock is tick constrained to 1 cent for non-penny stocks and exchange access fee can be as high as 0.3 cents. Angel, Harris, and Spatt (2011) argue the introduction of a maker rebate financed by a taker fee should wash-out and thus have no effect because the best bid and ask prices in competitive markets should narrow by the rebate

[3] Matthew Andersen, the chief executive of Island ECN, said that Island hoped to "Jump-start the market" and that turned out to be true (Spicer, 2009).

[4] The historical growth of Island is available at http://www.hofstra.edu/pdf/biz_mlc_concannon1.pdf.

amount if all non-marketable limit orders are subsidized by an equivalent tax on market orders, with the tax precisely offsetting the narrowed spread.

Colliard and Foucault (2012) distinguish the change in the components of the exchange access fee (maker rebate and taker fee) and the change in the net exchange access fee (difference between maker rebate and taker fee), and find that only changes in the net exchange access fee matters in the absence of market frictions. If an exchange introduces a maker rebate and finances it by an increased taker fee, and keep the net exchange access fee constant, then, ceteris paribus, placing a marketable order becomes relatively more expensive than trading with a non-marketable order. As some traders switch from marketable orders to non-marketable orders, the execution probability of each non-marketable order declines and thus traders will improve quotes to attract matches for their limit orders. The benefit from maker rebates will be exactly offset by the narrowed bid-ask spread in the absent of frictions. Depending on the parameters, changes in the net exchange access fee affect a trader's choice of order type, and an increase in the net fee can cause an increase or decrease in trading volume.

Foucault, Kadan, and Kandel (2013) build on the Colliard and Foucault (2012) model to show that trading volume may either increase or decrease even in the absence of a change in the net total fee, as a fixed minimum tick prevents prices from neutralizing the effect of the maker rebate. Similarly, Chao, Yao, and Ye (2017a) extend the model of Foucault, Kadan, and Kandel (2013) to ascertain that a minimum tick size is sufficient to establish an equilibrium in which venues can co-exist with different fee structures. These models all share a common feature, namely the absence of informed trading. Foucault, Kadan, and Kandel (2013) examine the relative importance of the net fee and the levels of the maker and taker fee and claim that the breakdown of the total exchange access fee between makers and takers only becomes economically meaningful when the tick size restricts adjustments to quote prices. They model the degree of monitoring intensity by makers and takers such that an imbalance in latency and monitoring efficiency by HFT and algorithmic traders can induce an exchange to vary its fee and rebate structure. For example, if taker monitoring intensity is higher than maker, then the exchange can improve efficiency by reducing the maker fee with an offsetting increase in the taker fee. Hence, their model implies rigidity in the maker-taker decision.

In the maker-taker pricing model, the liquidity supplier has two main sources of revenue, namely bid-ask spreads and rebates, while liquidity takers profit from price movement minus the taker fee. Foucault, Kadan, and Kandel (2013) theorize that the make-take liquidity cycle process appears repeatedly in electronic markets. Large market orders consume the available liquidity and widen the quoted spread. This decrease in liquidity creates a profit opportunity for liquidity makers, which react by posting new quotes (make liquidity phase) that in turn create a new trading opportunity for liquidity takers (take liquidity phase).

Empirical

Malinova and Park (2015) examine whether and why the breakdown of exchange fees between liquidity takers and makers matters by using a change in trading fees on the Toronto Stock Exchange (TSX) which operates in a non-competitive or fragmented market during their sample period. They find that posted quotes adjust after the change in exchange trading fee composition, but that the trading costs for liquidity takers remain unaffected, and hence washes-out, once fees are considered. In the absence of competition from competing exchanges, one does not expect any meaningful change in the composition of the order flow. On a monopoly exchange, the variation in order flow informativeness is likely to be exceedingly small. Chao, Yao, and Ye (2017b) demonstrate empirically using reverse split events for leveraged exchange-traded funds (ETFs) that even a modest tick size of 1 cent for stocks priced at $2 or more can help explain market fragmentation.

Exchange pricing is increasingly important in a high frequency trading environment and deterministic of the recent shift from lit markets to off-exchange venues. Battalio, Corwin, and Jennings (2016) study the impact of differential exchange access fee schedules on broker routing decisions to find evidence that four of ten examined national retail brokers sell orders to capture the maximum make rebates. But high rebate venues experience lower fill rates due to the length of the queue. Consequently, on this measure of execution quality, some clients of retail brokers who do not receive the broker rebate may suffer from a conflict of interest.

The effect of changes in the breakdown of the total exchange fee into a maker rebate and taker fee is not neutral if some traders only pay them on average, such as via a flat commission to their brokers (Brolley and Malinova, 2012). Only a fraction of traders receives maker rebates for executed non-marketable orders in their model. As the maker rebate increases, these traders offer better quotes, and the raw quoted spread thus decreases. Ceteris paribus, traders who pay the flat fee based their order choices on the raw quoted spread, thereby submitting relatively more marketable orders as the raw quoted spread decreases. Thus, the change in trader behavior causes marketable orders to become less informative in the presence of asymmetric information. Similar to Kaniel and Liu (2006) and Rosu (2016), their prediction is driven by the monotonic equilibrium behavior of traders in their model whereby traders with a sufficiently large informational advantage use marketable orders and those with weaker information use non-marketable orders[5].

Barclay, Kandel, and Marx (1998) examine the effect of changes in bid-ask spreads on trading volume and prices of stocks that switched from Nasdaq to the New York Stock Exchange (NYSE) or American Stock Exchange (AMEX)[6] and stocks that move from AMEX to Nasdaq. They find that higher transaction costs reduce trading volume, but do not have a significant effect on prices. Lutat (2010) empirically investigate the impact of the maker-taker pricing on the Swiss Stock Exchange (SWX) after the introduction of Markets in Financial Instruments Directive (MiFID), and finds that the removal of a maker fee (without changing the taker fee) do not affect spreads, but leads to an increase in the best price level quoted depth. Cardella, Hao, and Kalcheva (2015) examine a number of maker-taker fee changes in the U.S. from 2008 to 2010 and find that both trading volume and revenue are not equally sensitive to changes in the maker or taker fees. An exchange's total fee relative to other exchanges in a fragmented trading environment affects the exchange's trading volume, and that a change in the taker fee has a stronger effect than a change in the maker rebate. However, Malinova and Park (2015) argue that changes in the maker-taker fees are accompanied by changes in the total fee in Lutat (2010) and Cardella, Hao, and Kalcheva (2015). Hence, the root of the impact cannot be differentiated precisely.

[5] Baruch, Panayides and Venkataraman (2017) show while extant literature views informed traders as using market orders, it might be optimal under some scenarios for informed traders to use limit orders.

[6] NYSE Euronext acquired AMEX on October 1, 2008, then renamed to NYSE MKT on May 10, 2012. After IEX become an exchange in 2016, NYSE MKT rebranded as NYSE American and introduced a 350-microsecond delay in trading ('speed bump').

Skjeltorp, Sojli, and Tham (2016) examine whether liquidity demand (supply) attracts or reduces liquidity supply (demand) using an increase of the rebate on the take side on the Nasdaq BX[7] which the market share is about 5 percent during their sample period. They document positive cross-sided liquidity externalities where liquidity taking begets liquidity making. However, the externality is negative in periods of high adverse selection.

Panayides, Rindi, and Werner (2016) study the effect of maker-taker fee reduction on market quality and market share and develop a model of an order book with make and take fees that faces competition from an alternative trading system. They find that simultaneous reduction in make and take fees results in a deterioration of market quality and significant spillovers between venues following fee changes in BATS Europe between 2013 and 2015.

Additionally, how the maker-taker fee reduction affects market efficiency is another important aspect of market quality. The notion of market efficiency hypothesis was introduced by Fama (1970) which points out the lack of return predictability as the efficiency criteria, while Chordia, Roll, and Subrahmanyam (2008) state that the microstructure literature emphasizes the amount of information reflected in prices (Lo and MacKinlay, 1988). They claim that return predictability can arise from order flows in at least two channels: 1) liquidity stimulates arbitrage activity which enhances market efficiency in turn; 2) market makers might fail to eliminate return predictability due to mis-reacting to the information content of order flow. The mispricing might create an incentive for traders to gather such information and trade on this information. In this case, market makers will face increased adverse selection cost, which may lead to a less liquid market. Meanwhile, stock price may

[7] Nasdaq completed the acquisition of the Boston Stock Exchange on August 29, 2008 and launched Nasdaq BX on January 16, 2009.

be more efficient as more information is reflected. Overall, Chordia, Roll, and Subrahmanyam (2008) find that liquidity enhances market efficiency and increased incorporation of private information into stock prices when liquidity is high.

2.2 Tick Size

Most equity markets around the globe have rules on minimum price variation, i.e. tick size or minimum tick. The tick size limits the minimum quoted spread, and no bid-ask spread may be smaller than the mandated minimum tick. Angels (1997b) points out the primary difference across markets is whether a trading venue uses a single absolute tick size that applies to most stocks (e.g. U.S.) or a tick size that is a step function of share price (e.g. Japan). The debate on the optimal tick size still continues and the impact of tick size change on different fee structure markets remains unclear.

Theoretical

Harris (1994) states that tick size creates discrete bid-ask spreads, and increased tick size will naturally widen quoted spreads and the quoted depth may be large. The cross-sectional discrete spread model of the U.S. market shows spreads would decrease 38%, quotation sizes would decrease 16%, and daily volume would increase for stocks priced under $10. Foucault, Kadan, and Kandel (2005) develop a limit order book (LOB) model populated by traders with varying patience level and find that a zero-minimum tick size is not optimal. As stated in Angel, Harris, and Spatt (2011), traders cannot fully neutralize maker-taker fees as quotes must be expressed as multiples of a minimum monetary unit in reality. Kadan (2006) argues that investors benefits from tick size reduction on markets with many dealers whereas they may suffer losses by small ticks when the number of dealers is small. Dayri and Rosenbaum (2015) discuss the notion of an 'implicit spread' for large tick LOBs which the value could be below the minimum tick and argue that the tick size is optimal when it equals the 'implicit spread'. The model is applied in Huang, Lehalle, and Rosenbaum (2015) to predict the impact on the Tokyo Stock Exchange when the tick size changes.

Harris (1991) and Angel (1997b) discuss the following advantages of a large tick size. First of all, it reduces the complexity of the trading environment as the number of possible price levels is limited which reduce bargaining costs. Secondly, it reduces the bandwidth requirements of a trading platform. A change in the single stock price can lead to a price update for hundreds of related derivative products. Thirdly, it sets the floor income for liquidity providers because bid-ask spreads cannot be quoted below the minimum tick. Last but not least, a finite tick size protects the time-priority rule. Time priority grows in importance when the tick size is large, and traders are incentivized to submit orders to the LOB early.

Empirical

Existing literature mainly focuses on the impact of tick size reductions. Harris (1994) predicts that if minimum tick decreases from one-eighth to one-sixteenth, exchange-designated market maker profits will remain unchanged for low-priced stocks, but it will decrease for high-priced stocks which assumes that volume supplied by market makers will increase proportionally with overall volume, and profits per share are proportional to quoted spreads. In U.S. markets, when the minimum tick decreased from one-eighth to one-sixteenth on June 24, 1997, consistent with predictions established in Harris (1994), tighter quoted spreads and lower trading volumes at the best quotes were generally reported. Specifically, using cross-sectional daily averages over a forty-trading day window centred on June 24, 1997, Bollen and Whaley (1998) find that the volume-weighted bid-ask spread decreases by more than 13 percent and quoted depth drops by about 38 percent on the NYSE. Goldstein and Kavajecz (2000) find quoted spreads and available depth decrease after the tick size reduction using NYSE LOB data. This combined effect has made liquidity takers executing small orders better off; however, traders submitting larger orders in low volume and low priced stocks do not benefit. Van Ness, Van Ness, and Pruitt (2000) study AMEX, Nasdaq and NYSE markets and find that the quoted spread and effective spread decrease. However, total quoted depth declines on the AMEX and NYSE but increases on the Nasdaq. In addition, Chordia, Roll, and Subrahmanyam (2001) find depth and spread declined around June 1997 on the NYSE.

Using U.S. 2016 tick size pilot as an exogenous shock which is similar to this thesis, Griffith and Roseman (2017) test the overall impact of tick size increase on market liquidity and find that the quoted and effective spreads widen; trading volume decrease; return volatility increase; and a wealth transfer from taking liquidity to adding liquidity although cumulative depth remains unchanged or decreases after widening the tick size. Moreover, Penalva and Tapia (2017) find both spreads and quoted depth increase, minimum price variation reduces, and trading volume and the level of market activity (messages and trades which are posted and subsequently cancelled in a short period) decreases for the stocks which quoted spread is similar to the new tick size which is five cents.

MacKinnon and Nemiro (2004) claim that the decrease in quoted spreads is partly caused by increased competition between exchange-designated market makers, while Chung and Chuwonganant (2004) report tighter quoted spreads on the Nasdaq after the change in order handling rules which allow competition between dealers and traders. Chakravarty, Wood, and Van Ness (2004) report spreads decline on the NYSE after decimalization. As tick size reduction increases the number of price levels available to liquidity suppliers, it effectively distributes liquidity onto a finer pricing grid. This can mechanically reduce quoted depth at the best level without reducing total liquidity or increase total liquidity in reality. Furthermore, Goldstein and Kavajecz (2000), Sie and McInish (1995), and Pavabutr and Prangwattananon (2009) find that the cumulative liquidity in the LOB decreases when the minimum tick is reduced. Chakravarty, Van Ness, and Van Ness (2005) examine adverse selection costs around NYSE decimalization and find a meaningful increase in the percentage adverse selection cost and a decrease in dollar adverse selection cost. Moreover, there are less stealth trading and less institutional trading following decimalization.

Besides U.S. equity markets, narrower quoted spreads and lower trading volume after tick size reductions are also documented for the Toronto Stock Exchange (Ahn, Cao, and Choe, 1998; Bacidore, 1997; MacKinnon and Nemiro, 1999; Porter and Weaver, 1997), the Singapore Exchange (Sie and McInish, 1995), the Taiwan Stock Exchange (Hsieh, Chuang, and Lin, 2008), the Hong Kong Stock Exchange (Chan and Hwang, 2001), the Jakarta Stock Exchange (Ekaputra and Ahmad, 2007), the Tokyo Stock Exchange (Ahn, Jun, Chan, and Hamao, 2007), and the Stock Exchange of Thailand (Pavabutr and Prangwattananon, 2009). On derivatives markets, similar pattern is also reported for the Chicago Mercantile Exchange (Kurov, 2008; Karagozoglu, Martell, and Wang, 2003), the Sydney Futures Exchange (Alampieski and Lepone, 2009).

In addition, Niemeyer and Sandas (1994), Chung, Kim, and Kitsabunnarat (2005) and Ke, Jiang, and Huang (2004) study the impact of the relative tick size for stocks listed on the Stockholm Stock Exchange, the Kuala Lumpur Stock Exchange, and the Taiwan Stock Exchange respectively, and all report larger quoted spreads for larger relative tick sizes. Jain (2003) tests 51 exchanges globally and report narrower spreads for smaller tick size. Cai, Hamao, and Ho (2008) and Bessembinder (2000) exploit threshold effects in the tick size bands of the Tokyo Stock Exchange and the Nasdaq respectively, and both observe larger spreads when the stock price entered a larger tick size band. Moreover, Hau (2006) finds a higher volatility for stocks with a larger tick size on the Paris Bourse.

Mixed results are also reported in examining the relationship between tick size and execution costs. Bollen and Whaley (1998), Alampieski and Lepone (2009), MacKinnon and Nemiro (1999) and Smith, Turnbull, and White (2006) report smaller transaction costs after tick size is reduced. However, Jones and Lipson (2001), and Bollen and Busse (2006) report that execution costs increased after tick size reductions. Angel (1997a) finds the amount of stock splits increases after tick size is reduced so as to keep the stock price in the 'optimal trading range' and maintain relative tick size. Schultz (2000) proposes that stock splits prompted brokers to promote the stock by increasing the spread and market making revenues, and this was critically discussed in Lipson and Mortal (2006) and Easley, O'Hara, and Saar (2001).

Some literature questioned the positive relationship between smaller minimum tick and higher liquidity. Bourghelle and Declerck (2004) find no change in quoted spreads after a tick size reduction on the Paris Bourse. Aitken and Comerton-Forde (2005) find that spreads increase for stocks whose relative tick size was already very small prior to the Australian

Stock Exchange (ASX) lower the minimum tick. Wu, Krehbiel, and Brorsen (2011) find similar results on the NYSE. Using splits/reverse splits as exogenous shocks, Yao and Ye (2015) argue that the uniform one cent tick size imposed by SEC Rule 612 harms liquidity. A large relative tick size (low-priced securities) constrains price competition and harms liquidity, and encourages HFTs to achieve speed advantage over non-HFTs at the constrained prices to earn the liquidity supply revenue. In addition, using NYSE order-level data, O'Hara, Saar, and Zhong (2015) find a larger relative tick size results in greater depth and more volume in a one-tick spread environment while the opposite outcome prevails in a multi-tick spread environment.

Tick Size and Dark Pools

Dark pools are equity trading systems that do not publicly display orders, and offer potential price improvements without execution guarantee (Zhu, 2014). Buti, Rindi and Werner (2017) model the interaction between a LOB and a dark pool. Their model shows that dark pool activity is likely to increase with market depth and tick size, and decrease when the inside spread widens. When market depth is high and the stock is trading at the minimum tick size, an order is either placed at the end of the queue of limit orders or the order may cross the spread to gain priority. Alternatively, the trade can be executed at the midpoint in dark pools. Thus, dark pools provide a cheaper alternative to trading on the LOB.

Zhu (2014) shows that dark pools can improve price discovery. Informed traders tend to trade in the same direction, crowd on the heavy side of the market, thus, facing a higher execution risk compared with uninformed traders. As a consequence, exchanges are more attractive to informed traders, while dark pools are more attractive to uninformed traders.

Adding a dark pool alongside an exchange may concentrates price-relevant information into the exchange and improves price discovery which coincides with the reduced exchange liquidity.

Jiang, McInish and Upson (2012) provide empirical support for Zhu's (2014) theoretical predictions. They find that the off-exchange order flow is significantly less informed which is consistent with the segmentation of uninformed traders into off-exchange venues, and conclude that price discovery and market quality is improved when informed order flow concentrates on the exchange.

Kwan, Masulis, and McInish (2015) investigate competition between lit markets and new dark trading venues using Reg NMS Rule 612 as a natural experiment.[8] They find that the limit order queues is larger as the minimum tick constrains certain stock spreads, and dark pools allow traders to bypass existing limit order queues with minimal price improvement. Also, increased orders migration to dark pools raises liquidity. They conclude that the ability to circumvent time priority of displayed limit orders is one reason of U.S. equity market fragmentation.

Foley and Putnins (2014) analyse the effects of dark trading using the minimum price improvement rules in Australia and Canada as natural experiments. They find dark limit order markets are beneficial to market quality as they encourage aggressive competition in liquidity provision. Also, tick size can affect the structure of dark trading when minimum price improvement rules are imposed. The level of dark trading decreases when the minimum price improvements is implemented on Australia and Canada.

---

[8] Rule 612 prohibits displaying, ranking, or accepting orders priced at more than two decimal places for stocks priced at or above $1 by broker-dealers and exchanges. When stock prices fall below $1, the required minimum pricing increment for exchange trades decreases from a penny, or $0.01, to $0.0001 (SEC, 2005).

2.3 Price Limit

Circuit breakers are procedures that halt trading temporarily or close the market for the rest of the trading day when the market price of a security or an index moves by a limit threshold percentage. It was first implemented at NYSE following the Black Monday of 1987, and have been widely adopted by equity and derivative markets around the world since then (Gomber, Clapham, Haferkorn, Panz, and Jentsch, 2016). Circuit breakers are triggered when prices move beyond predetermined price limits. Greenwald and Stein (1991) argue a system of continuous trading sacrifices some degree of informational efficiency for timeliness. The informational loss is severe when volume shocks are large. In this scenario, it is better off to switch temporarily to circuit breaker in spite of compromising the ability of market to provide immediacy. Subrahmanyam (2013) states that “the general notion is that rapidly falling prices may exacerbate panic amongst investors and cause limit orders to become unfairly stale. Circuit breakers that allow a cool-down period and a batching of trades can mitigate this problem.” Also, price limits could be an effective policy in “improving market efficiency and reducing the risks associated with financial instability” (The UK’s Government Office for Science Report, 2012).

The debate about whether the price limit rule is beneficial or harmful to the market continues in the existing academic literature. The advantages of circuit breakers are as follows. First, a circuit breaker could restrict upfront trading, or trading motivated by the fear of a future liquidity shock rather than true liquidity needs (Draus and Achter, 2015). Second, extreme order imbalances during rapid market movements might result in prices deviating from fundamentals; hence, allowing orders to accumulate and then batching them may lead to better execution quality by setting a ceiling price and providing a cooling off period (Kim and Rhee, 1997; Kim and Sweeney, 2002; Hsieh, Kim, and Yang, 2009; Subrahmanyam, 2013). Third, it counters overreaction and may preclude trade at prices that occur in response to automated executions of erroneous orders (Subrahmanyam, 2013). Fourth, regulators could use price limits to counter stock price manipulations by large investors (Kim and Park, 2010).

In contrast, Kim and Rhee (1997) argue that it may interfere with trading due to restrictions imposed by price limits (*trading interference hypothesis*); cause higher volatility levels on subsequent periods (*volatility spillover hypothesis*); delay prices from efficiently reaching their equilibrium level (*delayed price discovery hypothesis*); and accelerate price toward the limits as it gets closer to the limits (*magnet effect hypothesis*).

Theoretical

Subrahmanyam (1994) provides a formal theoretical model of the *magnet effect*, which is a generalization of Kyle (1985), engendering a line of empirical scrutiny and hypothesis testing. It posits that *uninformed* traders rush to trade in anticipation of market halts, thus increasing volatility, decreasing price efficiency, and increasing the probability of price limit hits. Subrahmanyam (1997) proposes the *hold back hypothesis* that *informed* traders will strategically change their aggressiveness to hold back their trading to avoid a price limit and continue trading on mispricing. This model predicts that a price limit will decrease volatility and the likelihood of an extreme price movement, but also decrease price efficiency. Brogaard and Roshak (2016) study the staggered introduction of the price limit rules in the U.S. and find that price limits reduce the frequency and severity of extreme price movements. Brogaard and Roshak (2016) test the two competing hypotheses from

Subrahmanyam (1994, 1997) and find they induce price under-reaction and cause informed traders to be less aggressive, which is consistent with the holding back hypothesis.

Kim and Park (2010) introduce the first model to propose a manipulation-based rationale for the existence of price limits in stock markets. They show that price limits may deter stock market manipulators and regulators impose price limit rules for markets where the likelihood of manipulation is high.

Chen, Petukhov, and Wang (2016) claim to have built the first dynamic model to examine the mechanism of market-wide circuit breakers with an optimistic and pessimistic agent and in the absence of market frictions. They conclude that circuit breakers give too much weight to pessimistic investors. This distortion generates both a magnet effect and an excess volatility effect which, surprisingly, is more severe the smaller the wealth of the irrational (pessimistic) investor.

Empirical

Lauterbach and Ben-Zion (1993) state that the interference of trading is cited as the ‘obvious cost’ to circuit breakers. If price limits interfere with trading, stocks may become less liquid, and this may lead to intensified trading activity in the subsequent trading periods (Fama, 1989; Lehmann, 1989). Lee, Ready, and Seguin (1994) find that both volume and volatility increase after trading halts on the NYSE, and the increase is even higher than on normal days.

Fama (1989) states the underlying volatility may increase if the price discovery process is interfered with. Kuhn, Kuserk, and Locke (1991) find that limits are ineffective in reducing volatility during the 1989 U.S. mini-crash. Lehmann (1989) suggests that trading imbalances between supply and demand actually induce prices to reach the limits, which implies a transfer of transactions to subsequent days. Ma, Ramesh, and Sears (1989b) conclude the volatility declines on days following limit days as a benefit of price limits; however, both Lehmann (1989) and Miller (1989) argue that such a finding is inevitable and trivial as volatility is biased to decrease on days after high volatility. Kim and Rhee (1997) report that volatility does not return to the normal level after reaching the price limits using the Tokyo Stock Exchange data. Therefore, instead of reducing volatility, price limits may cause volatility to spread out over longer periods as price limits prevent both large one-day price changes and immediate order imbalance corrections.

As the trade prices approach to upper or lower price boundaries, trading halts are usually triggered, which interfere the price discovery process (Fama, 1989; Lehmann, 1989; Lee, Ready, and Seguin, 1994). The delayed price discovery hypothesis suggests that price limits prevent prices from reacting to new information and reaching the new equilibrium level. Stocks may be prevented from reaching their equilibrium price if constraint is set on price movements; hence, true price can only be reached in the subsequent trading period (Kim and Rhee, 1997).

The magnet effect suggests that the stock price accelerates toward the price limits as it moves closer to the upper or lower price limits. Two channels are proposed for the magnet effect: rational anticipated illiquidity and behavioural finance. First, if traders are fearful of illiquidity, they would get involved in active trading moving the price closer to the limits. Subrahmanyam (1994) shows in an intertemporal one-market model that price limits can increase ex ante price variability and trading volume with the increased probability of the price crossing the limits. Lehmann (1989) contends that order imbalances and the consequent lack of trading induce prices to reach the limits. Second, Arak and Cook (1997) argue that behavioural investors who follow price trends can act in a way that produce the magnet effect. To avoid being shut out of a trend, traders who expect price limits to be reached might execute sooner. This behaviour will accelerate price changes as the trading price moves closer to the limits. On the other hand, the cooling off effect, which is the opposite of the magnet effect, is one of the major benefits of price limits. A trading halt will be triggered when the price reaches its limit for a certain period of time. The market will then have time to re-evaluate the 'true' value to counter the overreaction (Cho, Russell, Tiao, and Tsay, 2003).

There are mixed evidences in the literature concerning the magnet effect of price limits. Cho, Russell, Tiao, and Tsay (2003) estimate the return process under a tight (7%) price limit and find strong evidence that stock prices accelerate toward upper limits, but little evidence that prices accelerate toward lower limits on the Taiwan Stock Exchange. Goldstein and Kavajecz (2004) find that trading on the NYSE accelerated just before a trading halt in October 1997 with just one episode evidence. Chan, Kim, and Rhee (2005) examine how the magnet effect occurs via order imbalance on the Kuala Lumpur Stock Exchange, and find a magnet effect due to the order imbalance prior to the limit-hit, as well as a subsequent order imbalance reversal after the limit-hit, which supports the idea that magnet effects take place during the pre-hit period.

Hsieh, Kim, and Yang (2009) report that the magnet effect activates when the price falls within nine ticks of the upper price limits and about four ticks of the lower price limits using logit regression on the Taiwan Stock Exchange. In addition, Du, Liu, and Rhee (2009) find evidence consistent with the magnet effect in returns, trading volume, and volatility using a time-distanced quadratic model on the Korea Stock Exchange. Tooma (2011) finds that the conditional probability of reaching a limit increases after imposing the price limits on the Egyptian Stock Exchange.

On the other side, Nath (2003) finds mixed evidence that trading activity reduces when price approaches upper limit price and accelerates when price approaches lower price limit on the National Stock Exchange of India. Abad and Pascual (2007) find no support for the magnet effect on the Spanish Stock Exchange (SSE), and price limits do not cause traders to accelerate trading activities. Moreover, Arak and Cook (1997) and Berkman and Steenbeek (1998) find no evidence of the magnet effect on Treasury bond, commodity futures, and Nikkei 225 futures.

In addition, Kirilenko, Kyle, Samadi, and Tuzun (2017) find that HFT do not cause the May 2010 ‘Flash Crash’ via studying the intraday market intermediation in an electronic market. Madhavan (2012) finds that prices are more sensitive to liquidity shocks when markets are fragmented as fragmentation leads to a thinner LOB in each venue than in a consolidated market.

2.4 Conclusion

The importance and impact of exchange access fee, tick size and price limits remain unclear and further research is required to clarify and fill in the literature gap. First, whether the component of exchange access fee matters in a competitive trading environment remains unclear. Previous literature finds that only the net exchange access fee matters, as the impact of component of exchange access fee wash-out. Other studies find that the breakdown of the total exchange fee between makers and takers only matters when the minimum tick restricts adjustments to quoted spreads. Second, previous studies on tick size mainly focused on its overall impact on liquidity, and have mixed results. Moreover, extremely limited literature has examined the impact of tick size change across different exchange access fee venues in a fragmented trading environment. Third, there is a lack of empirical studies of the *dynamic* intraday price limit rule, i.e., Limit Up Limit Down. Previous empirical evidence about whether the *static* price limit is beneficial or harmful to trading mainly using low frequency data (often daily data) and lack of episodes. This thesis aims to address those questions and provide conclusive evidence.

Chapter 3 Why Maker-Taker Fee Matter

3.1 Introduction

Exchange venue fee structures and their impact remain unresolved. Angel, Harris, and Spatt (2011) propose that fee-rebate schemes wash out, while Colliard and Foucault (2012) and Malinova and Park (2015) underscore net fee changes since the maker-taker fee breakdown is irrelevant. In addition, Foucault, Kadan, and Kandel (2013) claim that maker-taker venues benefit from a rise in the minimum tick from one to twelve cents because a finite minimum tick prevents prices from fully neutralizing the maker rebate effect. Like Foucault, Kadan and Kandel (2013), Chao, Yao, and Ye (2017a) show without a minimum tick, the principle of tax neutrality requires that it does not matter whether the buyer or seller is liable for the tax. However, in the presence of a minimum tick constraint, fee structures could be used to constrain or segment the market if liquidity maker preferences are sufficiently heterogeneous. While this taxation principle is correct if no investors possess information, it is a mistake to conclude that a make tax with a corresponding take subsidy, or vice versa, could cancel out when there is a zero net-fee other than in the absence of informed trading.

We show that it is in the interests of informed traders, here takers, that are endowed with the true value of the security – high or low – to raise the information content of their trades by the full amount of the maker rebate to achieve the efficient corner solution in maker-taker markets. In inverted markets, the reverse occurs with information content falling by the full extent of the fee paid by limit orders, leaving the raw spread unchanged, together with deterioration in market efficiency. Thus, far from washing out, the maker-taker fee structure improves market efficiency and price discovery. Additionally, it raises the market share of trading volume to exchanges with the highest rebates by attracting more traders with information and more makers to act as counterparties to the newly recruited informed traders. Since liquidity traders who place limit orders earn the same raw spread plus an additional rebate when their trades are executed, and since a rebate raises informed trader profitability, fee rebates are Pareto-improving. Additionally, maker-taker venues provide price discovery enhancement, a positive externality. In contrast, inverted taker-maker markets with a fee on makers and rebate to takers are Pareto-inefficient relative to tax-neutral zero-fee regimes as informed trader profitability falls while leaving the net fee on uninformed traders the same.

This ongoing debate on the exchange fee structure in U.S. markets led Nasdaq to conduct a 'quasi-natural' fee-pilot experiment by lowering both the maker rebate and taker fee by an identical amount (25 cents per 100 shares traded) for a selected group of securities for a pre-specified period. This access fee pilot is described in greater detail in Section III below. While Nasdaq's fee-pilot experiment was motivated by competition from dark pools and other off-exchange venues[9], for us the real benefit lies in its ability to provide robust tests of our theoretical model that reveal for the first time why maker-taker fee structures are of critical importance for almost every aspect of market quality. In this chapter, we show that the maker-taker fee structure employed by Nasdaq has highly desirable features in that it improves price discovery and market efficiency while at the same time maximizing Nasdaq's trading volume and market share. How can that be? In short, maker-taker tax-subsidy schemes result in fundamental changes to the information content of equilibrium order flow across venues with no prospect of ever cancelling out, irrespective of the minimum tick size.

[9] Over the past decade, hundreds of exchange-listed securities have traded more volume in off-exchange markets than on exchange markets. In response to claims that public markets are too expensive to trade, Nasdaq wanted to know whether access fees may be discouraging the use of public markets (SEC, 2014).

Our findings from the fee pilot experiment support all the predictions from our theoretical model while rejecting the conclusions of all those models that find that either fees and rebate cancel out or that maker-taker venues benefit from an increase in the minimum tick.

Ours is not the only paper to address these issues. Panayides, Rindi, and Werner (2016) study the effects of BATS Europe maker-taker fee reduction on market quality and market share in a fragmented market. Their model has predictions that differ from ours. In their model, an increase in the taker fee and maker rebate reduces the spread and increases the depth, but inter-venue competition leads to a migration of order flow away to other venues, worsening both market quality and venue share. This is because in their model, as in all maker-taker models prior to ours, informed order flow is not modeled and a minimum tick constraint is relied on to prevent washing out. By contrast, in our model, the increased make rebate does not reduce the spread or increase depth because both informed order flow and makers prepared to counter their orders are attracted from other venues. Hence, both market share and pricing efficiency improve.

Barclay, Kandel, and Marx (1998) empirically study how changes in quoted spreads influence trading volume and prices, finding that higher transaction costs reduce trading volume. Lutat (2010) argues that the removal of a maker fee (without changing the taker fee) has not affected the quoted spreads on the Swiss Stock Exchange. In two empirical tests, Dosanjh (2013) examines the introduction of maker fee rebates to the Australian Stock Exchange (ASX) Electronic Traded Funds (ETFs) in 2010 to show that rebates improve depth and liquidity while Clapham, Gomber, Lausen, and Panz (2017) examine the 2016 Xetra Liquidity Provider Program at Deutsche Bourse, which provides maker rebates to find it increased the venue’s liquidity share without increasing that of the entire market.

Cardella, Hao, and Kalcheva (2015) study several maker-taker fee changes in the U.S. from 2008 to 2010. They find that an exchange's total fee relative to that of other exchanges affects its trading volume and that a change in the taker fee has a stronger effect than a change in the maker rebate. They also show that the breakdown of the total exchange fee into maker rebate and taker fee does not affect the quoted spreads.

Our theory predicts that inverted taker-maker markets would suffer an outflow of informed traders, thin markets, and reduced price discovery, as limit order book access fees increase. While Nasdaq's fee pilot experiment, with a take fee of 5 cents and a maker rebate of 4 cents per 100 shares, approximately modeled a neutral market, stopping short of an inverted market, we nevertheless see a severe departure of informed traders, a thinning of the limit order book, and increased pricing inefficiency as the rebate to non-marketable[10] limit orders is withdrawn. The reason the existing models of the maker-taker fee reach a conclusion that differs from ours is that none of them considers informed traders. Hence, they obtain the limiting equilibrium in our model in which informationless trading leads to a precise cancelling out of the cum-fee and matching rebate equilibrium. As predicted, we find empirically that non-High Frequency Traders (non-HFTs), with their less efficient monitoring of the order flow dynamics, switch from takers to makers as both the information content and the rebate are reduced, while HFTs do the reverse. This is not only because can HFT switch sides but because they focus on the most informed traders requiring more speed and expertise. Again, our theory explains why these outcomes occur in maker rebate-taker

[10] Non-marketable limit orders enter the limit order book, but do not generate immediate trades that remove existing orders from the book.

fee markets; higher access fees increase beneficial informed equilibrium order flow and lower access fees reduce it.

A. Our Contribution

Our model of maker-taker fees predicts that the introduction of a fee/rebate from an initial zero fee-subsidy regime will not narrow the raw (observed) effective spreads or quotes as wash-out requires, but will raise the expected market impact cost to the full extent of the make rebate due to increased market order information content. Hence, the raw realized spread, reflecting the new higher level of information content, must fall by the extent of the price impact to leave the effective raw spread unaltered and the cum-fee spread widened. This is because the higher information content of trades should fully absorb the rebate. Once the new take fee is factored in, market orders face a higher cost; that said, informed trader profit nevertheless rises due to the greater information content of order flow. Our model predicts that the cum-fee effective spread and quotes will rise by the entire extent of the fee rebate but the rise in raw market impact costs will be fully absorbed by the fee rebate, leaving a smaller alteration in cum-fee market impact costs. Hence, the rise in the cum-fee realized spread will precisely equal the rebate.

Our model also predicts that alterations to the rebate, or its abolition, by a lit venue will have little or no effect on competition between the altered venue and dark pools as the latter are designed to exclude informed traders. The sole effect of the rebate in lit markets is as a magnet to attract informed traders from other high-rebate venues. Rebate modifications therefore have no bearing on dark pools. Instead, the model indicates that the unilateral removal of the make rebate in a high-rebate venue will reduce market share as informed traders flee to the next-highest high-rebate venue. In this study, we employ difference-in-differences (DiD) methodology to test the predictions of our theoretical model when access fees are reduced. All these predictions that specify the consequences when a maker-taker fee venue is introduced will be reversed when the maker-taker fee structure is largely eliminated, as it was during the Nasdaq fee pilot.

Additionally, we test the Foucault, Kadan, and Kandel (2013) maker-taker model which predicts that in the presence of a finite minimum tick the success of maker-taker venues is driven by exogenous imbalances in the monitoring ability of specialized firms such as HFT between maker and taker markets. Since we show that HFT switches from adding to removing liquidity as the tax-subsidy is removed, we reject Foucault, Kadan, and Kandel's (2013) implicit hypothesis that monitoring imbalances lead to exogenous rigid maker-taker roles. Instead, our theory predicts these specialized monitors can freely and endogenously switch sides between maker and taker markets in response to imbalance in monitoring effectiveness. This novel finding indicates that HFT is a distinct trader-type, neither informed nor uninformed, and has both a substantial and flexible niche regardless of tax-subsidy schemes. Moreover, both the speed and expertise of HFT indicate their relative specialization in dealing with highly informed trades.

Finally, our model rejects as redundant the claim made by Chao, Yao, and Ye (2017a) that competing fee venues can be explained by heterogeneous limit order providers with the minimum tick size constraint such that fee structures would otherwise wash out. This model does not predict the alterations to asymmetric information and information content that occurred because of the Nasdaq fee experiment.

The strength of our analysis lies both in our novel theory and in the structure of the Nasdaq access fee pilot. First, Nasdaq introduced the maker-taker fee reduction only for a predefined subset of 14 securities, permitting the analysis of the impact using a DiD approach. Second, the pilot only lasted four months in 2015, enabling us to compare PilotOn with PilotOff events. Third, Nasdaq data allow us to analyze how HFT reacts to the maker-taker fee reduction compared with non-HFT across different resting order types, giving us new insights into the nature of HFT to address issues raised by Foucault, Kadan, and Kandel's (2013) assumed monitoring discrepancy between maker and taker markets.

Our results show that, as predicted, when Nasdaq reduces the maker-taker fee unilaterally without the cooperation of other exchanges, there are no significant changes in the market share of off-exchange trading venues (we classify these into dark pools and non-dark pools). Instead, we observe a redistribution of the traded volume between lit venues. Nasdaq's market share was reduced with trading activity shifting to exchanges with higher maker rebates. Hence, holding the exchange net fee relatively constant, when the make rebate is reduced, the queue of the Nasdaq limit orders shrinks as it shifts to the most closely competing venues, as evidenced by the drop in the percentage of time and depth at the National Best Bid and Offer (NBBO). Depth cannot be provided in the limit order book if there are no highly informed orders to hit orders placed away from the best bid and ask. Note that one cannot logically argue that this depth decline occurred because of the removal of the subsidy to limit orders as in the extant 'wash-out' models because subsidy removal has no effect on the depth in the LOB.

Our theoretical model predicts that the information content of trades will decline by the precise amount of the rebate decline. The limit order book becomes thinner because, deprived of informed traders who have switched to other high-rebate exchanges, limit order placements away from the best bid and ask lack the essential informed counterparties. Realized spreads should then increase to reflect the falling proportion of informed traders, but the cum-rebate (i.e., the net realized spread) decreases because of the smaller proportion of informed traders hitting limit orders. Hence, not only does the pilot reduce the net realized spread, but the effective spread and market impact costs are also lower; these are all indicators of the switch in the informed order flow toward other high-rebate venues. As a net result, the Nasdaq market share declines on average and is captured by the other high rebate-paying exchanges in close competition with Nasdaq. Our theoretical model and these findings indicate that the components of the exchange access fee do matter greatly in a competitive trading environment. Hatheway, Kwan, and Zheng (2016) reported that investors can trade on about 300 different venues, including thirteen registered exchanges, approximately forty or so active alternative trading systems (ATSs) and numerous broker-dealer platforms in the U.S. equity markets.

A natural objection to our claim that maker-taker markets are Pareto-superior to either inverse or neutral markets (no fees or rebates) is that while the raw spread remains constant as the rebate rises, uninformed market orders are still subject to the rebate-matching fee and hence these traders are apparently worse off. However, since these traders are free to place limit orders and receive the rebate when their orders are executed, in addition to the raw spread that has remained unchanged, they appear to be no worse off. However, a possible downside is the lower fill rate reflecting the deeper limit order book market. Linnainmaa (2010) shows that households that lack short-term information are more likely to place limit orders. Moreover, our model of maker-taker venues indicates that neither uninformed dealers (makers), nor uninformed liquidity traders, are any worse off.

B. Maker-Taker Fees

The International Organization of Securities Commissions defines maker-taker fees as "a pricing model whereby the maker of liquidity, or passive [limit] order, is paid a rebate and the taker of liquidity, or aggressive [market] order, is charged a fee." Maker-taker fees for protected quotes in the U.S. equity markets are bound by Rule 610 of Regulation NMS[11], which caps fees at 30 cents per 100 shares (CPS) traded, i.e., 0.3 cents per share. Maker rebates aim to both improve liquidity by rewarding its provision, and increase trading volume by raising execution quality.

The maker-taker payment model originated with electronic trading venues in the late 1990s (Harris, 2013). In 1997, the Island ECN was among the first markets to adopt maker-taker fees, which attract order flow through liquidity rebates[12]. These maker rebates provide traders with an additional source of income other than the quoted spread, incentivizing liquidity providers to post more competitive quotes to attract order flows from other markets. Thus, Island's market share of reported Nasdaq trades increased from approximately 3% in 1997 to almost 13% in 1999 (Cardella, Hao, and Kalcheva, 2015). This sizeable rise in market share is unlikely to have either 'wash-out' or a vital role for minimum tick as its explanation. Other ATSs soon followed Island's fee model to attract order flows and liquidity from lit exchanges (SEC, 2015a). Alternative versions of maker-taker fee structures were soon

[11] If the price of a protected quotation is less than $1, the fee cannot exceed 0.3% of the quotation price. See SEC Rule 610.

[12] The Island ECN historical growth is available at http://www.hofstra.edu/pdf/biz_mlc_concannon1.pdf.

introduced. For example, in 2008 the NYSE abolished its specialist system of market makers in favor of contractual Designated Market Makers (DMMs) that were rewarded with a maker-taker fee structure with a taker fee of 0.275 cents and a maker rebate of 0.27 cents in 2017. These contracts require DMMs to maintain competitive bids and offers for a fraction of the trading day that depends on the prior month trading volume. More stringent market making requirements are associated with increased depth, narrower bid-ask spreads, increased firm value, and improved price efficiency (Bessembinder, Hao, and Zheng, 2017).

In response to the ATSs competition, many exchanges adopted their own maker-taker fee model. Over the past decade, the maker-taker pricing model has thus gained widespread adoption in the U.S. equity markets, rewarding liquidity suppliers and charging liquidity demanders. Only three U.S. exchanges (BATS-Y, EDGA, Nasdaq/BX) adopt the inverted fee (taker-maker) model, offering rebates to remove liquidity accompanied by a higher fee to add liquidity. As predicted by our model, the market shares of these inverted venues are small in comparison to those of maker-taker venues.

Taker fees and maker rebates comprise a sizeable proportion of overall trading costs, given the typical bid-ask spread in a liquid stock is tick-constrained to one cent set by the SEC for stock prices above $1. Although most retail investors do not directly observe fees and rebates, all institutional investors and market makers who account for the majority of trading activity are quite cognizant. Brokers commonly sell their marketable orders to wholesale dealers to capture the quoted spread and to avoid exchange access fees, and send their non-marketable orders to exchanges for executions in order to receive maker rebates (Angel, Harris, and Spatt, 2015). Moreover, so-called 'smart order routers' that consider fee rebates along with real-time state information and formulate an order routing problem that considers various execution metrics to decide whether to place a limit order or market order and accordingly to which venue(s) to direct their order (Maglaras, Moallemi, and Zheng, 2012).

Policy questions are foremost in this maker-taker fee debate. The SEC has questioned whether "rebates [are] unfair to long-term investors because they necessarily will be paid primarily to [high-frequency] proprietary firms engaging in passive market making strategies. Or do they generally benefit long-term investors by promoting narrower spreads and more immediately accessible liquidity?" (SEC, 2010). Angel, Harris, and Spatt (2011, p. 39) further argue that the maker-taker fee model has "aggravated agency problems among brokers and their clients" because typical brokers do not forward the exchange fees to their clients on a trade-by-trade basis and may have a conflict of interest with their clients regarding the choice of trading venues.

C. Prior Literature

From an empirical perspective, Malinova and Park (2015) analyze whether and why the breakdown of trading fees between liquidity demanders and suppliers matters by using a change in trading fees on the Toronto Stock Exchange (TSX) which then operated in a non-competitive and thus unfragmented market. Hence, there is no alternative stock exchange for informed trades to flow. They find that posted quotes adjust after the change in fee composition, but that the transaction costs for liquidity demanders remain unaffected, and hence washes out, once fees are considered. In the absence of competition from competing exchanges, one does not expect any meaningful change in the composition of the order flow. On a monopoly exchange (TSX at the time), the variation in order flow informativeness, as measured by the proportion of informed to overall trades, is likely to be exceedingly small such that the introduction of a maker rebate has negligible effect. In contrast, when Nasdaq largely removed its maker rebate in the access fee experiment, there was a sizeable fall in order-flow informativeness.

Using reverse split events for leveraged ETFs, Chao, Yao, and Ye (2017b) demonstrate empirically that even a modest minimum tick size of 1 cent for stocks priced at $2 or more can help explain market fragmentation. However, in our framework, the minimum tick size is redundant as an explanation for fragmentation. Each venue specializes in a different type of order flow, with the highest rebate venue seeking the most informed order flow.

Most relevant to our research, Foucault, Kadan, and Kandel (2013) examine the relative importance of the exchange net fee and the levels of the maker and taker fees. They claim that the breakdown of the total fee between makers and takers only becomes economically meaningful when the minimum tick size restricts adjustments to bid and ask prices. However, they do not consider how the competition between different maker-taker venues and business models affects the information content of order flow. They model the degree of monitoring intensity by makers and takers such that an imbalance in latency and monitoring efficiency by HFT and algorithmic traders can induce an exchange to vary its fee and rebate structure. For example, if taker monitoring intensity is higher than maker monitoring intensity, then the exchange can improve efficiency by reducing the maker fee with an offsetting increase in the taker fee. Hence, their model implies rigidity in the maker-taker decision. We test this aspect of their model to show that HFT traders switch from making to taking liquidity for the duration of the Nasdaq's removal of the make rebate, while less effective non-HFT monitors do the reverse. They also use their model to simulate a sizeable benefit to maker-taker venues from a minimum tick rise from one to 12 cents, but when the SEC mandated tick pilot raised the minimum tick from one to only five cents, the market share and performance of maker-taker venues deteriorated significantly.

Battalio, Corwin, and Jennings (2016) examine the impact of differential exchange access fee schedules on broker routing decisions to find evidence that four of an examined ten national retail brokers sell orders to capture the maximum make rebates. However, high rebate venues experience lower fill rates due to the length of the queue. Consequently, on this measure of execution quality, some clients of retail brokers who do not receive the broker rebate may suffer from a conflict of interest.

Brolley and Malinova (2012) argue that the effect of changes in the breakdown of the total exchange fee into a maker rebate and a taker fee is not neutral if some traders (e.g., retail traders) only pay them on average. In their model, only a fraction of traders receives maker rebates for executed non-marketable orders. Those traders improve their quotes as the maker rebate increases, and the raw quoted spread thus declines. Ceteris paribus, traders who pay the flat fee decide their order choices based on the raw quoted spread instead of the cum fee quoted spread, thereby submitting relatively more marketable orders as the raw quoted spread decreases. They then predict that the change in trader behaviour causes market orders to become less informative with asymmetric information. Their prediction is driven by the monotonic equilibrium behavior of traders in their model whereby, like Kaniel and Liu (2006), Rosu (2016) and Baruch, Panayides and Venkataraman (2017), traders with a sufficiently large informational advantage may use market orders and those with weaker information use limit orders in certain scenarios.

Instead, we wish to emphasize that HFT traders who rely on speed and more effective monitoring of the informed order flow dynamics can place either limit orders or market orders. In our theoretical model, traders can switch instantaneously from one side of the market to the other, depending on relative profitability. In the maker-taker pricing model, the liquidity supplier has two main sources of revenue, namely, bid-ask spreads and rebates, while liquidity takers profit from price movement minus the taker fee. In electronic markets, large market orders consume the available liquidity and widen the quoted bid-ask spread. This fall in liquidity creates a profit opportunity for liquidity suppliers, which react by posting new quotes (make liquidity phase), which in turn create a new trading opportunity for liquidity demanders (take liquidity phase). Foucault, Kadan, and Kandel (2013) theorize this 'liquidity cycle' process appears repeatedly. In our theoretical analysis, we also assume that each trading opportunity is short-lived as it disappears once a trader exploits it.

3.2 Model

Following in the framework of the seminal analysis by Glosten and Milgrom (1985) and also Aitken, Garvey, and Swan (1995), we are the first to analyse access fees in the presence of informed order flow. There is one unit of a representative stock whose price can take on one of two values, V^H or V^L, with $V^H > V^L$. Setting the unconditional share price at $\left(V^H + V^L\right)/2 \equiv V$, namely, the valuation placed on a share by the uninformed trade, then $V^H \equiv (1+\alpha)V$ and $V^L \equiv (1-\alpha)V$, where $1 \geq \alpha \geq 0$ is the value of the information about the 'true' price revealed only to informed traders prior to placing their order. One can think of α as our measure of the degree of information content in an informed market order. At $\alpha = 1$, the informed trader's informational advantage is at its maximum and as $\alpha \to 0$ the informational advantage evaporates. Throughout this paper, information content is treated as exogenous, and thus the costs of acquiring information are irrelevant. In Section II.C below we appear to endogenize the degree of information content when modelling the informed trader's choice of venue. However, this must be understood as a situation in which information content is still exogenous but endogenous to a particular venue as it varies the size of its rebate since the informed trader has to decide on his choice of venue given his degree of information content. A risk-neutral liquidity trader who is a potential seller values the share at a fraction, $(1-\lambda)V$, of the unconditional value for liquidity or portfolio rebalancing reasons while an equivalent potential buyer values the share at $(1+\lambda)V$, where λV is a measure of the gains from a trade, with $(1-\lambda)V$ and $(1+\lambda)V$ representing private valuations of the liquidity seller and buyer, respectively, with $0 \le \lambda \le 1$. Liquidity traders are randomly assigned either a low or high valuation, while $\lambda = 0$ for both makers and informed traders alike. This difference in private valuation motivation for liquidity traders is identical to that employed by Colliard and Foucault (2012), Foucault, Kadan, and Kandel (2013) and Chao, Yao, and Ye (2017a), except there is no informed trading in those models, and hence $\alpha = 0$. Informed traders who know the true value of V prior to placing their order consist of a proportion $1/2 > \gamma \ge 0$ of the entire population of traders with this proportion known to dealers. Potential sellers with valuation $(1-\lambda)V$ are offered the 'bid, i.e., 'sell', price, denoted as $p_s \ge (1-\lambda)V$ in the limit order market, so that the price must equal or exceed his private valuation and potential buyers are offered the 'ask', i.e., 'buy', price $p_b \le (1+\lambda)V$ with the exogenously fixed maximum width of the 'inside' quotes, $p_b - p_s = 2\lambda V$, increasing in gains from trade, λV.

A. No Fees or Rebates

We begin with an examination of one of potentially many competing venues that charges no fee and offers no rebate. An informed trader with short-lived information about future prices will only buy when he knows the true value of the security is high at $(1+\alpha)V$ and sell when he knows it is low at $(1-\alpha)V$. Since his information is short-lived and only market orders execute with certainty, he will only place market orders. Hence, a dealer (maker) who encounters an informed seller with probability γ by placing a limit order to buy loses the amount $(1-\alpha)V-p_s$ and one who encounters an informed buyer loses $p_b-(1+\alpha)V$. The maker who places limit orders to buy breaks even if the expected profit from buying from uninformed sellers at the (low) sell price (i.e., the 'bid'), $p_s=(1-\lambda)V$, and selling to uninformed buyers at the unconditional value, V, with probability $1-\gamma$, makes up for his loss by buying from informed sellers at the (low) sell price and having to sell at the even lower 'true' price, $(1-\alpha)V$, known only to the informed seller prior to the trade.

The maker's expected profit, π_m, from placing an uninformed limit order to buy at the sell price, $p_s=(1-\lambda)V$, is given by:

$$\pi_m=(1-\gamma)\left[V-(1-\lambda)V\right]+\gamma\left[(1-\alpha)V-(1-\lambda)V\right]\geq 0\,, \tag{1}$$

consisting of his expected profit of $(1-\gamma)\lambda V$ obtained with probability $(1-\gamma)$ from uninformed liquidity traders that must at least compensate for his expected loss of $\gamma(\lambda-\alpha)V$, should $\alpha>\lambda$, if he encounters an informed trader with probability, γ.

When the maker posts a limit order to sell at the buy price (i.e., the ‘ask’), $p_b = (1+\lambda)V$, the maker breaks even, or makes a profit, π_m , if

$$\pi_m = (1-\gamma)\big[(1+\lambda)V - V\big] + \gamma\big[(1+\lambda)V - (1+\alpha)V\big] \geq 0\,, \tag{2}$$

depending on the inequality ≥ 0 . The maker receives the buy price $p_b = (1+\lambda)V$ from both informed and uninformed buyers when hit with a market order as he is selling and replenishes his inventory from uninformed traders at the unconditional price V with probability $(1-\gamma)$ and at the high price $(1+\alpha)V$ from informed traders with probability γ, as under this eventuality the market has turned against him.

Solving either equation (1) or (2) for the maker break-even, $\pi_m = 0$, zero-profit condition that sets the upper limit to the spread with free-entry into making, then the loss to the maker due to the information content of the informed market order given by the product of information content α and the likelihood of encountering an informed order, γ , must precisely exhaust the gains from trade, λ :

$$\alpha\gamma = \lambda\,, \tag{3}$$

Fixing the likelihood of encountering an informed trader at $\gamma = \gamma_n$, namely, our initial *neutral*, i.e., no fee or rebate, equilibrium, then the degree of information content in the informed order flow is capped by the gains from trade deflated by the probability of encountering an informed trader:

$$\alpha_n = \frac{\lambda}{\gamma_n}\,, \tag{4}$$

On solving for the buy price, $p_b = (1+\lambda)V = (1+\alpha_n\gamma_n)V$, and the sell price, $p_s = (1-\lambda)V = (1-\alpha_n\gamma_n)V$, with $p_b - p_s = 2\alpha_n\gamma_n = 2\lambda$, so that the spread is widening in the expected loss due to asymmetric information given by the expected informational product $\alpha_n\gamma_n$ that cannot exceed the exogenously given degree of gains from trade, λ.

Upon arrival, the investor can place either a limit order at the inside quotes that is not guaranteed execution, a limit order away from the inside quotes that faces a greater risk of non-execution, or a market order guaranteed to execute, so long as a posted limit order exists on the side of the market opposite to him. Since the information possessed by informed traders is short-lived, informed investors will only post market orders while uninformed investors could do either, unless they are impatient in which case they will have a preference for market orders, as in Colliard and Foucault (2012).

An informed trader who places a market order to sell will receive the competitively determined sell price, $p_s = (1-\lambda)V = (1-\alpha_n\gamma_n)V$, but the stock is only worth the low price, $(1-\alpha_n)V$, known only to informed sellers. Hence the informed seller's profit, π_{in}, is:

$$\pi_{in} = \alpha_n(1-\gamma_n)V, \tag{5}$$

with an identical expression for an informed purchase. Hence, profitability is determined by the product of the degree of information content, α_n, and the proportion of uninformed traders, $1-\gamma_n$. Thus, for a given prospect of encountering an uninformed trader, $1-\gamma_n$, informed trader profitability is always increasing in the degree of information in the order flow and the proportion of the order flow that is uninformed. However, in the absence of the maker-taker rebate, information content is limited by the exogenously given level of gains from trade, λ. Of course, if the proportion of informed traders, $\gamma_{nr} \to 1, \pi_{in} \to 0$, since none would have an informational advantage, but the upper bound, $\gamma_n < 1$, applies.

B. Numerical Example

Let $V = 1$, $\alpha_n = 1/2$, $V^H = (1+\alpha_n)V = 3/2$, $V^L = (1-\alpha_n)V = 1/2$, and $1-\lambda = 7/8$, with $\lambda = 1/8$ $\gamma_n = 1/4$, $p_b = (1+\lambda)V = (1+\alpha_n\gamma_n)V = 1+1/8$, $p_s = (1-\lambda)V = (1-\alpha_n\gamma_n)V = 7/8$, and gains from trade, λV, of for each uninformed investor making $2\lambda V = 1/4$, taking into account both sides of the market. Hence, the 'inside' spread $p_b - p_s = 1/4$ with a half-spread of $1/8$ and informed trader profit, $\pi_{in} = 3/8$. Keeping the initial informed trader probability, $\gamma_n = 1/4$, and increasing the insider's informational advantage by 50% ($1/4$) to $\alpha = 3/4 > \alpha_n$, then sustainability of the LOB market requires that gains from trade, λ, must rise by $1/16$ from $1/8$ to $3/16$, which generally will not be possible. The inside quotes now widen to $p_b = 1+3/16$ and $p_s = 1-3/16 = 13/16$ with $A - B = 3/8$. Here, the corresponding gain from trade has now increased to $2\lambda = 2\times 3/16 = 3/8$. Consequently, to support a rise in the information in the stock price while keeping the same proportion of informed investors, uninformed investors need to gain a higher level of benefit from trading. This higher benefit level means they willingly subsidize the higher level of losses incurred by market makers when facing more informed traders.

Since the gains from trade are exogenously fixed, it would normally be impossible to raise the degree of information content by this means. What we demonstrate in this chapter is that competition in so-called 'fragmented' markets has led to the discovery of a clever means of raising the information level, or 'information content', in trades without the trading venues necessarily being aware of precisely why they are achieving better outcomes, or even knowing that a fee rebate is precisely equivalent to a rise in the gains from trade and thus a Pareto improvement.

C. Revenue-Neutral Fee Scheme with Offsetting Rebate

The inside spread must notionally contract in the size of the subsidy to non-marketable limit orders when a revenue-neutral maker-taker fee structure with a fee for make orders and rebate for take orders is introduced. This rebate is applicable only to limit orders that are not certain of immediate execution, that is, lower than the best ask or higher than the best bid. Pseudo limit orders, i.e., disguised aggressive market orders, placed at or above the best-ask or at or below the best bid, are certain to execute and thus pay the fee applicable to take orders. To the extent that the subsidy is simply passed-through as a lower cost of the spread to offset the fee, nothing alters, i.e., the cum-prices and cum-spread is unaltered, as neither the likelihood of a market order nor the information content of the market order is affected. This is the case in Colliard and Foucault (2012) in which the tax-subsidy system always washes out, or in both Foucault, Kadan, and Kandel (2013) and Chao, Yao, and Ye (2017a) in which the fee washes out if the minimum tick size is zero.

For simplicity, the cost to the exchange of matching buyers and sellers is set to zero. Nevertheless, there is an exchange matching fee, given by f_t, on takers in the maker-taker market with a precisely offsetting rebate, $-f_t$, paid to makers representing limit order providers. This fee structure raises zero net revenue. In the inverted taker-maker market, the fee applied to makers is f_m and the rebate applied to takers is $-f_m$. In maker-taker markets, the take fee is applied to all 'take' trades, i.e., market orders, and marketable 'make' trades, i.e., limit orders at or outside the inside quotes, such that all buyers and sellers regardless of their information status pay fees on all trades certain to execute. In inverted markets, the reverse is true and in traditional neutral markets both the fee and rebate are zero.

Commencing with the initial neutral, no-rebate, regime and initially (counterfactually) assuming no alteration to the degree of information content, when the liquidity supplier (maker) places a limit order to buy at the sell price he receives the maker transaction rebate on a limit order not certain to execute of $-f_t$, which is payable regardless of the identity of the market order placer, uninformed or informed. The competitive zero expected profit condition, equation (1), now becomes:

$$\pi_m = (1-\gamma_n)[V + f_t - p_s] + \gamma_n\left[(1-\alpha_n)V + f_t - p_s\right] = 0, \tag{6}$$

so that the sell price in the *maker-taker* venue appears to be increased by the rebate to limit orders:

$$p_s^{mt} = (1-\alpha_n\gamma_n)V + f_t, \tag{7}$$

and, if the maker places a limit order to sell at the ask, the ask price appears to be reduced by the rebate:

$$p_b^{mt} = (1+\alpha_n\gamma_n)V - f_t, \tag{8}$$

Hence the inside quotes appear to narrow to $p_b^{mt} - p_s^{mt} = 2[\alpha_n\gamma_n V - f_t]$ under the maker-taker regime for a given initial information content, $\alpha_n\gamma_n$, while additional limit orders would appear to queue up behind the best bid and ask price, adding to the depth and reducing the fill rate, while seeking the maker rebate.

This gives rise to our first proposition, *Proposition 1: If the degree of informational advantage, α_n, is zero in the neutral regime, then $p_b^{mt} = V + f_t$ and $p_s^{mt} = V - f_t$ in equations (7) and (8) respectively, and hence a peculiar raw negative spread of $-2f_t$ is generated. However, since takers' pay a fee of f_t on one side of the market, the maker-taker market clears with fees and rebates netting out with the cum-spread remaining at zero.*

This is precisely as in the models of Colliard and Foucault (2012), Foucault, Kadan, and Kandel (2013) and Chao, Yao, and Ye (2017a) in which there is also no informational advantage. Thus, these authors are correct given their assumption that informed traders are absent.

Likewise, in the inverted taker-maker market, *Proposition 2* states: *If the degree of informational advantage, α_n, is zero in the neutral regime, then $p_b^{tm} = V - f_m$ and $p_s^{tm} = V + f_m$ in equations (7) and (8) respectively, and hence a raw spread of $2f_m$ is generated which is precisely matched by the subsidy to market orders such that the cum-fee spread is effectively zero once again.*

If there exists both informed traders and maker-taker fee competition between venues, as was the case during the 2015 Nasdaq fee experiment, the maker-taker regime is different. For example, when an individual venue such as Nasdaq with a pre-existing maker-taker fee unilaterally removes or raises the fee, existing informed traders can either flee to the next highest fee-rebate venue or shift to Nasdaq from a lower rebate venue, respectively. Hence, it will no longer be the case that the information content remains unaltered. We model this inter-venue rebate competition by endogenizing the informed trader's choice of his degree of informed trading information content, α_t, under the new maker-taker fee regime, commencing from the initial competitive equilibrium for this venue in the absence of a fee rebate in which the degree of taker order flow information content, given by the product of information content and the proportion of informed trades, is initially at its neutral equilibrium level in the absence of a fee rebate, $\alpha_n \gamma_n$. Recall from the discussion above that the degree of information content for a particular trade always remains exogenous but now the venue where that trade is executed is determined endogenously.

An informed buyer could use a market order to buy at the buy-price market clearing condition, $p_b = (1+\alpha_n \gamma_n)V - f_t$, discounted by the rebate, and sell at the same high-price as before, $p_s = (1+\alpha_n)V - f_t$, with receipts reduced by the same fee of f_t to achieve wash-out, i.e., the cum-spread is unaltered. It will not pay him to do so, however, because by raising information content by precisely the amount of the rebate,

$$\alpha_t - \alpha_n = \Delta\alpha^{mt} = \frac{f_t}{\gamma_n V}, \tag{9}$$

the cum-fee buy-price, $p_b^{f_t}$, remains unaltered at the pre-rebate level once the maker rebate is considered:

$$p_b^{f_t} = (1+\alpha_t \gamma_n)V = (1+\alpha_n \gamma_n)V + f_t, \tag{10}$$

What we have derived here is the standard 'complete washout' in which nothing alters.

We now allow informed takers with a predetermined level of information in their potential trade to decide on their venue of choice. Hence making information content (quasi) endogenous. With the purchase price unaltered but with higher information content (but with still the same proportion of informed traders), the informed buyer can now sell at the high net price, $(1+\alpha_t)V - f_t$, which has raised the profit in the absence of the rebate, $\pi_{in} = \alpha_n (1-\gamma_n) V$ from equation (5) above, by the dollar amount of positive profit gain,

$$\Delta\pi_{in}^{mt} = \Delta\alpha(1-\gamma_n)V - f_t = f_t(1-\gamma_n)/\gamma_n - f_t = \left[(1-2\gamma_n)/\gamma_n\right] f_t > 0 \ , \qquad (11)$$

due to the information content increase, after paying the levy of f_t on his market order. Profit is thus maximized when the expected incremental cost of information content, $(\alpha_t - \alpha_n)\gamma_n V$, fully absorbs the fee rebate.

Hence, we have *Proposition 3*: *The raw effective spread and quotes must remain unaltered when the rebate is introduced with the cum-fee effective spread and cum-fee quotes rising precisely by the rebated fee amount. Nonetheless, the rebate will lower the raw realized spread and raise the raw market impact by the same amount due to greater information (higher order flow information content) leaving the raw effective spread unaltered.*

A maker-taker fee rebate induces a Pareto efficiency improvement since the rebate precisely compensates uninformed traders placing limit orders for the greater information content of taker order flow while informed traders receive enhanced trading profits and the market generally benefits from better price discovery and pricing efficiency.

Conversely, if the take fee and make rebate are replaced by a make fee, f_m, and equivalent take rebate, $-f_m$ in an inverted *taker-maker* regime,

$$\alpha_n - \alpha_m = -\Delta\alpha^{tm} = \frac{f_m}{\gamma_n V}, \qquad (12)$$

the information content of the order flow falls by the amount of the tax, f_m, on make trades (limit orders) as the LOB thins, with a lower bound of zero, i.e., no information content, in the LOB market. An inverted market worsens Pareto efficiency as uninformed traders receive no benefit while the profitability of informed trades falls by the amount of the fee on make orders:

$$\Delta\pi_{in}^{tm} = -\left[\left(1-2\gamma_n\right)/\gamma_n\right] f_m < 0 , \tag{13}$$

What the model establishes is that in competitive (i.e., fragmented) markets there can exist simultaneously potentially an infinite number of apparently similar venues ranging from inverted markets to maker-taker markets, each with its own unique fee (rebate) structure. Each venue differs according to the degree of information, α_t , in the market orders of informed traders, with taker-maker (inverted) markets having the lowest information content and the maker-taker venue with the highest maker rebate benefitting from the highest information-content trades. The highest rebate venue attracts the highest volume as the maker rebate attracts more market makers which in turn provides liquidity to the most informed traders who are attracted to the venue. This concentration of informed traders also makes pricing in this venue more efficient.

D. Maker-Taker Numerical Example

Commencing with the same numerical example as in Section 3.2.D above in the absence of a taker rebate, now introduce a rebate given by $f_t/V = 1/16$. Keeping the initial informed trader probability, $\gamma_n = 1/4$, the maker rebate increases the insider's informational advantage by 50% (0.25) from $\alpha_n = 1/2$ to $\alpha_t = 3/4$ as before, but this time without the need to raise gains from trade, λ , which remain at $1/8$. The 50% rise in information content, α_n to α_t, raises the bid price by $1/16$, and hence leaves the bid price unaltered once the rebate is considered. With the informed buyer's initial information content level maintained he could replenish his take sale at price $(1-\alpha_n)V = 1/2$ but at the new higher information content level the purchase price falls by half to only $1/4$. After paying the fee of $1/16$, informed trader profitability has improved by $\pi_{in}^{mt} - \pi_{in}^{n} = 3/16$, as can be shown by evaluating equation (13). Precisely the same profit gain is achieved if an informed trader buys at the same ask price as prior to the maker-taker fee structure after raising the degree of information content and then sells at the high price. This price has been enhanced by the rise in information content net of the make fee. The opposite process occurs in the inverted taker-maker market with the equilibrium level of information content falling relative to the neutral benchmark.

E. Minimum Tick Size Constraint

So far, a continuously variable price has been assumed but at the time of the Nasdaq experiment the SEC imposed a minimum one-cent (penny) tick size for stocks priced at one dollar or more. Hence, it is important to incorporate this constraint into our model as neglect could bias our testable implications.

Proposition 4: If fee structures are flexible, then both maker-taker and taker-maker inverted market fee structures can be utilized to remove minimum tick size constraints in the absence of informed traders, the standard assumption in the literature to date.

Proof: Denote the minimum tick size constraint by θ and set the maker-taker fee, $f_t = (1/2)\theta$, and, likewise, the taker-maker fee, $f_m = (1/2)\theta$. Thus, from *Proposition 1*, the raw (observed) maker-taker spread meets the minimum tick size constraint of $-2f_t = -\theta$.

Likewise, from *Proposition 2*, the spread in the inverted market becomes $2f_m = \theta$, which is also equal to the required minimum tick size of θ.

Since fee structures are not fully flexible, with the SEC imposed limit, $f_t, f_m \leq 0.3$ of one cent per share under current rules, fee arrangements can neutralize a minimum tick size of 2×0.3 = 0.6 of one cent, but not fully the standard minimum tick size of a cent or the five cent minimum tick size currently in force for 1,200 small stocks subject to the SEC's current (2016-2018) minimum tick size experiment. Of course, this finding puts to one side the role of asymmetric information and informed trades in generating positive cum-fee spreads.[13]

Since spreads generally exist even in a neutral world in the absence of fee structures due to asymmetric information and the presence of informed traders, the introduction of informed traders lessens any concern over minimum tick size constraints since the cum-fee spread will now exceed zero. Moreover, the cum-fee spread will be higher in maker-taker venues and lower in inverted venues than in neutral venues due to altered information content levels. In none of our empirical tests of the fee experiment do we find that the minimum tick size of one cent for stocks priced at a dollar or more played any role.

This thesis extends the current model to incorporate changes to the minimum tick and to explain why lit markets in general, and maker-taker venues in particular, are being substantially harmed during the SEC's ongoing tick size pilot.

[13] Even if fee structures were sufficiently flexible, it is not necessarily the case that venues would simply use fee structures to neutralize minimum tick size constraints in the absence of informed trades. For example, Chao, Yao, and Ye (2017a) demonstrate that, under these conditions, if liquidity makers are drawn from a distribution such that they have heterogeneous preferences, then exchange fees could be used to break-up what would otherwise be a single market if the minimum tick constraint was neutralized into two or more segments.

3.3 Nasdaq Fee Pilot

Nasdaq has adopted a maker-taker fee model that charges a take fee for removing liquidity by submitting marketable orders, and provides a make rebate for adding liquidity by submitting non-marketable orders (i.e., limit orders that cannot be executed immediately). On February 2, 2015, Nasdaq implemented a maker-taker fee pilot for 14 traded stocks on Nasdaq where the take fee was lowered to 5 cents per 100 shares (CPS) from 30 CPS to remove displayed liquidity; the make rebate for adding displayed liquidity was lowered to 4 CPS from an indicative 29 CPS[14]. Seven of these stocks were listed on the NYSE and seven on Nasdaq.

Essentially, the participating stocks were chosen non-randomly by Nasdaq to improve the quality of the experiment from its perspective. To improve the effectiveness of the pilot, stocks had to be very liquid with a high volume traded, especially in off-exchange and dark pools, as Nasdaq was primarily concerned about whether fees and rebates discouraged its lit market. The high liquidity of the chosen stocks turns out to be beneficial in terms of observing significant changes to the trading pattern over the course of the experiment. It is not inconceivable, but seems unlikely to us, that Nasdaq's process induced some sort of bias to our results. For example, since we are largely interested in the quotes and off-exchange venues do not provide quotes, we chose not to match on exposure to off-market trading. For simplicity, the indicative rebate for the PilotOff period is summarized in Table 3.1. The pilot ended on May 31, 2015, when the fee reverted to its pre-pilot level.

[14] The maker rebate scheme is much more complicated, available at http://nasdaqtrader.com/Trader.aspx?id=PriceListTrading2. Further, see ITG Takeaways from the NASDAQ Pilot Program report, available at http://www.itg.com/marketing/ITG_Pearson_WP_20150602.pdf.

Table 3.1: Nasdaq Maker-Taker Fee Structure

This table reports the Nasdaq pricing measured in cents per 100 shares (CPS) traded during (PilotOn) and pre/post (PilotOff) the Nasdaq access fee pilot implemented on February 2, 2015 and ended on May 31, 2015. The sample period is from December 1, 2014 to July 31, 2015. Net fee is defined as the sum of the take fee and the make rebate, which is the exchange revenue per 100 shares traded. The take fee is the highest rate Nasdaq can charge and make rebate is the most indicative rate Nasdaq provides in the Nasdaq pricing table.

Fee/ Rebate	**PilotOn (C*PS*)**	**PilotOff (C*PS*)**	**Difference**
Take Fee to Remove Liquidity:	5	30	-25
Make Rebate to Add Liquidity:			
Displayed Liquidity	4	29	-25
Non-Displayed Midpoint	2	25	-23
Other Non-Displayed Liquidity	0	10	-10
Net Fee:			
Displayed Liquidity	1	1	0
Non-Displayed Midpoint	3	5	-2
Other Non-Displayed Liquidity	5	20	-15

Table 3.2 identifies the following 11 lit exchanges and its indicative fee structure during our sample period. Among those lit exchanges, there are eight maker-taker markets and three taker-maker markets. Those markets in aggregate account for approximately 65% of the market share in U.S. equities[15]: BATS Exchange (BATS), BATS Y Exchange (BATS Y), Chicago Stock Exchange (CHX), EDGA Exchange (EDGA), EDGX Exchange (EDGX), NASDAQ BX (BX), NASDAQ PHLX (PSX), Nasdaq Stock Market (NASDAQ), New York Stock Exchange (NYSE), NYSE MKT (AMEX), and NYSE Arca (ARCA). Among these 11 lit exchanges, NASDAQ, BX, and PSX are within the NASDAQ group (18.2% market share); NYSE, ARCA, and AMEX are within the Intercontinental Exchange group (24.6%); and BATS, BATS Y, EDGA, and EDGX are within BATS Global Markets (21.6%) and Chicago Stock Exchange (0.6% of total market share)[16]. At the exchange level,

[15] The CBOE Stock Exchange and National Stock Exchange, Inc. ceased market operations on April 30, 2014 and May 30, 2014, respectively.

[16] Nasdaq Trader website in December 2015: http://nasdaqtrader.com/trader.aspx?id=FullVolumeSummary.

NASDAQ has the largest market share (approximately 15%) followed by NYSE and ARCA. NASDAQ follows the price, display type, and time execution priority model.

Table 3.2: Indicative U.S. Exchange Fee Structure

This table reports the U.S. stock exchanges pricing, measured in cents per 100 shares (CPS) traded. Fee is the exchange charge (pay) in the maker-taker (taker-maker) market. Rebate is the exchange pay (charge) in the maker-taker (taker-maker) market. Net fee is defined as the sum of the take fee and the maker rebate, which is the exchange revenue per 100 shares traded. For simplicity, the fee is the highest rate the exchange can charge and rebate is the highest rate below the fee given its pre-determined fee structure in their pricing table.

Exchange	Fee Model	Fee (CPS)	Rebate (CPS)	Net Fee (CPS)
NASDAQ	Maker-Taker	30	-29	1
ARCA	Maker-Taker	30	-29	1
BATS Z	Maker-Taker	30	-29	1
NYSE	Maker-Taker	27.5	-26	1.5
AMEX	Maker-Taker	30	-25	5
PHLX	Maker-Taker	29	-23	6
EDGX	Maker-Taker	29	-20	9
CHX	Maker-Taker	30	-20	10
EDGA	Taker-Maker	-2	5	3
BATS Y	Taker-Maker	-15	18	3
BX	Taker-Maker	-17	19	2

Off-exchange trades for NMS stocks, which account for approximately 35% of total volume, must be reported to FINRA trade reporting facility (TRF) by members for which the Participant ID (Pid) is D in centralized SIP data. FINRA has established two TRFs in conjunction with NASDAQ and NYSE[17]. U.S. equities trade execution venues can be classified into three main categories: exchanges, dark pools, non-dark pools off-exchange venues, which consist of ECNs[18], voice-brokered trades, and broker-dealer internalization. Dark pools and ECNs are also called ATSs. The primary difference between an ATS (typically operated by broker-dealers) and an exchange is that the former includes less regulatory scrutiny, fewer reporting requirements, and restricted access.

[17] FINRA Trade Reporting FAQ: http://www.finra.org/industry/trade-reporting-faq.

[18] ECNs are prohibited from listing stocks and are not self-regulating organizations. The only remaining ECN in the United States, LavaFlow owned by Citi, ceased market operations on January 30, 2015.

3.4 Data, Sample Selection, and Methodology

A. Data source

Our analysis is based on trader-level data and U.S. SIP data. Our HFT data are identified by Nasdaq based on the method described in Brogaard, Hendershott, and Riordan (2014). We analyse the effect of Nasdaq's fee pilot for an eight-month window (December 1, 2014 to July 31, 2015), which is two months before to two months after the introduction of the maker-taker fee pilot, which ran from February 2, 2015 to May 31, 2015. We exclude the half-day trading on Christmas Eve.

Our data include all information on order submission and trades, including price, volume, and a unique identifier for the trader that submitted the order, which allow us to construct the HFT data. We restrict our attention to transactions that occur in the LOB and trades during regular trading hours. For each LOB transaction, the data contain identifiers for buyer- or seller-initiated trade, adding or removing liquidity, and types of liquidity (such as displayed, non-displayed midpoint, and non-displayed non-midpoint). Furthermore, our ATS data are provided via www.FINRA.org/ATS, copyrighted by FINRA 2015. ATS data is reported weekly. Our 38 dark pools list is taken from the SEC report on the "Regulation of NMS Stock Alternative Trading Systems."

B. Sample selection and Methodology

Nasdaq introduced a maker-taker fee reduction for 14 stocks that they split equally between their own listings and those of the NYSE (seven listed on Nasdaq—AAL, MU, FEYE, GPRO, GPRN, SIRI, and ZNGA—and seven listed on the NYSE—BAC, GE, KMI, RAD, RIG, S, and TWTR). We use the remaining companies on each corresponding listing exchange to find a one-to-one control group without replacement to ensure that our results are not driven by market-wide, exchange-wide, or industry-wide fluctuations. Moreover, I exclude securities that had stock splits, switched listing exchange, or had days with a stock price below \$1. In the sample, the minimum trade count per stock per day is 450.

Our control sample matches listing exchange, closing price, market capitalization, and average daily trading volume (ADV) based on one month prior to our sample period data. Davies and Kim (2009) argue that one-to-one matching without replacement based on closing price and market capitalization is the most appropriate method to test for differences in trade execution costs. O'Hara and Ye (2011) followed their approach and matched on closing price, market capitalization and listing exchange. We add ADV as a matching criterion since this study focuses not only on trade execution costs but also on trader behaviour.

In addition, we randomize the matching order by sorting the stocks in the treatment group alphabetically by ticker symbol. The match for each treatment group security *i* is then defined to be the control group security *j* that minimizes the following matching error:

$$matcherror_{i,j} = \left|\frac{CP_i - CP_j}{CP_i + CP_j}\right| + \left|\frac{MC_i - MC_j}{MC_i + MC_j}\right| + \left|\frac{ADV_i - ADV_j}{ADV_i + ADV_j}\right|, \qquad (14)$$

where CP, MC, and ADV denote the security's closing price, market capitalization as of the end of November 2014, and average November 2014 ADV on its corresponding listing exchange (i.e., Nasdaq or NYSE), respectively. Our panel regression analysis employs a DiD approach to account for market-wide fluctuations. The estimation is based on the following DiD regression specification:

$$Y_{i,t} = \alpha_0 + \alpha_1 Treat_i + \alpha_2 PilotOn_t + \beta_1 Treat_i * PilotOn_t + \gamma VIX_t + \varphi X_i + \epsilon_{it}, \qquad (15)$$

where Y_{it} is the dependent variable; α_0 is the intercept; $Treat_i$ is the dummy variable if security i is a pilot stock; PilotOn is the dummy variable that is one if date t is during the pilot period and zero otherwise; VIX_t is the closing value of CBOE's volatility index for day t; and X_i is the vector of security-level control variables including the log of the average closing price during the sample period, the log of the average market capitalization during the sample period, and average volatility, measured by daily high price minus daily low price over closing price, during the sample period. We estimate the specification with and without stock-fixed effects.

Table 3.3 reports the summary statistics across U.S. equity trading venues for our sample of 28 stocks between December 1, 2014 and July 31, 2015. The sample period is divided into two parts: (i) the pilot period between February 2, 2015 and May 31, 2015 and (ii) the two months before and after the pilot from December 1, 2014 to February 1, 2015 and June 1, 2015 to July 31, 2015.

Table 3.3: Summary Statistics

The table reports summary statistics for the selected variables for Nasdaq fee pilot stocks and its matching stocks. Nasdaq access fee pilot is implemented on February 2, 2015 and ended on May 31, 2015. The sample period is from December 1, 2014 to July 31, 2015. PilotOn is the period when Nasdaq implemented the fee pilot; and PilotOff refers to two months prior and post the fee pilot.

	Treatment				**Control**				**Diff in Diffs**	
	PilotOn		**PilotOff**		**PilotOn**		**PilotOff**			
Number of Daily Obs.	1148		1190		1148		1190			
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Diff	t-stats
Panel A: Trading Stats and Market Share										
Closing Price	24.21	17.28	23.78	17.93	31.08	18.66	30.11	18.70	-0.54	-0.44
Market Cap (millions)	47.63	73.32	46.78	74.40	43.56	63.50	42.84	62.93	0.13	0.03
Nasdaq Volume (millions)	2.27	2.29	2.62	2.11	1.71	1.36	1.86	1.56	-0.19	-1.77
Nasdaq MarketShare (%)	12.77	4.72	14.72	5.77	18.73	6.07	19.22	6.69	-1.46	-4.25
Panel B: Quote Quality and Fill Ratio										
Nasdaq Time at NBBO (%)	88.25	13.19	94.00	9.30	93.02	10.64	93.80	9.80	-4.98	-6.86
Nasdaq Depth at NBBO (millions)	1.69	3.51	2.33	5.26	0.41	0.47	0.41	0.46	-0.64	-3.67
Nasdaq Depth Share	15.68	5.88	19.31	7.29	21.64	6.58	21.25	7.58	-4.02	-10.15
Fill Rate (%)	3.91	1.46	3.65	1.48	2.99	1.27	2.96	1.14	0.24	3.02
Fill Time (seconds)	189	197	236	265	134	119	123	127	-58	-5.59
Panel C: Transaction Cost_cents										
Quoted Spread (cents)	1.35	0.83	1.59	1.28	1.63	1.86	2.15	3.23	0.28	2.33
Effective Spread (cents)	1.14	0.58	1.18	0.70	1.30	1.13	1.40	1.34	0.06	1.02
Realised Spread 1s (cents)	-0.12	0.48	-0.38	0.35	-0.39	0.51	-0.36	0.68	0.29	9.97
Realised Spread 5s (cents)	-0.30	0.59	-0.62	0.50	-0.56	0.66	-0.51	0.68	0.37	10.18
Price Impact 1s (cents)	1.26	0.74	1.56	0.70	1.69	1.09	1.77	1.13	-0.22	-3.95
Price Impact 5s (cents)	1.44	0.89	1.81	1.08	1.86	1.45	1.92	1.41	-0.31	-3.95
Panel D: HFT vs Non-HFT Trading Behaviour										
HFT Volume (millions)	1.69	1.71	1.96	1.59	1.26	0.98	1.40	1.11	-0.13	-1.59
Non-HFT Volume (millions)	2.86	2.96	3.29	2.74	2.15	1.81	2.33	2.11	-0.25	-1.78
HFT Adding Liquidity (%)	22.91	7.67	33.59	10.15	33.15	13.12	33.80	12.18	-10.04	-13.60
HFT Taking Liquidity (%)	50.68	9.88	41.05	8.80	41.18	9.83	41.48	10.09	9.94	17.20
Non-HFT Adding Liquidity (%)	77.09	7.67	66.41	10.15	66.85	13.12	66.20	12.18	10.04	13.60
Non-HFT Taking Liquidity (%)	49.32	9.88	58.95	8.80	58.82	9.83	58.52	10.09	-9.94	-17.20

The Nasdaq market share dropped by 1.45% on average during the pilot period, while the average price and market capitalization remained relatively stable. The Nasdaq depth at the NBBO and depth share declined, while the fill ratio and speed of fill improved due to the reduced queue of limit orders following the removal of the rebate. None of these changes should have occurred if fees and rebates wash out. The raw effective spread increased, while the realized spread decreased, and the price impact decreased. By using Nasdaq data, we found that HFT changed trading behaviour from adding liquidity (dropped by 10.04%) to taking liquidity (increased by 9.94%) after the reduction of the make rebate and take fee. The decline in HFT adding liquidity mainly came from the displayed liquidity type. Non-HFT traders moved in the opposite direction to HFT.

3.5 Empirical Results

A. Market quality and the information content of trades

A1. How does the Nasdaq percentage of time at the NBBO, market depth, NBBO quoted spread, fill rate and speed change?

Since these five measures affecting the NBBO are at the national market level, they are not expected to be precisely the same as our model predictions or what we find when we examine just Nasdaq alone. If our model was precisely applicable at the national level, there should be no change in either the quoted or effective spread since the predicted decrease in information content should fully offset the rebate removal. With the shift of informed trades to the remaining high-rebate venues, limit order providers must also desert Nasdaq to chase these more informative orders, leaving the limit order book much thinner and lower-rated in the NBBO. By contrast, the extant theoretical and empirical literature predicts that the quoted and effective spread will rise by the full amount of the rebate reduction to achieve a wash-out, with no alteration to the depth or the fill rate in the LOB as the rise in the spread should match the fall in the fee on market (take) orders. Note that the quoted spread is the difference between the ask price and bid price, measured in cents. Specifically,

$$qspread_{i,t} = a_{i,t} - b_{i,t}, \tag{16}$$

Table 3.4 shows that the Nasdaq percentage time and depth at the NBBO declined. Our model predicts that this decline was due to the flight of both limit and (relatively informed) market orders to the remaining high-rebate venues. The drop in the depth share at the NBBO (coefficient of -4.010*** in Table 3.4) is much larger compared with the relatively slight decrease in market share (coefficient of -1.451*** in Table 3.7), as makers joined the more informed takers on competing venues. The NBBO quoted spread increased (coefficient of 0.285*** in Table 3.4), but by far less than the drop in the rebate of 50 cents (0.500) for 100 round-trip trades required for wash-out. Our model predicts this mitigated quoted spread increase due to the reduced information content of take order flow. Furthermore, when the take fee was reduced, the fill rate increased and time to fill decreased (i.e., speed of fill increased) due to the thinness of Nasdaq's limit order book with orders now clustered at the best bid and ask, matching the lower information content of orders. Hence, after the make rebate was reduced, the queue of Nasdaq orders in the LOB declined, contrary to the extant literature predicting a wash-out.

Table 3.4: Quote Quality, Fill Ratio and Speed

The table tests the impact of trading fee change on the percentage of time and market depth when Nasdaq at national best bid or offer (NBBO), as well as the fill rate and speed. Time% NQ at NBBO is the average percentage time Nasdaq bid at NBB and Nasdaq ask at NBO. NQ Depth at NBBO is the average quote size when Nasdaq bid at NBB and Nasdaq ask at NBO. NQ Depth Share at NBBO is the average percentage of quote size when Nasdaq bid at NBB and Nasdaq ask at NBO. Fill rate is the ratio of volume of executed orders to the volume of resting orders. Quoted spread_raw is calculated as the time weighted difference between the bid and ask using NBBO quote data. Fill time represents the average time (measured in seconds) that it takes executed to receive their first (and perhaps only) execution. All the variables are measured on a daily basis per security. The control variables (Price, Mkt Cap and Volatility) are computed as the average value over the sample period. Nasdaq access fee pilot is implemented on February 2, 2015 and ended on May 31, 2015. The sample period is from December 1, 2014 to July 31, 2015. Standard errors are in parentheses and are robust to both heteroscedasticity and correlation within stocks. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**Time% NQ at NBBO**		**NQ Depth at NBBO**		**NQ Depth Share at NBBO**		**Quoted Spread_raw**		**Fill Rate**		**Log (Fill Time)**	
Stock Fixed Effect	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Intercept	70.375***	63.922***	10.392***	13.492***	107.689***	75.317***	8.322***	26.152***	7.352***	9.663***	8.469***	11.026***
	(2.907)	(3.987)	(0.226)	(0.267)	(2.929)	(3.689)	(0.794)	(1.213)	(0.603)	(1.056)	(0.239)	(0.360)
Treat	-3.064***	-2.870***	0.194***	-0.281***	-1.034***	1.944***	-0.148**	-1.277***	0.843***	0.415***	0.203***	-0.180***
	(0.261)	(0.357)	(0.020)	(0.024)	(0.263)	(0.330)	(0.071)	(0.109)	(0.054)	(0.094)	(0.022)	(0.032)
PilotOn	-0.962***	-0.962***	-0.047**	-0.047***	0.262	0.262	-0.407***	-0.407***	-0.002	-0.002	0.038*	0.038**
	(0.265)	(0.180)	(0.021)	(0.012)	(0.267)	(0.167)	(0.072)	(0.055)	(0.055)	(0.048)	(0.022)	(0.016)
Treat*PilotOn	-4.977***	-4.977***	-0.337***	-0.337***	-4.010***	-4.010***	0.285***	0.285***	0.237***	0.237***	-0.261***	-0.261***
	(0.365)	(0.248)	(0.028)	(0.017)	(0.368)	(0.229)	(0.100)	(0.075)	(0.076)	(0.066)	(0.030)	(0.022)
VIX	-0.115***	-0.115***	-0.032***	-0.032***	-0.077**	-0.077***	0.072***	0.072***	-0.015**	-0.015**	-0.063***	-0.063***
	(0.038)	(0.026)	(0.003)	(0.002)	(0.038)	(0.024)	(0.010)	(0.008)	(0.008)	(0.007)	(0.003)	(0.002)
Log (Price)	-9.495***	-10.259***	-1.729***	-1.628***	0.206*	-0.399***	1.109***	1.669***	0.373***	0.273***	-0.822***	-0.712***
	(0.108)	(0.131)	(0.008)	(0.009)	(0.109)	(0.122)	(0.029)	(0.040)	(0.022)	(0.035)	(0.009)	(0.012)
Log (Mkt Cap)	2.656***	3.596***	0.372***	0.236***	-3.271***	-1.196***	-0.484***	-1.411***	-0.223***	-0.199***	-0.009	-0.137***
	(0.110)	(0.146)	(0.009)	(0.010)	(0.110)	(0.135)	(0.030)	(0.044)	(0.023)	(0.039)	(0.009)	(0.013)
Volatility	-249.743***	-661.217***	-31.073***	-27.563***	-311.181***	-701.937***	19.009***	102.189***	-2.274	-59.183***	-7.885***	1.084
	(13.732)	(23.146)	(1.068)	(1.548)	(13.836)	(21.416)	(3.749)	(7.041)	(2.849)	(6.127)	(1.130)	(2.092)
Adjusted R^2	0.682	0.853	0.911	0.97	0.251	0.708	0.307	0.602	0.152	0.362	0.693	0.829
Observations	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676

A2. How do the effective spread, realized spread, and price impact change?

Since the Nasdaq fee pilot removes the make fee rebate, the effective spread should rise by the fall in the rebate but, as shown in our theoretical model (*Proposition 3*), this should be precisely offset by a fall in market impact costs as the order flow becomes less informative. Hence, there should be no significant change in the raw effective spread. However, the raw realized spread should rise due to the removal of the rebate to make orders and market impacts fall by the same amount due to the lowered information content of the taker order flow. On a cum-fee basis the effective spread should fall to the extent of the rebate reduction with both the realized spread and market impact declining by a significantly smaller amount.

The effective spread is twice the signed difference between the transaction price ($p_{i,t}$) and the midpoint of the bid and offer quotes ($m_{i,t}$) at the time of the transaction. $q_{i,t}$ is a dummy variable which is 1 for buyer initiated trade and 0 for seller initiated trade. Specifically,

$$espread_{i,t} = 2 * q_{i,t} * \left(p_{i,t} - m_{i,t} \right), \tag{17}$$

The realized spread is a measure of the profit to market makers, based on a given horizon for them to adjust their inventory. Previous studies have set this time horizon difference, τ, to five minutes after the trade. The choice of this time horizon should be sufficiently long to incorporate the permanent impact of the trade and thus to ensure that quotes are subsequently stabilized, temporary effects are dissipated, and there is a sufficiently long period for liquidity providers to close their positions (Conrad, Wahal, and Xiang, 2015). In today's ultra-high frequency trading environment, which has upgraded trading systems with an accuracy of mere nanoseconds, five minutes is excessively long. Like Conrad, Wahal, and Xiang (2015), I estimate realized spreads from one second to five seconds after each trade. The realized spread is then calculated as twice the signed difference between the transaction price and the midpoint of the bid and offer quotes one second and five seconds after the transaction. Specifically,

$$rspread_{i,xt} = 2 * q_{i,t} * (p_{i,t} - m_{i,t+\tau}), \tag{18}$$

Price impact is defined as the signed change between the midpoint of the quote one second and five seconds after the trade and the midpoint of the prevailing quote at the time of the trade. It captures the information that is revealed by the trade. A decline in the price impact indicates a decline in adverse selection costs. Specifically,

$$priceimpact_{i,t} = 2 * q_{i,t} * \left(m_{i,t+\tau} - m_{i,t}\right), \tag{19}$$

To test whether the fee change truly matters, we also compute the cum-fee, cum-fee effective spread, and cum-rebate realized spread as follows:

$$cum\ fee\,espread_{i,t} = 2 * q_{i,t} * \left(p_{i,t} - m_{i,t}\right) + 2 * f_{i,t}, \tag{20}$$

$$cum\ rebate\ rspread_{i,xt} = 2 * q_{i,t} * \left(p_{i,t} - m_{i,t+\tau}\right) + 2 * r_{i,t}, \tag{21}$$

where $f_{i,t}$ is the taker fee for security *i* at time t, $r_{i,t}$ is the maker rebate for security *i* at time t, $a_{i,t}$ is the ask price of the quote, where $b_{i,t}$ is the bid price of the quote, $p_{i,t}$ is the trade price for security i at time t, $m_{i,t}$ is the midpoint of the prevailing quote at the time of the trade, $m_{i,t+\tau}$ is the midpoint of the quotes at one-second and five-second intervals after the trade, and $q_{i,t}$ is an indicator variable that equals one if the trade is buyer-initiated and minus one if the trade is seller-initiated. Our data report the prevailing quotes and contain a flag that signs each trade as buyer- or seller-initiated.

Table 3.5a shows that, as predicted by our model (*Proposition 3*), the raw effective spread, either remained the same or rose slightly based on stock-fixed effects. Per the extant literature, it should have risen by the full 50 cent rebate reduction per 100 round-trip trades. Table 3.5b shows the cum-fee effective spread decrease (coefficients of -0.438***) which corresponds to the rebate reduction and is as predicted by our model. Moreover, it is inconsistent with the findings of Malinova and Park (2015) and the remaining extant literature which says there should be no change in the cum-fee effective spread due to wash-out. The cum-fee effective spreads are adjusted by the take fee, while the cum-rebate realized spread is adjusted by the make rebate. Moreover, we find that the cum-fee price impact declined (-0.232***) for the one second price impact and (-0.304***) for the five second price impact after the maker-taker fee reduction. These falls in price impact are as predicted by our model and indicate the sizeable extent to which the Nasdaq taker order flow became less informative as both informed traders and makers departed to other competing venues.

Table 3.5a: Raw Bid-Ask Spreads and Price Impact_Cents

The table tests the impact of trading fee change on bid-ask spreads and price impact. Effective spread_raw is calculated as twice the volume weighted signed difference between the transaction price and the midpoint of the prevailing bid and ask quotes. Realized spread_raw is calculated as twice the volume weighted signed difference between the transaction price and the midpoint of the bid and ask quotes 1 and 5 seconds after the trade. Price impact_raw is twice the volume weighted signed difference between the quote midpoint and the quote midpoint 1 and 5 seconds after the trade. All spreads and price impact measures are measured in cents on a daily basis per security. The control variables (Price, Mkt Cap and Volatility) are computed as the average value over the sample period. Nasdaq access fee pilot is implemented on February 2, 2015 and ended on May 31, 2015. The sample period is from December 1, 2014 to July 31, 2015. Standard errors are in parentheses and are robust to both heteroscedasticity and correlation within stocks. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**Effective Spread_raw**		**Realized Spread 1s_raw**		**Realized Spread 5s_raw**		**Price Impact 1s_raw**		**Price Impact 5s_raw**	
Stock Fixed Effect	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Intercept	4.639***	16.241***	2.119***	6.235***	0.750***	2.497***	2.520***	10.006***	3.889***	13.744***
	(0.364)	(0.447)	(0.237)	(0.432)	(0.272)	(0.539)	(0.337)	(0.555)	(0.440)	(0.697)
Treat	-0.011	-0.737***	0.003	-0.227***	-0.152***	-0.175***	-0.014	-0.510***	0.141***	-0.562***
	(0.033)	(0.040)	(0.021)	(0.039)	(0.024)	(0.048)	(0.030)	(0.050)	(0.040)	(0.062)
PilotOn	-0.084**	-0.084***	-0.039*	-0.039**	-0.055**	-0.055**	-0.045	-0.045*	-0.029	-0.029
	(0.033)	(0.020)	(0.022)	(0.020)	(0.025)	(0.024)	(0.031)	(0.025)	(0.040)	(0.032)
Treat*PilotOn	0.062	0.062**	0.294***	0.294***	0.366***	0.366***	-0.232***	-0.232***	-0.304***	-0.304***
	(0.046)	(0.028)	(0.030)	(0.027)	(0.034)	(0.034)	(0.042)	(0.035)	(0.055)	(0.043)
VIX	0.012**	0.012***	-0.006*	-0.006**	-0.006*	-0.006*	0.017***	0.017***	0.018***	0.018***
	(0.005)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.006)	(0.004)
Log (Price)	0.597***	0.989***	0.0001	0.175***	-0.182***	-0.096***	0.597***	0.813***	0.780***	1.084***
	(0.013)	(0.015)	(0.009)	(0.014)	(0.010)	(0.018)	(0.012)	(0.018)	(0.016)	(0.023)
Log (Mkt Cap)	-0.246***	-0.832***	-0.092***	-0.297***	-0.009	-0.090***	-0.154***	-0.534***	-0.237***	-0.741***
	(0.014)	(0.016)	(0.009)	(0.016)	(0.010)	(0.020)	(0.013)	(0.020)	(0.017)	(0.026)
Volatility	16.565***	57.462***	-8.139***	3.055	-13.501***	-13.663***	24.704***	54.407***	30.066***	71.125***
	(1.718)	(2.596)	(1.122)	(2.509)	(1.283)	(3.131)	(1.592)	(3.225)	(2.081)	(4.048)
Adjusted R^2	0.385	0.772	0.068	0.242	0.128	0.155	0.429	0.619	0.423	0.645
Observations	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676

Table 3.5b: Cum Fee Bid-Ask Spreads and Price Impact_cents

The table tests the impact of trading fee change on bid-ask spreads and price impact. Effective spread_cum is calculated as twice the volume weighted signed difference between the transaction price and the midpoint of the prevailing bid and ask quotes. Realized spread_cum is calculated as twice the volume weighted signed difference between the transaction price and the midpoint of the bid and ask quotes 1 and 5 seconds after the trade. Price impact_cum is twice the volume weighted signed difference between the quote midpoint and the quote midpoint 1 and 5 seconds after the trade. All spreads and price impact measures are measured in cents on a daily basis per security. The control variables (Price, Mkt Cap and Volatility) are computed as the average value over the sample period. Nasdaq access fee pilot is implemented on February 2, 2015 and ended on May 31, 2015. The sample period is from December 1, 2014 to July 31, 2015. Standard errors are in parentheses and are robust to both heteroscedasticity and correlation within stocks. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**Effective Spread_cum**		**Realized Spread 1s_cum**		**Realized Spread 5s_cum**		**Price Impact 1s_cum**		**Price Impact 5s_cum**	
Stock Fixed Effect	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Intercept	5.239***	16.841***	2.699***	6.815***	1.330***	3.077***	2.520***	10.006***	3.889***	13.744***
	(0.364)	(0.447)	(0.237)	(0.432)	(0.272)	(0.539)	(0.337)	(0.555)	(0.440)	(0.697)
Treat	-0.011	-0.737***	0.003	-0.227***	-0.152***	-0.175***	-0.014	-0.510***	0.141***	-0.562***
	(0.033)	(0.040)	(0.021)	(0.039)	(0.024)	(0.048)	(0.030)	(0.050)	(0.040)	(0.062)
PilotOn	-0.084**	-0.084***	-0.039*	-0.039**	-0.055**	-0.055**	-0.045	-0.045*	-0.029	-0.029
	(0.033)	(0.020)	(0.022)	(0.020)	(0.025)	(0.024)	(0.031)	(0.025)	(0.040)	(0.032)
Treat*PilotOn	-0.438***	-0.438***	-0.206***	-0.206***	-0.134***	-0.134***	-0.232***	-0.232***	-0.304***	-0.304***
	(0.046)	(0.028)	(0.030)	(0.027)	(0.034)	(0.034)	(0.042)	(0.035)	(0.055)	(0.043)
VIX	0.012**	0.012***	-0.006*	-0.006**	-0.006*	-0.006*	0.017***	0.017***	0.018***	0.018***
	(0.005)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.006)	(0.004)
Log (Price)	0.597***	0.989***	0.0001	0.175***	-0.182***	-0.096***	0.597***	0.813***	0.780***	1.084***
	(0.013)	(0.015)	(0.009)	(0.014)	(0.010)	(0.018)	(0.012)	(0.018)	(0.016)	(0.023)
Log (Mkt Cap)	-0.246***	-0.832***	-0.092***	-0.297***	-0.009	-0.090***	-0.154***	-0.534***	-0.237***	-0.741***
	(0.014)	(0.016)	(0.009)	(0.016)	(0.010)	(0.020)	(0.013)	(0.020)	(0.017)	(0.026)
Volatility	16.565***	57.462***	-8.139***	3.055	-13.501***	-13.663***	24.704***	54.407***	30.066***	71.125***
	(1.718)	(2.596)	(1.122)	(2.509)	(1.283)	(3.131)	(1.592)	(3.225)	(2.081)	(4.048)
Adjusted R^2	0.428	0.788	0.061	0.236	0.123	0.15	0.429	0.619	0.423	0.645
Observations	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676

A3. How does market efficiency change?

To test the informational efficiency of prices, we use variance ratio tests of the random walk hypothesis (Lo and MacKinlay, 1988) and autocorrelation, which are identified as the two information-associated measures in Chordia, Roll, and Subrahmanyam (2008), in multiple time intervals:

$$Variance\,ratio = var\left(r_{i,t}\right) * x \,/\, var\left(r_{i,\,x*t}\right), \tag{22}$$

$$Autocorrelation = Corr(r_{i,t}, r_{i,t-1}), \tag{23}$$

where $var\left(r_{i,t}\right)$ refers to the variance of the return during the t^{th} time interval for *i*, $var(r_{i,x*t})$ refers to the variance of the return during the *x*t* time interval for *i* and $Corr(r_{i,t}, r_{i,t-1})$ refers to the autocorrelation of the midpoint return during the t^{th} time period for *i*. Specifically, when a stock's price follows a random walk, the returns variance is a linear function of the measurement frequency, i.e., $var(r_{i,x*t})$ is *k* times larger than $var\left(r_{i,t}\right)$.

When the maker-taker fee is reduced, we find that liquidity worsens (see the results presented above) and, as predicted by our model, there is less information in the order flow. Thus, we expect the informational efficiency of prices to decrease. When Nasdaq's maker-taker fee was reduced, our model predicts reduced liquidity and less information in the order flow. Thus, for both reasons, we expect the informational efficiency of prices to decrease.

Table 3.6 shows that in response to the maker-taker fee reduction, the variance ratios in Nasdaq generally increased, while first-order return autocorrelations declined when one would have expected it to increase.[19] This pattern suggests that the variance ratios rise

[19] These inconsistent patterns have previously been noted; see Tarun, Roll and Subrahmanyam (2008). They explain: "the higher variance ratios can be attributed either to an increase in mispricing or to an increase in privately informed trading that results in the incorporation of more information into prices……Regardless of

because of increased mispricing, with less private information being reflected in prices following the maker-taker fee reduction. Compared with Nasdaq, the market efficiency of the two highest rebate-paying exchanges (Arca and BATS) also declined slightly but much less than Nasdaq.

the cause, if such mispricing were driving the increase in variance ratios across time, autocorrelations should have increased along with variance ratios as the tick size decrease; but there is no evidence of this increase. In fact, there is reliable evidence that the opposite transpired for smaller firms." (p. 266).

Table 3.6: Market Efficiency Univariate Analysis

This table reports variance ratio (panel A) and autocorrelation (panel B) univariate analysis for Nasdaq fee pilot stocks and its matching stocks. Nasdaq access fee pilot is implemented on February 2, 2015 and ended on May 31, 2015. The sample period is from December 1, 2014 to July 31, 2015. PilotOn is the period when Nasdaq implemented the fee pilot; and PilotOff refers to two months prior and post the fee pilot. Variance ratio are computed in 1 to 10 second, 10 to 60 second, and 1 to 300 second time intervals; and autocorrelation are computed in 1 seconds, 10 seconds and 300 seconds. Chicago Stock Exchange and NYSE Amex are not reported in this table due to small market share which both are smaller than 0.5%.

	Treatment						**Control**											
	PilotOn			**PilotOff**			**PilotOn**			**PilotOff**			**DiD**			**t-Stats**		
Panel A: Variance Ratio																		
Exchange	1t10s	10t60s	1t300s	1t10s	10t60s	1t300s	1t10s	10t60s	1t300s	1t10s	10t60s	1t300s	1t10s	10t60s	1t300s	1t10s	10t60s	1t300s
NASDAQ	1.083	1.146	1.509	1.003	1.044	1.221	1.108	1.099	1.466	1.117	1.091	1.413	0.089	0.094	0.234	8.590	6.763	5.819
ARCA	1.125	1.169	1.585	1.087	1.109	1.425	1.175	1.153	1.701	1.195	1.167	1.732	0.058	0.075	0.190	5.687	6.204	4.214
BATS Z	1.214	1.236	1.933	1.172	1.172	1.723	1.226	1.188	1.912	1.235	1.187	1.865	0.051	0.063	0.163	3.959	4.260	2.327
NYSE	1.128	1.153	1.521	1.092	1.108	1.401	1.101	1.086	1.389	1.128	1.120	1.558	0.063	0.080	0.290	6.830	4.772	3.497
PHLX	3.375	3.098	40.301	2.647	2.358	22.506	3.453	3.214	53.811	3.189	2.904	40.701	0.465	0.431	4.684	3.493	4.478	1.436
EDGX	1.167	1.192	1.699	1.152	1.162	1.705	1.268	1.263	2.379	1.353	1.358	3.076	0.100	0.125	0.692	6.168	6.282	4.802
EDGA	1.796	1.575	5.772	1.753	1.550	5.522	1.928	1.660	6.781	2.048	1.775	8.464	0.163	0.141	1.934	2.350	3.351	2.532
BATS Y	1.790	1.551	5.604	1.786	1.543	5.934	1.942	1.663	8.752	1.895	1.643	7.722	-0.044	-0.011	-1.361	-0.691	-0.260	-1.519
BX	2.478	2.055	19.667	2.711	2.187	30.773	2.452	2.095	20.437	2.956	2.459	34.489	0.271	0.231	2.946	2.130	2.875	0.975
Consolidated	1.057	1.128	1.416	1.605	1.302	5.290	1.044	1.059	1.294	1.764	1.386	9.881	0.172	0.152	4.713	3.155	4.300	4.455
Panel B: Autocorrelation																		
Exchange	ac1s	ac10s	ac300s	ac1s	ac10s	ac300s	ac1s	ac10s	ac300s	ac1s	ac10s	ac300s	ac1s	ac10s	ac300s	ac1s	ac10s	ac300s
NASDAQ	-0.016	0.892	0.995	-0.002	0.900	0.996	-0.024	0.890	0.995	-0.028	0.889	0.996	-0.017	-0.009	-0.001	-6.538	-8.800	-6.108
ARCA	-0.029	0.888	0.995	-0.021	0.892	0.996	-0.038	0.884	0.995	-0.040	0.882	0.995	-0.011	-0.006	-0.001	-4.996	-6.147	-4.286
BATS Z	-0.045	0.879	0.994	-0.037	0.884	0.995	-0.047	0.879	0.994	-0.048	0.878	0.994	-0.009	-0.005	0.000	-3.485	-4.163	-2.204
NYSE	-0.029	0.887	0.995	-0.023	0.891	0.996	-0.029	0.890	0.996	-0.029	0.888	0.995	-0.005	-0.007	-0.001	-2.563	-7.031	-3.418
PHLX	-0.234	0.675	0.890	-0.174	0.754	0.940	-0.206	0.676	0.869	-0.193	0.703	0.904	-0.047	-0.051	-0.015	-5.589	-4.460	-1.987
EDGX	-0.041	0.884	0.995	-0.035	0.885	0.995	-0.053	0.875	0.993	-0.064	0.867	0.991	-0.017	-0.010	-0.002	-6.070	-6.329	-5.223
EDGA	-0.112	0.827	0.982	-0.110	0.831	0.984	-0.117	0.817	0.980	-0.120	0.810	0.978	-0.004	-0.010	-0.003	-0.703	-1.826	-1.505
BATS Y	-0.119	0.827	0.983	-0.128	0.824	0.981	-0.125	0.816	0.972	-0.130	0.815	0.976	0.003	0.002	0.005	0.493	0.325	1.610
BX	-0.127	0.771	0.952	-0.150	0.739	0.922	-0.120	0.774	0.950	-0.169	0.719	0.917	-0.026	-0.023	-0.002	-3.538	-2.083	-0.311
Consolidated	-0.009	0.894	0.995	-0.061	0.840	0.983	-0.010	0.896	0.996	-0.080	0.825	0.970	-0.018	-0.017	-0.013	-4.026	-3.145	-4.203

B. Trading volume and market share

B1. Does a reduced exchange access fee attract trading volume from off-exchange venues?

The Nasdaq access fee pilot aimed to test whether a lower exchange access fee can raise the market share of off-exchange venues. In this regard, the SEC (2014) filing states the following: "*Off-exchange orders do not generate quotes on public markets, do not interact with orders on public markets and consequently do not promote or contribute to price discovery to the same extent as do orders posted and executed on exchanges. Economic studies from markets spanning the world conclude that as more orders migrate away from exchanges, the price discovery process weakens, trading spreads widen, and overall investor trading costs increase… Nasdaq believes that proposed changes may improve price discovery in the select securities.*" An early study, Barclay, Hendershot, and McCormick (2003), find that ECNs can compete with lit markets due to an ability to utilize sub-penny tick sizes denied lit markets and can attract informed order flow. Kwan, Masulis, and McInish (2015) find that the U.S. minimum tick constrains some stock spreads, causing large limit order queues and that dark pools allow traders to bypass existing limit order queues with minimal price improvement. Moreover, Foley and Putnins(2014) find that the level of dark trading decreases when Canada and Australia implemented minimum price improvement. We expect no change in dark pools trading volume after the maker-taker fee reduction since their ability to circumvent the time priority of displayed limit orders is not affected and they are not required to provide price improvement.

Table 3.7: Trading Volume and Market Share across each Trading Venue

This table reports the daily average trading volume and market share in each U.S. trading venue before, during and after the Nasdaq access fee pilot implemented on February 2, 2015 and ended on May 31, 2015. During the sample period, there are 11 lit exchanges in the U.S. equity markets, and TRF captures all the off-exchange trades including Dark Pools, ECNs, voice-brokered trades and Broder/Dealer Internalization. The sample period is from December 1, 2014 to July 31, 2015.

	Treatment		Control			
Trading Venue (Pid)	**PilotOn**	**PilotOff**	**PilotOn**	**PilotOff**	**Diff in Diffs**	**t-stats**
Panel A: Market Share Per Trading Venue						
NASDAQ (Q/T)	12.77	14.72	18.73	19.22	-1.46	-4.25
ARCA (P)	9.91	8.73	10.58	10.59	1.19	6.75
BATS Z (Z)	8.05	6.82	10.10	10.02	1.15	5.54
NYSE (N)	8.81	8.31	9.50	9.82	0.81	1.55
EDGX (K)	8.71	8.09	7.06	6.79	0.35	2.13
AMEX (A)	0.32	0.26	0.28	0.22	0.01	0.16
CHX (M)	0.58	0.52	0.65	0.56	-0.02	-0.17
EDGA (J)	2.74	2.92	3.14	3.11	-0.21	-2.83
PHLX (X)	1.14	1.20	1.32	1.14	-0.24	-5.26
BX (B)	1.68	2.16	2.15	2.29	-0.34	-6.01
BATS Y (Y)	4.39	4.49	4.99	4.67	-0.41	-3.40
TRF (D)	40.89	41.78	31.51	31.57	-0.83	-2.09
Panel B: Trading Volume Per Trading Venue						
Consolidated Volume	20,349,535	21,114,357	11,012,865	11,585,448	-192,240	-0.72
NASDAQ (Q/T)	2,272,872	2,622,836	1,705,936	1,862,962	-192,937	-1.77
ARCA (P)	1,958,653	1,812,350	1,076,705	1,125,105	194,704	2.00
BATS Z (Z)	1,568,537	1,468,514	975,009	1,070,096	195,110	2.69
NYSE (N)	2,518,804	2,488,645	1,395,893	1,577,664	211,930	0.94
EDGX (K)	1,514,448	1,557,424	682,305	690,438	-34,843	-0.49
AMEX (A)	64,420	51,374	23,448	19,026	8,624	1.20
CHX (M)	126,856	125,932	124,190	103,060	-20,206	-0.73
EDGA (J)	619,019	689,642	350,232	355,551	-65,303	-1.79
PHLX (X)	250,213	297,040	151,272	138,851	-59,247	-3.77
BX (B)	386,755	524,416	239,178	270,556	-106,284	-3.91
BATS Y (Y)	946,024	1,038,663	574,679	533,566	-133,752	-2.53
TRF (D)	8,122,933	8,437,522	3,714,019	3,838,572	-190,035	-0.49

Table 3.8: Market Share (weekly) across Types of Trading Venue

The table tests the impact of trading fee change on the market share for Nasdaq, other lit exchange, dark pools and non-dark pools (other off exchange trades other than the dark pools). The NQ_MktShare_weekly is the percentage weekly total Nasdaq trading volume over the weekly total trading volume on all the U.S. markets. The OtherExch_MktShare_weekly is the percentage of weekly other 10 lit exchanges total trading volume over the weekly total trading volume on all the U.S. markets. The DarkPools_MktShare_weekly is the percentage of weekly dark pools trading volume over the weekly total trading volume on all the U.S. markets. The NonDarkPools_MktShare_weekly is the percentage of weekly non-dark pools off exchange trading volume over the weekly total trading volume on all the U.S. markets. The control variables (Price, Mkt Cap and Volatility) are computed as the average value over the sample period. Nasdaq access fee pilot is implemented on February 2, 2015 and ended on May 31, 2015. The sample period is from December 1, 2014 to July 31, 2015. Standard errors are in parentheses and are robust to both heteroscedasticity and correlation within stocks. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**NQ_ MktShare_weekly**		**OtherExch_ MktShare_weekly**		**DarkPools_ MktShare_weekly**		**NonDarkPools_ MktShare_weekly**	
Stock Fixed Effect	No	Yes	No	Yes	No	Yes	No	Yes
Intercept	136.286***	96.538***	-106.988***	-85.781***	47.669***	57.004***	23.033***	32.239***
	(5.255)	(5.353)	(6.697)	(7.109)	(3.335)	(5.158)	(5.466)	(7.296)
Treat	-1.611***	0.304	-7.646***	-8.796***	0.116	-1.533***	9.141***	10.025***
	(0.468)	(0.478)	(0.597)	(0.635)	(0.297)	(0.460)	(0.487)	(0.651)
PilotOn	-0.057	-0.057	0.374	0.374	-0.053	-0.053	-0.264	-0.264
	(0.476)	(0.241)	(0.607)	(0.321)	(0.302)	(0.233)	(0.496)	(0.329)
Treat*PilotOn	-1.451**	-1.451***	1.809**	1.809***	0.17	0.17	-0.528	-0.528
	(0.657)	(0.333)	(0.837)	(0.442)	(0.417)	(0.321)	(0.683)	(0.453)
VIX	0.304***	0.304***	0.059	0.059	-0.125***	-0.125***	-0.238***	-0.238***
	(0.074)	(0.038)	(0.094)	(0.050)	(0.047)	(0.036)	(0.077)	(0.051)
Log (Price)	3.905***	4.829***	-3.635***	-3.156***	-1.695***	-1.669***	1.425***	-0.004
	(0.194)	(0.176)	(0.247)	(0.234)	(0.123)	(0.170)	(0.202)	(0.240)
Log (Mkt Cap)	-5.203***	-3.321***	6.507***	5.043***	-0.804***	-1.399***	-0.500**	-0.323
	(0.197)	(0.196)	(0.251)	(0.26)	(0.125)	(0.189)	(0.205)	(0.267)
Volatility	-385.575***	-497.086***	454.371***	652.067***	-226.512***	-31.693	157.716***	-123.289***
	(24.709)	(31.038)	(31.491)	(41.221)	(15.681)	(29.907)	(25.705)	(42.306)
Adjusted R^2	0.543	0.883	0.478	0.855	0.306	0.590	0.458	0.761
Observations	980	980	980	980	980	980	980	980

Table 3.7 shows that after the Nasdaq fee reduction, the Nasdaq market share declined on average (see also Hatheway, 2015a, 2015b), while the consolidated trading volume remained stable. The Nasdaq share loss was captured by the remaining two highest rebate-paying stock exchanges (ARCA and BATS Z). Table 3.8 shows that when Nasdaq reduced the maker rebate and taker fee, there was no significant market share drop from the off-exchange trading venues of dark pools and other non-dark pools. Instead, we observe a redistribution of market share among lit exchanges (i.e., the lost market share shifted to those lit exchanges with the highest liquidity provision rebate).

Nasdaq alone reduced the access fee unilaterally during the pilot in an otherwise competitive trading environment. Dark pools rely on sub-tick price improvement relative to midpoint prices of the lit exchanges to attract uninformed order flow and are protected by the inability of lit exchanges to reduce the minimum tick because of the SEC's regulations. Instead, of volume migrating to the off-exchange venues, we observe a redistribution effect among lit exchanges. Even with a uniform fee reduction across the entire lit market, a movement of trading volume toward the lit market would remain unlikely as dark pools could still undercut the 1 cent tick rule.

B2. How does the lit exchange trading volume and market share change in response to the maker-taker fee reduction?

Our theoretical model predicts a decline in Nasdaq's market share during the pilot as the elimination of the rebate to limit orders destroys the Pareto-efficiency of the market by removing at least some of the profit earned by informed traders and does not improve the position of liquidity, i.e., uninformed traders. Since it eliminated its rebate unilaterally, these harmed informed traders, together with the limit order providers who have lost their counterparties, flee to the maker-taker venue with the next highest rebate relative to Nasdaq prior to the rebate. Consequently, Nasdaq's market share falls and its limit order book thins out.

The DiD regression model in Table 3.9 shows that the Nasdaq market share decreases in response to the maker-taker fee reduction, as our theory predicts. To verify whether this reduced market share is caused by a reduction in liquidity supply, we compare the Nasdaq depth at the NBBO change with the Nasdaq trading volume change. We find that the Nasdaq depth at the NBBO (coefficient of -0.337*** in Table 3.4) dropped about three times more than the Nasdaq trading volume drop (coefficient of -0.131*** in Table 3.9). This finding indicates that the reduced Nasdaq market share is also associated with the drop in the depth at the NBBO which in turn is due to the departure of limit order providers who are now redundant due to the departure of informed traders.

Since our model shows that maker-taker fee arrangements are Pareto-superior to the outcome without fee rebate, during the Nasdaq pilot, informed traders suffered trading profit reductions while uninformed traders gained nothing. Hence, many informed traders, together with associated limit order providers, withdrew from Nasdaq during the pilot, thus accounting for the decline in trading volume and market share.

Table 3.9: Trading Volume, Market Share and Routing Dynamics

The table tests the impact of trading fee change on the trading volume, market share, as well as the Nasdaq routing dynamics. Consolidated_Volume is the consolidated daily trading volume, and NQ_Volume is the Nasdaq daily trading volume. The NQ_MktShare is the percentage of daily total Nasdaq trading volume over the daily total trading volume on all the U.S. markets. Routing to NQ% is the percentage of orders routed to Nasdaq, and Routing to OtherExch% is the percentage of orders routed to other U.S. markets. The control variables (Price, Mkt Cap and Volatility) are computed as the average value over the sample period. Nasdaq access fee pilot is implemented on February 2, 2015 and ended on May 31, 2015. The sample period is from December 1, 2014 to July 31, 2015. Standard errors are in parentheses and are robust to both heteroscedasticity and correlation within stocks. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**Log (Consolidated _Volume)**		**Log (NQ_Volume)**		**NQ_MktShare**		**Routing to NQ%**		**Routing to OtherExch%**	
Stock Fixed Effect	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Intercept	0.745***	1.738***	3.723***	2.975***	98.210***	75.012***	131.561***	99.373***	-31.561***	0.627
	(0.264)	(0.422)	(0.284)	(0.449)	(1.913)	(2.817)	(1.631)	(2.471)	(1.631)	(2.471)
Treat	0.454***	0.104***	0.308***	0.152***	-2.368***	-0.583**	-1.659***	1.739***	1.659***	-1.739***
	(0.024)	(0.038)	(0.026)	(0.040)	(0.172)	(0.252)	(0.147)	(0.221)	(0.147)	(0.221)
PilotOn	0.011	0.011	0.034	0.034*	0.134	0.134	-0.637***	-0.637***	0.637***	0.637***
	(0.024)	(0.019)	(0.026)	(0.020)	(0.174)	(0.127)	(0.149)	(0.112)	(0.149)	(0.112)
Treat*PilotOn	-0.024	-0.024	-0.131***	-0.131***	-1.463***	-1.463***	-0.381*	-0.381**	0.381*	0.381**
	(0.033)	(0.026)	(0.036)	(0.028)	(0.240)	(0.175)	(0.205)	(0.154)	(0.205)	(0.154)
VIX	0.031***	0.031***	0.058***	0.058***	0.388***	0.388***	0.02	0.02	-0.02	-0.02
	(0.003)	(0.003)	(0.004)	(0.003)	(0.025)	(0.018)	(0.021)	(0.016)	(0.021)	(0.016)
Log (Price)	-0.466***	-0.674***	-0.227***	-0.404***	3.710***	4.460***	-0.214***	-0.877***	0.214***	0.877***
	(0.010)	(0.014)	(0.011)	(0.015)	(0.071)	(0.093)	(0.060)	(0.081)	(0.060)	(0.081)
Log (Mkt Cap)	0.660***	0.687***	0.422***	0.527***	-3.824***	-2.803***	-1.318***	0.178**	1.318***	-0.178**
	(0.01)	(0.015)	(0.011)	(0.016)	(0.072)	(0.103)	(0.062)	(0.091)	(0.062)	(0.091)
Volatility	20.708***	-2.343	8.572***	-18.238***	-243.160***	-284.998***	-90.175***	-119.650***	90.175***	119.650***
	(1.247)	(2.451)	(1.342)	(2.605)	(9.038)	(16.350)	(7.705)	(14.343)	(7.705)	(14.343)
Adjusted R^2	0.679	0.798	0.453	0.665	0.595	0.785	0.195	0.546	0.195	0.546
Observations	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676

In contrast, Cardella, Hao, and Kalcheva (2015) find that reductions in relative taker fees in U.S. equity markets are associated with increased market share, which is inconsistent with our finding and our theory. However, Malinova and Park (2015) argue that the change in the maker-taker fees in Cardella, Hao, and Kalcheva (2015) are accompanied by changes in the total access fee. Thus, the Cardella, Hao, and Kalcheva result is suspect and not congruent with the Nasdaq pilot fee experiment which held the total access fee relatively constant and instead varied the fee components.

B3. How do the Nasdaq routing dynamics change?

Reg NMS Rule 611[20], also known as the 'Order Protection Rule' or 'Trade-through' rule, restricts trading, either as agent or principal, on one venue at prices inferior to the displayed quotations on another trading venue during regular trading hours (9:30 am to 4:00 pm ET). Thus, if one exchange does not have the NBBO when it receives incoming orders and if it is eligible for immediate electronic execution, it is obligated to route the incoming order to another trading venue with the NBBO. After the Nasdaq rebate is reduced, the incentive to submit non-marketable orders on Nasdaq decreases, but only because of the flight of informed orders to competing venues. Without the possibility of being hit by a market order, there is no longer any point in adding depth to the Nasdaq market. Thus, the Nasdaq percentage of time at the NBBO is likely to decline and outbound routing is likely to increase.

[20] "Rule 611 does not affirmatively require the routing of orders to trading venues displaying the best prices. Rather, it only restricts trades at prices worse than a protected quotation. Any trading venue is free to execute trades at prices equal to or better than a protected quotation, regardless of whether such a trading center is currently quoting at that price or is a dark venue that never displays quotations." (SEC, 2015b).

Table 3.9 shows that the Nasdaq incoming orders routed to Nasdaq declined and that the proportion routed away to other exchanges increased, consistent with our model's prediction away from the NBBO. Moreover, the percentage of volume routed to off-exchange trading venues remained stable.

C. HFT vs. non-HFT trading behaviour

C1: HFT vs. non-HFT adding/taking liquidity_volume

HFT is distinguished by more efficient monitoring of the informed order flow dynamics. Any trader who sees a signal containing information before a slower trader will take liquidity before that information advantage is exploited by other traders. Nevertheless, as holding inventory is expensive, Rosu (2015) argues a risk-averse trader reverses part of his order to make money supplying liquidity to slower traders who receive a delayed signal. The fast trader could even post some of these reversed trades as limit orders. Hence, HFT traders are likely to have information and thus take liquidity but can also make liquidity. As both the rebate and the tax on liquidity makers fell during the Nasdaq fee pilot, HFT traders are likely to shift from making to taking liquidity.

Table 3.10 shows that in response to the maker-taker fee reduction, HFT traders switched trading behaviour on Nasdaq from posting liquidity (coefficient of -10.043***) to taking liquidity (coefficient of 9.936***), as predicted, while the overall Nasdaq HFT volume remained stable (coefficient of -0.054 with no statistical significance).

Table 3.10: HFT vs non-HFT Add or Take Liquidity

The table tests the impact of trading fee change on the HFT vs Non-HFT trading behaviour. HFT is identified using Nasdaq proprietary data. HFT% is the ratio of HFT trading volume over total trading volume, while non-HFT% is the ratio of non-HFT trading volume over total trading volume. HFT Add (Take)_Liquidity% is the ratio of HFT adding (taking) liquidity volume over total trading volume. Non-HFT Add (Take)_Liquidity% is the ratio of non-HFT adding (taking) liquidity volume over total trading volume. All the variables are measured on a daily basis per security. Nasdaq access fee pilot is implemented on February 2, 2015 and ended on May 31, 2015. The control variables (Price, Mkt Cap and Volatility) are computed as the average value over the sample period. The sample period is from December 1, 2014 to July 31, 2015. Standard errors are in parentheses and are robust to both heteroscedasticity and correlation within stocks. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**HFT Add_Liquidity%**		**HFT Take_Liquidity%**		**HFT Volume%**		**non-HFT Add_Liquidity%**		**non-HFT Take_Liquidity%**		**non-HFT Volume%**	
Stock Fixed Effect	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Intercept	-48.69***	-80.89***	-36.67***	-29.90***	-42.68***	-55.40***	148.69***	180.89***	136.67***	129.90***	142.68***	155.40***
	(3.287)	(5.586)	(4.172)	(7.239)	(2.675)	(4.421)	(3.287)	(5.586)	(4.172)	(7.239)	(2.675)	(4.421)
Treat	-4.757***	-3.991***	-1.350***	2.329***	-3.053***	-0.831**	4.757***	3.991***	1.350***	-2.329***	3.053***	0.831**
	(0.296)	(0.500)	(0.375)	(0.648)	(0.241)	(0.396)	(0.296)	(0.500)	(0.375)	(0.648)	(0.241)	(0.396)
PilotOn	-0.366	-0.366	0.272	0.272	-0.047	-0.047	0.366	0.366	-0.272	-0.272	0.047	0.047
	(0.300)	(0.252)	(0.380)	(0.327)	(0.244)	(0.200)	(0.300)	(0.252)	(0.380)	(0.327)	(0.244)	(0.200)
Treat*PilotOn	-10.04***	-10.04***	9.936***	9.936***	-0.054	-0.054	10.04***	10.04***	-9.936***	-9.936***	0.054	0.054
	(0.413)	(0.347)	(0.524)	(0.450)	(0.336)	(0.275)	(0.413)	(0.347)	(0.524)	(0.450)	(0.336)	(0.275)
VIX	0.171***	0.171***	0.359***	0.359***	0.265***	0.265***	-0.171***	-0.171***	-0.359***	-0.359***	-0.265***	-0.265***
	(0.042)	(0.036)	(0.054)	(0.046)	(0.035)	(0.028)	(0.042)	(0.036)	(0.054)	(0.046)	(0.035)	(0.028)
Log(Price)	-9.714***	-10.99***	0.587***	2.935***	-4.564***	-4.027***	9.714***	10.989***	-0.587***	-2.935***	4.564***	4.027***
	(0.122)	(0.184)	(0.155)	(0.239)	(0.099)	(0.146)	(0.122)	(0.184)	(0.155)	(0.239)	(0.099)	(0.146)
Log (Mkt Cap)	4.299***	6.108***	2.383***	1.320***	3.341***	3.714***	-4.299***	-6.108***	-2.383***	-1.320***	-3.341***	-3.714***
	(0.124)	(0.205)	(0.157)	(0.266)	(0.101)	(0.162)	(0.124)	(0.205)	(0.157)	(0.266)	(0.101)	(0.162)
Volatility	324.35***	101.60***	504.70***	678.40***	414.53***	390.00***	-324.35***	-101.60***	-504.70***	-678.40***	-414.53***	-390.00***
	(15.525)	(32.426)	(19.708)	(42.020)	(12.635)	(25.664)	(15.525)	(32.426)	(19.708)	(42.020)	(12.635)	(25.664)
Adjusted R^2	0.649	0.751	0.270	0.460	0.402	0.599	0.649	0.751	0.270	0.460	0.402	0.599
Observations	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676

C2: HFT vs. non-HFT adding and taking liquidity: spreads and price impact

Carrion (2013) finds that spreads are wider (tighter) when HFT traders provide (take) liquidity, which suggests HFTs provide liquidity when it is scarce and consume liquidity when it is plentiful. Because of their effective real-time forecasting of the state of the market, HFT traders face lower adverse selection costs than non-HFT traders when supplying liquidity in larger trades, suggesting they have an informational advantage when demanding liquidity and avoiding the supply of liquidity to informed traders. SEC (2010) thus questions whether “rebates generally benefit long-term investors by promoting narrower spreads and more immediately accessible liquidity?”

Focusing on the cum fee/rebate column in Table 3.11, we analyse the net effect of the participation of the two trader types, after considering the fee/rebate alteration, the effective spread for HFT is lower compared with non-HFT when placing market orders, with the spread is sizably reduced for both types because of the removal of the fee on market orders. This is because HFT market orders are less informed than non-HFT market orders as the market impact over both the one-second and the-five second intervals falls by a larger amount than that for non-HFT, such that the fall in the realized spread is less for HFT market orders.

HFT is therefore more inclined to add to liquidity relative to non-HFT when the effective spread is narrower because of its relatively less informed order flow, as shown by both the one-second and the five-second delay price impacts being relatively low. Therefore, the realized spread is higher when HFT traders provide liquidity relative to non-HFT traders. These findings imply that HFT is more likely to take liquidity when it is relatively plentiful (spreads are narrow and consistent with Carrion (2013)) and provide liquidity when it is relatively cheap (the effective spread is relatively narrow), as it uses its informational advantage to supply liquidity to less informed counterparts.

Subsidized limit orders and taxed market orders benefit relatively informed short-term traders, as seen when Nasdaq's unilateral withdrawal of the subsidy and tax during the pilot led to the movement of such traders to exchange with higher subsidy/tax regimes. Insofar as long-term investors are also informed, high subsidy/tax regimes might also benefit such investors. However, if they are not informed, then they should be no worse off under the subsidy/tax regime. This is because they still trade on the same terms as other uninformed traders whose terms remain the same in the limit order book market. Of course, if long-term traders are uninformed in the short-term and persist in using taxed market orders in preference to subsidized limit orders, then they are worse off. On the other hand, to the extent that all market participants gain from improved price discovery and pricing efficiency, long-term traders benefit along with the entire market.

Table 3.11: Coefficient of HFT vs Non-HFT Bid-Ask Spread and Price Impact (cents)

This table reports the regression coefficient of Treat*PilotOn from the model (1) in the chapter for stock fixed effect. HFT is identified using Nasdaq proprietary data. HFT Take (Add) is when HFT is the liquidity taker (maker) in the trade. Non-HFT Take (Add) is when non-HFT is the liquidity taker (maker) in the trade.

	Cents				Cum-fee/Rebate Cents			
Regression Coefficient	**HFT Take**	**HFT Add**	**Non-HFT Take**	**Non-HFT Add**	**HFT Take**	**HFT Add**	**Non-HFT Take**	**Non-HFT Add**
Effective Spread	0.041*	0.450***	0.071**	0.628***	-0.459***	-0.05	-0.429***	0.128
Realized Spread 1s	0.425***	0.111**	0.255***	-0.018	-0.075**	-0.389***	-0.245***	-0.518***
Realized Spread 5s	0.530***	0.067	0.328***	-0.131***	0.03	-0.433***	-0.172***	-0.631***
Price Impact 1s	-0.384***	0.339***	-0.184***	0.647***	-0.384***	0.339***	-0.184***	0.647***
Price Impact 5s	-0.489***	0.383***	-0.257***	0.760***	-0.489***	0.383***	-0.257***	0.760***

3.6 Robustness Test

To explore the robustness of our strikingly different results to those presented by previous studies, we change the matching variables used to identify the control sample. An obvious alternative was to construct the control sample by using only the price and ADV selection criteria in line with the sample in the Nasdaq fee pilot report. The results are qualitatively consistent.

A second robustness test involved normalizing our dependent variables. We compute the quote spread, effective spread, realized spread, and price impact (both one second and five second) in bps rather than cents (see Table 3.12). Again, the results are generally consistent.

Table 3.12: Relative Bid-Ask Spread and Price Impact_bps

The table tests the impact of trading fee change on bid-ask spreads and price impact. Quoted spread_rel is calculated as the time weighted difference between the bid and ask over the quote midpoint using NBBO (Nasdaq) quote data. Effective spread_rel is calculated as twice the volume weighted signed difference between the transaction price and the midpoint of the prevailing bid and ask quotes. Realized spread_rel is calculated as twice the volume weighted signed difference between the transaction price and the midpoint of the bid and ask quotes 1 and 5 seconds after the trade. Price impact is twice the volume weighted signed difference between the quote midpoint and the quote midpoint 1 and 5 seconds after the trade. All spreads and price impact measures are expressed in basis points (bps) of the prevailing midpoint, and measured on a daily basis per security. The control variables (Price, Mkt Cap and Volatility) are computed as the average value over the sample period. Nasdaq access fee pilot is implemented on February 2, 2015 and ended on May 31, 2015. The sample period is from December 1, 2014 to July 31, 2015. Standard errors are in parentheses and are robust to both heteroscedasticity and correlation within stocks. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**Quoted Spread_rel**		**Effective Spread_rel**		**Realized Spread 1s_rel**		**Realized Spread 5s_rel**		**Price Impact 1s_rel**		**Price Impact 5s_rel**	
Stock Fixed Effect	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Intercept	73.71***	169.31***	58.87***	139.04***	25.91***	43.07***	12.83***	21.29***	32.96***	95.96***	46.04***	117.75***
	(2.273)	(2.565)	(1.625)	(1.608)	(1.566)	(2.861)	(1.574)	(3.059)	(2.117)	(3.429)	(2.304)	(3.684)
Treat	1.280***	1.500***	1.025***	1.798***	0.682***	-0.252	-0.075	-0.685**	0.342*	2.049***	1.099***	2.483***
	(0.204)	(0.230)	(0.146)	(0.144)	(0.141)	(0.256)	(0.142)	(0.274)	(0.190)	(0.307)	(0.207)	(0.330)
PilotOn	-1.14***	-1.138***	-0.260*	-0.260***	0.038	0.038	0.005	0.005	-0.298	-0.298*	-0.264	-0.264
	(0.207)	(0.116)	(0.148)	(0.073)	(0.143)	(0.129)	(0.144)	(0.138)	(0.193)	(0.155)	(0.210)	(0.166)
Treat*PilotOn	0.284	0.284*	-0.003	-0.003	3.615***	3.615***	3.644***	3.644***	-3.618***	-3.62***	-3.647***	-3.647***
	(0.285)	(0.160)	(0.204)	(0.100)	(0.197)	(0.178)	(0.198)	(0.190)	(0.266)	(0.213)	(0.289)	(0.229)
VIX	0.193***	0.193***	0.050**	0.050***	-0.118***	-0.118***	-0.113***	-0.113***	0.168***	0.168***	0.163***	0.163***
	(0.029)	(0.016)	(0.021)	(0.010)	(0.020)	(0.018)	(0.020)	(0.020)	(0.027)	(0.022)	(0.030)	(0.024)
Log (Price)	-6.60***	-6.768***	-6.54***	-6.907***	1.437***	1.837***	2.021***	2.189***	-7.972***	-8.74***	-8.556***	-9.095***
	(0.084)	(0.085)	(0.060)	(0.053)	(0.058)	(0.094)	(0.058)	(0.101)	(0.078)	(0.113)	(0.085)	(0.121)
Log (Mkt Cap)	-1.95***	-5.735***	-1.34***	-4.381***	-1.083***	-1.856***	-0.671***	-1.049***	-0.255***	-2.53***	-0.667***	-3.332***
	(0.086)	(0.094)	(0.061)	(0.059)	(0.059)	(0.105)	(0.059)	(0.112)	(0.08)	(0.126)	(0.087)	(0.135)
Volatility	-54.2***	-165.8***	-31.3***	-194.6***	-207.1***	-181.4***	-174.5***	-157.1***	175.86***	-13.14	143.20***	-37.51*
	(10.737)	(14.890)	(7.675)	(9.335)	(7.396)	(16.608)	(7.436)	(17.758)	(9.998)	(19.905)	(10.885)	(21.388)
Adjusted R^2	0.699	0.906	0.796	0.951	0.355	0.471	0.399	0.443	0.771	0.853	0.770	0.855
Observations	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676

3.7 Conclusion

This study provides a new theoretical model of maker-taker fees to show that the entire fee rebate must be consumed by an increase in the information content of informed trades in equilibrium, motivated by a *rise* in the trading profitability of informed market orders paying the taker fee. Uninformed traders neither benefit from the fee rebate nor suffer from it. We then empirically investigate how a reduction in the maker-taker fee affects market competition, liquidity, and HFT/non-HFT make-or-take decisions during the Nasdaq pilot that was constructed as a natural experiment.

The current literature holds that only changes in the net exchange fee matter for market quality and transaction cost efficiency. Our findings dispute that result and instead support our theoretical implications that although the raw effective spreads were largely unaffected by the rebate reduction, the information content of the taker order flow fell substantially. Holding the net fee relatively constant, the change in component fees and rebates does matter, but not for competition between lit and dark venues because the latter predominantly exclude informed traders. Rather, when Nasdaq's fee pilot reduced both the fee and the rebate, its market share decline benefitted other high rebate-paying lit exchanges. O'Hara and Ye (2011) shows that more fragmented stocks have lower execution costs and faster execution speeds while Gresse (2017) uses European evidence to show that market fragmentation in lit markets improves liquidity. In this chapter, we have shown how competition between venues differentiated by differences in maker-taker fee structures can lead to greater depth, higher trade volume, more informed trading, and better price discovery, especially in venues with the highest permitted levels of limit order book fee rebate. Hence, we question the capping of the rebate by the SEC.

We find that the equivalently reduced fee and rebate lowers quote quality and routed volume to Nasdaq while enhancing the fill rate and speed of fill because of the reduced taker fee and a thinner market. With their relative routing position improved, adverse selection costs decline on Nasdaq and liquidity supplier profit increases. The decline in informed equilibrium order flow and reduced liquidity means that Nasdaq's existing strategy of providing the highest possible subsidy to liquidity makers funded by its tax on liquidity takers encourages better price discovery and higher market efficiency. Profits of informed trades are enhanced with liquidity providers and liquidity traders no worse off and probably better off. Hence, and as shown by our simple model, fee rebates to liquidity suppliers can improve Pareto efficiency.

Nasdaq designed the access fee pilot as a natural experiment to address the question of whether high exchange access fees cause trading to shift from exchanges to dark pools. We find no evidence for such a shift, perhaps because only one exchange reduced its access fee unilaterally in a competitive trading environment. However, even if all exchanges had participated, the outcome may have remained the same, as the net fee did not substantially alter. Something fundamental changed in the competition between exchanges (the informed equilibrium order flow decreased on Nasdaq) but nothing fundamental changed between exchange and off-exchange venues.

Finally, we have shown that as exchange access fees and rebates decrease, HFT traders tend to switch from adding to removing liquidity, contrary to the liquidity cycles model of Foucault, Kadan, and Kandel (2013) in which imbalances in monitoring efficiency by HFT versus other non-HFT lead to rigidities in shifting from one side of the market to the other. While standard quality measures such as the effective spread, realized spread, and market impact costs all appear to improve on a cum-rebate basis, these apparent improvements reflect the lower information content in Nasdaq's order flow during the experiment as informed order flows shift to exchanges that maintain high rebates. Given their relative inability to monitor the now less informed order flow dynamics, non-HFT firms increased liquidity making and decreased their liquidity taking while HFT firms did just the reverse, albeit with a less informed equilibrium order flow.

Chapter 4 Tick Size, Exchange Access Fee and Market Liquidity[21]

4.1 Introduction

Do exchange access fees matter when there is an exogenous alteration in the minimum tick size? To compete for liquidity, three major exchange access fee structures operate in global equity markets: *taker-taker*, which charge both liquidity suppliers and consumers; *maker-taker*, which reward liquidity suppliers and charge liquidity consumers; and *taker-maker (‘inverted’),* which reward liquidity consumers and charge liquidity suppliers.[22] Angel, Harris, and Spatt (2011, 2015) and Colliard and Foucault (2012) argue that the tax-subsidy arrangements simply wash-out in competitive markets, without affecting transaction prices or quotes. For example, Colliard and Foucault (2012, p.3392) claim that “holding the total fee constant, any change in make and take fees is neutralized by an adjustment in the raw bid-ask spread.” Taking it further, Foucault, Kadan, and Kandel (2013) present trader monitoring model in which tax-subsidy arrangements wash-out in a zero-minimum tick regime but become more efficacious for maker-taker venues and traders the higher the tick size. Likewise, Chao, Yao, and Ye (2017a) prove that fee structures wash-out with a zero-minimum tick or continuous pricing in the absence of informed traders, to show that a monopoly exchange can survive competition between exchange operators with different fee structures venues provided price changes are

[21] The Internet appendix that accompanies this chapter may be found at https://goo.gl/Lz227q.

[22] In the remainder of the chapter, when we refer to ‘make’ or ‘limit’ orders, we mean specifically ‘non-marketable’ limit orders that add liquidity by improving on the best ‘bid’ by offering the seller a higher price, or on the best ‘ask’ by offering the buyer a lower price, and thus are not guaranteed automatic executions. Similarly, when we refer to ‘take’ or ‘market’ orders, we mean ‘marketable’ limit orders that are equal to or above the best ‘ask’ price or equal to or below the best ‘bid’ price that remove liquidity (see Interactive Broker Knowledge Base, 2017).

discrete. However, previous chapter of this thesis model and empirically report that the conventional 'wash-out' view with continuous pricing is only correct in the empirically unlikely case in which there is no information in order flow (Hasbrouck, 1991; Easley, Kiefer, and O'Hara, 1996; and Corwin and Lipson, 2000). Using the 2015 Nasdaq fee experiment, this thesis shows a maker-taker fee and rebate set at the (capped) maximum level is most beneficial for market efficiency and the market share of a venue. In the present chapter, we show that, far from making existing fee structures work better, the minimum tick size itself operates as an excessively-severe tax on liquidity, i.e., uninformed, traders imposed in conjunction with existing venue fee structures.

We find that an exogenous rise in the tick size for small capitalization firms relatively benefits investors trading in the taker-maker (inverted) fee venues, crossing networks and dark pools while harming those in the maker-taker fee venues and at the expense of the well-being of investors in lit markets generally. The SEC mandated market-wide tick size pilot[23] increased the minimum tick increment on October 3, 2016 from 1 cent (penny) to 5 cents (nickel) which supposedly improve liquidity and trading of small capitalization securities (SEC, 2015c). The pilot will last for 18 months and complete in October 2018. This carefully designed pilot provides for an exogenous rise in the tick size while retaining comparable control samples to unravel the role of the minimum tick in affecting market execution quality and pricing.

[23] The tick size pilot was introduced under the auspices of the Jumpstart Our Business Startups Act ("JOBS Act") in 2002, which sought to improve access to the public capital markets for emerging growth companies. Ritter (2014) reports that the small-company initial public offerings (IPOs) declined 83 percent from 165 IPOs a year during 1980-2000 to only 28 a year during 2001-2012 in spite of the doubling of real gross domestic product (GDP) during this 33-year period. This should be alarmed as it is the conventional wisdom that companies going public create many jobs.

The tick size pilot sets forth four groups of securities, treatment and control, selected by the three stocks listing markets (Nasdaq, NYSE, and NYSE Amex). The control group includes small capitalization firms that experience no change in regulatory policy and tick size. The treatment firms are categorized into three subsets of securities: those that quote in 5 cents increments but continue to trade at 1 cent increments; those that trade and quote in 5 cents increments; and those that trade and quote in 5 cents increment and subject to the additional 'trade-at' protection rule.[24]

Our findings show the 400% rise in the minimum tick from 1 cent to 5 cents relatively benefits traders in inverted venues, as it is essentially a tax on uninformed traders whose liquidity demands become even more price sensitive as trading costs rise. This increase in price sensitivity motivates flight from higher-cost maker-taker markets to both inverted markets and off-exchange venues that are affected by SEC mandated tick size pilot. Since informed trades are less price sensitive, the rise in information content caused by the departure of liquidity traders is most pronounced in maker-taker venues with the highest make rebate.

Both our theory and findings are in contradiction to the conventional 'wash-out' theory of exchange fee models and to the view that tick size constraints make exchange fee structures efficacious. While this chapter agree with Foucault, Kadan, and Kandel (2013) and Chao, Yao, and Ye (2017a, b) that exchange fee structures matter, it does not support that fee structures transfer from takers to makers (more than $1 billion p.a. in the case of NYSE-Arca), or that these arrangements need to work as depicted in their models. For example, if it is minimum tick size constraints that make it possible for competing

[24] More detail is available here: http://www.finra.org/investors/tick-size-pilot-program. The 'trade-at' protection rule prevents dark pools from competing with lit venues by providing minimal sub-penny price improvement.

maker-taker venues to survive with different maker-taker fees then higher minimum tick size constraints should improve their competitive position, as contended by Foucault, Kadan, and Kandel (2013). By contrast, we find that a higher tick size worsens the competitive position of maker-taker venues.

The maker-taker rebate (subsidy) acts directly to both narrow the inside quotes and increase the queue of orders in the LOB, but as the model shows, this narrowing is offset by informed traders who increase trading profit by raising the information content of their trades. This increase can come about by the migration of informed traders from venues with a lower level of subsidy to make orders. Hence, the subsidy mechanism can never wash-out, even with a zero-tick size, and nor can it transfer billions of dollars annually from takers to makers as discussed in Foucault, Kadan, and Kandel (2013). To the contrary, its leading role is in encouraging better terms for asymmetrically informed market orders, thus promoting price discovery. Consequently, the maker-taker model improves price discovery and market efficiency by encouraging more informed trading to the overall betterment of the market.

Foucault, Kadan, and Kandel (2013) is the first to find that the fee structure matters, assuming that tax-subsidy effects wash-out in a zero-tick size regime in the absence of any informed traders. In their model, participants differ with respect to their monitoring capacity as reflected in latency, algorithmic, and high frequency trading skills in the two sides of the market, taker or maker. For example, given their model, if makers are relatively poor monitors, the tick size is relatively tight, and makers are few relative to takers, takers obtain only limited gains from trade[25]. Under these circumstances, it is optimal to increase the taker fee and reduce the maker fee to enhance the speed of execution. They provide a numerical example in which the welfare gains from a maker-

[25] See pp. 318-319 and Proposition 3 of Foucault, Kadan, and Kandel (2013).

taker fee over a uniform fee structure rise from \$33 million to \$294 million per annum as the minimum tick rises from one cent to one-eighth of a dollar.[26]

There are obviously several possible flaws in the Foucault, Kadan, and Kandel (2013) model, in addition to the problem that the fee structure does matter even in a zero-minimum tick regime. First, market participants with varying latency, algorithmic, and high frequency trading skills are equally free to both make and take liquidity and can instantly switch from one side of the market to the other. Hence, the assumption of permanent differences in monitoring skills between the two sides of the market that enable differing fee structures lacks both evidence and plausibility. For example, as shown in previous chapter of this thesis, HFT shifted from one side of the market to the other on average during Nasdaq's fee experiment. Second, even if there were differences in monitoring ability between the two sides of the market, a higher tick size, such as in the 2016 SEC mandated tick size experiment, will act as a tax on liquidity trades and thus favour the inverse market, as argued above, worsening rather than improving the welfare of participants in the maker-taker markets.

An alternative and perhaps more plausible hypothesis to that of Foucault, Kadan, and Kandel (2013) is that they have incorrectly specified the effects of the maker-taker fee structure and its inverse by ignoring a major constituent of the order flow and trading volume, namely informed trades. Because the existing literature assumes either the absence or the irrelevance of informed traders hiding in the liquidity-based order flow, fee structures can never wash-out, even without tick size. This is because it is profitable for these informed traders to raise the information content of their trades to absorb the subsidy to make orders such that the *raw* best bid and ask does not contract by the amount of the fee rebate in maker-taker regimes or widen by the fee in inverted regimes–a

[26] In Table 3 (pp. 333) of Foucault, Kadan, and Kandel (2013).

necessary condition for wash-out. This information content offset to the make subsidy is accomplished by an influx of highly informed traders.

In this chapter, we use the 2016 U.S. tick size pilot natural experiment to test the Foucault, Kadan, and Kandel (2013) hypothesis that the welfare of maker-taker participants should improve with a higher minimum tick size. We split our sample into three tick-constrained groups, based on the NBBO quoted spreads prior to the pilot. The first tick-constrained group experiences a 'true' increase in the minimum tick size (previous NBBO quoted spread is between 1 cent to 5 cents). The two additional subsets of non-minimum tick-constrained securities may not experience an increase in tick size (5 cents to 10 cents and greater than 10 cents) as its average quoted spread prior the pilot is already higher than 5 cents.

We find that inverted venue participant's gain markedly in every investigated respect.[27] Moreover, consolidated trading volume for the small stocks included in the tick size pilot diminishes as trading and market share shifts off-exchange to crossing networks and dark pools for all but the sub-sample subject to the 'trade-at' rule. These findings are in support of our contention that exchange fees can never wash-out as their effect is limited to altering the best quotes and depth of the LOB and thus altering the treatment of informed versus uninformed orders. Also, the effective spread in inverted venues for the most tick-constrained stocks rose relative to maker-taker venues (as distinct from the exchange level), consistent with a much higher proportionate rise in the cost of take orders in response to the minimum tick size rise in inverted markets. Moreover, our contention that increases in the tick size for tick-constrained stocks are simply a disguised tax on trading, especially liquidity trading, is also borne out by the overall decline in the lit

[27] The inverted venue, EDGA, showed a slight fall but with its fee and rebate close to zero, its limit order book improvement was very modest.

market for stocks affected by the pilot, as well as the shift from maker-takers to the inverted venue.

It might seem puzzling that trade volume and market share increased, in absolute and relative terms, in inverted venues following a minimum tick size increase that widened the effective spread. Moreover, the composition of trading altered with a relative increase in uninformed compared to informed order flow in the inverted market, as indicated by a relative fall in market impact costs and order flow information content compared to maker-taker venues.

Our explanation is based on the altered composition of order flow, with price-sensitive uninformed order flow fleeing lit markets generally when confronting a 400% increase in trading costs, admittedly with most of this taking the form of an increased subsidy to the LOB. Relatively price-inelastic informed order flow is far less severely affected and, in any case, benefits from the increased depth in all lit order books.

Despite the higher proportionate rise in transaction costs (cum-fee spreads) in inverted venues, these venues remained substantially cheaper for uninformed make orders not requiring the added depth in maker-taker markets to the extent of the sum of the maker-taker make fee plus the inverted market take rebate. This sum specifies the degree to which take (market) orders are relatively cheaper on the inverted venue compared to the maker-taker venue. As the cost of transacting rose enormously and uninformed order flow became more price sensitive, it was now worthwhile for uninformed order flow to shift from maker-taker to inverted venues, whereas in the pre-event low transaction cost regime these differences were regarded as immaterial.

More generally, we find that market-wide consolidated trading volume decreased and effective spreads increased, suggesting the economic welfare deteriorated for small capitalization firms subject to the tick size pilot. This was despite an improvement in the percentage of time these markets quote at the NBBO and the NBBO quoted depth together with associated decreases in exchange-level quoted spreads and volatility. In this chapter, we devise a new NBBO metric that credits multiple markets displaying liquidity at the NBBO in the case of ties to describe the underlying quote quality that exists across fragmented markets. In addition, lit market share generally decreased for treatment firms and improved in off-market and dark pools, except for those subjects to the 'trade-at' protection rule. Market-wide NBBO quoted spreads and effective spreads increased, especially for the most tick-constrained group of securities. As transaction costs increased significantly, only the highly informed orders can now afford to cross the widened spread; hence, the price impact increases. However, the price discovery process is enhanced.

Like the present chapter, Comerton-Forde, Gregoire, and Zhong (2017) find that the SEC tick size pilot harms the lit market to benefit the dark market, and within the lit market, the inverted venue gains at the expense of the maker-taker venue. However, their explanation for the better performance of the inverted venue is that it offers a 'finer pricing grid' with sub-penny pricing that allows participants to gain price improvement relative to the NBBO. Yes, the net fee on inverted venues was lower than on maker-taker venues following the 400% increase in the minimum tick size, and in this sense, there is a 'finer pricing grid'. But on this peculiar definition, the 'pricing grid' was much finer still prior to the tick-pilot experiment as a percentage of transaction costs in the inverted venue. Despite this, the inverted venue share was exceedingly small. Rather, the increase in the minimum tick size acted as a severe tax on relatively uninformed traders, forcing them to become far more price sensitive and hence flee the higher-cost maker-taker market to either the relatively cheaper inverted market or to the dark market.

The present chapter shows that both maker-taker and inverted markets can offer a 'finer pricing grid' in the sense of undoing an imposed minimum tick that can help to neutralize very modest minimum tick rules, but the imposed cap on fees limits the ability of all fee structures to neutralize sizeable minimum tick rules. The fact remains that all venues were subject to a one-cent pricing grid prior to the tick size pilot and a five-cent grid during the pilot. Hence, it is not the case that fee structures are capable of altering the pricing grid, per se, as contended by Comerton-Forde, Gregoire, and Zhong (2017). At best, they can ameliorate some of the adverse effects of a minimum tick. In addition, Griffith and Roseman (2017) find the widening tick size fails to improve market liquidity in small-cap stocks, and a wealth transfer from taking liquidity to adding liquidity although cumulative depth remains unchanged or decreases. Penalva and Tapia (2017) find the spreads and depth increased, and trading volume and the level of market activity decreased for the stocks which the quoted spread is similar to the new tick size.

4.2 Literature Review

A. Exchange Access Fee and Off-Exchange Trading

In 1997, the Island ECN was the first venues to adopt maker-taker fee model to attract order flow through liquidity provision rebates, incentivizing liquidity providers to post competitive quotes and attract order flow, particularly informed order flow, from other markets. Consequently, Island's market share of reported Nasdaq trades increased by approximately 10% from 1997 to 1999 (Cardella, Hao, and Kalcheva, 2015).

In response to competition from ATSs, many exchanges also adopted maker-taker fee model with varying component fees over the past decade. The maker-taker pricing model has gained widespread adoption, notionally rewarding liquidity providers and charging liquidity consumers but in reality, encouraging informed traders and better price discovery. Three U.S. exchanges adopt an inverted fee model, offering rebates to liquidity takers and conversely charging a higher fee to liquidity suppliers. On August 19, 2016, IEX became the latest exchange in the U.S. trading landscape and adopted the third type of exchange fee model (taker-taker) other than the maker-taker and the taker-maker fee model, which charges fees for non-displayed liquidity for both adding and removing liquidity.[28] These inverted fee structure venues remain small in comparison with maker-taker venues, although their market share rose for stocks subject to the SEC tick size pilot.

Taker fees and maker rebates comprise a significant proportion of overall trading costs, especially for tick-constrained stocks and relatively uninformed trades that do not require a deep LOB market. Angel, Harris, and Spatt (2015) find brokers avoid such large marginal costs by routing their marketable orders to wholesale dealers to capture the bid-ask spread and avoid access fees; conversely, they send their non-marketable orders to exchanges for executions to gain maker rebates. They employ so-called 'smart order routers' that use real-time state information to solve the order routing problem (Maglaras, Moallemi, and Zheng, 2012).

Angel, Harris, and Spatt (2011) argue the introduction of a maker rebate financed by a taker fee should wash-out and thus have no effect because the best bid and ask prices in competitive markets should narrow by the rebate amount if all non-marketable limit orders are subsidized by an equivalent tax on market orders, with the tax precisely offsetting the narrowed spread. This argument would be valid if there were, counterfactually, no information in the order flow, or alternatively, if the information content in order flows is the same across venues and is unresponsive to make rebates designed to attract more informed order flow. This higher informed order flow is then matched with greater depth in the LOB without any of the postulated narrowing of the inside quotes. This postulated washing-out effect breaks down when different competing venues have different fee structures, each with its own specific information content in

[28] IEX fee schedule details are available at: https://www.iextrading.com/trading/alerts/2016/036/.

their order flow. When fees/rebates are altered, the venue's order flow information content adjusts accordingly. The maker-taker fee structure, by adding to the depth in the LOB, induces more informed trading that precisely offsets the tendency for the quotes to narrow due to the rebate. This alteration to the order flow induces better price discovery and a redistribution of informed orders across the set of venues.

Colliard and Foucault (2012) attempt to prove that, without market frictions, only changes in the net fee retained by the exchange affect liquidity and trading volume. While this would be the case if there were no informed trades, or if order flow information content were not subject to inter-venue competition, this thesis use the Nasdaq fee-rebate experiment to show that neither supposition is correct. Foucault, Kadan, and Kandel (2013) build on the Colliard and Foucault (2012) model to show that trading volume may either increase or decrease, even without a change in the net exchange fee, because a fixed tick size prevents prices from neutralizing the effect of the maker rebate. Similarly, Chao, Yao, and Ye (2017a) extend the model of Foucault, Kadan, and Kandel (2013) to establish that a minimum tick is sufficient to establish an equilibrium in which venues can co-exist with different fee structures. These models all share a common feature, namely the absence of informed trading.

From an empirical perspective, Malinova and Park (2015) examine whether and why the breakdown of trading fees between liquidity demanders and suppliers matters by using a change in trading fees on the TSX during a sample period when Canada operated under a monopoly market setting. They find that the requirements for wash-out were met as the quoted spreads decreased with trading volume unaffected, while holding the net exchange fee constant. In the absence of competing venues, it remained impossible for there to be a redistribution of informed orders across venues. By contrast, this thesis study the impact of unilateral fee reduction under a fragmented trading environment in U.S. and find that Nasdaq's market share reduced, and such reduced volume shifted to other lit exchanges with competitive maker rebates while the consolidated volume and off-exchange trading volume remained stable. Adverse selection costs also declined in line with the improved relative position of a market in the routing table. That is, informed traders fled Nasdaq to competing maker-taker exchanges that did not participate in Nasdaq's own experiment and thus price discovery was reduced on the Nasdaq exchange.

Exchange pricing model is increasingly important in a high frequency trading environment and deterministic of the recent shift from traditional exchanges to off-exchange venues. Battalio, Corwin, and Jennings (2016) study the impact of differential exchange fee models on broker routing decisions to find evidence that four of an examined ten national retail brokers sell orders to capture the maximum make rebates. But high rebate venues experience lower fill rates due to the length of the queue. Consequently, on this measure of execution quality, some clients of retail brokers who do not receive the broker rebate may suffer from a conflict of interest.

B. Minimum tick Increment

Foucault, Kadan, and Kandel (2005) present a LOB model populated by traders with varying patience level and find that a zero-minimum tick is not optimal. Kadan (2006) argues that markets with many dealers benefit from a small tick size whereas investors may suffer when the number of dealers is small. Finally, Dayri and Rosenbaum (2015) introduce the 'implicit spread' for large tick LOBs, which can take values below the tick size. They argue that the tick size is optimal when it equals the 'implicit spread'. The model is adopted in Huang, Lehalle, and Rosenbaum (2015) to predict the consequences of the tick size change on the Tokyo Stock Exchange.

Harris (1991) and Angel (1997b) discuss the following advantages of a large tick size. First, it reduces the complexity of the trading environment as the number of possible price levels is limited which reduce bargaining costs. Second, it reduces the bandwidth requirements of a trading platform. A change in the single stock price can lead to a price update for hundreds of related derivative products. Third, it sets the floor income for liquidity providers because bid-ask spreads cannot be quoted below the minimum tick. Finally, a finite tick size protects the time-priority rule. Time priority grows in importance when the tick size is large, and traders are incentivized to submit orders to the LOB early. As we find no benefit from the increased tick and sizeable costs, this suggests that the benefits from a sizeable tick are largely illusory and that time priority is not so important.

Most of the existing literature examined the impact of tick size reductions. Bollen and Whaley (1998), Goldstein and Kavajecz (2000), Van Ness, Van Ness, and Pruitt (2000), and Chordia, Roll, and Subrahmanyam (2001) report lower spreads and lower volumes at the best quotes after the change from eighths to sixteenths in U.S. exchanges, consistent with predictions established in Harris (1994).[29] The decline in spreads is partially due to increased competition between market makers (MacKinnon and Nemiro, 2004), while Chung and Chuwonganant (2004) report tighter spreads on the Nasdaq after the change in order handling rules allowing competition between designated dealers and traders. Decimalization also see declines in spreads on the NYSE (Chakravarty, Wood,

---

[29] When tick size is reduced, narrower spreads and lower quoted depth are also reported for the Chicago Mercantile Exchange (Kurov, 2008; Karagozoglu, Martell, and Wang, 2003), the Hong Kong Stock Exchange (Chan and Hwang, 2001), the Jakarta Stock Exchange (Ekaputra and Ahmad, 2007), the Sydney Futures Exchange (Alampieski and Lepone, 2009), the Taiwan Stock Exchange (Hsieh, Chuang, and Lin, 2008), the Tokyo Stock Exchange (Ahn, Jun, Chan, and Hamao, 2007), the Toronto Stock Exchange (Ahn, Cao, and Choe, 1998; Bacidore, 1997; MacKinnon and Nemiro, 1999; Porter and Weaver, 1997), the Stock Exchange of Singapore (Sie and McInish, 1995), and the Stock Exchange of Thailand (Pavabutr and Prangwattananon, 2009).

and Van Ness, 2004). A tick size reduction increases the number of price levels available to liquidity providers, thus, it effectively distributes liquidity onto a finer pricing grid. This can mechanically decrease market level 1 quoted depth without reducing total liquidity, or increase total liquidity in reality. Furthermore, Goldstein and Kavajecz (2000), Sie and McInish (1995), and Pavabutr and Prangwattananon (2009) find that the LOB cumulative liquidity decreases after tick size reductions.

Niemeyer and Sandas (1994), Ke, Jiang, and Huang (2004), and Chung, Kim, and Kitsabunnarat (2005) examine the effects of the relative tick size for stocks listed on the Stockholm Stock Exchange, the Taiwan Stock Exchange, and the Kuala Lumpur Stock Exchange respectively. They all find larger spreads for larger relative tick sizes. Jain (2003) examined 51 exchanges globally and report tighter spreads for narrower minimum ticks. Cai, Hamao, and Ho (2008) and Bessembinder (2000) exploited threshold effects in the tick size bands of the Tokyo Stock Exchange and the Nasdaq, and report larger spreads when the stock price entered a larger tick size band. Also, Hau (2006) reported higher volatility for stocks with a larger tick size on the Paris Bourse.

There are mixed results examining the relationship between tick size and execution costs. Bollen and Whaley (1998), Alampieski and Lepone (2009), MacKinnon and Nemiro (1999), and Smith, Turnbull, and White (2006) report lower transaction costs when tick size is smaller. However, Jones and Lipson (2001) and Bollen and Busse (2006) find the opposite that the execution costs are higher following tick size reductions. Angel (1997a) finds the incidence of stock splits increase after tick size reduction to keep the stock price in the 'optimal trading range' and maintain relative tick size. Schultz (2000) argue that brokers promote stocks using stock splits by increasing the spread and market making revenues (see also Lipson and Mortal, 2006; and Easley, O'Hara, and Saar, 2001).

Later academic literature question the positive relationship between smaller tick sizes and higher liquidity. Bourghelle and Declerck (2004) find no change in bid-ask spreads after a tick reduction on the Paris Bourse. Furthermore, Aitken and Comerton-Forde (2005) find that spreads increase for stocks with small relative tick size before the Australian Stock Exchange reduced the tick size, and Wu, Krehbiel, and Brorsen (2011) find similar results on the NYSE. Fears that tick sizes could have become too small finally prompted the SEC to implement the 2016 tick size pilot which widened the minimum tick size from 1 cent to 5 cents for small stocks.

Using relative tick size change, Yao and Ye (2015) find a large relative tick size harms liquidity and drives speed competition in liquidity provision in low-priced securities using splits/reverse splits as exogenous shocks. O'Hara, Saar, and Zhong (2015) find a larger relative tick size benefits HFT in market making on the NYSE. It results in greater depth and more volume in a one-tick spread environment, and the opposite outcome in a multi-tick spread environment.

4.3 Model

Following in the framework of Glosten and Milgrom (1985), Aitken, Garvey, and Swan (1995), and the maker-taker fee model presented in chapter 3 of this thesis, there is one unit of a representative stock whose price can take on one of two values, V^H or V^L, with $V^H > V^L$. Setting the unconditional share price at $\left(V^H + V^L\right)/2 \equiv V$, namely the valuation placed on a share by the uninformed, then $V^H \equiv (1+\alpha)V$ and $V^L \equiv (1-\alpha)V$, where $1 \geq \alpha \geq 0$ is the value of the information about the 'true' price revealed only to informed traders prior to placing their order. One can think of α as the measure of the degree of information content of an informed market order. At $\alpha = 1$, the informed trader's informational advantage is at its maximum and as $\alpha \to 0$, the informational advantage evaporates. A liquidity, i.e., uninformed, trader who is a potential seller values the share at a fraction, $(1-\lambda)V$, of the unconditional value for liquidity or portfolio rebalancing reasons while an equivalent potential buyer values the share at $(1+\lambda)V$, where λV is a measure of the gains from trade, with $(1-\lambda)V$ and $(1+\lambda)V$ representing private valuations of the liquidity seller and buyer, respectively, with $0 \leq \lambda \leq 1$. Liquidity traders are randomly assigned either a low or high valuation, while $\lambda = 0$ for both makers and informed traders alike. Moreover, some liquidity traders may be assigned more extreme λ values than others, such that when lit-market trading costs rise, those with less extreme values either migrate to crossing networks or dark venues, or cease to trade altogether. As trading costs rise, the remaining liquidity traders with more extreme λ values display greater price sensitivity, consistent, for example, with a conventional linear downward sloping demand schedule.

The difference in private valuation motivation for liquidity traders is similar to that employed by Colliard and Foucault (2012), Foucault, Kadan, and Kandel (2013), and Chao, Yao, and Ye (2017a), except in these models there is no informed trading, and hence $\alpha = 0$. Informed traders who know the true value of V prior to placing their order consist of a proportion $1 > \gamma \geq 0$ of the entire population of traders. Potential sellers with valuation $(1-\lambda)V$ are offered the 'bid', i.e., sell, price, denoted $p_s \geq (1-\lambda)V$ in the limit order market, so that the price must equal or exceed his private valuation and potential buyers are offered the 'ask', i.e., buy, price $p_b \leq (1+\lambda)V$ with the exogenously fixed maximum width of the 'inside' quotes, $p_a - p_b = 2\lambda V$, increasing in the gains from trade, λV, enjoyed exclusively by uninformed traders. By contrast, both professional limit order providers, other than liquidity traders, and informed traders are motivated purely by profit.

A. Fee Scheme with Offsetting Rebate to Liquidity Suppliers

For simplicity, the cost to the exchange of matching buyer and seller is set to zero. Nonetheless, and in common with the conventional literature, there is an exchange matching fee, given by f_t, on *takers* in the maker-taker venue with an offsetting rebate, $-f_t$, to makers representing non-marketable limit order providers such that execution is not certain. In the inverted taker-maker market, the fee applicable to *makers* is f_m and the rebate to non-marketable takers is $-f_m$. In maker-taker venues, the take fee is applied to all 'take' trades, i.e., market orders, and marketable 'make' trades, i.e., limit orders at the inside quotes, such that all buyers and sellers regardless of their information status pay fees on all trades certain to execute. In inverted venues, the reverse is true with non-marketable limit orders that provide additional liquidity paying the fee of f_m and take trades in receipt of the rebate, $-f_m$, and in traditional neutral taker-taker markets both the fee and rebate are zero.

Chapter 3 of this thesis proves that with the introduction of a maker-taker fee-rebate scheme the information content in venue order flow must increase from the *neutral*, no rebate (n), regime level, α_n, to the higher maker-taker level, α_t, with the information content increment:

$$\alpha_t - \alpha_n = \Delta\alpha = \frac{f_t}{\gamma_n V}, \tag{24}$$

Intuitively, informed traders can improve profitability by raising information content by the precise amount of the rebate to make trades without making liquidity traders any worse off. Liquidity traders can still place limit orders themselves and, having received the rebate, are no worse off.

Informed traders (*in*) gain a profit increment of:

$$\Delta \pi_{in}^{mt} = (1-\gamma_n)\left(\frac{f_t}{\gamma_n}\right)V > 0, \tag{25}$$

which increases in the proportion of uninformed order flow in the initial no-rebate regime, $(1-\gamma_n)$, the magnitude of the rebate deflated by the proportion of informed traders, $\frac{f_t}{\gamma_n}$, and the unconditional stock value, V. Intuitively, informed traders gain more, the greater the exclusivity of their informational advantage, the more sizeable is the rebate relative to the size of the pool of informed traders, and the more dollars that are involved, i.e., the price of the stock. Profit is maximized when the expected incremental cost of information content, $(\alpha_t - \alpha_n)\gamma_n V$, fully absorbs the fee rebate. The fee-rebate 'wash-out' models of Colliard and Foucault (2012), Foucault, Kadan, and Kandel (2013), and Chao, Yao, and Ye (2017a) represent the limiting case of this fee-rebate model when the information level or degree of information content is zero such that there is no informed trading.

So far, this model of fees and matching rebates has ignored price discreetness in the form of the tick size, with the minimum tick implicitly set at zero. In the SEC mandated tick size experiment, which is applied to 1,200 relatively small firms, the minimum tick size for a single trade, denoted θ, is raised from one cent to five cents, a huge increase of 400%. This increased minimum tick size can be compared to the permissible maximum fee of 30 cents per hundred shares traded or 0.3 cents per share traded. Since the minimum tick increase acts in many ways just as a maker-taker fee-rebate does, it is remarkable that the SEC had mandated this minimum tick experiment with a fee structure that is (5/0.3) = 16.67 times higher than its own maximum permitted fee of 0.3 cents. Essentially, this minimum tick experiment will have the largest impact on fees and matching rebates when the minimum tick is binding. These will typically be lower priced/higher volume stocks. For example, if the market-determined spread exceeds five cents such that the constraint is never binding then the experiment may tell us little. We now assume that the minimum tick is binding and ask how it modifies the maker-taker analysis of chapter 3 of this thesis.

As a thought experiment, suppose there are no informed trades and there are two venues with a maker-taker fee, $f_t = 1/2 \times \theta$, in one and an inverted fee structure, $f_m = 1/2 \times \theta$, in the other venue, where θ is the minimum tick size. If the minimum tick size, θ, increases from a binding 1 cent to a binding 5 cents, then *Proposition 1* of chapter 3 of this thesis shows that the raw spread will increase from $-(p_a - p_b) = 1$ to $-(p_a - p_b) = 5$ cents in the maker-taker venue and *Proposition 2* shows that the spread will increase, likewise, from $(p_a - p_b) = 1$ to $(p_a - p_b) = 5$ cents in the inverted venue. In both venues, the cum-fee spread remains constant at zero cents due to the absence of informed traders when the venue fee alters to match the minimum tick size rule. Hence, with a flexible fee structure and no informed trades, not only do diverse fee structures net out but, in addition, the minimum tick size constraint also washes out. Unfortunately, it is impossible for fee structures to wash-out the 5 cents minimum tick size because the SEC caps fees at 0.3 cents per share, whereas a much higher fee of 2.5 cents would be required to wash-out the 5 cents minimum tick size requirement by artificially widening the spread prior to applying the correcting fee-rebate.

Now, examine a maker-taker venue in the presence of informed trading, such that, the buy-price becomes:

$$p_b = (1+\lambda)V \geq (1+\alpha_n \gamma_n)V + f_t = \left(1+\alpha_t^{mt}\gamma_n\right)V\ , \tag{26}$$

and the sell-price:

$$p_s = (1-\lambda)V \leq (1-\alpha_n \gamma_n)V + f_t = \left(1-\alpha_t^{mt}\gamma_n\right)V\ , \tag{27}$$

where α_n denotes the level of information content of informed market orders in the initial zero-minimum tick regime and γ_n denotes the likelihood of encountering an informed trader, i.e., the proportion of informed traders, in the zero-minimum tick regime. Note that all participants are assumed risk-neutral. Hence, the binding maker-taker spread condition becomes:

$$2\alpha_t^{mt}\gamma_n V = 2\left(\alpha_n\gamma_n V + f_t\right) \leq p_b - p_s = 2\lambda V = \theta, \tag{28}$$

and, similarly, the inverted taker-maker spread requirement becomes:

$$2\alpha_m^{tm}\gamma_n V = 2\left(\alpha_n\gamma_n V - f_m\right) \leq p_b - p_s = 2\lambda V = \theta, \tag{29}$$

Taking the difference in spread composition between the maker-taker (Equation (28)) and inverted market (Equation (29)), taker-maker venues are more heavily concentrated on informed traders by the amount, i.e., twice the sum of the maker-taker rebate plus the inverted market fee for providing liquidity.

If there is a sizeable increase in the minimum spread, as was mandated by the SEC tick size pilot, the degree of information asymmetry in the purely market-driven spread specified by the two LHS terms in Equations (28) and (29), representing the market-determined spread, may not be sufficient to match the required minimum spread, indicated by the three RHS terms in Equations (28) and (29). As when the constraint is met, gains from trade, λ, are equal to $\theta/2V$, the constraint can only be satisfied by the departure of enough marginal liquidity traders from the lit market until investor gains from trade for remaining investors rise sufficiently to satisfy the constraint.

The remaining vital issue to address is: will these fleeing liquidity investors come from the maker-taker market, the inverted market, or both? Market (take) orders pay a fee-inclusive price of the raw spread plus the fee, $\theta + f_t$, on maker-taker venues and the raw spread minus the fee, $\theta - f_m$, on inverted venues, making inverted venues particularly attractive to increasingly price-sensitive uninformed order flow, following what is effectively a 400% increase in the fee due to the spread rise. The more inelastic nature of informed order flow is indicated by the reduction in informed order flow to the maker-taker Nasdaq venue when fees were largely removed during the Nasdaq fee pilot. Moreover, while there is greater depth in the LOB for both inverted and maker-taker markets due to the 400% rise in the tick size from one cent to five cents, the relative depth is even greater on the maker-taker venue due to the rebate of f_t to limit orders and fee of f_m on limit orders placed on inverted venues. This greater relative depth, with corresponding lower fill rate, makes maker-taker venues particularly attractive to informed traders placing take orders when the minimum tick size is raised, helping to account for the lower price sensitivity of informed orders.

Proposition 1: *When the minimum tick size,* θ, *increases from one cent to five cents, liquidity traders with less extreme, either high or low gains from trade,* λ, *and increasingly price sensitive, will switch from maker-taker venues with a high make fee of* $\theta + f_t$ *to either inverted venues with the lower fee of* $\theta - f_m$, *off-market, or cease trading altogether. Similarly, more price sensitive liquidity traders will depart the inverted market for off-market venues or cease trading altogether. Since cum-fee take trading costs are higher in the maker-taker venue due to the fee structure, more highly price sensitive liquidity traders will depart this venue for the inverted venue, ensuring that the relative information content of trading in the inverted market falls. Despite this relative fall, the departure of liquidity traders implies that information content levels must increase in both markets while market share rises in both inverted and dark venues and falls in maker-taker venues.*

4.4 Institutional Details

The SEC approved the National Market System (NMS) Plan to implement the tick size pilot proposed by the exchanges and FINRA on May 6, 2015. The pilot commenced on October 3, 2016. The control group consists of approximately 1,400 securities and three treatment groups, each with approximately 400 securities selected by a stratified sampling. The groups are defined as follows: *treatment group one (G1)* must quote in 5 cents increments, but continue to trade at 1 cent increment; *treatment group two (G2)* must quote and trade in 5 cents increments; and *treatment group three (G3)* will adhere to the requirements of the G2, but will also be subject to a 'trade-at' prohibition. Under the trade-at prohibition, it will prevent a trading venue that was not quoting from the price-matching NBBO which are protected quotations for NMS stocks, and permit a trading venue that was quoting at a protected quotation to execute orders at that level, but only up to the amount of its displayed size. This would require brokerages to route trades to public exchanges at the NBBO, unless they can execute the trades at a meaningfully better price than what is available in the public market. The control group will be quoted and trade at the existing 1 cent tick size increment.

Table 4.1 identifies the 12 lit exchanges in U.S. equity markets during our sample period. Eight are maker-taker venues (NASDAQ, PSX, NYSE, ARCA, AMEX, BATS, EDGX, and CHX) and three are the inverted venues (BX, BATS Y, and EDGA). IEX applies the taker-taker fee model for non-displayed liquidity only. Table 4.1 reports the indicative exchange access fee for all U.S. markets during our sample period as well as the percentage change in incentive to add liquidity to the LOB (Incentive %). Incentive% is measured as the post pilot minimum tick (500 CPS) adjusted by the rebate and then divided by the pre-pilot minimum tick amount (100 CPS) adjusted by the rebate. BATS Y and BX are the two inverted venues that experience the highest percentage increase in posting liquidity while maker-taker venues generally experience the lowest percentage increase.

Table 4.1: Indicative U.S. Exchange Fee Structure

This table reports exchanges pricing, measured in cents per 100 shares (CPS) traded for stocks price above $1 and the percentage incentive change (Incentive%). Fee is the exchange charge (pay) in the maker-taker (taker-maker) market. Rebate is the exchange pay (charge) in the maker-taker (taker-maker) market. Net fee is defined as the sum of the taker fee and the maker rebate, which is the exchange revenue per 100 shares traded. For simplicity, the fee is the highest rate the exchange can charge and rebate is the highest rate below the fee given its pre-determined fee structure in their pricing table. Incentive% is measured as the post pilot minimum tick (500 CPS) adjusted by the rebate divided by the pre pilot minimum tick (100 CPS) adjusted by the rebate. * IEX only charge fees for non-displayed liquidity for both adding and removing liquidity, please see more details here: https://www.iextrading.com/trading/alerts/2016/036/ .

Exchange	Fee Model	Fee (CPS)	Rebate (CPS)	Net Fee (CPS)	Incentive%
BX	Taker-Maker	-17	19	2	5.94
BATS Y	Taker-Maker	-15	18	3	5.88
EDGA	Taker-Maker	-2	5	3	5.21
NASDAQ	Maker-Taker	30	-29	1	4.10
ARCA	Maker-Taker	30	-29	1	4.10
BATS Z	Maker-Taker	30	-29	1	4.10
NYSE	Maker-Taker	27.5	-26	1.5	4.17
AMEX	Maker-Taker	30	-25	5	4.20
PHLX	Maker-Taker	29	-23	6	4.25
EDGX	Maker-Taker	29	-20	9	4.33
CHX	Maker-Taker	30	-20	10	4.33
IEX	Taker-Taker*	9	9	18	5.40

4.5 Data, Sample Selection, and Methodology

A. Data Source

The data examined in this study are processed by the Market Quality Dashboard (MQD)[30] developed and managed by Capital Markets CRC. The data includes end-of-day security level metrics, calculated from Thomson Reuters Tick History (TRTH) data. The TRTH data includes level one bid and ask quotes, and trade details for all stocks in U.S. markets. The trading statistics include trading price, trading volume and qualifiers. For each quote and trade, TRTH reports time stamps to the nearest millisecond. In contrast to SIP NBBO data that credits one market with the NBBO quote, we construct and evaluate the NBBO in a way that multiple exchanges are credited with the best price in the market, in the event of ties, to evaluate quote quality in LOBs.

B. Sample Selection

Consistent with the criteria set out by FINRA, the sample of tick size pilot securities are selected based on the following criteria during the measurement period[31]: market capitalization less than \$3 billion; closing price greater than \$2 on the last day of the measurement period; each daily closing price greater than \$1.5; consolidated average daily trading volume less than one million shares and volume-weighted average price greater than \$2. We exclude stocks with corporate actions such as mergers and acquisitions, switches in listing market, and prices below \$1 during the sample period since the minimum-tick size for such stocks is less than 1 cent.

[30] MQD website: https://www.mqdashboard.com.

[31] Three-month period ending at least 30 days prior to the effective start date of the pilot on October 3, 2016.

Table 4.2: Summary Statistics

This table summarizes the number of securities, average NBBO quoted spread, and firm characteristics for each tick constrained group established prior to the commencement of the tick size pilot. Panel A reports the number of securities, Panel B reports the average national NBBO quoted spread measured in dollars and Panel C reports the average market capitalization, high-low volatility and average closing price of each pilot groups prior our sample period. The high-low volatility is measured by the difference of daily highest price minus lowest price over the closing price. The sample period is two months pre and post the U.S. tick size pilot implemented on October 3, 2016. Pre period is from August 1, 2016 to September 30, 2016. The half trading day has been excluded. Due to the three treatment groups were implemented gradually in October, the post period starts from November 1, 2016 to December 31, 2016. C stands for tick size pilot control group while G1, G2 and G3 represents treatment group 1, 2, and 3 respectively. Most, Medium and Least Tick Constrained group refers to average quoted spread in pre-period is less than 5 cents, between 5 cents and 10 cents, and greater than 10 cents respectively.

Tick-constrained Groups	**C**	**G1**	**G2**	**G3**	**Total**
Panel A: Number of Securities					
Most Tick Constrained	534	177	166	180	1057
Medium Tick Constrained	260	82	82	78	502
Least Tick Constrained	369	124	136	128	757
Total	1163	383	384	386	2316
Panel B: Average NBBO quoted spread ($)					
Most Tick Constrained	0.03	0.03	0.03	0.03	0.03
Medium Tick Constrained	0.07	0.07	0.07	0.07	0.07
Least Tick Constrained	0.32	0.33	0.33	0.29	0.32
Average	0.13	0.14	0.14	0.12	0.13
Panel C: Firm Characteristics					
Average Market Cap	724,680,553	719,574,515	701,899,309	717,250,460	718,819,542
Average High-Low Volatility	0.0285	0.0301	0.0307	0.0292	0.0293
Average Closing Price	22.68	22.72	21.98	23.81	22.76

Table 4.2 summarizes the number of securities and the average NBBO quoted spread for each of the stock groups included in the tick size pilot across three tick classifications: most, medium, and least tick constrained. Most (medium or least) tick constrained refers to the NBBO quoted spread below 5 cents (5 to 10 cents or above 10 cents) prior to our sample period. Panel A reports the number of securities for each tick classification under each control and treatment groups. Around 45 percent of tick size pilot stocks fall into the most tick constrained group. Panels B and C reports the average NBBO quoted spread, market capitalization, volatility and closing price. The NBBO quoted spread is computed as the difference between the national best bid and offer quotes over the national level midpoint. The results indicate the NBBO quoted spread and firm characteristics are identical across control and treatment groups.

C. Methodology

To examine the relation of exchange access fees and tick size and market quality, we conduct univariate and multivariate analyses. The implementation of the tick size pilot lends itself to the use of the difference-in-differences (DID) framework. In the univariate analysis, we test the pre- and post- mean market share for each treatment group, and conduct DID tests for each treatment group relative to the control group.

We apply the following DID framework to assess the interaction between access fees, tick size and market quality:

$$Y_{i,t} = \alpha_0 + \beta_1 \times factor\left(Treat_i \times PilotOn_t \times FeeDummy_m\right) + \gamma VIX_t + \varepsilon_{it}, \quad (30)$$

and the following DID regression specification to assess the market-wide impact of the tick size change:

$$Y_{i,t} = \alpha_0 + \beta_1 \times factor\left(Treat_i \times PilotOn_t\right) + \gamma VIX_t + \varepsilon_{it}, \quad (31)$$

where $Y_{i,t}$ is a series of dependent variables including trading volume, market share, relative effective spread, relative realized spread, relative price impact, intraday volatility, variance ratio, percentage of time at NBBO, and NBBO quoted depth on day *i* and security *t*; α_0 is the intercept; and $Treat_i$ is the factor variable which is 0 (1, 2, or 3) if the security *i* is in the control group (treatment group 1, 2, or 3 respectively). The main effects of these factors are included, together with the interaction effects, but are not reported in the tables for parsimony reasons; $PilotOn_t$ is a dummy variable set to one for day *t* post the pilot period and zero otherwise; $FeeDummy_m$ is a dummy variable that is one if the market fee structure is the inverted venue and zero for the maker-taker venues; and VIX_t is the closing value of the Chicago Board Options Exchange (CBOE) volatility index for day *t* to control for the market-wide effect.[32]

4.6 Empirical Results

A. Tick Size and Exchange Access Fee Structure

This section assesses the interaction between exchange fees structures of market-taker and inverted venues, as firms experience an increase in the minimum tick.

A1. Does an increase in minimum tick size alter the relative market share of venues with different exchange access fee structures?

Prediction: Proposition 1 in Section 4.3 above shows that the information content of the order flow in the inverted venue must fall relative to the maker-taker venue as the minimum tick size, θ, increases. A higher proportion of liquidity traders either depart

[32] We also have run regressions without controlling for *VIX* and results are consistent and robust. For parsimony, we only report regression coefficients of interest.

maker-taker venues for cheaper inverted venues or are forced to depart the lit market altogether due to the increasing price sensitivity of liquidity, i.e., uninformed order flow. The increase is greater, the higher is the minimum tick size and the more different are the fee structures in the two competing venues as given by the sum of the two fee structures, $f_t + f_m$, the maker-taker make rebate plus the inverted make fee.

Result: Overall, Figure 1 shows that market share on the inverted venues for treatment firms increased following the minimum tick increase from 1 cent to 5 cents. Table 4.3 reports our univariate results and tests of difference in market share across each of the 12 exchanges and all off-exchange trading activities including dark pools and internalizations (i.e., *OffExch*). Table 4.3 reports that market share increased for the treatment groups relative to the control group on inverted taker-maker markets associated with high exchange access fees (BATS Y and BX[33] by 4.16% and 2.88% in G1, 4.08% and 2.87% in G2, and 6.66% and 5.91% in G3 respectively), and maker-taker venues' share decreased significantly with Nasdaq experiencing the largest drop (-4.38% in G1, -4.16% in G2 and -2.16% in G3). Market share mainly shifted from the maker-taker venues experiencing the smallest increase in incentive to post liquidity, that is, those with the highest maker-taker fee structure to begin with, and diverted to inverted venues which provided the largest increase in such incentive, i.e., the highest sum of $f_t + f_m$, the maker-taker maker rebate plus the inverted maker fee.

[33] Although it is the case that EGDA is classified as one of the inverted venues. The components of its access fee are close to zero (see Table 4.2).

Figure 1: Average Daily Trading Volume on Maker-Taker vs Inverted Markets across Control and Treatment Groups

This figure shows the average daily trading volume on maker-taker and inverted markets across control and three pilot treatment groups represented by different lines from August 1, 2016 to December 31, 2016. The tick size pilot started on October 3, 2016.

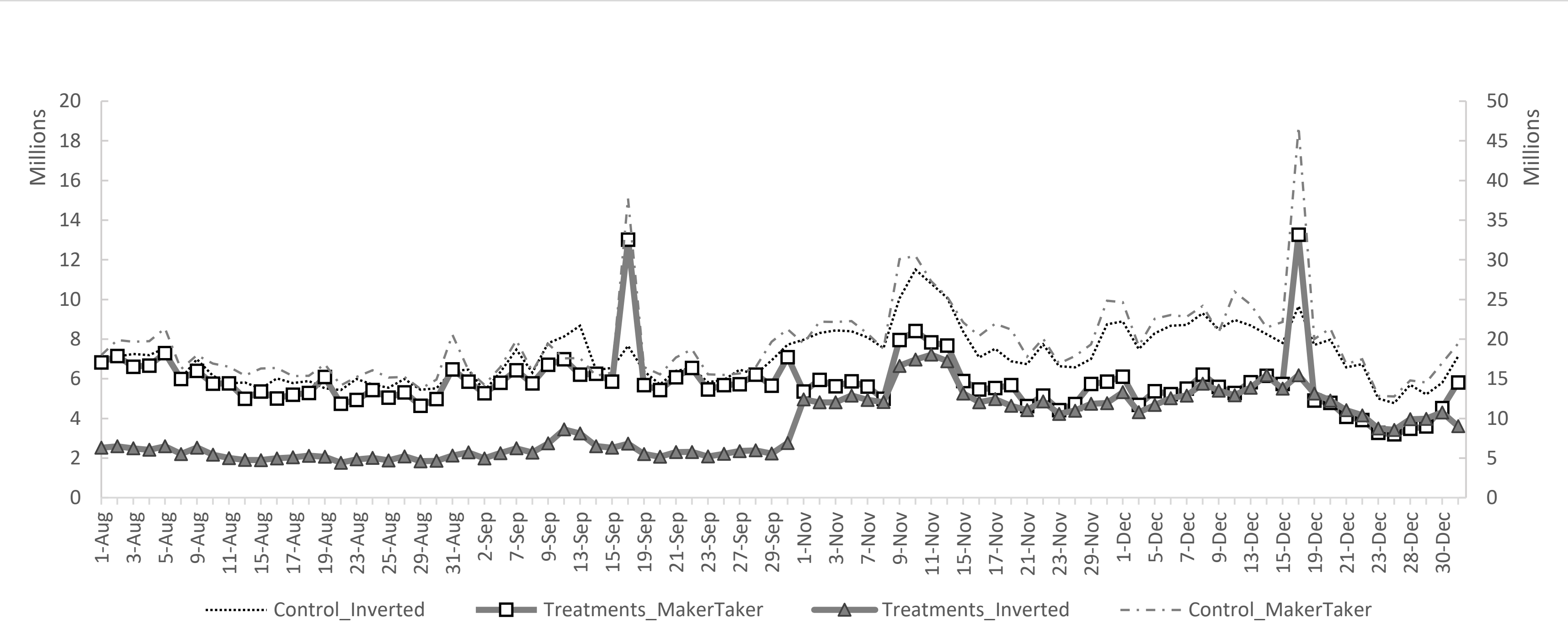

Table 4.3: Market Share Change across Tick Size Pilot Groups

This table reports trading venue market share change pre- and post- the tick size pilot for three treatment groups, and the difference in differences (DID) change pre and post for each treatment group compared with the control group. The sample period is two months pre and post the U.S. tick size pilot implemented on October 3, 2016. Pre period is from August 1, 2016 to September 30, 2016. The half trading day has been excluded. Due to the three treatment groups were implemented gradually in October, the post period starts from November 1, 2016 to December 31, 2016.

	Treatment Group 1				Treatment Group 2				Treatment Group 3			
Markets	**Pre**	**Post**	**DID**	**t.stats**	**Pre**	**Post**	**DID**	**t.stats**	**Pre**	**Post**	**DID**	**t.stats**
BATS Y	3.81	7.85	4.16	25.54	3.88	7.85	4.08	23.26	3.48	10.02	6.66	35.76
BX	2.57	5.67	2.88	23.39	2.62	5.73	2.87	23.00	2.46	8.60	5.91	32.51
EDGA	2.24	1.86	-0.08	-1.68	2.39	1.86	-0.24	-3.84	2.23	2.06	0.12	1.87
NASDAQ	20.32	16.45	-4.38	-7.32	21.56	17.91	-4.16	-7.03	20.64	18.99	-2.16	-3.71
EDGX	6.39	5.22	-1.53	-7.28	6.68	5.16	-1.88	-9.75	6.75	6.02	-1.09	-4.79
NYSE	11.20	8.04	-3.19	-6.84	9.78	8.34	-1.46	-3.45	10.22	9.26	-0.98	-2.07
ARCA	7.49	5.64	-1.49	-8.35	7.48	5.68	-1.45	-7.92	7.35	6.15	-0.85	-3.17
BATS	5.94	4.57	-0.33	-1.97	6.34	4.56	-0.75	-4.51	5.94	5.21	0.31	2.08
CHX	0.10	0.08	-0.10	-2.30	0.44	0.09	-0.44	-2.04	0.05	0.17	0.03	0.65
PHX	0.66	0.70	0.01	0.41	0.68	0.70	-0.02	-0.43	0.63	1.00	0.32	7.11
AMEX	0.39	0.22	-0.11	-1.78	0.17	0.19	0.08	1.27	0.07	0.09	0.07	1.33
IEX	1.30	2.17	0.42	2.70	1.48	2.21	0.28	1.52	1.66	2.50	0.38	2.08
OffExch	38.11	41.52	3.89	5.05	37.17	39.74	3.05	4.34	39.35	29.95	-8.92	-12.28

Table 4.4: Trading Volume and Market Share Across Exchange Fee Structure and Tick Groups_DDD

This table shows key coefficients of the impact of the U.S. tick size pilot across exchange fee structures on trading volume and market share for different tick groups using difference-in-difference-in-differences (DDD). The most (medium/least) tick constrained refers to stock whose NBBO quoted spread is lower than 5 cents (5 to 10 cents/greater than 10 cents) prior to the tick size pilot and the result is displayed in panels A, B, and C, respectively. The sample period is two months pre and post the U.S. tick size pilot implemented on October 3, 2016. Pre period is from August 1, 2016 to September 30, 2016. The half trading day has been excluded. Due to the three treatment groups were implemented gradually in October, the post period starts from November 1, 2016 to December 31, 2016. Post is the dummy variable which is 1 for dates after October 3, 2016 and zero otherwise. TakerMaker is a dummy variable that is 1 if the exchange fee structure is an inverted taker-maker market and 0 for a maker-taker market. G1, G2 and G3 refer to tick size pilot treatment group 1, 2, and 3, respectively. Individual dummy variables are included for Post, TakerMaker and the three treatment groups but not reported. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**Log(Volume)**	**MktShare**	**Log(Value)**	**MktShare$**
Panel A: Most Tick Constrained				
Post*TakerMaker*G1	0.716***	4.697***	0.686***	4.697***
	(0.022)	(0.126)	(0.028)	(0.126)
Post*TakerMaker*G2	0.688***	4.559***	0.643***	4.559***
	(0.023)	(0.130)	(0.028)	(0.130)
Post*TakerMaker*G3	0.845***	5.552***	0.807***	5.552***
	(0.022)	(0.127)	(0.028)	(0.127)
Observations	717,857	717,857	717,857	717,857
Adjusted R^2	0.04	0.101	0.026	0.101
Panel B: Medium Tick Constrained				
Post*TakerMaker*G1	0.541***	3.967***	0.400***	3.968***
	(0.037)	(0.233)	(0.049)	(0.233)
Post*TakerMaker*G2	0.559***	3.618***	0.477***	3.619***
	(0.037)	(0.231)	(0.049)	(0.231)
Post*TakerMaker*G3	0.616***	3.238***	0.549***	3.240***
	(0.038)	(0.234)	(0.049)	(0.234)
Observations	306,789	306,789	306,789	306,789
Adjusted R^2	0.037	0.1	0.017	0.1
Panel C: Least Tick Constrained				
Post*TakerMaker*G1	0.204***	2.864***	-0.008	2.865***
	(0.040)	(0.305)	(0.050)	(0.305)
Post*TakerMaker*G2	0.117***	2.713***	-0.123**	2.713***
	(0.039)	(0.299)	(0.049)	(0.299)
Post*TakerMaker*G3	0.236***	1.702***	0.012	1.704***
	(0.040)	(0.304)	(0.050)	(0.304)
Observations	344,397	344,397	344,397	344,397
Adjusted R^2	0.015	0.093	0.006	0.093

Table 4.3 also shows the off-exchange trading volume increased for treatment groups 1 and 2 (by 3.89% and 3.05%, respectively) but decreased for treatment group 3 (-8.92%) which is subject to the ‘trade-at’ rule. The decline in the overall lit market to the benefit of off-exchange trading is consistent with *Proposition 1* due to the rise in the minimum tick acting as a tax on all uninformed, i.e., liquidity, trading in lit markets, combined with a relative transfer of uninformed trading from maker-taker to inverted markets.

Using the difference-in-difference-in-differences (DDD) methodology, coefficient estimates of Equation (30), reported in Table 4.4, show that post the tick size change, inverted venues, relative to maker-taker venues, experienced a significant increase in trade volume and market shares, in both volume and value terms. The findings are consistent across tick- constrained groups reported in Panels A-C. These relative improvements in value and volume in inverted markets represent the greater relative flight of uninformed trades from the highest make-subsidy maker-taker venues to the highest take-subsidy venues, alternatively off-market or trade cessation, as predicted by *Proposition 1*.

A2. Does quote quality and price discovery alter in the different venues in response to the minimum tick size increase?

Prediction: Proposition 1 in Section 4.3 indicates that, as the minimum tick size widens, the percentage of time that inverted venues match the NBBO, together with the NBBO quoted depth, should increase as these venues relatively benefit as uninformed traders relocate from maker-taker venues or flee the lit market entirely. Many proxies for quote quality are assessed, including percentage time at the NBBO, NBBO quoted depth, exchange quoted spread, and exchange intraday volatility. An advantage of our data is that, in contrast to SIP NBBO data that credits one market with the NBBO quote, we construct the NBBO such that multiple exchanges are credited with the best price in the market, in the event of ties. Exchange intraday day volatility (*ExchIntraVola*) is defined as the standard deviation of the exchange quote midpoint return.

Variance ratios across multiple time intervals are examined as proxies for informational efficiency consistent with Lo and MacKinlay (1988) and Chordia, Roll, and Subrahmanyam (2008), specifically, 1 to 10 seconds, 10 to 60 seconds, and 1 to 5 minutes as follows:

$$ExchVR_{i,t} = abs\left(\operatorname{var}\left(r_{i,t}\right)\times x \big/ \operatorname{var}\left(r_{i,x\times t}\right)-1\right), \tag{32}$$

where $\operatorname{var}\left(r_{i,t}\right)$ refers to the variance of the return during the t^{th} time interval for i, $\operatorname{var}\left(r_{i,x\times t}\right)$ refers to the variance of the return during the $x\times t^{th}$ time interval for i. Specifically, when a stock's price follows a random walk, the variance of its returns is a linear function of the measurement frequency, i.e., $\operatorname{var}\left(r_{i,x\times t}\right)$ is k times larger than $\operatorname{var}\left(r_{i,t}\right)$, and the specified relative variance ratio has been normalized to zero since 1 is subtracted.

The exchange-level quoted spread (*ExchQuoSpread*) is the difference between each venue's best bid and offer quotes over the midpoint. *ExchQuoSpread* differs from the NBBO quoted spread as the former uses each venue's best bid and offer quotes while the latter use the national best bid and offer. The exchange best bid and offer in a venue does not necessarily represent the national best bid and/or offer quotes. Specifically,

$$ExchQuoSpread_{i,t} = \left(a_{i,t}-b_{i,t}\right)\big/b_{i,t}\,, \tag{33}$$

where $a_{i,t}$ is the exchange's best ask (buy) price and $b_{i,t}$ is the exchange's best bid (sell) price.

Table 4.5 reports that, as predicted, and relative to control group firms and maker-taker venues, the percentage of time exchanges quote at the NBBO and NBBO quoted depth on the inverted venues has increased for the treatment groups that experienced a widening in the minimum tick to 5 cents. This is in contrast with Foucault, Kadan, and Kandel (2013), who conclude that the welfare of maker-taker participants should improve with a higher minimum tick size since a higher minimum tick size enables maker-taker markets to work better. In their framework, which neglects the role of informed traders, both taker-maker and inverted markets fee structures wash-out to the detriment of traders when the minimum tick is set at a suboptimal level. In fact, maker-taker venues have experienced a sizeable drop in market share, trade volume and value, and depth. The variance ratios in inverted venues are closer to zero which suggests the price discovery has been enhanced, consistent with our theoretical prediction of a rise in information content in the order flow for inverted venues. Notwithstanding the fact that, as also predicted, the information content increase in maker-taker markets was even greater due to the shift in trading to inverted markets and off-market, especially the loss of uninformed traders from the maker-taker market. The exchange quoted spread fell significantly more in the inverted venues relative to maker-taker venues due to the increased depth in the LOB that came about from the significantly greater rise in the implicit spread in the inverted market because of increased information content. It is of interest that quoted and implicit spreads move in opposite directions. Similarly, exchange intraday volatility decreased in the inverted markets, due to the rise in depth and implicit spread, after controlling for VIX, and the changes are generally consistent across each of the tick-constrained groups.

Table 4.5: NBBO, Volatility, and Price Discovery Across Exchange Fee Structure and Tick Groups_DDD

This table shows key coefficients of the impact of the U.S. tick size pilot across exchange fee structures on NBBO, volatility and price discovery for different tick groups using difference-in-difference-in-differences (DDD). The most (medium/least) tick constrained refers to stock whose NBBO quoted spread is lower than 5 cents (5 to 10 cents/greater than 10 cents) prior to the tick size pilot and the result is displayed in panels A, B, and C, respectively. The sample period is two months pre and post the U.S. tick size pilot implemented on October 3, 2016. Pre-period is from August 1, 2016 to September 30, 2016. The half trading day has been excluded. Due to the three treatment groups were implemented gradually in October, the post period starts from November 1, 2016 to December 31, 2016. Post is the dummy variable which is 1 for date after October 3, 2016 and 0 otherwise. TakerMaker is the dummy variable that is 1 if the exchange fee structure is an inverted taker-maker venue and 0 for a maker-taker venue. G1, G2, and G3 refer to tick size pilot treatment group 1, 2, and 3, respectively. Individual dummy variables are included for Post, TakerMaker and the three treatment groups but not reported. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**NBBO%**	**Log(NBBODepth)**	**ExchQuoSpread**	**ExchIntraVola**	**ExchVR1t10s**	**ExchVR10t60s**	**ExchVR1t5m**
Panel A: Most Tick Constrained							
Post*TakerMaker*G1	15.723***	0.705***	-513.956***	-1.477***	-0.060***	-0.086***	-0.089***
	(0.334)	(0.022)	(13.201)	(0.032)	(0.003)	(0.003)	(0.003)
Post*TakerMaker*G2	16.303***	0.685***	-521.619***	-1.616***	-0.070***	-0.088***	-0.081***
	(0.344)	(0.022)	(13.625)	(0.033)	(0.003)	(0.003)	(0.003)
Post*TakerMaker*G3	13.977***	0.744***	-454.065***	-1.435***	-0.083***	-0.086***	-0.073***
	(0.334)	(0.022)	(13.206)	(0.032)	(0.003)	(0.003)	(0.003)
Observations	717,857	717,857	709,130	716,860	716,551	716,459	716,230
Adjusted R^2	0.311	0.228	0.054	0.137	0.13	0.159	0.114
Panel B: Medium Tick Constrained							
Post*TakerMaker*G1	12.423***	0.832***	-506.571***	-0.568***	-0.042***	-0.064***	-0.038***
	(0.467)	(0.038)	(18.654)	(0.052)	(0.005)	(0.005)	(0.005)
Post*TakerMaker*G2	12.704***	0.709***	-626.815***	-0.889***	-0.021***	-0.056***	-0.052***
	(0.461)	(0.038)	(18.440)	(0.051)	(0.005)	(0.005)	(0.005)
Post*TakerMaker*G3	10.968***	0.769***	-389.250***	-0.421***	-0.038***	-0.037***	-0.025***
	(0.468)	(0.038)	(18.697)	(0.052)	(0.005)	(0.005)	(0.005)
Observations	306,789	306,789	289,255	305,110	304,731	304,677	304,564
Adjusted R^2	0.214	0.128	0.132	0.175	0.027	0.088	0.082
Panel C: Least Tick Constrained							
Post*TakerMaker*G1	12.733***	0.906***	-313.949***	0.008	-0.037***	-0.038***	-0.032***
	(0.476)	(0.041)	(14.182)	(0.047)	(0.005)	(0.005)	(0.005)
Post*TakerMaker*G2	11.186***	0.856***	-266.111***	-0.029	-0.042***	-0.046***	-0.034***
	(0.467)	(0.040)	(13.780)	(0.046)	(0.005)	(0.005)	(0.005)
Post*TakerMaker*G3	14.031***	0.984***	-255.519***	0.280***	-0.026***	-0.025***	-0.024***
	(0.473)	(0.041)	(13.995)	(0.047)	(0.005)	(0.005)	(0.005)
Observations	344,397	344,397	308,136	339,570	338,539	338,401	338,154
Adjusted R^2	0.124	0.079	0.178	0.143	0.009	0.07	0.062

A3. Do transaction costs and price impact on markets vary with exchange access fee structures and minimum tick size?

Prediction: The effective spread should rise by more in inverted markets because the spread is lower to begin with, prior to the increase in the minimum tick size, due to the subsidy paid to market orders. For example, consider the indicative maker rebate on Nasdaq for displayed liquidity which is 29 cents per 100 shares traded (CPS) and taker fee, f_t, is 30 CPS[34], while on Nasdaq/BX, an inverted market, the maker fee, f_m, is 19 CPS and taker rebate is 17 CPS. A change in the minimum tick size from 1 cent to 5 cents, and hence 100 to 500 CPS, results in a percentage change in the take, i.e., market, order cost increase on the maker-taker market of 308 %[35] and 482%[36] on the inverted market. Thus, the rise in the minimum tick size represents an increase of 57%[37] in the relative cost of *take* trades on the inverted market. Moreover, *Proposition 1* in Section 4.3 implies that, compared with maker-taker venues, the relative price impact should decrease on inverted venues for the most tick-constrained group because the proportion of informed trades must rise by less than in maker-taker venues.

The relative effective spread is twice the signed difference between the transaction price and the midpoint of the *national* bid and offer quotes at the time of the transaction. Specifically,

$$EffSpread_{i,t} = 2 \times q_{i,t} \times \left(p_{i,t} - m_{i,t} \right) / m_{i,t}, \tag{34}$$

[34]Exchange access fees for protected quotes in the equities markets are bound by Rule 610 of Regulation NMS; fees are capped at 30 CPS traded.

[35] [(500 + 30) - (100 + 30)] / (100 + 30) = 308%.

[36] [(500 - 17) - (100 - 17)] / (100-17) = 482%.

[37] (4.8193-3.0769)/ 3.0769 = 57%.

where $p_{i,t}$ is the transaction price for security i at time t, $m_t = (a_{i,t} + b_{i,t})/2$ is the exchange midpoint quote of the best bid and ask price, and $q_{i,t}$ is an indicator variable that equals 1 if the trade is buyer-initiated and -1 if the trade is seller-initiated.

The relative realized spread is a measure of profits earned by market makers. Previous studies have set the time lag, τ, to five minutes after the trade. The choice of this time horizon should be sufficiently long to incorporate the permanent impact of the trade and thus to ensure that quotes are subsequently stabilized, temporary effects are dissipated, and there is a sufficiently extended period for liquidity providers to close their positions (Conrad, Wahal, and Xiang, 2015). In today's ultra-high frequency trading environment, which has upgraded trading systems with an accuracy within 100 microseconds[38], five minutes is excessively long. Like Conrad, Wahal, and Xiang (2015), we estimate realized spreads from one second to ten minutes (i.e., 1, 10, 30, 60, 300, and 600 seconds) after each trade for a robustness check.[39] The relative realized spread is then calculated as twice the signed difference between the transaction price and the midpoint of the national bid and offer quotes one second and five seconds after the transaction. Specifically,

$$ReaSpread_{i,t} = 2 \times q_{i,t} \times \left(p_{i,t} - m_{i,t+\tau}\right)/m_{i,t}. \tag{35}$$

The relative price impact is defined as the signed change between the NBBO midpoint of the quote one second to ten minutes (i.e., 1, 10, 30, 60, 300, and 600 seconds)[40] after the

[38] SEC Approves Plan to Create Consolidated Audit Trail, available at https://www.sec.gov/news/pressrelease/2016-240.html

[39] Realized spread results for different time interval are consistent, to save space, we only report 30 seconds, 1 minute, and 5 minutes' interval results in the regression table.

[40] Similarly, price impact results for different time interval are consistent, so to save space, we only report 30-second, 1-minute and 5-minute interval results in the regression table.

trade and the NBBO midpoint of the prevailing quote at the time of the trade. It captures the information that is revealed by the trade. A decline in the price impact indicates a decline in adverse selection costs. Specifically,

$$PrImpact = 2 \times q_{i,t} \times \left(m_{i,t+\tau} - m_{i,t} \right) / m_{i,t}, \tag{36}$$

where $m_{i,t+\tau}$ is the midpoint at one second and five seconds after the trade. We follow the Lee and Ready (1991) approach to mark each trade as buyer- or seller-initiated.

As shown in Table 4.6, compared with maker-taker venues, both the inverted venue's NBBO effective spread and NBBO realized spreads increased for the most tick-constrained group (Panel A), and price impact decreased in the inverted taker-maker markets using the DDD method. Panels B and C report the effective spreads reduced for the medium and least tick-constrained groups, which do not experience the 'true' tick size increase as their NBBO quoted spreads were higher than 5 cents prior to the tick size pilot. The price impact also dropped for those two groups.

Table 4.6: Transaction Cost and Price Impact Across Exchange Fee Structure and Tick Groups_DDD

This table shows key coefficients of the impact of the U.S. tick size pilot across exchange fee structures on the relative transaction cost and price impacts for different inverted-market tick groups using difference-in-difference-in-differences (DDD). The most (medium/least) tick constrained refers to stock whose NBBO quoted spread is lower than 5 cents (5 to 10 cents/greater than 10 cents) prior to the tick size pilot and the results are displayed in panels A, B, and C, respectively. The sample period is two months pre and post the U.S. tick size pilot implemented on October 3, 2016. Pre-period is from August 1, 2016 to September 30, 2016. The half trading day has been excluded. Due to the fact that the three treatment groups were implemented gradually in October, the post-period starts from November 1, 2016 to December 31, 2016. Post is the dummy variable which is 1 for date after October 3, 2016 and 0 otherwise. TakerMaker is a dummy variable that is 1 if the exchange fee structure is an inverted taker-maker market and 0 for maker-taker market. G1, G2, and G3 refer to tick size pilot treatment group 1, 2, and 3, respectively. Individual dummy variables are included for Post, TakerMaker and the three treatment groups but not reported. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**EffSpread**	**ReaSpread30s**	**ReaSpread1m**	**ReaSpread5m**	**PrImpact30s**	**PrImpact1m**	**PrImpact5m**
Panel A: Most Tick Constrained							
Post*TakerMaker*G1	3.416***	7.536***	7.418***	6.566***	-4.188***	-4.071***	-3.219***
	(0.573)	(0.581)	(0.601)	(0.705)	(0.582)	(0.606)	(0.729)
Post*TakerMaker*G2	3.310***	7.673***	7.262***	6.616***	-4.208***	-3.798***	-3.153***
	(0.590)	(0.597)	(0.618)	(0.725)	(0.599)	(0.624)	(0.750)
Post*TakerMaker*G3	5.457***	10.932***	10.602***	10.607***	-5.361***	-5.032***	-5.038***
	(0.573)	(0.580)	(0.600)	(0.704)	(0.581)	(0.606)	(0.729)
Observations	709,819	701,761	701,761	701,761	701,747	701,747	701,747
Adjusted R^2	0.055	0.021	0.018	0.012	0.014	0.014	0.012
Panel B: Medium Tick Constrained							
Post*TakerMaker*G1	-3.967**	1.611	0.924	1.875	-5.440***	-4.752***	-5.705***
	(1.826)	(1.406)	(1.444)	(1.640)	(1.710)	(1.757)	(1.954)
Post*TakerMaker*G2	-1.92	2.152	0.986	1.146	-3.264*	-2.099	-2.258
	(1.800)	(1.382)	(1.420)	(1.612)	(1.681)	(1.727)	(1.921)
Post*TakerMaker*G3	-6.263***	-2.760*	-3.592**	-3.258**	-2.980*	-2.149	-2.486
	(1.837)	(1.411)	(1.449)	(1.645)	(1.716)	(1.763)	(1.960)
Observations	297,465	289,895	289,895	289,895	289,871	289,871	289,871
Adjusted R^2	0.006	0.003	0.002	0.001	0.002	0.003	0.003
Panel C: Least Tick Constrained							
Post*TakerMaker*G1	-6.827*	1.387	2.05	4.703	-10.370***	-11.007***	-13.795***
	(4.075)	(3.316)	(3.301)	(3.322)	(3.710)	(3.709)	(3.892)
Post*TakerMaker*G2	-8.991**	4.315	3.905	4.851	-13.533***	-13.121***	-14.184***
	(4.016)	(3.275)	(3.261)	(3.281)	(3.664)	(3.664)	(3.845)
Post*TakerMaker*G3	-10.004**	-3.449	-3.333	-1.964	-7.583**	-7.498**	-9.368**
	(4.054)	(3.299)	(3.285)	(3.305)	(3.691)	(3.691)	(3.873)
Observations	324,346	309,805	309,805	309,805	309,351	309,351	309,351
Adjusted R^2	0.006	0.002	0.002	0.002	0.002	0.002	0.003

B: Tick Size and Overall Economic Impact

Turning to market-wide effects of the tick size pilot on treatment and control groups[41], exclusive of fees, we examine the following three questions:

B1. Does tick size affect consolidated and off-exchange trading volume in small capitalization securities?

Prediction: We use the consolidated trading volume as a proxy for the net economic welfare. As the tick size increased by 400% from a penny to a nickel, this increase in transaction cost should lead to a fall in the consolidated trading volume, as described in Section 4.3 above. As the tick size increased in lit markets, traders can still trade at the midpoint for stocks in treatment groups 1 and 2 in both crossing networks and dark pools; hence, the off-market share should tend to increase. However, treatment group 3, which is subject to the 'trade-at' rule, is expected to decrease. Kwan, Masulis, and McInish (2015) find that the uniform minimum tick-constrained bid-ask spreads resulted in large limit order queues, and dark pool allow traders to bypass existing limit order queues with minimal price improvement. Also, Foley and Putnins (2014) find that when Canada and Australia implemented minimum price improvements, the level of dark trading decreased.

As shown in Table 4.7, displaying the interaction effects on the post tick size pilot and treatment groups using difference-in-differences (DID) methodology, the consolidated trading volume decreased for all treatment groups for the most tick-constrained stocks as tick size widens. This evidence indicates the tick size pilot reduced

[41] To test the overall impact of the tick size pilot for all stocks in aggregate, we undertook further analysis using daily volume-weighted exchange-level data rather than security level data as a robustness check. The result is generally qualitatively consistent, especially with the most tick-constrained group. The Internet appendix that accompanies this chapter may be found at https://goo.gl/Lz227q.

trading volume for small stocks, rather than improving it, contrary to the suggested outcome for the pilot.

Table 4.7 also reports that the off-exchange market share increased for treatment group 1 and 2 but decreased for treatment group 3. The decrease in off-exchange trading for treatment group 3 is economically and statistically significant across all tick groups. This is consistent with expectations, as the treatment group 3 is subject to 'trade-at', which requires brokerages to route trades to public exchanges, unless they can execute the trades at a meaningfully better price than what is available in the public lit markets.

Table 4.7: Trading Volume and Market Share Across Tick Groups_DID

This table shows key coefficients of the impact of the U.S. tick size pilot on trading volume and market share for different tick groups using difference-in-differences (DID). The most (medium/less) tick constrained refers to stock whose NBBO quoted spread is lower than 5 cents (5 to 10 cents/greater than 10 cents) prior to the tick size pilot and the result is displayed in panels A, B and C, respectively. The sample period is two months pre and post the U.S. tick size pilot implemented on October 3, 2016. Pre period is from August 1, 2016 to September 30, 2016. The half trading day has been excluded. Due to the three treatment groups were implemented gradually in October, the post period starts from November 1, 2016 to December 31, 2016. Post is the dummy variable which is 1 for date after October 3, 2016 and 0 otherwise. G1, G2, and G3 refer to tick size pilot treatment group 1, 2, and 3, respectively. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**Log(ConsVol)**	**Log(LitConsVol)**	**LitShare**	**Log(OffExchVol)**	**OffExchShare**
Panel A: Most Tick Constrained					
Post*G1	-0.092***	-0.157***	-4.224***	0.073***	4.511***
	(0.023)	(0.023)	(0.270)	(0.022)	(0.263)
Post*G2	-0.182***	-0.224***	-3.613***	0.034	4.391***
	(0.023)	(0.024)	(0.278)	(0.023)	(0.272)
Post*G3	-0.062***	0.102***	9.229***	-0.322***	-9.097***
	(0.023)	(0.023)	(0.269)	(0.022)	(0.263)
Observations	89,444	89,157	89,157	88,415	88,415
Adjusted R^2	0.003	0.006	0.037	0.01	0.04
Panel B: Medium Tick Constrained					
Post*G1	-0.037	-0.103**	-3.376***	0.056	3.626***
	(0.039)	(0.042)	(0.470)	(0.038)	(0.462)
Post*G2	-0.119***	-0.128***	-1.102**	-0.025	1.777***
	(0.039)	(0.042)	(0.470)	(0.038)	(0.461)
Post*G3	-0.150***	0.004	9.664***	-0.504***	-9.628***
	(0.040)	(0.043)	(0.478)	(0.039)	(0.468)
Observations	42,289	42,156	42,156	41,751	41,751
Adjusted R^2	0.008	0.01	0.024	0.014	0.027
Panel C: Least Tick Constrained					
Post*G1	-0.006	-0.073	-0.461	-0.025	0.895*
	(0.051)	(0.053)	(0.494)	(0.048)	(0.501)
Post*G2	0.066	0.039	0.523	0.049	-0.226
	(0.049)	(0.051)	(0.480)	(0.047)	(0.487)
Post*G3	-0.120**	0.092*	11.711***	-0.433***	-10.818***
	(0.050)	(0.053)	(0.492)	(0.049)	(0.505)
Observations	62,229	60,834	60,834	59,350	59,350
Adjusted R^2	0.002	0.004	0.02	0.004	0.017

B2. Does market wide quote quality and price discovery improve?

Prediction: If the tick size increases, the incentive to provide liquidity increases. Hence, the proportion of time a market quotes at the NBBO and quoted depth should increase. Furthermore, the exchange-level quoted spread and volatility should decrease as there is greater incentive to place meaningful quotes on each venue for small cap stocks. However, a mandated increase in the minimum tick will give rise to an increase in the national level of quoted spreads. With fewer negotiation levels, the percentage of time at NBBO and NBBO quoted depth should increase as competition amongst liquidity providers increases. Correspondingly, price discovery should improve as the tick size widens, as deeper limit order queues benefit informed traders.

Parameter estimates reported in Table 4.8 confirm exchange-level quoted spreads, intraday volatility and variance ratio decreases, and the national-level quoted spread, the percentage of time at the NBBO and NBBO quoted depth increase as the minimum tick size widens. These results are consistent with either an increased incentive to quote and given fewer negotiation price points at a nickel tick, the likelihood of being at the NBBO increases naturally.

Table 4.8: NBBO, Volatility, and Price Discovery Across Tick Groups_DID

This table shows key coefficients of the impact of U.S. tick size pilot on NBBO, volatility and price discovery for different tick groups using difference-in-differences (DID). The most (medium/least) tick constrained refers to stock whose NBBO quoted spread is lower than five cents (five to ten cents/greater than ten cents) prior the tick size pilot and the result is displayed in panel A, B and C respectively. The sample period is 2 months pre and post the U.S. tick size pilot implemented on October 3, 2016. Pre period is from August 1, 2016 to September 30, 2016. The half trading day has been excluded. Due to the three treatment groups were implemented gradually in October, the post period starts from November 1, 2016 to December 31, 2016. Post is the dummy variable which is 1 for date after October 3, 2016 and 0 otherwise. G1, G2 and G3 refer to tick size pilot treatment group 1, 2 and 3 respectively. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**NBBO%**	**Log(NBBODepth)**	**NBBOQuoSpread**	**ExchQuoSpread**	**ExchIntraVola**	**ExchVR1t10s**	**ExchVR10t60s**	**ExchVR1t5m**
Panel A: Most Tick Constrained								
Post*G1	30.874***	1.672***	0.410***	-202.590***	-0.497***	-0.060***	-0.064***	-0.059***
	(0.161)	(0.010)	(0.010)	(6.022)	(0.015)	(0.002)	(0.001)	(0.002)
Post*G2	31.098***	1.674***	0.424***	-178.808***	-0.519***	-0.068***	-0.068***	-0.055***
	(0.166)	(0.010)	(0.012)	(6.210)	(0.016)	(0.002)	(0.001)	(0.002)
Post*G3	34.726***	1.861***	0.436***	-187.637***	-0.577***	-0.060***	-0.066***	-0.058***
	(0.161)	(0.010)	(0.012)	(6.020)	(0.015)	(0.002)	(0.001)	(0.002)
Observations	777,660	777,660	89,432	765,830	775,431	774,765	774,586	774,201
Adjusted R^2	0.224	0.178	0.132	0.011	0.009	0.016	0.025	0.016
Panel B: Medium Tick Constrained								
Post*G1	23.012***	1.343***	0.155***	-173.405***	-0.114***	-0.060***	-0.068***	-0.042***
	(0.218)	(0.018)	(0.024)	(8.889)	(0.025)	(0.002)	(0.002)	(0.002)
Post*G2	22.904***	1.268***	0.121***	-188.955***	-0.191***	-0.055***	-0.069***	-0.040***
	(0.216)	(0.018)	(0.021)	(8.821)	(0.025)	(0.002)	(0.002)	(0.002)
Post*G3	21.451***	1.231***	0.125***	-142.896***	-0.125***	-0.051***	-0.064***	-0.034***
	(0.218)	(0.018)	(0.019)	(8.935)	(0.026)	(0.002)	(0.002)	(0.002)
Observations	331,980	331,980	42,425	311,039	328,835	328,146	328,049	327,881
Adjusted R^2	0.148	0.086	0.010	0.008	0.002	0.012	0.024	0.012
Panel C: Least Tick Constrained								
Post*G1	13.058***	0.726***	-0.095*	-81.067***	0.100***	-0.041***	-0.057***	-0.028***
	(0.218)	(0.019)	(0.046)	(6.717)	(0.022)	(0.002)	(0.002)	(0.002)
Post*G2	11.941***	0.691***	-0.028	-60.282***	0.058***	-0.035***	-0.052***	-0.036***
	(0.214)	(0.018)	(0.049)	(6.532)	(0.022)	(0.002)	(0.002)	(0.002)
Post*G3	12.925***	0.711***	-0.047	-42.160***	0.189***	-0.039***	-0.046***	-0.026***
	(0.218)	(0.019)	(0.046)	(6.696)	(0.022)	(0.002)	(0.002)	(0.002)
Observations	370,719	370,719	61,017	328,335	362,725	361,223	361,048	360,727
Adjusted R^2	0.046	0.025	0.002	0.002	0.001	0.01	0.017	0.009

B3. How does an increase in the minimum tick impact market wide transaction cost and price impact change?

Prediction: As tick size widens, transaction costs will increase, especially when the stocks were previously the most tick constrained. The realized spread should be higher, as the posting of non-marketable limit orders is encouraged by the implicit subsidy. A nickel tick has created a higher 'barrier' to cross the spread; consequently, only the highly informed orders can now afford to cross the widened spread. Hence, the price impact should increase, as shown in *Proposition 1*.

As shown in Table 4.9, for the most tick-constrained group, the effective spread increased, indicating a higher transaction cost, as well as a higher realized spread. Also, the price impact at different time intervals increased significantly for all three treatment groups suggesting that the information content of trades is higher.

Table 4.9: Transaction Cost, Price Impact, and NBBO Across Tick Groups_DID

This table shows key coefficients of the impact of the U.S. tick size pilot on transaction cost and price impact for different tick groups using difference-in-differences (DID). The most (medium/least) tick constrained refers to stock whose NBBO quoted spread is lower than 5 cents (5 to 10 cents/greater than 10 cents) prior the tick size pilot and the result is displayed in panels A, B, and C, respectively. The sample period is two months pre and post the U.S. tick size pilot implemented on October 3, 2016. Pre period is from August 1, 2016 to September 30, 2016. The half trading day has been excluded. Due to the three treatment groups were implemented gradually in October, the post period starts from November 1, 2016 to December 31, 2016. Post is the dummy variable which is 1 for date after October 3, 2016 and 0 otherwise. G1, G2, and G3 refer to tick size pilot treatment group 1, 2, and 3, respectively. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**EffSpread**	**ReaSpread30s**	**ReaSpread1m**	**ReaSpread5m**	**Primpact30s**	**Primpact1m**	**Primpact5m**
Panel A: Most Tick Constrained							
Post*G1	15.460***	10.374***	10.166***	9.692***	5.196***	5.405***	5.878***
	(0.287)	(0.270)	(0.279)	(0.328)	(0.289)	(0.300)	(0.354)
Post*G2	14.983***	9.894***	9.725***	8.325***	5.105***	5.274***	6.674***
	(0.295)	(0.277)	(0.287)	(0.337)	(0.297)	(0.308)	(0.364)
Post*G3	19.548***	13.278***	12.867***	12.114***	6.318***	6.730***	7.483***
	(0.287)	(0.269)	(0.279)	(0.327)	(0.288)	(0.299)	(0.353)
Observations	750,426	738,936	738,936	738,936	738,922	738,922	738,922
Adjusted R^2	0.036	0.017	0.015	0.009	0.006	0.006	0.005
Panel B: Medium Tick Constrained							
Post*G1	-0.231	-0.401	-0.475	-0.814	0.341	0.414	0.754
	(0.870)	(0.647)	(0.665)	(0.755)	(0.816)	(0.837)	(0.924)
Post*G2	-3.491***	-2.110***	-2.003***	-2.437***	-1.419*	-1.525*	-1.092
	(0.861)	(0.639)	(0.657)	(0.745)	(0.806)	(0.827)	(0.913)
Post*G3	1.013	1.964***	2.357***	2.888***	-0.956	-1.349	-1.878**
	(0.877)	(0.651)	(0.669)	(0.759)	(0.820)	(0.841)	(0.929)
Observations	309,923	300,824	300,824	300,824	300,800	300,800	300,800
Adjusted R^2	0.002	0.002	0.001	0.001	0.0003	0.0003	0.0003
Panel C: Least Tick Constrained							
Post*G1	-7.115***	-8.733***	-9.041***	-6.538***	3.164*	3.470**	1.049
	(1.839)	(1.489)	(1.483)	(1.493)	(1.674)	(1.674)	(1.756)
Post*G2	-1.775	-6.801***	-6.174***	-4.815***	5.877***	5.263***	3.972**
	(1.809)	(1.465)	(1.458)	(1.468)	(1.647)	(1.647)	(1.728)
Post*G3	-1.815	-2.299	-1.584	0.873	1.473	0.636	-1.478
	(1.840)	(1.490)	(1.483)	(1.494)	(1.674)	(1.674)	(1.757)
Observations	334,667	318,554	318,554	318,554	318,098	318,098	318,098
Adjusted R^2	0.0003	0.0003	0.0003	0.0002	0.0003	0.0003	0.0003

4.7 Conclusion

In this chapter, we examine whether exchange access fees matter when tick size changes, using the nickel tick size pilot as an exogenous shock. We extend the fee-structure model provided in chapter 3 to incorporate minimum tick size constraints to show that the limit on fees set by the SEC is too low to enable fees to neutralize the harmful effect of the 400% increase in the minimum tick size during the SEC tick size pilot. Since there is a downward sloping demand schedule for liquidity trades, our model shows that maker-taker venues with the highest cost of make orders, must lose the most market share to both inverted venues and off-market as the minimum tick size is raised. Moreover, both venue types should experience a relative increase in informed order flow as uninformed traders depart, with maker-taker markets experiencing the largest relative increase.

We show, as expected, that inverted venues, which experience a substantial increase in incentive to post non-marketable orders, gain in market share while the maker-taker venues experience a decline. The intuition is that as the tick size widens, the cost of crossing the spread increases, however, the cost is even higher on maker-taker venues since there is an additional fee for taking liquidity. Consequently, one is better off by executing market orders in an inverted fee market since consumers of liquidity receive a rebate to offset a material portion of the widened tick. The increase is higher for the most tick-constrained group. Price discovery becomes more efficient while relative price impact decreases on the inverted venue as its quote quality improves further.

In addition, we find that the consolidated trading volume decreased overall suggesting small capitalization stocks have not attracted increased trading interest despite an increase in the quote quality. The overall transaction cost increased significantly, which leads to a higher price impact as only the highly informed orders can afford to cross the spread. However, the price discovery process is enhanced and less volatile for the most tick-constrained group. In addition, lit market share increased for treatment firms, which are subject to the ‘trade-at’ rule. This indicates that imposing minimum price improvement in opaque venues will restrict off-exchange trading activity.

Chapter 5 Limit Up Limit Down, Exchange Access Fee, and High Frequency Trading

5.1 Introduction

Circuit breakers are the commonly-used mechanism that monitor the market continuously and trigger a trading halt when the price of a security or an index goes beyond a predetermined level, aiming to restrict extreme daily price movement. The market crash in October 1987 triggered the introduction of a market circuit breaker system. Today, many stock exchanges impose price limits on daily price movement, such as the markets in China, Egypt, France, Japan, India, South Korea, Spain, Taiwan, Thailand and the United States.[42] The aim of this chapter is to analyse the impact of the dynamic price limit rule named Limit Up Limit Down (LULD) on U.S. markets and the role of HFT in pushing the price to the limits in maker-taker and taker-maker markets.

After the May 2010 'Flash Crash', the SEC implemented the security-level price limits that halt trading after a security's price experiences a sudden large movement. Effective from April 8, 2013, the LULD, replacing the single stock circuit breakers (SSCBs) [43] that started on June 14, 2010, began to address extraordinary market volatility on a pilot

[42] Gomber, Clapham, Haferkorn, Panz, Jentsch (2016) surveyed the circuit breakers among international trading venues.

[43] Moise and Flaherty (2017) state that some trading pauses were triggered by clearly erroneous trades under the SSCB, which can create the perception of greater volatility. This may discourage trading during high volatility periods as counterparties can lose money if a trade is eventually cancelled. Unlike SSCBs, the LULD aims to prevent all trades in individual securities outside of a specified price band. However, the LULD design may still permit clearly erroneous trades due to LULD parameters being larger than the clearly erroneous guidelines in certain scenarios (for example, for Tier 1 stocks trading above $50 or for Tier 2 stocks trading above $25), or due to a slow LULD reference price update.

basis. The role of circuit breakers is emphasized again during the four-day China stock markets crash in 2016. On January 4, 2016, The China Securities Regulatory Commission (CSRC) started implementing a market-wide circuit breaker that triggered a 15-minute trading halt if the CSI 300 Index falls by 5% (Level 1) from the previous closing price, and the market would be closed after a 7% decline (Level 2). Both thresholds were reached on the first day, and it took only seven minutes to reach Level 2 halt from the re-opening of the Level 1 halt. On January 7, two level thresholds were triggered again in 30 minutes from opening. This leads the CSRC to suspend the circuit breaker rule on the same day.

Subrahmanyam (2013) states that "the general notion is that rapidly falling prices may exacerbate panic amongst investors and cause limit orders to become unfairly stale." There are debates about price limit rules in academic literature. The advocates claim that a circuit breaker could restrict upfront trading, which is trading motivated by the fear of a future liquidity shock rather than true liquidity needs (Draus and Achter, 2015). In addition, large order imbalances during rapid market movements might result in prices deviating from fundamentals, hence, allowing orders to accumulate and then batching them may lead to better execution quality by setting a ceiling price and providing a cooling off period (Greenwald and Stein, 1991; Kim and Rhee, 1997; Kim and Sweeney, 2002; Hsieh, Kim, and Yang, 2009; Subrahmanyam, 2013). It is also supposed to counter overreaction, and not interfere with trading activity. Moreover, regulators can use price limits to overcome stock price manipulation by large investors (Kim and Park, 2010). However, the critics argue that price limits can have several adverse effects: interfere with trading due to restrictions imposed by price limits (*trading interference hypothesis*) (Kim and Rhee, 1997); cause higher volatility levels on subsequent periods (*volatility spillover hypothesis*); delay prices from reaching their equilibrium level efficiently (*delayed price discovery hypothesis*); and accelerate price toward the limits as it gets closer to the limits (*magnet effect hypothesis*).

Subrahmanyam (1994) introduced a *magnet effect hypothesis,* which posits that *uninformed* traders rush to trade in anticipation of market halt, thus increasing volatility, decreasing price efficiency, and increasing the probability of price limit hits. Subrahmanyam (1997) also puts forth the *hold back hypothesis* that *informed* traders will strategically change their aggressiveness to hold back their trading to avoid a price limit and continue trading on mispricing. This model predicts that a price limit will decrease volatility and the likelihood of an extreme price movement, but also decrease price efficiency. Brogaard and Roshak (2016) studies the staggered introduction of the price limit rules in the U.S. markets and find that price limits reduce the frequency and severity of extreme price movements. Brogaard and Roshak (2016) tested the two competing hypotheses from Subrahmanyam (1994, 1997) and find they induce price under-reaction and cause informed traders to be less aggressive, which is consistent with the holding back hypothesis. Chen, Petukhov, and Wang (2016) build the first dynamic model to examine the mechanism of market-wide circuit breakers with an optimistic and pessimistic agent and in the absence of market frictions. They conclude that circuit breakers give too much weight to pessimistic investors. This distortion generates both a magnet effect and an excess volatility effect which, surprisingly, is more severe the smaller is the wealth of the irrational (pessimistic) investor. Kirilenko, Kyle, Samadi, and Tuzun (2017) study intraday market intermediation in an electronic market before and during the ‘Flash Crash’, a period of large and temporary selling pressure. They find that the trading pattern of the HFT did not change when prices fell during the Flash Crash, suggesting that HFT traders did not cause the crash. Madhavan (2012) finds that prices are more sensitive to liquidity shocks when markets are fragmented because fragmentation leads to a thinner limit order book (LOB) in each venue than in a consolidated market.

To extend the discussion of each of the hypotheses, first, the interference of trading is cited as the 'obvious cost' to circuit breakers (Lauterbach and Ben-Zion, 1993). If price limits interfere with trading, then stocks become less liquid, which may cause intensified trading activity in the subsequent trading periods (Fama, 1989; Lehmann, 1989). Lee, Ready, and Seguin (1994) found that trading halts increase both the volume and volatility for the NYSE, and volume and volatility after a trading halt are higher than on normal days.

Second, Fama (1989) states the underlying volatility may increase if the price discovery process is interfered with. Kuhn, Kuserk, and Locke (1991) find that limits were ineffective in reducing volatility during the 1989 U.S. mini-crash. Lehmann (1989) also suggests that trading imbalances between supply and demand may induce prices to reach their limits, which causes shift of transactions to subsequent days. Ma, Ramesh, and Sears (1989b) state that the volatility declines on days following limit days, which provides favourable evidence for price limits; however, both Lehmann (1989) and Miller (1989) argue that such a finding is inevitable and trivial because volatility is biased to decrease on days after high volatility. Kim and Rhee (1997) find that volatility does not return to the normal level after reaching the price limits on Tokyo Stock Exchange. Thus, price limits may cause volatility to spread out over a longer period instead of reducing volatility because it prevents both large one-day price changes and immediate order imbalance corrections.

Third, as trading prices approach upper or lower price boundaries, trading halts are usually triggered, which create an interference with the price discovery process (Fama, 1989; Lehmann, 1989; Lee, Ready, and Seguin, 1994). The delayed price discovery hypothesis suggests that price limits prevent prices from reacting to new information and reaching the new equilibrium level. By putting constraints on price movements, stocks may be prevented from reaching their equilibrium price; hence, they must wait until a subsequent trading period to adjust to the true price (Kim and Rhee, 1997).

Fourth, the magnet effect refers to the stock price accelerates toward the limits as it moves closer to the upper or lower price limits. Subrahmanyam (1994) provides a formal theoretical model of the magnet effect which is a generalization of Kyle (1985), engendering a line of empirical scrutiny and hypothesis testing. Two reasons are suggested for the magnet effect: rational anticipated illiquidity and behavioural finance. First, if traders are fearful of illiquidity, they would get involved in active trading that pushes the price closer to the price boundary. Subrahmanyam (1994) shows if the price gets close to the limits, the limits can increase ex-ante price variability and trading volume with the probability of the price crossing the limits increasing. Rather differently and more behaviourally, Lehmann (1989) suggests that order imbalances and the consequent lack of trading induce prices to reach the limits. Second, Arak and Cook (1997) argued that behavioural investors who follow price trends can act in a way that produces the magnet effect. To avoid being shut out of a trend, traders who expect price limits to be reached may execute trades earlier. This behaviour will accelerate price movements as the trading price gets closer to the limits. On the other hand, the cooling off effect, in contrast to the magnet effect, is claimed to be one of the major benefits of price limits. A trading halt will be triggered when the price reaches its limit for a predetermined period of time. The market will then have time to re-evaluate the 'true' value to overcome the overreaction (Cho, Russell, Tiao, and Tsay, 2003). Thus, answering definitively if the price limit rule creates a magnet effect is of importance.

The UK's Government Office for Science report (2012) recommends that price limits could be an effective policy in "improving market efficiency and reducing the risks associated with financial instability." The existing academic literature has offered inconclusive evidence concerning the magnet effect of price limits (Subrahmanyam, 2013). In support of the magnet effect, first, Cho, Russell, Tiao, and Tsay (2003) estimate the return process under a tight (7%) price limit, and only find strong evidence that prices accelerate toward upper limits on the Taiwan Stock Exchange. Second, Goldstein and Kavajecz (2004) find that trading on the NYSE accelerated just before a trading halt in October 1997 given it is just one episode evidence. Third, Chan, Kim, and Rhee (2005) examine how the magnet effect occurs through order imbalance on the Kuala Lumpur Stock Exchange. They find a magnet effect due to the order imbalance before the limit-hit, as well as a subsequent order imbalance reversal after the limit-hit. Fourth, Hsieh, Kim, and Yang (2009) provide evidence of when magnet effects start to emerge on the Taiwan Stock Exchange. Using logit regressions, they show that the magnet effect starts when the price is within nine ticks of the upper price limits and approximately four ticks of the lower price limits. In addition, Du, Liu, and Rhee (2009) also present evident supporting magnet effect in returns, trading volume, and volatility on the Korea Stock Exchange using a time-distanced quadratic model. Tooma (2011) find magnet effect using a logit model of the conditional probability of reaching a limit on the Egyptian Stock Exchange.

In contrast, Arak and Cook (1997) find no magnet effect evidence on Treasury bonds and commodity futures. Berkman and Steenbeek (1998) analysed Nikkei 225 futures trading and also did not find a magnet effect. The lack of magnet effects evidence in futures markets may not generalize to stock markets because futures contracts[44] typically have close substitutes, while individual stocks may not. Subrahmanyam (1994) showed that having a second substitute market could reverse the magnet effect as traders flee in anticipation of closure. In equity markets, Nath (2003) examine the National Stock Exchange of India and find mixed evidence that trading activity reduces when price approaches upper limit price and accelerates when price approaches lower price limit. Abad and Pascual (2007) find that price limits do not cause traders to accelerate trades on the Spanish Stock Exchange (SSE), and conclude there is no magnet effect. However, this can be attributed to the specific trading halt mechanism dealing with the price limit hit in the SSE, which results in five-minute call auctions and then resumption of trading. Thus, the magnet effect may be unobservable in a rule design where investors know that trade can continue after limit hits or trading halts, creating little incentive to advance their trades. The LULD in the U.S. financial market has a similar design, where the LULD trading halt is usually 5 minutes and can be extended to a maximum of 10 minutes each time, but a stock can be halted multiple times in a day. When a listing market announces LULD halt for a stock, all other markets will stop trading during the halt time. We examine whether the magnet effect exists under LULD using causal methodology, and we differentiate the magnet effect and momentum effect in our analysis. Also, we study how does HFT trading behaviour changes around price limits.

Although the price limit rule has already been implemented in numerous financial markets, empirical research remains relatively scarce. The main reason is the difficulty of

---

[44] Brennan (1986) states the price limits in certain futures markets may act as a partial substitute for margin requirements in ensuring contract performance, and its effectiveness is a decreasing function of available information.

obtaining intraday data on hitting limits with a reasonable sample size. Also, the existing empirical literature on the price limits mainly concentrates on the first three hypotheses - trading interference, volatility spillover, and delayed price discovery hypotheses - which typically use daily prices, such as the close and open prices, and trading volume in testing them. To test the magnet effect, we examine intraday price and volume changes to see how the price reacts as it gets closer to the limits and whether different percentage price limits and exchange fee structure influence the magnet effect. In U.S. equity markets, a securities information processor (SIP) consolidates quote and trade data for U.S. stocks across all markets ('SIP data').

To compete for liquidity, most U.S. equity trading venues have introduced a maker-taker model which charges liquidity-demanding orders a fee that exceeds the rebate offered to liquidity supplying orders. More recently, three taker-maker ('inverted') exchanges charge liquidity suppliers a fee that exceeds the rebate paid to liquidity demanders. The fill ratio tends to be higher on the inverted market due to removing liquidity is rewarded. When the price approaches the limit, investors may rush into the inverted market. We examine how does HFT trading behaviour changes for adding or removing liquidity, and how does trading behaviour changes in maker-taker market (Nasdaq) versus the inverted market (Nasdaq/BX) around the price limit.

In this chapter, we test the impact of LULD across different security types (stocks vs. ETFs) and examine the role of HFT when price approaches the price limit. This study is of importance for several reasons. First, this chapter tests the causal impact of the dynamic intraday price limit rule (LULD) using Nasdaq, Nasdaq/BX proprietary intraday data, and SIP data. Second, this is the first paper to investigate who (HFT vs. non-HFT) pushed the price to the limits under both maker-taker market and inverted markets. Third, this chapter uses high frequency intraday trading data to separate the magnet effect from the intraday momentum effect, distinguishing between the magnet effects associated with hitting upper versus lower limits. Fourth, further policy evaluation of price limits is encouraged by the SEC, and this chapter attempts to address the void.

5.2 Limit Up Limit Down Rule

The LULD mechanism is designed to prevent trades in the NMS stocks from occurring outside of predetermined price bands, coupled with trading halts in the event of extreme price movements. The implementation of LULD will prohibit trades from exceeding a percentage (up or down) of the reference price. The reference price is calculated as the arithmetic mean price of eligible trades for the security over the preceding five minutes. The reference price is updated only if the new reference price is at least 1% away in either direction from the current reference price. The SIP will republish the current price bands every 30 seconds. Table 5.1 reports the LULD implementation schedule which is implemented in multiple stages. During Phase I, LULD was in effect from 9:45 am to 3:30 pm for all securities in the S&P 500 Index, the Russell 1000 Index, and selected exchange trade products (ETPs), collectively called Tier 1 NMS stocks. LULD Plan Amendments 4 and 6 split the implementation of Phase II into two separate parts. During part 1 of Phase II, it was rolled out to all Tier 2 securities (all NMS stocks not in Tier 1) with bands in effect from 9:30 am to 3:45 pm. In part 2 of Phase II, LULD bands will be extended to the close (3:45 pm - 4:00 pm). Only warrants and rights are exempted.

Table 5.1: LULD Rule Implementation Schedule

This table reports the detailed Limit Up Limit Down (LULD) rule implementation dates for two phases, and the affected trading hours and affected securities.

LULD	**Start Date**	**Effective Date**	**Effective Time**	**Affected Securities**
Phase I	April 8, 2013	May 3, 2013	9:45am-3:30pm	Tier 1 NMS securities only
Phase II part 1	August 5, 2013	December 8, 2013	9:30am-3:45pm	Both Tier 1 and 2 NMS securities (except rights and warrants)
Phase II part 2	December 8, 2013	February 24, 2014	9:30am-4:00pm	

Table 5.2 illustrates the thresholds of LULD, which is based on the previous closing price, time of the day, and type of security. LULD price bands are a certain percentage away from the reference price. The percentage parameter of a security is based on the previous day's closing price on the primary listing exchange and does not change intraday regardless of intraday price changes. Those stocks in the S&P 500 Index, Russell Index, and certain ETFs (Tier 1 stocks) will have a 5% band, and there will be a 10% band for other listed securities (Tier 2 stocks). Securities with prices less than $3 will have wider bands. The percentages are doubled during the opening and closing periods which are defined as the first and last 15 minutes of continuous trading hours.

Table 5.2: LULD Rule Thresholds and Number of Halts

This table reports the LULD threshold percentage and number of halts for all stocks, and stocks switching above and below $3 and $0.75 under different price, time, and stock type categories. The LULD threshold differs based on the type of the stocks, time of the day and previous closing price of the stocks. Threshold percentage is doubled during the first and last 15 minutes.

Previous Closing Price	9:45am-3:35pm	9:30-9:45am /3:35-4:00pm*
Panel A: LULD Threshold		
<$0.75	Min ($0.15, 75%)	Min ($0.3, 150%)
$0.75-$3	20%	40%
>$3 (Tier 1)	5%	10%
>$3 (Tier 2)	10%	20%
Panel B: Halts for all stocks		
1.<0.75	52	6
2.0.75-3	145	21
3.>3Tier1	554	322
4.>3Tier2	2453	611
Panel C: Halts for stocks switching above and below $3		
$0.75-$3	67	9
>$3 (Tier 1)	0	0
>$3 (Tier 2)	166	6
Panel D: Halts for stocks switching above and below $0.75		
<$0.75	14	2
$0.75-$3	19	1

A LULD halt is triggered if the *limit state* continues for more than 15 seconds, then the primary exchange will call a 5-minute LULD halt which can be extended to maximum of 10 minutes for each halt; however, multiple halts can be triggered for a security in a given day. A limit state is defined as the national best offer (NBO) equals to the lower price band, or the national best bid (NBB) equals to the upper price band, and is not crossed. Table 5.3 provides numerical examples for limit states and straddle states. During a limit state, the SIP will not disseminate a new reference price or new price bands. If a limit state is ended within 15 seconds, the NBBO is no longer resting at a price band. Quotes outside of the price band will not be executable.

Table 5.3: Numerical Examples of Limit and Straddle States

Ref Price	Lower Band	NBB	NBO	Upper Band	State	Explanation
$100	**$95**	$92	**$95**	$105	Limit	NBO equal to Lower Band
$100	$95	**$105**	$107	**$105**	Limit	NBB equal to Upper Band
$100	$95	$104	**$106**	**$105**	Straddle	NBO is higher than Upper Band
$100	**$95**	**$94**	$96	$105	Straddle	NBB is lower than Lower Band

A straddle state exists when the NBB is below the lower price band while the NBO is inside the price band or when the NBO is above the upper price band and the NBB is within the band. The primary listing exchange has discretion to declare a trading halt when a security is in a straddle state. However, no LULD halt is triggered because of a straddle state during our sample period. If the listing exchange cannot open trading and has not declared a regulatory halt after 10 minutes, other market centres are free to commence trading. During the LULD halt, no trades in the halted stocks will be executed, but all bids and offers may be displayed. Options exchanges will halt trading in contracts if the underlying symbol is subject to a LULD halt. Any trade that does not update the last sale price or excluded from Order Protection Rule 611 of Reg NMS (Trade-Through rule) would be exempt from LULD restrictions.

5.3 Data, Sample Selection, and Methodology

A. Data Source

Our analysis is based on intraday trader-level data and U.S. SIP data, as well as Nasdaq and Nasdaq/BX proprietary data. Data is collected for all securities with limit states and LULD halts and its corresponding matched sample from April 8, 2013 to December 31, 2015. Our HFT data is identified based on the method described in Brogaard, Hendershott and Riordan (2014).

Our data include all information on order submission and trades, including timestamp, price, volume, and a unique identifier for the trader that submitted the order, which allows us to construct the HFT data. We restrict attention to transactions that occur in the LOB and trades during regular trading hours. For each LOB transaction, the data contains identifiers for buyer- or seller-initiated trade, adding or removing liquidity, and halt reason. We also construct the intraday dynamic upper and lower limit prices and matched with trade and quote data.

B. Sample Selection

We use *propensity matching score* methodology to match 'Treatment' and 'Control' stocks using one-to-one matching without replacement nearest neighbour matching. 'Treatment' stocks refer to stocks with limit states and LULD halts, while 'Control' stocks refer to the rest of stocks without limit states and LULD halts during the sample period. The 'Control' stocks are on the same listing exchange, the same security type (stocks or ETFs) with the 'Treatment' stocks, and share very similar characteristics in terms of the security's average closing price, average daily trading volume based on one month prior to our sample period data.

We place the pseudo halt time on matched 'Control' stocks on the same day as the halted treatment stock. We exclude securities that had stock splits, switched listing exchange, or had days with a stock price below 20 cents. Also, each sample stock has at least 20 trades on lit markets each day. Davies and Kim (2009) argue that one-to-one matching without replacement based on the closing price and market capitalization maybe the most appropriate method to test for differences in trade execution costs. O'Hara and Ye (2011) followed their approach and matched on closing price, market capitalization and listing exchange. Because our sample also includes ETFs, we replace market capitalization with the average daily trade volume one month before the LULD starts.

Figure 2: Intraday Pattern of LULD Halts

This figure depicts the total number of LULD halts for 30-minute time buckets between continuous trading hours 9:30am – 4:00pm across different listing markets. We first split the first and last half hour into 15-minute time buckets. Tape A refers to NYSE-listed stocks; Tape B refers to other mixed securities listing exchanges consisting of NYSE MKT, NYSE Arca and BATS; and Tape C refers to Nasdaq-listed stocks.

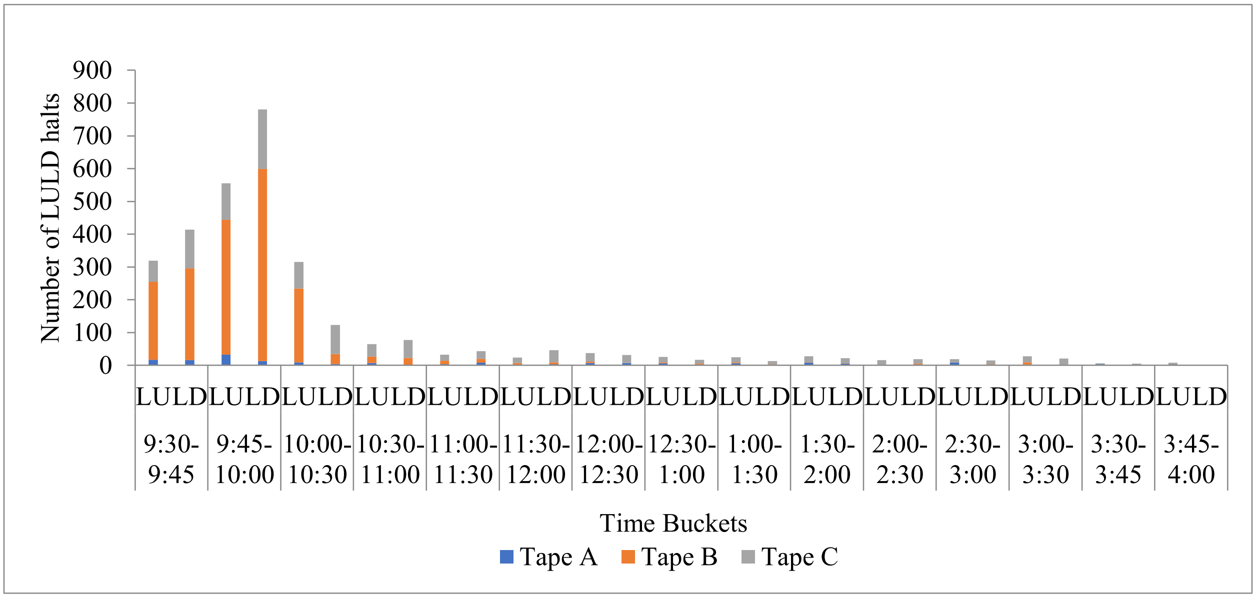

Table 5.4: Summary Statistics

This table reports the number of securities with limit states and LULD halts (Halt) per month per listing exchange, as well as total number of limit up states (LU), limit down states (LD), total limit states (LS) and LULD halts from April 2014 to December 2015. In this sample, only securities with at least 20 trades on lit markets and a closing price is greater than 20 cents are included. A, N, P, Q and Z refers to AMEX, NYSE, NYSE_Arca, Nasdaq, and BATS listed securities respectively. We also report the conditional probability of LULD halts given LU, LD, and LS.

	Limit States					**Halt**					**Security Count**		**Total LS/Halt Count**			
Month	**A**	**N**	**P**	**Q**	**Z**	**A**	**N**	**P**	**Q**	**Z**	**LS**	**Halt**	**LU**	**LD**	**LS**	**Halt**
201304	0	0	2	0	0	0	0	0	0	0	2	0	3,576	0	3,576	0
201305	0	4	1	6	0	0	2	1	2	0	11	5	4,053	366	4,419	8
201306	0	2	1	2	0	0	1	1	2	0	5	4	550	210	760	4
201307	0	3	0	1	0	0	2	0	1	0	4	3	816	109	925	5
201308	1	10	5	13	0	0	6	4	6	0	29	16	525	595	1,120	22
201309	4	6	3	33	0	2	3	3	16	0	46	24	2,043	646	2,689	50
201310	4	8	7	42	0	3	5	4	21	0	61	33	3,651	3,206	6,857	68
201311	3	11	8	36	0	1	2	8	17	0	58	28	379	1,100	1,479	32
201312	5	7	10	46	0	1	2	10	25	0	68	38	4,330	2,153	6,483	52
201401	8	14	7	46	0	4	6	6	30	0	75	46	2,452	2,344	4,796	93
201402	3	9	5	42	0	2	4	4	20	0	59	30	1,149	5,361	6,510	44
201403	6	6	14	57	0	3	4	11	32	0	83	50	2,659	2,051	4,710	79
201404	4	9	10	50	0	0	2	10	36	0	73	48	916	2,932	3,848	92
201405	1	5	6	53	0	1	1	6	36	0	65	44	1,206	3,958	5,164	75
201406	5	4	27	40	0	3	2	26	21	0	76	52	3,838	2,345	6,183	72
201407	1	11	12	50	0	0	9	11	30	0	74	50	9,598	4,046	13,644	109
201408	5	14	13	58	0	3	3	11	42	0	90	59	2,912	2,489	5,401	86
201409	2	8	11	71	0	1	2	10	48	0	92	61	2,927	1,978	4,905	122
201410	7	22	18	99	1	3	10	12	68	0	147	93	5,804	7,554	13,358	163
201411	3	6	10	47	0	1	4	10	29	0	66	44	4,364	1,224	5,588	69
201412	6	12	17	77	1	3	6	13	45	1	113	68	10,619	3,075	13,694	117
201501	3	13	21	55	0	3	8	17	34	0	92	62	3,558	1,972	5,530	95
201502	7	12	16	38	0	3	5	15	27	0	73	50	3,165	1,105	4,270	87
201503	1	7	22	54	0	1	4	22	35	0	84	62	3,020	4,469	7,489	119
201504	2	14	19	48	0	0	8	18	32	0	83	58	4,042	1,748	5,790	88
201505	1	11	12	51	0	0	7	10	33	0	75	50	6,282	6,050	12,332	98
201506	3	14	10	56	0	0	6	9	27	0	83	42	7,235	440	7,675	67
201507	4	19	10	80	0	2	13	8	53	0	113	76	3,231	2,650	5,881	173
201508	10	152	287	193	3	6	70	253	126	2	645	457	27,527	47,773	75,300	1,309
201509	14	12	9	95	0	12	5	9	71	0	130	97	8,841	1,750	10,591	217
201510	9	26	10	102	0	5	15	8	66	0	147	94	6,266	4,845	11,111	158
201511	9	29	5	94	0	3	16	4	57	0	137	80	3,976	3,328	7,304	208
201512	6	18	12	122	2	3	8	11	75	1	160	98	2,212	4,026	6,238	199
Total	137	498	620	1857	7	69	241	545	1163	4	3119	2022	147,722	127,898	275,620	4,180
Unique	83	391	411	967	7	46	196	365	681	4	1859	1292				
Probability of LULD Halts given Limit States													2.8%	3.3%	1.5%	

Figure 2 reports the intraday pattern of LULD halts across different listing exchanges, showing that most of halts occurred during the first continuous trading hour of the day. Table 5.4 reports the monthly number of securities with limit states and LULD halts during our sample period from April 2014 to December 2015. The difference-in-differences (DID) can be used given the staggered implementation of the LULD rule, especially during Phase I, which only applies to Tier 1 National Market System (NMS) securities. However, there is an insufficient sample amount of LULD halts that were triggered during Phase I implementation stages. As shown in Table 5.4, only 5 LULD halts were trigger between April and May 2013. Two parts implementation of Phase II only extend the LULD coverage among the continuous trading session from 3:30pm to 4:00pm. As shown in Figure 2, the majority of LULD halts were triggered during the first half hour of the continuous trading session. The sample may not be homogenous in certain respects and further research is necessary. In addition, prior to the LULD rule, there was another SSCB rule, which does not provide a clean window comparing pre and post adoption of the LULD. To overcome this difficulty, we use the pseudo 'control stocks'. In addition, we construct a subset of stocks switched above and below the $3 LULD threshold.

5.4 Empirical Results

H1: The trading interference hypothesis

Does trading volume increase after trading halts? Fama (1989) and Lehmann (1989) point out that if price limits prevent trading, then stocks will become less liquid causing intensified trading activities in the subsequent trading periods. We compare the percentage of trading volume pattern over a 5-minute interval before and after a LULD trading halt between the treatment group and control group.

The estimation is based on the following regression specification:

$$Y_{i,t} = \alpha_0 + \beta_1 Halt_i + \beta_2 Post_i + \beta_3 Halt_i * Post_t + \beta_4 ETF_i + \gamma VIX_T + \varepsilon_{i,t}, \quad (37)$$

where $Y_{i,t}$ is the trading volume and market share in a time interval; α_0 is the intercept; $Halt_i$ is a dummy variable, which is 1 if the security *i* has at least one LULD halt on a particular day and 0 if it does not have a LULD halt; $Post_t$ is a dummy variable that is 1 if the time interval is after the first halt and 0 otherwise; and VIX_T is the closing value of CBOE's volatility index for day *T*. We estimate the specification with and without stock fixed effects.

Table 5.5: Trading Volume, Number of Trades and Volatility

The table tests the impact of LULD on national-level trading volume, number of trades, and volatility during five-minute windows pre and post LULD halts. ConsolVol stands for the consolidated trading volume during the time window. #Trades stands for the total number of trades during the time window. VWAP refers to the volume-weighted average trade price during the time window. Volatility is defined as the highest price minus lowest price over the VWAP during the time window. Standard errors are in parentheses and are robust to both heteroscedasticity and correlation within stocks. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	Log(ConsolVol)	Log(#Trades)	Log(VWAP)	Log(Volatility)
Intercept	9.532***	4.111***	2.524***	0.0003
	(0.078)	(0.073)	(0.031)	(0.002)
Post	-0.125*	0.102	0.001	-0.002
	(0.073)	(0.068)	(0.029)	(0.002)
Halt	1.421***	1.290***	-0.175***	0.117***
	(0.085)	(0.080)	(0.034)	(0.003)
Post*Halt	0.393***	0.179*	0.015	-0.013***
	(0.117)	(0.109)	(0.047)	(0.004)
ETF	0.291***	0.09	1.005***	-0.005*
	(0.078)	(0.073)	(0.031)	(0.002)
VIX	-0.040***	-0.039***	0.008***	0.0004***
	(0.003)	(0.003)	(0.001)	(0.000)
Observations	6,020	6,020	6,020	6,020
Adjusted R^2	0.107	0.101	0.287	0.411

Results: As shown in Table 5.5, the five-minute trading volume and number of trades increased after LULD halts (coefficients of 0.393 and 0.179, respectively). This indicates that LULD interferes with trading activity.

H_2: The volatility spillover hypothesis

Does volatility spillover to the subsequent trading period for stocks vs. ETFs? Chen, Petukhov, and Wang (2016), Kyle (1985), Fama (1989), and Kuhn, Kuserk, and Locke (1991) state underlying volatility may increase if there is interference with the price discovery process. ETFs tracks a basket of assets and are more diversified, thus, ETFs are less volatile than stocks.

The estimation is based on the following regression specification:

$$Y_{i,t} = \alpha_0 + \beta_1 Halt_i + \beta_2 Post_i + \beta_3 Halt_i * Post_t + \beta_4 ETF_i + \gamma VIX_T + \varepsilon_{i,t}, \qquad (38)$$

where $Y_{i,t}$ is the trading volume and market share in a time interval; α_0 is the intercept; $Halt_i$ is a dummy variable, which is 1 if the security *i* has at least one LULD halt on a particular day and 0 if it does not have an LULD halt; $Post_t$ is a dummy variable, which is 1 if the time interval is after the first halt and 0 otherwise; ETF_t is a dummy variable, which is 1 if the security is ETF and 0 for stock; and VIX_T is the closing value of CBOE's volatility index for day *T*. We estimate the specification with and without stock fixed effects.

Results: As shown in Table 5.5, the five-minute interval volatility decreased (coefficient of -0.013) after the LULD halts. This indicates that a LULD halt curbs short-term over-reactive action in the stock market.

H_3: The delayed price discovery hypothesis

Does the dynamic price limit LULD delay price discovery? Kim and Rhee (1997) find price limits prevent price adjustments to their equilibrium prices on the halt days, without curbing over-reactive behaviour. The price continuation behaviour suggests that price limits prevent rational or informed trading (Roll, 1989), otherwise, we would find price reversals in the context of over-reactive behaviour (Ma, Ramesh, and Sears (1989a, 1989b)).

To identify intraday price continuations and reversals, we examine the following two return series: open-to-halt return measured by $\ln(H_t/O_t)$ and halt-to-close return measured by $\ln(C_t/H_t)$. If the halt time is ended at the end of the continuous trading hour, the close-to-open measured by $\ln(O_{t+1}/C_t)$ will be used instead. Following the intuition of Kim and Rhee (1997), we compare the sign of price return series before and after a LULD halt. Stock return can be either positive (+), negative (-), or zero (0). Consequently, nine returns series are possible: [+, +], [+, 0], [+, -], [0, +], [0, 0], [0, -], [-, +], [-, 0], and [-, -], where the first return represents an open-to-halt return and the second return represents a halt-to-close return. *For upper limit hits*, we classify [+, +] and [0, +] as price continuations; [+, -], [0, -], [-, +], [-, 0], and [-, -] as price reversals; and [+, 0] and [0, 0] as no change. *For lower limit hits*, we classify [-, -] and [0, -] as price continuations; [-, +], [0, +], [+, -], [+, 0], and [+, +] as price reversals; and [-, 0] and [0, 0] as no change. If there are multiple halts for a stock-day, the return after the first halt will be selected. Then, we count the number of price continuations, price reversals, and no change, and compare between the treatment group and control group using mean tests.

Table 5.6: Delayed Price Discovery

This table summarizes the number and proportion of price continuation, price reversal and no change when the price reached the lower price limit (Limit Down) or upper price limit (Limit Up). To identify intraday price continuations and reversals, we examine the following two return series: open-to-halt return measured by ln(Ht/Ot) and halt-to-close return measured by ln(Ct/Ht). If the halt time is ended at the end of the continuous trading hour, the close-to-open measured by ln(Ot+1/Ct) will be used instead. Following Kim and Rhee's (1997) intuition, we compare the sign of the price return series before and after the LULD halt. Stock returns can be either positive (+), negative (-), or zero (0). Consequently, nine returns series are possible: [+, +], [+, 0], [+, -], [0, +], [0, 0], [0, -], [-, +], [-, 0], and [-, -], where the first return represents open-to-halt return and the second return represents halt-to-close return. For upper limit hits, we classify [+, +] and [0, +] as price continuations; [+, -], [0, -], [-, +], [-, 0], and [-, -] as price reversals; and [+, 0] and [0, 0] as no change. For lower limit hits, we classify [-, -] and [0, -] as price continuations; [-, +], [0, +], [+, -], [+, 0], and [+, +] as price reversals; and [-, 0] and [0, 0] as no change. If there are multiple halts for a stock-day, the return after the first halt will be selected. The last column reports the difference between the Treatment and Control groups. The z-values, based on a binomial test statistic, are given in parenthesis.

Price Pattern	**Treatment**	**Control**	**Difference**	**(z-value)**
Panel A: Limit Up				
Sample Size	219	219		
Price Continuation	0.30	0.26	0.04	(0.956)
Price Reversal	0.68	0.71	-0.03	(-0.519)
No Change	0.01	0.03	-0.01	(-1.243)
Panel B: Limit Down				
Sample Size	272	272		
Price Continuation (%)	0.10	0.17	-0.10	(-2.34)
Price Reversal (%)	0.90	0.81	0.13	(2.866)
No Change (%)	0.00	0.02	-0.03	(-0.01)

Results: As shown in Table 5.6, when the price reaches its LULD upper limit, there is no statistically significant change in price pattern. Interestingly, when the price reaches its lower limit, the price tends to revert after LULD trading halts. This suggests the dynamic price limit rule LULD curbs over-reactive action when the price reaches its lower limit, which is different than Kim and Rhee's (1997) analysis of the traditional price limit rule. There is no evidence showing that LULD delayed price discovery.

H_4: The magnet effect hypothesis

Prior to the anticipated price limit hit, traders would try to rush in and cover their positions, which would increase trading activity, raise *ex-ante* price variability, and accelerate price to its permitted limit for that day. Subrahmanyam (1994) engenders a line of empirical inquiry of the magnet effect based on an extension of the Kyle (1985) model with risk-averse market makers. Illiquidity and behavioural finance are two suggested reasons for the magnet effect. First, if traders are fearful of illiquidity, they would adopt active trading strategies that pull the price closer to the limit. Subrahmanyam (1994) presents that if the price is close to the limits, the limits can increase ex-ante price variability and the probability of the price crossing the limits as discretionary liquidity traders inefficiently concentrate their orders, rather than split them over time, in anticipation of closure. That is, trading halts due to price limits can perversely exacerbate the very volatility they are supposed to ameliorate while at the same time temporarily increasing trading volume and liquidity. Rather differently, Lehmann (1989) finds order imbalances and lack of trading induce prices to move towards the limits. Arak and Cook (1997) argue that behavioural investors who believe in the price trends can act in a way that produces a magnet effect. Traders who think that the ceiling will be reached might execute their trades faster. This behaviour will accelerate price movements as price gets closer to the ceiling. However, the magnet effect and the intraday momentum effect need to be differentiated (Du, Liu, and Rhee, 2009). In the cleanest theoretical justification for the magnet effect that does not rely on either market frictions or behavioural effects, Chen, Petukhov, and Wang (2016) show that this effect can arise from giving too much weight to pessimistic beliefs when the market is shut down. To test this hypothesis, actual and quasi-limit hit (90% of the price limit bands) will be compared. Actual limit hits can be driven by both magnet and momentum effects whereas the quasi-limit hits are driven by the momentum effect alone since they represent large price changes but not large enough yet to trigger the LULD halts. Thus, the difference will contribute to measuring the impact of the magnet effect. When price approaches the price limit, the average time to the next trade will decrease.

We estimate this hypothesis using the following two models:

When price reaches the upper limit:

$$Y_{i,t} = \alpha_0 + \beta_1 du.rel_i + \beta_2 Halt_i + \beta_3 du.rel_i * Halt_t + \beta_4 ETF_i + \gamma VIX_T + \varepsilon_{i,t}, \quad (39)$$

and when price reaches the lower limit:

$$Y_{i,t} = \alpha_0 + \beta_1 dl.rel_i + \beta_2 Halt_i + \beta_3 dl.rel_i * Halt_t + \beta_4 ETF_i + \gamma VIX_T + \varepsilon_{i,t}, \quad (40)$$

where $Y_{i,t}$ refers to the time to the next trade (speed) or trading volume (magnitude) for the trade price between 90% and 100%, and 80% to 90% of the price limit bands. $Halt_i$ is a dummy variable, which is 1 if the security *i* has at least one LULD halt on a particular day and 0 if it does not have an LULD halt; $du.rel_i$ is the percentage of price difference away from the upper band; and $dl.rel_i$ is the percentage of price difference away from the lower band. We run the regression when the price difference is between 95%-100%, 90%-95%, and 85%-90%. If there is a magnet effect, the time to next trade (trade volume) is expected to be smaller (larger) for an actual hit compared with a quasi-limit hit.

Table 5.7: Magnet Effect_Trading Speed

The table tests the magnet effect in terms of trading speed, measured by time to next trade, when price approaches the upper/lower price limits. We test using four subsets of data for both price approach upper limit price (LU) and lower limit price (LD): the whole price range, greater than 95%, 90%-95%, and 85%-90% of the price limit. The variable du.rel(dl.rel) refers to the trade price relative to the upper(lower) price limit. Standard errors are in parentheses and are robust to both heteroscedasticity and correlation within stocks. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**LU**	**LU >0.95**	**LU 0.9-0.95**	**LU 0.85-0.9**	**LD**	**LD >0.95**	**LD 0.9-0.95**	**LD 0.85-0.9**
Intercept	8.576***	8.233***	8.742***	10.149***	7.021***	23.790***	8.700***	-6.855***
	(0.029)	(0.326)	(0.078)	(0.546)	(0.021)	(0.601)	(0.077)	(0.430)
du.rel	-5.327***	-5.243***	-5.359***	-6.872***				
	(0.031)	(0.342)	(0.085)	(0.613)				
dl.rel					-3.671***	-21.652***	-5.436***	12.166***
					(0.023)	(0.630)	(0.083)	(0.480)
Halt	-3.316***	24.843***	0.536***	-12.427***	-1.630***	10.532***	-6.186***	-11.160***
	(0.041)	(0.378)	(0.131)	(0.594)	(0.032)	(0.637)	(0.130)	(0.494)
du.rel*Halt	2.502***	-26.505***	-1.780***	12.370***				
	(0.044)	(0.394)	(0.141)	(0.669)				
dl.rel*Halt					0.691***	-11.514***	5.553***	11.252***
					(0.035)	(0.667)	(0.139)	(0.554)
ETF	-0.770***	-0.436***	-1.047***	-1.350***	-0.785***	-0.396***	-0.824***	-1.357***
	(0.002)	(0.004)	(0.003)	(0.019)	(0.002)	(0.004)	(0.003)	(0.006)
VIX	0.033***	0.032***	0.034***	0.025***	0.032***	0.034***	0.031***	0.032***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	14,425,886	6,114,696	6,812,200	752,489	14,425,886	4,611,553	6,791,418	2,213,083
Adjusted R^2	0.026	0.017	0.032	0.052	0.026	0.019	0.023	0.046

Table 5.8: Magnet Effect_Trading Volume

The table tests the magnet effect in terms of trading volume when price approaches the upper/lower price limits. We test using four subsets of data for both price approach upper limit price (LU) and lower limit price (LD): the whole price range, greater than 95%, 90%-95%, and 85%-90% of the price limit. The variable du.rel(dl.rel) refers to the trade price relative to the upper(lower) price limit. Standard errors are in parentheses and are robust to both heteroscedasticity and correlation within stocks. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**LU**	**LU >0.95**	**LU 0.9-0.95**	**LU 0.85-0.9**	**LD**	**LD >0.95**	**LD 0.9-0.95**	**LD 0.85-0.9**
Intercept	6.243***	5.194***	6.029***	7.622***	5.855***	3.681***	6.029***	6.650***
	(0.007)	(0.073)	(0.018)	(0.157)	(0.005)	(0.137)	(0.017)	(0.099)
du.rel	-1.508***	-0.410***	-1.283***	-3.005***				
	(0.007)	(0.077)	(0.019)	(0.176)				
dl.rel					-1.095***	1.176***	-1.293***	-1.962***
					(0.005)	(0.144)	(0.019)	(0.111)
Halt	-0.199***	-2.686***	-0.838***	0.602***	0.094***	-2.143***	-0.768***	0.361***
	(0.009)	(0.085)	(0.030)	(0.171)	(0.007)	(0.145)	(0.029)	(0.114)
du.rel*Halt	0.466***	3.057***	1.137***	-0.456**				
	(0.010)	(0.088)	(0.032)	(0.192)				
dl.rel*Halt					0.152***	2.471***	1.077***	-0.185
					(0.008)	(0.152)	(0.032)	(0.128)
ETF	0.135***	0.137***	0.129***	0.207***	0.132***	0.125***	0.134***	0.146***
	(0.001)	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)	(0.001)	(0.001)
VIX	0.00001	0.0002***	0.0002***	-0.002***	-0.0001***	0.001***	0.0001***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	14,425,886	6,114,696	6,812,200	752,489	14,425,886	4,611,553	6,791,418	2,213,083
Adjusted R^2	0.018	0.016	0.014	0.017	0.019	0.015	0.016	0.018

Results: As shown in Tables 5.7 and 5.8, when price moves closer to the upper or lower price limit, the time to the next trade is reduced and trading volume increases. Time to next trade refers to the trading timestamp difference between current trade price and next trade price, and trading volume refers to the trading volume as the price approaches the upper and lower price limits. If a magnet effect exists, in which price accelerates toward the limits as it gets closer to the limits, the time to next trade tends to decrease as price approaches the price limits. Also, we extend the analysis by not only examining the trading speed but also the magnitude, or trading volume. The trading volume tends to increase as price approaches the price limits. To differentiate between magnet effect and momentum effect, we select three different sets of price ranges: 95% to the upper or lower price limit, 90% to 95%, and 85% to 90%. When the price moves from 85% to 90% of the price limit, the time to next trade (trading volume) increases (decreases). However, the time to next trade (trading volume) decreases (increases) when price moves from 90% to the price limit. Consistent with previous literature, the results suggest there is a magnet effect when price approaches the upper or lower limit under the LULD framework.

H_5: What is the impact of the price percentage band and exchange fee structure on the magnet effect?

To further examine the magnet effect, we constructed two subsets of stocks that switched above and below the \$3 and \$0.75 LULD thresholds. As shown in Table 5.2 Panel D, given the smaller subset of the group around \$0.75, and different tick size regime above and below \$1 issues, the \$0.75 group is excluded from this analysis. We focus on the subset of stocks around \$3 to test the impact of the price band percentage and exchange fee structures. Again, to differentiate the magnet effect and momentum effect, we construct three slightly different groups to balance the sample size in each: 90% to the price limits, 80% to 90%, and 70% to 80%. We estimate those hypotheses using the following models:

When price reaches the upper limit:

$$Y_{i,t} = \alpha_0 + \beta_1 du.rel_i + \gamma VIX_T + \varepsilon_{i,t}, \quad (41)$$

$$Y_{i,t} = \alpha_0 + \beta_1 du.rel_i + \beta_2 PriceBand_i + \beta_3 du.rel_i * PriceBand_t + \gamma VIX_T + \varepsilon_{i,t}, \quad (42)$$

$$Y_{i,t} = \alpha_0 + \beta_1 du.rel_i + \beta_2 Inverted_i + \beta_3 du.rel_i * Inverted_t + \gamma VIX_T + \varepsilon_{i,t}, \quad (43)$$

and when price reaches the lower limit:

$$Y_{i,t} = \alpha_0 + \beta_1 dl.rel_i + \gamma VIX_T + \varepsilon_{i,t}, \quad (44)$$

$$Y_{i,t} = \alpha_0 + \beta_1 dl.rel_i + \beta_2 PriceBand_i + \beta_3 dl.rel_i * PriceBand_t + \gamma VIX_T + \varepsilon_{i,t}, \quad (45)$$

$$Y_{i,t} = \alpha_0 + \beta_1 dl.rel_i + \beta_2 Inverted_i + \beta_3 dl.rel_i * Inverted_t + \gamma VIX_T + \varepsilon_{i,t}, \quad (46)$$

where $PriceBand_i$ is a dummy variable, which is 1 for stocks with an LULD halt and whose previous closing price is above \$3 and 0 for below \$3. All the stocks are above \$1 and, as shown in Table 2 Panel C, all the stocks above \$3 are Tier 2 NMS securities. $Inverted_i$ is a dummy variable which is 1 for the inverted fee markets and 0 for the maker-taker markets. Off-exchange venues are not included for the regression model involved with the $Inverted_i$ dummy variable.

Table 5.9: The Impact of Price Band on Magnet Effect_Trading Speed

This table reports the impact of LULD price bands on the magnet effect, specifically, the pattern of time to next trade (TTNT) when price approaches the upper/lower price limits using the same stocks with different price categories on different days. We examine the impact on three sub-samples: trades with price greater than 90% of price limits (>0.9); trades with price between 80% and 90% of price limits (0.8-0.9); and trades with price between 70% and 80% of price limits (0.7-0.8). The variable du.rel (dl.rel) is a percentage distance between the current trade price and upper (lower) price limits. PriceBand is a dummy variable which is 1 for a stock priced above \$3 in a day and 0 for the same stock when its price is below \$3 but above \$0.75 in another day. The percentage limit on the \$0.75-\$3 group is twice higher than the above \$3 group for Tier 2 stocks (majority of the cross-price sample), and four times higher for Tier 1 stocks. Inverted is a dummy variable, which is 1 for markets with an inverted fee structure and 0 for the maker-taker markets (i.e. off-exchange venues are not included in the regression model involved with the Inverted dummy variable).

Panel A: TTNT when Price Approaches Upper Price Limit

Sample Group	>0.9			0.8-0.9			0.7-0.8		
Model	4	5	6	7	8	9	10	11	12
Constant	19.402***	3.631**	18.001***	-0.619*	5.656***	-1.297***	6.164***	7.692***	2.453***
	(0.476)	(1.583)	(0.624)	(0.361)	(0.610)	(0.503)	(0.643)	(0.648)	(0.899)
du.rel	-15.672***	0.042	-14.157***	4.718***	-2.976***	5.335***	-6.238***	-8.614***	-2.618**
	(0.522)	(1.668)	(0.680)	(0.422)	(0.728)	(0.586)	(0.821)	(0.830)	(1.142)
PriceBand		15.636***			-0.861			1.173	
		(1.667)			(0.859)			(4.550)	
Inverted			-1.542			8.586***			2.838
			(1.791)			(1.619)			(2.568)
du.rel:PriceBand		-15.489***			1.823*			0.85	
		(1.761)			(1.008)			(5.861)	
du.rel:Inverted			1.102			-10.709***			-4.358
			(1.915)			(1.907)			(3.390)
VIX.Close	-0.012***	-0.013***	-0.024***	0.062***	0.060***	0.063***	0.144***	0.154***	0.184***
	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)	(0.004)	(0.006)	(0.006)	(0.009)
Observations	62,414	62,414	37,998	78,230	78,230	44,767	18,325	18,325	10,928
Adjusted R^2	0.016	0.022	0.018	0.009	0.016	0.01	0.033	0.052	0.042

Panel B: TTNT when Price Approaches Lower Price Limit

Sample Group	>0.9			0.8-0.9			0.7-0.8		
Model	16	17	18	19	20	21	22	23	24
Constant	14.879***	7.022***	12.768***	-3.154***	9.230***	-3.550***	-1.199**	0.07	-3.681***
	(0.713)	(2.039)	(0.929)	(0.303)	(0.638)	(0.423)	(0.547)	(0.571)	(0.805)
dl.rel	-10.920***	-2.964	-8.995***	8.190***	-7.045***	8.677***	5.356***	3.692***	8.398***
	(0.775)	(2.155)	(1.009)	(0.344)	(0.769)	(0.482)	(0.703)	(0.736)	(1.032)
PriceBand		7.941***			-11.524***			-12.702***	
		(2.188)			(0.791)			(2.218)	
Inverted			-0.566			4.767***			3.963*
			(3.061)			(1.222)			(2.105)
dl.rel:PriceBand		-8.067***			14.426***			16.555***	
		(2.322)			(0.940)			(2.834)	
dl.rel:Inverted			-0.047			-6.151***			-5.632**
			(3.312)			(1.435)			(2.716)
VIX.Close	0.005	0.007*	0.016***	0.033***	0.032***	0.021***	0.079***	0.078***	0.077***
	(0.004)	(0.004)	(0.006)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.006)
Observations	39,276	39,276	23,922	80,534	80,534	47,243	36,607	36,607	20,249
Adjusted R^2	0.005	0.006	0.006	0.008	0.016	0.009	0.013	0.015	0.013

Table 5.10: The Impact of Price Band on Magnet Effect_Trading Volume

This table reports the impact of LULD price band on the magnet effect, specifically, the pattern of trading volume when price approaches the upper/lower price limits using the same stocks with different price categories on different days. We examine the impact on three sub-samples: trades with price greater than 90% of price limits (>0.9); trades with price between 80% and 90% of price limits (0.8-0.9); and trades with price between 70% and 80% of price limits (0.7-0.8). The variable du.rel (dl.rel) is a percentage distance between the current trade price and upper (lower) price limits. PriceBand is a dummy variable, which is 1 for a stock priced above $3 in a day and 0 for the same stock when its price is below $3 but above $0.75 in another day. The percentage limit on the $0.75-$3 group is twice higher than the above $3 group for Tier 2 stocks (majority of the cross-price sample), and four times higher for Tier 1 stocks. Inverted is a dummy variable, which is 1 for markets with an inverted fee structure and 0 for the maker-taker markets (i.e. off-exchange venues are not included in the regression model involved with the Inverted dummy variable).

Panel A: Trading Volume when Price Approaches Upper Price Limit

Sample Group	>0.9			0.8-0.9			0.7-0.8		
Model	4	5	6	7	8	9	10	11	12
Constant	4.267***	7.742***	3.949***	7.613***	5.067***	7.675***	5.211***	4.965***	4.692***
	(0.117)	(0.388)	(0.136)	(0.108)	(0.182)	(0.132)	(0.220)	(0.223)	(0.272)
du.rel	0.977***	-2.393***	1.331***	-2.364***	0.763***	-2.535***	0.398	0.787***	0.786**
	(0.128)	(0.408)	(0.149)	(0.126)	(0.217)	(0.153)	(0.281)	(0.286)	(0.346)
PriceBand		-3.251***			-0.047			-1.582	
		(0.408)			(0.256)			(1.566)	
Inverted			-0.921**			-1.204***			-1.888**
			(0.391)			(0.423)			(0.777)
du.rel:PriceBand		3.095***			-0.303			1.618	
		(0.431)			(0.300)			(2.018)	
du.rel:Inverted			0.655			1.070**			2.170**
			(0.419)			(0.499)			(1.026)
VIX.Close	-0.003***	-0.002**	-0.007***	-0.018***	-0.017***	-0.019***	-0.009***	-0.010***	-0.004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
Observations	62,414	62,414	37,998	78,230	78,230	44,767	18,325	18,325	10,928
Adjusted R^2	0.001	0.011	0.02	0.011	0.027	0.025	0.001	0.006	0.008

Panel B: Trading Volume when Price Approaches Upper Price Limit

Sample Group	>0.9			0.8-0.9			0.7-0.8		
Model	16	17	18	19	20	21	22	23	24
Constant	3.776***	6.115***	3.386***	8.215***	4.522***	7.944***	6.896***	6.356***	5.811***
	(0.174)	(0.497)	(0.204)	(0.085)	(0.178)	(0.105)	(0.173)	(0.180)	(0.218)
dl.rel	1.495***	-0.689	1.902***	-3.203***	1.372***	-2.939***	-1.515***	-0.816***	-0.287
	(0.190)	(0.526)	(0.221)	(0.097)	(0.215)	(0.120)	(0.223)	(0.232)	(0.280)
PriceBand		-1.762***			1.531***			1.664**	
		(0.534)			(0.221)			(0.700)	
Inverted			-0.801			-0.594*			-1.017*
			(0.671)			(0.304)			(0.571)
dl.rel:PriceBand		1.569***			-2.221***			-2.441***	
		(0.566)			(0.263)			(0.894)	
dl.rel:Inverted			0.533			0.333			0.945
			(0.726)			(0.356)			(0.737)
VIX.Close	-0.0002	-0.002*	-0.003***	-0.015***	-0.013***	-0.017***	-0.021***	-0.019***	-0.021***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Observations	39,276	39,276	23,922	80,534	80,534	47,243	36,607	36,607	20,249
Adjusted R^2	0.002	0.009	0.017	0.016	0.036	0.032	0.01	0.017	0.021

Results: As shown in Tables 5.9 and 5.10, using a subset of stocks switching above and below \$3, when price approaches upper limit (> 90%), the time to the next trade reduces (coefficient of -15.672***) and trading volume increases (coefficient of 0.977***). Similarly, when price approaches the lower limit (> 90%), the time to the next trade decreases (coefficient of -10.920***) and trading volume increases (coefficient of 1.495***). Those are consistent with previous findings.

The price percentage band for Tier 2 stocks above \$3 is twice smaller than the below \$3 percentage band (as shown in Table 5.2 Panel A). A narrower price percentage band does encourage the magnet effect when price approaches the price limits and the impact is stronger when price approaches upper limit. Specifically, when price approaches the upper limit, the narrower price band causes time to next trade to decrease further (coefficient of -15.489***) and trading volume to increase (coefficient of 3.095***). On the other hand, when price approaches the lower limit, the narrower price band also cause time to next trade to decrease (coefficient of -8.067***) and trading volume to increase (coefficient of 1.569***) but at a smaller magnitude compared with upper limits. In addition, the 80%-90% and 70%-80% of price limits groups shows a different pattern compared with the >90% group, which rules out the momentum effect. However, we did not observe evidence that the exchange fee structure plays a role for the magnet effect.

H_6: How does HFT trading behaviour change around a price limit in a maker-taker Fee market?

Using Nasdaq proprietary data, we compare the percentage of trading volume for HFT and non-HFT prior to the LULD halt between the treatment group and control group in the Nasdaq (maker-taker). The estimation is based on the following regression specification:

$$Y_{i,t} = \alpha_0 + \beta_1 Halt_i + \beta_2 Post_i + \beta_3 Halt_i * Post_t + \gamma VIX_T + \varepsilon_{i,t}, \tag{47}$$

where $Y_{i,t}$ is the market liquidity measures such as trading volume, which add/remove liquidity in a time interval prior and post the LULD halts for HFT and non-HFT in the maker-taker market; α_0 is the intercept; $Halt_i$ is a dummy variable, which is 1 if the security *i* has at least one LULD halt on a particular day and 0 if it does not have an LULD halt; $Post_t$ is a dummy variable, which is 1 if the time interval is after the first halt and 0 otherwise; and VIX_T is the closing value of CBOE's volatility index for day *T*. We estimate the specification with and without stock fixed effects.

Results: Table 5.11 shows that the percentage of HFT trading decreases while the total trading volume increases (coefficient of 0.265*) after the LULD halts on Nasdaq, a maker-taker market. The decrease (coefficient of -0.04***) in HFT trading on Nasdaq is caused by the drop in HFT taking liquidity (coefficient of -0.062***) after the LULD halt.

Table 5.11: Nasdaq HFT vs. Non-HFT Trading around the LULD Halts

The table tests the HFT trading activity during five-minute windows pre and post LULD halts using Nasdaq proprietary trading data. HFT% stands for the percentage of HFT trading on Nasdaq over the Nasdaq trading volume during the time window. HFT_Add% stands for the percentage of HFT adding liquidity on Nasdaq over the Nasdaq trading volume during the time window. HFT_Take% stands for the percentage of HFT taking liquidity on Nasdaq over the Nasdaq trading volume during the time window. NQVolume stands for the trading volume on Nasdaq during the time window. Standard errors are in parentheses and are robust to both heteroscedasticity and correlation within stocks. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**HFT%**	**HFT_Add%**	**HFT_Take%**	**Log(NQVolume)**
Intercept	0.310***	0.257***	0.364***	8.922***
	(0.008)	(0.011)	(0.012)	(0.090)
Post	0.008	-0.002	0.019*	0.039
	(0.008)	(0.011)	(0.011)	(0.086)
Halt	-0.169***	-0.131***	-0.206***	1.030***
	(0.009)	(0.012)	(0.013)	(0.100)
Halt *Post	-0.040***	-0.019	-0.062***	0.265*
	(0.012)	(0.017)	(0.018)	(0.138)
VIX	0.001***	0.001***	0.001***	-0.057***
	(0.000)	(0.000)	(0.000)	(0.003)
Observations	4,498	4,498	4,498	4,498
Adjusted R^2	0.165	0.054	0.133	0.109

H_7: How does HFT trading behaviour change around a price limit in an inverted fee market?

Using Nasdaq/BX proprietary data, we compare the percentage of trading volume for HFT and non-HFT prior to the LULD halt between the treatment group and control group in the Nasdaq/BX market (inverted fee structure). The estimation is based on the following regression specification:

$$Y_{i,t} = \alpha_0 + \beta_1 Halt_i + \beta_2 Post_i + \beta_3 Halt_i * Post_t + \gamma VIX_T + \varepsilon_{i,t}, \qquad (48)$$

where $Y_{i,t}$ is the market liquidity measures such as trading volume, which add/remove liquidity in a time interval prior and post the LULD halts for HFT and non-HFT in the taker-maker market; α_0 is the intercept; $Halt_i$ is a dummy variable, which is 1 if the security *i* has at least one LULD halt on a particular day and 0 if it does not have an LULD halt; $Post_t$ is a dummy variable, which is 1 if the time interval is after the first halt and 0 otherwise; and VIX_T is the closing value of CBOE's volatility index for day *T*. We estimate the specification with and without stock fixed effects.

Table 5.12: Nasdaq/BX HFT vs Non-HFT Trading around the LULD Halts

The table tests the HFT trading activity during five-minute windows pre and post LULD halts using Nasdaq/BX proprietary trading data. HFT% stands for the percentage of HFT trading on Nasdaq/BX over the Nasdaq trading volume during the time window. HFT_Add% stands for the percentage of HFT adding liquidity on Nasdaq/BX over the Nasdaq/BX trading volume during the time window. HFT_Take% stands for the percentage of HFT taking liquidity on Nasdaq/BX over the Nasdaq/BX trading volume during the time window. BXVolume stands for the trading volume on Nasdaq/BX during the time window. Standard errors are in parentheses and are robust to both heteroscedasticity and correlation within stocks. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Dependent Variable	**HFT%**	**HFT_Add%**	**HFT_Take%**	**Log(BXVolume)**
Intercept	0.451***	0.476***	0.427***	7.516***
	(0.015)	(0.021)	(0.019)	(0.107)
Post	-0.016	-0.037*	0.004	-0.043
	(0.014)	(0.020)	(0.019)	(0.098)
Halt	-0.181***	-0.117***	-0.245***	1.034***
	(0.015)	(0.021)	(0.019)	(0.119)
Halt *Post	0.004	0.046	-0.038	0.509***
	(0.020)	(0.029)	(0.026)	(0.174)
VIX	-0.001***	-0.004***	0.001	-0.037***
	(0.000)	(0.001)	(0.001)	(0.003)
Observations	2,246	2,246	2,246	2,246
Adjusted R^2	0.102	0.032	0.122	0.141

Results: Table 5.12 shows that the percentage of HFT trading remains unchanged (coefficient of 0.004) after the LULD halt on Nasdaq/BX, which is a taker-maker market, while the trading volume on Nasdaq/BX after the LULD halt increases (coefficient of 0.509***).

5.5 Conclusion

This chapter analyses the impact of the LULD price limit rule in U.S. equity markets and HFT trading behaviour around price limits in maker-taker and taker-maker markets. We examine the three classic hypotheses - volatility spillover, delayed price discovery, and trading interference - by comparing volatility levels, price pattern change, and trading activity change. We examine the magnet effect using time to next trade when the price approaches the limit prices. In addition, we test how HFT trading behavior changes around price limits in maker-taker market and the inverted markets.

I find that LULD does interfere with trading activity but also curbs short-term volatility without delaying the price discovery process. Also, the magnet effect exists when trade price approaches the price limit and such an effect is stronger when approaching the upper limit. HFT taking liquidity drops significantly while adding liquidity but remains stable after the LULD trading halt on Nasdaq. However, there is no HFT trading behaviour change in the inverted market.

This suggests that LULD still needs subsequent improvement, as it still causes the magnet effect because, theoretically, a stock can be halted for most of the day for multiple five-minute halts. Moreover, Chen, Petukhov, and Wang (2016) indicate that any kind of circuit breaker design may result in the magnet effect simply because the market is closed. Also, LULD appears to be only affecting HFT trading activities on the maker-taker market compared with the inverted markets.

Chapter 6 Conclusion

This thesis examines the impact of three of the most important stock market designs: exchange access fee, tick size, and dynamic price limit. *First of all*, this thesis provides a new theoretical model of maker-taker fees to show that the entire fee rebate must be consumed by an increase in the information content of informed trades in equilibrium, motivated by a *rise* in the trading profitability of informed market orders paying the taker fee. Uninformed traders neither benefit from the fee rebate nor suffer from it. We then use the Nasdaq access fee pilot as a natural experiment to empirically investigate how a reduction in the maker-taker fee affects market competition, liquidity, and HFT/non-HFT make-or-take decisions.

The current literature holds that only changes in the net exchange fee matter for market quality and transaction cost efficiency. Our findings disagree and instead support our theoretical implications that although the raw effective spreads were largely unaffected by the rebate reduction, the information content of the taker order flow fell substantially. Holding the exchange net fee constant, the change in component fees and rebates does matter, but not for competition between lit and dark venues because the latter predominantly exclude informed traders. Rather, when both the fee and the rebate were reduced during the Nasdaq access fee pilot, Nasdaq market share declines which benefitted other high rebate-paying lit exchanges.

The equivalently reduced fee and rebate lowers quote quality and routed volume to Nasdaq while enhancing the fill rate and speed of fill because of the reduced taker fee and a thinner market. With their relative routing position improved, adverse selection costs decline on Nasdaq and liquidity supplier profit increases. The decline in informed equilibrium order flow and reduced liquidity means that Nasdaq's existing strategy of providing the highest possible subsidy to liquidity makers funded by its tax on liquidity takers encourages better price discovery and higher market efficiency. Profits of informed trades are enhanced with liquidity providers and liquidity traders no worse off and probably better off. Hence, fee rebates to liquidity suppliers can improve Pareto efficiency as predicted by our model.

In addition, we find no supporting evidence to the public assertion that high exchange access fees cause trading to shift away from exchanges to dark pools. Regarding to the impact on HFT, HFT traders tend to switch from adding to removing liquidity when exchange access fees and rebates decrease. While standard quality measures all appear to improve on a cum-rebate basis, these apparent improvements reflect the lower information content in Nasdaq's order flow during the experiment as informed order flows shift to exchanges that maintain high rebates. Given their relative inability to monitor the now less informed order flow dynamics, non-HFT firms increased liquidity making and decreased their liquidity taking while HFT firms did just the reverse, albeit with a less informed equilibrium order flow.

Secondly, using the SEC mandated nickel tick size pilot as an exogenous shock, this thesis examines whether exchange access fees matter when tick size changes. We further extend the exchange fee-structure model to incorporate minimum tick size constraints to show that the limit on fees set by the SEC is too low to enable fees to neutralize the harmful effect of the 400% increase in the minimum tick size. Because there is a downward sloping demand schedule for liquidity trades, our model shows that maker-taker venues with the highest cost of make orders, must lose the most market share to both inverted venues and off-market as the minimum tick size is raised. Moreover, both venue types should experience a relative increase in informed order flow as uninformed traders depart, with maker-taker markets experiencing the largest relative increase.

Inverted venues gain market share due to a substantial increase in incentive to post non-marketable orders while the maker-taker venues experience a decline. The intuition is that the cost of crossing the spread increases as the tick size widens, however, the cost is even higher on maker-taker venues because there is an additional fee for taking liquidity. Consequently, one is better off by executing market orders in an inverted fee market since consumers of liquidity receive a rebate to offset a material portion of the widened tick, especially for the most tick-constrained stocks. Price discovery becomes more efficient while relative price impact decreases on the inverted venue as its quote quality improves further.

Moreover, the consolidated trading volume decreased overall suggesting small capitalization stocks have not attracted increased trading interest despite an increase in the quote quality. The overall transaction cost increased significantly, which leads to a higher price impact as only the highly informed orders can afford to cross the spread. However, the price discovery process is enhanced and less volatile for the most tick-constrained group. Furthermore, lit market share increased for treatment firms which are subject to the 'trade-at' rule indicating that imposing minimum price improvement in opaque venues will restrict off-exchange trading activity.

Thirdly, this thesis analyses the impact of the LULD in U.S. equity markets and HFT trading behaviour around price limits in maker-taker and taker-maker markets. We examine the three classic hypotheses-volatility spillover, delayed price discovery, and trading interference-by comparing volatility levels, price pattern change and trading activity change. We examine the magnet effect using time to next trade when price approaches the limit prices;

In addition, we test how HFT trading behaviour changes around price limits in maker-taker market and the inverted markets.

Using difference-in-difference with propensity score matching methodology, I find that LULD does interfere with trading activity but also curbs short-term volatility without delaying the price discovery process. Also, the magnet effect exists when trade price approaches the price limit and such an effect is stronger when approaching the upper limit. HFT taking liquidity drops significantly while adding liquidity but remains stable after the LULD trading halt on Nasdaq. However, there is no HFT trading behaviour change in the inverted market.

This suggests that LULD still needs subsequent improvement, as it still causes the magnet effect because, theoretically, a stock can be halted for most of the day for multiple five-minute halts. Also, LULD appears to be only affecting HFT trading activities on the maker-taker market compared with the inverted market.

References

Abad, D., and R. Pascual, 2007, On the magnet effect of price limits, European Financial Management 13:833-852.

Ahn, H.J., C.Q. Cao, and H. Choe, 1998, Decimalization and competition among stock markets: Evidence from Toronto Stock Exchange cross-listed securities, Journal of Financial Markets 1:51-87.

Ahn, H.J., C. Jun, K. Chan, and Y. Hamao, 2007, Tick size change and liquidity provision on the Tokyo Stock Exchange, Journal of the Japanese and International Economies 21:173-194.

Aitken, M., and C. Comerton-Forde, 2005, Do reductions in tick sizes influence liquidity? Accounting and Finance 45:171-184.

Aitken, M, G.T. Garvey, and P.L. Swan, 1995, How brokers facilitate trade for long-term clients in competitive securities markets, Journal of Business 68:1-33.

Alampieski, K., and A. Lepone, 2009, Impact of tick size reductions on liquidity: Evidence from the Sydney Futures Exchange, Accounting and Finance 49:1-20.

Angel, J.J., 1997a, Picking your tick: Towards a new theory of stock splits, Journal of Applied Corporate Finance 10:58-68.

Angel, J.J., 1997b, Tick Size, share prices, and stock splits, Journal of Finance 52:655-681.

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

Angel, J., L. Harris, and C.S. Spatt, 2015, Equity trading in the 21st century: An update, Quarterly Journal of Finance 5:1-39.

Arak, M., and R.E. Cook, 1997, Do daily price limits act as magnets? The case of treasury bond futures, Journal of Financial Services Research 12:5-20.

Baruch, S., M. Panayides, and K. Venkataraman, 2017, Informed trading and price discovery before corporate events, Journal of Financial Economics 125:561-588.

Bacidore, J.M., 1997, The impact of decimalization on market quality: An empirical investigation of the Toronto Stock Exchange, Journal of Financial Intermediation 6:92-120.

Barclay, M.J., T. Hendershot, and D.T. McCormick, 2003, Competition among trading venues: Information and trading on electronic communications networks, Journal of Finance 58:2637-2665.

Barclay, M.J., E. Kandel, and L.M. Marx, 1998, The effects of transaction costs on stock prices and trading volume, Journal of Financial Intermediation 7:130-150.

Battalio, R.H., S.A. Corwin, and R. Jennings, 2016, Can brokers have it all? On the relation between make-take fees and limit order execution quality, Journal of Finance 71:2193-2238.

Berkman, H., and O.W. Steenbeek, 1998, The influence of daily price limits on trading in Nikkei futures, Journal of Futures Markets 18:265-279.

Bessembinder, H., 2000, Tick size, spreads, and liquidity: An analysis of Nasdaq securities trading near ten dollars, Journal of Financial Intermediation 9:213-239.

Bollen, N.P.B., and J.A. Busse, 2006, Tick size and institutional trading costs: Evidence from mutual funds, Journal of Financial and Quantitative Analysis, 41:915-937.

Bollen, N.P.B., and R.E. Whaley, 1998, Are 'tennies' better? Journal of Portfolio Management 25:10-24.

Bollen, N.P.B., T. Smith, and R.E. Whaley, 2003, Optimal contract design: For whom? Journal of Futures Markets 23:719-750.

Bourghelle, D., and F. Declerck, 2004, Why markets should not necessarily reduce the tick size, Journal of Banking & Finance 28:373-398.

Brennan, M.E., 1986, A theory of price limits in future markets, Journal of Financial Economics 16:213-233.

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

Brogaard, J., and K. Roshak, 2016, Prices and price limits, Working Paper.

Brolley, M., and K. Malinova, 2012, Informed trading and maker-taker fees in a low latency limit order market, Working Chapter, University of Toronto.

Buti, S., B. Rindi, and I.M. Werner, 2017, Dark pool trading strategies, Journal of Financial Economics 124:244–265.

Cai, J., Y. Hamao, and R.Y.K. Ho, 2008, Tick size change and liquidity provision for Japanese stocks trading near ¥1000, Japan and the World Economy 20:19-39.

Cardella, L., J. Hao, and I. Kalcheva, 2015, Make and take fees in the U.S. equity market, Working Paper, Texas Tech University.

Carrion, A., 2013, Very fast money: High-frequency trading on the NASDAQ, Journal of Financial Markets 16:680-771.

Chakravarty, S., B.F. Van Ness, and R.A. Van Ness, 2005, The effect of decimalization on trade size and adverse selection costs, Journal of Business Finance & Accounting 32:1063-1081.

Chakravarty, S., R.A. Wood, and R.A. Van Ness, 2004, Decimals and liquidity: A study of the NYSE, Journal of Financial Research 27:75-94.

Chan. K.C. and C.Y. Hwang, 2001, The impact of tick size on the quality of a pure order-driven market: Evidence from the Stock Exchange of Hong Kong, Working Paper, Hong Kong University of Science & Technology.

Chan, S.H., K.A. Kim, and S.G. Rhee, 2005, Price limit performance: Evidence from transactions data and the limit order book, Journal of Empirical Finance 12:269-290.

Chao, Y., C. Yao, and M. Ye, 2017a, Why discrete price fragments U.S. stock exchanges and disperses their fee structures, Working Paper, University of Illinois at Urbana-Champaign.

Chao, Y., C. Yao, and M. Ye, 2017b, Discrete pricing and market fragmentation: a tale of two-sided markets, American Economic Review: Papers & Proceedings 107:196–199.

Chen, H., A. Petukhov, and J. Wang, 2016, A Dynamic Model of Circuit Breakers, Working Paper, MIT.

Cho, D., J. Russell, G.C. Tiao, and R. Tsay, 2003, The magnet effect of price limits: Evidence from high-frequency data on Taiwan Stock Exchange, Journal of Empirical Finance 10:133-168.

Chordia, T., R. Roll, and A. Subrahmanyam, 2001, Market liquidity and trading activity, Journal of Finance 56:501-530.

Chordia, T, R. Roll, and A. Subrahmanyam, 2008, Liquidity and market efficiency, Journal of Financial Economics 87:249-268.

Chung, K.H., and C. Chuwonganant, 2004, Tick size, order handling rules, and trading costs, Financial Management 33:47-62.

Chung, K.H., K.A. Kim, and P. Kitsabunnarat, 2005, Liquidity and quote clustering in a market with multiple tick sizes, Journal of Financial Research 28:177-195.

Clapham, B., P. Gomber, J. Lausen, and S. Panz, 2017, Liquidity provider incentives in fragmented securities markets, Working Paper, Goethe University.

Colliard, J., and T. Foucault, 2012, Trading fees and efficiency in limit order markets, Review of Financial Studies 25:3389-3421.

Comerton-Forde, C., V. Gregoire, and Z. Zhong, 2017, Inverted fee venues and market quality, Working Paper, University of Melbourne.

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

Corwin, S.A., and M.L. Lipson, 2000, Order flow and liquidity around NYSE trading halts, Journal of Finance 55:1771-1801.

Davies, R.J., and S.S. Kim, 2009, Using matched samples to test for differences in trade execution costs, Journal of Financial Markets, 12:173-202.

Dayri, K., and M. Rosenbaum, 2015, Large tick assets: Implicit spread and optimal tick size, Market Microstructure and Liquidity 1:1550003.

Dosanjh, J.S., 2013, Market maker incentives and market efficiency: Evidence from the Australian ETF market, Working Paper, CMCRC.

Draus, S., and M.V. Achter, 2015, Circuit breakers and market runs, Working Paper, Erasmus University, the Netherlands.

Du, Y., Q. Liu, and S.G. Rhee, 2009, An analysis of the magnet effect under price limits, International Review of Finance 9:83-110.

Easley, D., N.M. Kiefer and M. O'Hara, 1996, Cream-skimming or profit-sharing? The curious role of purchased order flow, Journal of Finance 51:811-33.

Easley, D., M. O'Hara, and G. Saar, 2001, How stock splits affect trading: A microstructure approach, Journal of Financial and Quantitative Analysis 36:25-51.

Ekaputra, I.A., and B. Ahmad, 2007, The impact of tick size reduction on liquidity and order strategy: Evidence from the Jakarta Stock Exchange, Economics and Finance in Indonesia 55:89-104.

Fama, E., 1970, Efficient capital markets: A review of theory and empirical work, Journal of Finance 25:383-417.

Fama, E.F., 1989. Perspectives on October 1987, or What did we learn from the crash? In: Kamphuis Jr., R.W., R.C. Kormendi, and J.W. Henry Watson (Eds.), Black Monday and the Future of the Financial Markets. Homewood, IL:Irwin.

Foley, S., and T. Putnins, 2014, Regulatory efforts to reduce dark trading in Canada and Australia: How have they worked? Report prepared for the CFA Institute.

Foucault, T., O. Kadan, and E. Kandel, 2005, Limit order book as a market for liquidity, Review of Financial Studies18:1171-1217.

Foucault, T., O. Kadan, and E. Kandel, 2013, Liquidity cycles and make/take fees in electronic markets, Journal of Finance 68:299-341.

Glosten, L., and P. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, Journal of Financial Economics 14:71-100.

Goldstein, M.A., and K.A. Kavajecz, 2004, Trading strategies during circuit breakers and extreme market movements, Journal of Financial Markets 7:301-333.

Goldstein, M.A., and K.A. Kavajecz, 2000, Eights, sixteenths, and market depth: Changes in tick size and liquidity provision on the NYSE, Journal of Financial Economics 56:125-149.

Gomber, P., B. Clapham, M. Haferkorn, S. Panz, and P. Jentsch, 2016, Circuit breakers – A survey among international trading venue, Goethe University Frankfurt.

Government Office for Science, 2012, Tick size regulation: Costs, benefits and risks.

Greenwald, B. C., and J.C. Stein, 1991, Transactional risk, market crashes, and the role of circuit breakers, Journal of Business 64:443-462.

Gresse, C., 2017, Effects of lit and dark market fragmentation on liquidity, Journal of Financial Markets, forthcoming.

Griffith, T.G., and B.S. Roseman, 2017, Making cents of tick sizes: An investigation of the 2016 U.S. SEC tick size pilot, Working Paper.

Harris, L., 1991, Stock price clustering and discreteness, Review of Financial Studies 4:389-415.

Harris, L., 1994, Minimum price variations, discrete bid-ask spreads, and quotation sizes, Review of Financial Studies 7:149-178.

Harris, L., 2013, Maker-taker pricing effects on market quotations, Working Paper, University of Southern California.

Hasbrouck, J., 1991, Measuring the information content of stocks trades, Journal of Finance 46:179-207.

Hatheway, F., 2015a, Nasdaq access fee experiment: Report I, Nasdaq Report.

Hatheway, F., 2015b, Nasdaq access fee experiment: Report II, Nasdaq Report.

Hatheway, F., A. Kwan and H. Zheng, 2016, An empirical analysis of market segmentation on U.S. equities markets, Journal of Financial and Quantitative Analysis, forthcoming.

Hau. H., 2006, The role of transaction costs for financial volatility: Evidence from the Paris Bourse, Journal of the European Economic Association 4:862-890.

Hsieh, P. H., Y.H. Kim, and J.J. Yang, 2009, The magnet effect of price limits: A logit approach, Journal of Empirical Finance 16:830-837.

Hsieh, T.Y., S.S. Chuang, and C.C. Lin, 2008, Impact of tick-size reduction on the market liquidity-evidence from the emerging order driven market, Review of Pacific Basin Financial Markets and Policies 11:591-616.

Huang, W., C.A. Lehalle, and M. Rosenbaum, 2015, How to predict the consequences of a tick value change: Evidence from the Tokyo Stock Exchange pilot program.

Interactive Broker Knowledge Base, 2017, Add remove liquidity, available at http://ibkb.interactivebrokers.com/node/201

Jain, P.K., 2003, Institutional design and liquidity at stock exchanges around the world, Working Paper, University of Memphis.

Jiang, C.X., T.H. McInish, and J. Upson, 2012, Market fragmentation and information quality: The role of TRF trades, SSRN eLibrary.

Jones, C.M., and M.L. Lipson, 2001, Sixteenths: Direct evidence on institutional execution costs, Journal of Financial Economics 59:253-278.

Kadan, O., 2006, So who gains from a small tick size, Journal of Financial Intermediation, 15:32-66.

Kaniel, R., and H. Liu, 2006, So what orders do informed traders use? Journal of Business 79:1867-1913.

Karagozoglu, A.K., T. F. Martell, and G.H.K. Wang, 2003, The split of the S&P 500 futures contracts: Effects on liquidity and market dynamics, Review of Quantitative Finance and Accounting 21:323-348.

Ke, M.C., C.H. Jiang, and Y.S. Huang, 2004, The impact of tick size on intraday stock price behaviour: Evidence from the Taiwan stock exchange, Pacific-Basin Finance Journal 12:19-39.

Kim, K. A., and J. Park, 2010, Why do price limits exist in stock markets? A manipulation-based explanation, European Financial Management, 16:296-318.

Kim, K., and S.G. Rhee, 1997, Price limit performance: evidence from the Tokyo stock exchange, Journal of Finance 52:885-901.

Kim, K.A., and R.J. Sweeney, 2002, The effects of price limits on information revelation: Theory and evidence, Working Paper, University of Wisconsin, Milwaukee, WI.

Kirilenko, A., A.S. Kyle, M. Samadi, and T. Tuzun, 2017, The Flash Crash: High-frequency trading in an electronic market, Journal of Finance 72:967-998.

Kuhn, B.A., G.J. Kuserk, and P. Locke, 1991, Do circuit breakers moderate volatility? Evidence from October 1989, Review of Futures Markets 10:136-179.

Kurov, A., 2008, Tick size reduction, execution costs, and informational efficiency in the regular and E-mini Nasdaq-100 index Futures Markets, Journal of Futures Markets 28:871-888.

Kwan, A., R. Masulis, and T.H. McInish, 2015, Trading rules, competition for order flow and market fragmentation, Journal of Financial Economics 115:330-348.

Kyle, A.S., 1985, Continuous auctions and insider trading, Econometrica 53:1315-1335.

Lauterbach, B., and U. Ben-Zion, 1993, Stock Market Crashes and the Performance of Circuit Breakers: Empirical Evidence, Journal of Finance 48:1909-1925.

Lee, M.C., and M.J. Ready, 1991, Inferring Trade Direction from Intraday Data, Journal of Finance 46:733-746.

Lee, M.C., M.J. Ready, and P.J. Seguin, 1994, Volume, volatility, and New York Stock Exchange trading halts, Journal of Finance 49:183-214.

Lehmann, B.N., 1989, Commentary: volatility, price resolution, and the effectiveness of price limits, Journal of Financial Services Research 3:205-209.

Linnainmaa, J.T., 2010, Do limit orders alter inferences about investor performance and behavior? Journal of Finance 65:1473–1506.

Lipson, M.L., and S. Mortal, 2006, The effect of stock splits on clientele: Is tick size relevant? Journal of Corporate Finance 12:878-896.

Lo, A.W., and C. MacKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, Review of Financial Studies 1:41-66.

Lutat, M., 2010, The effect of maker-taker pricing on market liquidity in electronic trading systems—Empirical evidence from European equity trading, Working Paper, University of Frankfurt.

Ma, C.K., P.R. Ramesh, and R.S. Sears, 1989a, Limit moves and price resolution: The case of the treasury bond futures market, Journal of Futures Markets 9:321-335.

Ma, C.K., P.R. Ramesh, and R.S. Sears, 1989b, Volatility, price resolution, and the effectiveness of price limits, Journal of Financial Services Research 3:165-199.

MacKinnon, G., and H. Nemiro, 1999, Liquidity and tick size: Does decimalization matter? Journal of Financial Research 22:287-299.

MacKinnon, G., and H. Nemiro, 2004, Tick size and the returns to providing liquidity, International Review of Economics & Finance 13:57-73.

Madhavan, A., 2012, Exchange-traded funds, market structure and the flash crash, Financial Analyst Journal 68:20-35.

Maglaras, C., C.C. Moallemi, and H. Zheng, 2012, Optimal order routing in a fragmented market, Working Paper, Columbia University.

Malinova, K., and A. Park, 2015, Subsidizing liquidity: The impact of make/take fees on market quality, Journal of Finance 70:509-536.

Miller, M.H., 1989, Commentary: Volatility, price resolution, and the effectiveness of price limits, Journal of Financial Services Research 3:201-203.

Moise, C., and P. Flaherty, 2017, Limit Up-Limit Down pilot plan and associated events, SEC White Paper.

Nath, P., 2003, Do price limits behave like magnets? Working Paper, London Business School.

Niemeyer, J., and P. Sandas, 1994, Tick size, market liquidity and trading volume: Evidence from the Stockholm Stock Exchange, Working Paper.

O'Hara, M., G. Saar, and Z. Zhong, 2015, Relative tick size and the trading environment, Working Paper.

O'Hara, M., and M. Ye, 2011, Is market fragmentation harming market quality? Journal of Financial Economics 100:459-474.

Panayides, M., B. Rindi, and I.M. Werner, 2016, Trading fees and intermarket competition, Working Paper.

Pavabutr, P., and S. Prangwattananon, 2009, Tick size change on the Stock Exchange of Thailand, Review of Quantitative Finance and Accounting 32:351-371.

Penalva, J.S., and M. Tapia, 2017, Revisiting tick size: Implications from the SEC tick size pilot, Working Paper.

Porter, D.C., and D.G. Weaver, 1997, Tick size and market quality, Financial Management 26:5-26.

Ritter, J.R., 2014, Reenergizing the IPO market, Journal of Applied Finance 24:37-48.

Roll, R., 1989, Price volatility, international market links, and their implications for regulatory policies, Journal of Financial Services Research 3:211-246.

Rosu, I., 2015, Fast and slow informed trading, Working Paper, HEC Paris.

Rosu, I., 2016, Liquidity and information in order driven markets, Working Paper, HEC Paris.

Schultz, P., 2000, Stock splits, tick size and sponsorship, Journal of Finance 55:429-450.

SEC, 2005, Final Rule: Regulation NMS, available at https://www.sec.gov/rules/final/34-51808.pdf

SEC, 2010, Concept release on market structure, Release No. 34-61358, File No. S7-02-10, Discussion Paper.

SEC, 2014, Nasdaq Form 19b-4 filing on proposed access fee pilot, available at http://nasdaq.cchwallstreet.com/NASDAQ/pdf/nasdaq-filings/2014/SR-NASDAQ-2014-128.pdf.

SEC, 2015a, Memo on maker-taker fees on equities exchange, available at https://www.sec.gov/spotlight/emsac/memo-maker-taker-fees-on-equities-exchanges.pdf.

SEC, 2015b, Memo on Rule 611 of Regulation NMS, available at https://www.sec.gov/spotlight/emsac/memo-rule-611-regulation-nms.pdf

SEC, Commissioner Luis A. Aguilar, 2015c, The need for greater secondary market liquidity for small businesses, available at https://www.sec.gov/news/statement/need-for-greater-secondary-market-liquidity-for-small-businesses.html.

Sie, T.L., and T.H. McInish, 1995, Reducing tick size on the Stock Exchange of Singapore, Pacific-Basin Financial Journal 3:485-496.

Skjeltorp, J.A., E. Sojli, and W.W. Tham, 2016, Identifying cross-sided liquidity externalities, Management Science, forthcoming.

Smith, B.F., D.A.S. Turnbull, and R.W. White, 2006, The impact of pennies on the market quality of the Toronto stock exchange, Financial Review 41:273-288.

Spicer, J., 2009, Lure of rebates drives US trading despite selloff, Reuters April 25, 2009, available at http://www.reuters.com/article/us-trading-volumes-analysis/lure-of-rebates-drives-us-trading-despite-selloff-idUSTRE53N68S20090424.

Subrahmanyam, A., 1994, Circuit breakers and market volatility: A theoretical perspective, Journal of Finance 49:527-543.

Subrahmanyam, A., 1997, The ex-ante effects of trade halting rules on informed trading strategies and market liquidity, Review of Financial Economics 6:1-14.

Subrahmanyam, A., 2013, Algorithmic trading, the Flash Crash, and coordinated circuit breakers, Borsa Istanbul Review 13:4-9.

The UK's Government Office for Science, 2012, The future of computer trading in financial markets: An international perspective, Final Project Report.

Tooma, E.A., 2011, The magnetic attraction of price limits, International Journal of Business 16:35-50.

Van Ness, B.F., R.A. Van Ness, and S.W. Pruitt, 2000, The impact of the reduction in tick increments in major us markets on spreads, depth, and volatility, Review of Quantitative Finance and Accounting 15:153-167.

Wu, Y., T. Krehbiel, and B. W. Brorsen, 2011, Impacts of tick size reduction on transaction costs, International Journal of Economics and Finance 3:57-65.

Yao, C., and M. Ye, 2015, Tick size constraints, high-frequency trading, and liquidity, Working Paper, University of Illinois.

Zhu, H., 2014, Do dark pools harm price discovery? Review of Financial Studies, 27:747–789.