Essays on

Financial Markets



Tito Bastianello

Merton College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2019

Abstract

This thesis is a collection of three stand-alone essays.

The first essay, *Modelling the Maker-Taker Pricing in Financial Markets*, investigates the trading fees charged by stock exchanges to brokers. In particular, it studies the maker-taker pricing structure which is a form of second degree price discrimination. This is a theoretical work which develops a model of a limit order book with a closed-form solution that proxies a modern electronic trading platform. The model builds on the current literature by introducing competition among exchanges and allowing depth to build, thus providing predictions for market quality (spread, depth, and volume) and trading activity (limit and market orders) following a change in the trading fees. Maker-taker fees are the most prominent pricing scheme amongst stock exchanges.

The second essay, *Assessing the Attributes and Implications of Mini Flash Crashes*, investigates all mini flash crashes (MFCs) by U.S. equities between January 3, 2006 and February 3, 2011. A mini flash crash is defined as a change in the stock price that is larger than 0.8%, takes place in less than 1.5 seconds for at least 10 consecutive trades, and recovers to within 0.1% of the starting price. This is an empirical work that explores both the causes (or attributes) and implications of MFCs, and it is concluded that MFCs negatively affect market quality (measured as inside spread and BBO depth) over a 10-minute window around the MFC. Besides, the phenomenon of MFCs has deep repercussions on market agents, e.g. brokers trading at the volume-weighted average price.

The third essay, *Leadership of the Price Discovery Process at Outer Levels of the Book*, investigates price discovery, defined as the impounding of newly formed fundamental information into the observed price of the asset, and measured by the contribution different trading venues make to the common efficient price. This work contributes to the literature by showing that the market which leads the price discovery process at the BBO needs not enjoy leadership at other levels of the book, particularly at outer levels, i.e. levels that are distant from the BBO. This is of relevance since less than 50% of liquidity is posted to the BBO so focusing exclusively on the first level of the book does not provide a full understanding of which market truly leads the price discovery process.

Acknowledgements: I would like to thank my supervisor Jerry Tsai for his availability, feedback, and general advice, and without whose continuous support this thesis simply would not have been possible. I would also like to thank Francesco Zanetti and Kevin Sheppard for their helpful contribution to this work, and Lesley Darcy, Julie Minns, Michalis Rousakis and Bassel Tarbush for their guidance. Last, but not least, I should also mention Guus and Sebastian for clarifying that ipsa scientia potestas est.

Word Count: Using page 1 as a typical page of text, this thesis is approximately 58,000 words long.

List of Contents

Abstract i

List of contents iii

List of figures v

List of tables vii

Introduction 1

Chapter:

(i) Modelling the maker-taker pricing in financial markets 7

(1) Introduction . . . 8

(2) Literature review . . . 12

(3) Model . . . 13

(4) Empirical predictions . . . 22

(5) Empirical analysis . . . 29

(6) Conclusions . . . 39

(7) Appendix . . . 40

(ii) Assessing the attributes and implications of mini flash crashes 51

(1) Introduction . . . 52

(2) Literature review . . . 54

(3) MFC overview – relevance, definition and dataset . . . 55

(4) MFC attributes – risk and return . . . 59

(5) MFC implications – market quality . . . 70

(6) Conclusions . . . 76

(7) Appendix 78

(iii) Leadership of the price discovery process at outer levels of the book 95

(1) Introduction 96

(2) Literature review 98

(3) Model 101

(4) Information share 105

(5) Empirical evidence 110

(6) Conclusions 118

(7) Appendix 119

References 137

List of Figures

Chapter 1

1.1 Canadian Trading Centres 11
1.2 Generalising the Increase in the LOB Make Fee 45
1.3 Generalising the Reduction in the LOB Take Fee 46
1.4 Market Quality Metrics – Qualitative Descriptive Statistics 47

Chapter 2

2.1 Mini Flash Crashes over Sampling Period 62
2.2 Mini Flash Crashes by Stock Exchange 63
2.3 Frequency of MFCs – Daily and Intraday Distributions 64
2.4 Amplitude of MFCs – Daily and Intraday Distributions 66
2.5 Distribution of MFCs – Daily Persistence 69
2.6 Test 1 – MFC against 10-minute window 73
2.7 The Great Financial Crisis 89
2.8 Regulation NMS, High Frequency Trading, and Mini Flash Crashes 89
2.9 Test 1 by Stock Exchange I/II 90
2.10 Test 1 by Stock Exchange II/II 91

Chapter 3

3.1 Qualitative Descriptive Statistics by Book Level 114
3.2 The Best Bid-Offer of the Limit Order Book 130
3.3 The Best Bid-Offer of the Limit Order Book by Tercile 132
3.4 The Best Bid-Offer of the LOB by Market 133

List of Tables

Chapter 1

1.1 Make-Take Fee Changes applied to Canadian Equities on June 1, 2016 11

1.2 Increase in the LOB Make Fee (Reduction in the LOB Rebate) 24

1.3 Reduction in the LOB Take Fee 26

1.4 Increase in LOB/ATS Make Fee & Reduction in LOB/ATS Take Fee 28

1.5 Panel Data Regression 34

1.6 Robustness of Control Group – Parallel Trend Assumption 38

1.7 Daily Executed Volume for Canadian Equities 48

1.8 Market Quality Metrics – Definition 48

1.9 Sample 49

1.10 Market Quality Metrics – Quantitative Descriptive Statistics 50

Chapter 2

2.1 Sample I/II 60

2.2 Sample II/II 61

2.3 Mini Flash Crashes by Tabulated Frequency 62

2.4 Frequency of MFCs – Risk and Return 65

2.5 Amplitude of MFCs – Risk and Return 68

2.6 Distribution of MFCs – Risk and Return 70

2.7 Test 2 – pre-MFC against post-MFC 74

2.8 Market Quality Metrics – Definition 92

2.9 Test 2 by Stock Exchange 92

2.10 Tabulation of Figure 2.6 93

Chapter 3

3.1 Quantitative Descriptive Statistics by Book Level 113
3.2 Information Share Ratio 117
3.3 Sample I/II 134
3.4 Sample II/II 135
3.5 p-Values of Information Share Ratio 136

Introduction

The term "financial market" broadly describes any marketplace where the trading of securities including equities, bonds, currencies and derivatives occurs. Gathering spaces where goods and services are bought and sold date as far back as the 10th-8th centuries BC when free-born Greek citizens would gather in the city's Agora where merchants operated stalls to sell their goods. However, contrary to markets for goods and services which have been part of the economy since the very beginning of our history, capital markets are much more recent and it was not until the XVII$^{\text{th}}$ century that the first stock exchange was formed. Indeed, the Amsterdam Stock Exchange was established in 1602 by the Dutch East India Company which issued the first shares, as well as the first bonds. For the first four centuries since their initial inception stock exchanges were few and the overall value of the assets that were exchanged on them was modest, and so was the role of financial exchanges within the global economy as at the time they "only" played a subsidiary role to the real economy. Over time the role of financial markets grew steadily, as companies looked for new capital and investors for assets to exchange. Then, towards the end of the XX$^{\text{th}}$ century, technology enabled remote trading giving rise to a complex system of interlinked computers capable of trading at the speed of the light. This phenomenon coincided with the tumultuous growth of algorithmic trading which uses a computer program following a defined set of instructions (an algorithm) to place a trade, to the extent that the Wall Street Journal recently estimated that in 2018 approximately 85% of all stock trading was controlled by algorithms. Furthermore, remote trading was supported by a change in legislation and this led to a proliferation of trading centres, a phenomenon known as fragmentation. Financial markets today are therefore enabled by technology and are made up by a complex system of trading venues competing among each other for volumes, where agents can operate simultaneously on more than one venue and automated trading plays a dominant role.

Alongside their evolution in size, complexity and speed, financial markets have also grown in importance. To that extent, while they once played a subsidiary role to the real economy, today

they are a dominant force, an important driver of economic growth, and the way they perform often dictates the fate of companies, governments and real economies. A recent example of the role played nowadays by financial markets was the 2007/08 Global Financial Crisis which saw a crisis originating in the financial markets spread to the banking system and through this to the real economy: these are different crises all interconnected with one another, the firewalls having been removed by new regulation such as the Gramm-Leach-Bliley Act of 1999 which repealed the Glass-Steagall Act of 1933. The outcome was a virtual paralysis of the real economy as investments dried up in view of failing capital markets. This brought about a global crisis that some countries ten years on have not yet fully recovered from.

Driven by the increasing and determining role of financial markets for the overall performance of the real economy, and consequently by their direct impact on the well-being of billions of people, a growing interest developed on how capital markets work, what their effects are, how they should be framed and if necessary regulated to ensure they best serve their end purpose of efficiently allocating resources. And every time financial markets cause damage to the real economy there is a renewed drive to better understand their dynamics, based on the now accepted principle that even Adam Smith's "invisible hand" must somehow be guided and if necessary regulated. With this is mind, it is of no surprise that a new branch of literature called Market Microstructure was born to investigate how financial markets work and how they compete against each other, how they should be framed and regulated to make them more efficient, what sort of actions will maximise their result with what sort of impact on the welfare of agents, as well as the role played by different types of agents in determining outcomes. Market Microstructure deals with the study of the trading mechanisms used for financial securities, and while historical antecedents to the field can probably be found going as far back as the beginning of the written language, the field eventually acquired a distinct identity. As good a starting point as any is the minting of the term "market microstructure" in the paper of the same title by Garman (1976). This thesis sits within and contributes to Market Microstructure by investigating three different aspects pertaining to the modern functioning of financial markets.

The first essay, *Modelling the Maker-Taker Pricing in Financial Markets*, investigates the effect

of maker-taker fees on financial markets. Notably, the increased importance of maker-taker fees stems from the proliferation of exchanges on which an asset is traded, a phenomenon known as "fragmentation". This dynamic, enabled both by technology and remote trading as well as by the relaxation of regulations, introduced a new factor that was previously basically absent, namely competition among trading venues. Since trading centres are private enterprises run for profit, they compete among each other to attract volumes at remunerative terms, as this is key to their success. Maker-taker fees are a competitive lever that is widely used by exchanges to attract volumes and their structure significantly affects the behaviour and the choices of agents, which in turn impact market quality. To analyse them, this paper builds a model of a limit order book that has a closed-form solution, found by backward-induction on the set of all possible decisions a trader could make. The advantage of this approach is that it does not require a steady state, thus allowing depth to build on the book. This paper concludes that an increase of the make fee in a market with competition leads to a fall in liquidity while the effects on market quality are mixed: inside spread widens, BBO depth deepens and volumes fall. In summary, this paper contributes to the literature by building a model that allows to predict the effect of a change in fees on market quality while allowing for competition, a discrete tick size, and depth to build.

The second essay, *Assessing the Attributes and Implications of Mini Flash Crashes*, investigates a phenomenon that did not exist when prices were set by humans, namely spikes in the volatility of prices that are invisible to the human eye. Flash crashes are sharp movements in prices that quickly recover close to their previous levels within minutes, and they are presumably driven or at least enhanced by electronic and algorithmic trading. They constitute a burgeoning field of research with clear implications on traders' welfare and have attracted much attention both from academics and regulators. As an example, the flash crash of May 6th, 2010 was caused by one unusually large sell order of S&P futures by Waddel & Reed Financial. The size of the order significantly affected the price of the asset and this triggered a sale spree by high frequency algorithmic traders that, as the price dropped, immediately tried to cut their losses by selling too. This quickly spread to the rest of the market and as a result both the S&P and the Dow Jones dropped 9% within minutes, the largest such drop in history until that moment. Following this incident the Securities and

Exchange Commission introduced the so called "circuit breakers", measures aimed at preventing panic selling and the price volatility it causes by stopping trading when certain tripwires are crossed. A similar but lesser known phenomenon which also takes place on financial markets is that of the so called "mini" flash crashes. These are sharp oscillations in the price of an asset that are reversed within seconds. They occur on a daily basis (an average of three a day is detected in our sample) whenever the price significantly drops/surges and then recovers within a time frame that cannot possibly allow for a similar variation of the fundamentals, nor can it be driven by human actions. These oscillations have so far drawn little attention as most of the literature seems to have focussed on the bigger, wider and louder crashes. However, they are both significant and relevant. They are significant because they determine a substantial redistribution of wealth among agents which may or may not be random; and it is relevant because if one assumes that fundamentals cannot possibly show such wide and frequent swings within such a short period of time (it is so short it would not allow for news to reach the market and be embedded into the price of the asset), then it proves that the price of an asset does not always correctly reflect the value of the underlying fundamental. In summary, this paper concludes that around a mini flash crash the inside spread widens, the BBO depth dries up and the average trade size increases. It contributes to the literature by studying a phenomenon that has been widely neglected by both academics and regulators despite its deep repercussions on market quality, its impact on futures and derivatives, and above all its continued effect on brokers trading at the VWAP.

Finally, the third essay, *Leadership of the Price Discovery Process at Outer Levels of the Book*, investigates the impact trading across multiple venues has on the price discovery process. In 2006-2007 regulation was introduced in both the U.S. and Europe, through the Regulation National Market System and MiFID I directives respectively, in an attempt to reduce barriers to entry and increase transparency on trading venues. This, in conjunction with digitalised trading which made remote trading possible, led to the proliferation of alternative trading systems, and with it markets where the same security could be traded simultaneously. Therefore, financial markets nowadays operate in a paradigm where the same stock is simultaneously traded on multiple markets and consequently the same security has multiple prices. Whilst it is true that arbitrage opportunities

ensure that the various prices do not diverge, it is also true that in the very short run, i.e. in the immediate aftermath of the release of a new piece of information, some prices adjust faster to the news and embed said news at greater speed than others, and this inevitably has repercussions on investors. It should be of no surprise that the phenomenon examining the speed at which prices update to new information and the process through which this happens, known as price discovery, has attracted the attention of both academics and regulators. Price discovery is relevant for many applications. To give an example, price discovery can be used as a competitive lever as shown by Ibikunle (2015) who argues that BATS Chi-X Europe was able to successfully enter the European market at the expense of the incumbent, the LSE, thanks to its superior architecture which gave it price leadership over the LSE, thus allowing it to take market share away from the incumbent. From the trader's perspective, price discovery leadership signals which market provides the informationally best quotes and, since agents can simultaneously trade the same asset on more than one venue, price discovery as well as representing a competitive lever also presents traders with an opportunity for inter-market arbitrage or simply with the informationally best quotes. Furthermore, from the regulator's perspective price discovery informs which type of investor and which type of order is responsible for the inclusion of fundamental information in the current price, allowing for the introduction of regulation that enhances market quality. For example, Brogaard, Hendershott, and Riordan (2014) show that on NASDAQ between 2008 and 2009 the market orders of high frequency traders play a positive role in the price discovery process since they are posted in the direction of permanent price changes. So far the existing literature has studied price discovery by focusing on the best bid-offer (BBO) and this paper's contribution to the literature lies in emphasising the importance of the levels of the book other than the BBO by showing that price leadership may change across levels of the limit order book and therefore that one cannot assess price discovery purely on the basis of the BBO which accounts for less than 50% of total liquidity.

Chapter 1

Modelling the Maker-Taker Pricing in Financial Markets

Abstract

In this paper we study the trading fees of a stock exchange, namely the maker-taker fees. In particular, we show that an increase of the make fee (which in our setting is a negative fee) leads to a worsening of inside spread and executed volume, and to an improvement of BBO depth. This is done by first developing a theoretical model of a limit order book with a closed-form solution that proxies a modern electronic trading platform. Second, we draw empirical predictions to study the impact a change in make-take pricing has on trading activity and market quality, and to explain the dynamics behind the changes in market quality that we observe empirically on the Toronto Stock Exchange on June 1, 2016 when the make-take pricing scheme was altered.

1 Introduction

Financial markets were created to provide a bridge between the real economy and (financial) capital. Over the years financial markets have grown in size and have become increasingly sophisticated, taking a life of their own, dominating rather than serving the real economy and heavily influencing the economic cycle. The Global Financial Crisis provides ample evidence of this trend. One phenomenon that recently attracted the attention of regulators is competition among markets. Different markets are run by different institutions and are now actively competing to attract volume. Indeed over the past fifteen years financial markets both in the U.S. and in Europe have been characterised by an increase in the number of trading platforms which compete with traditional exchanges, and one of the most important strategic tools that managers of trading platforms use to attract volume is the set of trading fees that they impose on market participants, also called "pricing". These fees are a major source of revenues for exchanges, as they account for more than 60% of their total income.[1] They are also one of their main competitive tools, as managers use them to attract volume to their platform. The growing competition among trading platforms brought about a sharp reduction in the level of trading fees, but in an attempt to attract more volume and thus generate revenues it also led to the development of different pricing structures, which raised questions about their impact on market quality.

One particular and controversial set of trading fees is the so called "make-take" fee structure, that imposes second degree price discrimination by offering different fees to traders that demand liquidity and traders that supply liquidity.[2] In particular, a trader is said to be supplying liquidity

[1] This estimate is taken from NYSE's 2012 annual report where transaction fees constituted 64% of total revenue (listing fees formed 12%, market data 9%, technology services 9%, and the remaining 6% comes from other sectors) and transaction fees are generally believed to have increased since then. Subsequent reports for NYSE are not available since the NYSE was acquired by the Intercontinental Exchange (ICE) in November 2013, and the ICE consolidates all data from the multiple exchanges owned by the group in its annual report.

[2] Recall that Price Discrimination is a microeconomic pricing strategy where identical or largely similar goods or services are sold at different prices by the same provider in different markets. In particular, Second Degree Price Discrimination occurs when the price varies with the quantity demanded as the seller is not able to differentiate between different types of consumers, and Third Degree Price Discrimination occurs when different prices are charged to different consumer groups. Most of the literature classifies the make-take fee structure as an instance of Second Degree Price Discrimination since trading centres cannot tell liquidity makers and liquidity takers apart, and adopt different fees (namely the make fee and the take fee) to price discriminate between the two. However, others argue make-take pricing is in fact an instance of Third Degree Price Discrimination given there is a clear distinction between the liquidity makers and the liquidity takers. In this manuscript we sympathise with the latter school of thought but adopt the former terminology since it is more widely endorsed by fellow researchers.

if she submits an order (whether buy or sell) that rests on the book, as it cannot be immediately executed (usually because there is no counterpart willing to trade at the indicated, or limit, price); whereas a trader is said to be demanding liquidity if her order hits an existing limit order, thus taking/consuming liquidity on the book. Make-take fees generally include negative fees (make fees), which are effectively rebates and are offered to market participants who supply liquidity and de facto get paid for doing this; and positive fees (take fees) for market participants who demand liquidity and are effectively charged for doing this. There are however some stock exchanges, such as Alpha Canada, Omega Canada, or Nasdaq BX, that use inverted make-take fees, whereby make fees are positive fees and take fees are negative fees or rebates. We notice that whilst make fees provide a concrete incentive for market participants to submit more limit orders, hence increasing liquidity supply, take fees are a cost borne by liquidity demanders. Hence, the effects of maker-taker pricing on traders' welfare and market quality are widely debated and their assessment is still a relatively unexplored issue in market design. As a matter of fact, the make-take pricing structure has been recently criticised both by US and by European regulators.

The main criticisms raised about the maker-taker pricing structure, as summarised by Angel, Harris, and Spatt (2011) relate to: i) the fact that high rebates attract high frequency traders (HFTs), who gain a high proportion of rebates to the detriment of other investors who may therefore have an incentive to move away from "lit" and into "dark" markets. It is still being debated what the effect of dark markets is on lit ones, but for sure dark markets are less transparent; ii) as brokers do not pass the rebates they receive on to their clients, they may have an incentive to maximise rebates rather than search for the best possible execution price for their clients, thus exacerbating the inherent agency problem; iii) asymmetric fees may have a negative impact on price discovery as quoted prices do not include make-take fees, thus making cross market comparison of best execution difficult. The maker-taker pricing structure has therefore raised concerns about its effects both on the welfare of market participants and on the quality of the trading platforms.

Whilst this paper makes both a theoretical and an empirical contribution to the literature, its primary objective is to develop a theoretical model with a closed-form solution that proxies a modern electronic trading platform and to then use the model to draw empirical predictions on

make and take fees. This allows us to study the impact a change in fees has on trading activity and market quality. The model consists of a standard trading platform that resembles a limit order book (LOB) and a competing alternative trading system (ATS), executing at the mid-quote of the LOB, that by competing with the LOB for liquidity embeds the effects of competition in our model. By parametrising the model to an increase in the make fee on the LOB only (that is, leaving the take fee on the LOB, as well as the fees on the ATS, unchanged), results show that this leads to traders (both liquidity makers and liquidity takers) moving away from the LOB and into the now more profitable ATS. The reason being that increasing the make fee when this is a negative fee corresponds to a reduction in the rebate (defined as the absolute value of the make fee) offered to liquidity makers, and this reduces the number of limit orders posted to the LOB and with it the market depth for liquidity takers who thus follow the liquidity makers into the ATS. This result differs from a model without a competing trading venue, as in the latter, the liquidity makers's only outside option is to consume liquidity instead, thus leading to an increase in market orders submitted to the LOB. Finally, the outcome on market quality is that inside spread widens and BBO depth deepens, whilst executed volume falls. In a similar vein to what just described, parametrising the model to a reduction in the LOB take fee only, leads to opposite dynamics to the ones just described.

We test the model's predictions by means of a (high frequency) dataset centred around June 1, 2016 - it was on this day that the Toronto Stock Exchange (TSX) changed its pricing scheme by increasing the make fee and reducing the take fee on all inter-listed stocks (Inter-listed stocks are those stocks that are cross-listed on both the TSX and an American Exchange, and Non-inter-listed stocks are those stocks which are not Inter-listed. This is further discussed in Appendix 7.A). The key advantage of this dataset is that the fees were left unchanged on non-inter-listed stocks traded on the same platform and we can use the latter stocks as a control for other exogenous changes in market conditions (Section 5.4 gives a robustness assessment of the validity of the control group). For the sample period we take two months before and two months after the event, for a total of four months ranging from April 1, 2016 to July 31, 2016. The fee changes undertaken by TSX on June 1, 2016 are reported in Table 1.1, where one can also find the changes implemented by the

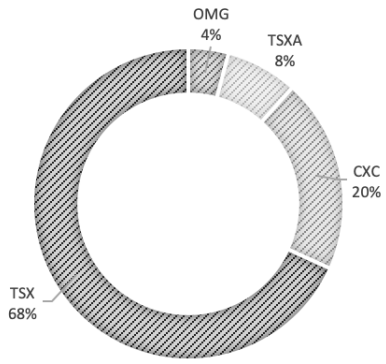
other major Canadian trading venues on the same day. A breakdown of the daily executed volume by stock category and time of trading is reported in Appendix 7.E Table 1.7, which is summarised by Figure 1.1 where we observe that the TSX is the largest Canadian platform followed by Chi-X Canada (CXC) which we use as a proxy for the ATS in our model.

Table 1.1: Make-Take Fee Changes applied to Canadian Equities on June 1, 2016

	Security	Prior to June 1, 2016		Effective June 1, 2016	
		Take Fee	Make Fee	Take Fee	Make Fee
TSX	**Inter-listed**	\$ 0.0030	(\$ 0.0026)	\$ 0.0027	(\$ 0.0023)
	Non-inter-listed	\$ 0.0023	(\$ 0.0019)	\$ 0.0023	(\$ 0.0019)
CXC	**Inter-listed**	\$ 0.0029	(\$ 0.0025)	\$ 0.0026	(\$ 0.0022)
	Non-inter-listed	\$ 0.0029	(\$ 0.0025)	\$ 0.0029	(\$ 0.0025)
TSXA	**All**	\$ 0.0018	(\$ 0.0014)	(\$ 0.0010)	\$ 0.0014
OMG	**0 $\leq$ EOV $<$ \$ 100MM**	(\$ 0.0014)	\$ 0.0019	(\$ 0.0016)	\$ 0.0020
	\$ 100MM $\leq$ EOV $<$ \$ 200MM	(\$ 0.0014)	\$ 0.0018	(\$ 0.0016)	\$ 0.0019
	\$ 200MM $\leq$ EOV	(\$ 0.0014)	\$ 0.0017	(\$ 0.0016)	\$ 0.0018

A detailed explanation of this table is given in Appendix 7.A.

Figure 1.1: Canadian Trading Centres



This figure exhibits the market share of the major Canadian trading centres as of January 2018: Toronto Stock Exchange, Chi-X Canada, TSX Alpha, and Omega ATS.

Empirically we find that over our sample period on TSX inside spread and BBO depth fall (albeit the latter not significantly) and that the average trade size increases. When we compare these results to the empirical predictions drawn from our model calibrated consistently with TSX before and after the fee change, we conclude that the empirics match the theory. This corroborates the model and allows us to use the model to explain the dynamics behind the changes in market

quality that we observe empirically, namely an injection of liquidity into the TSX accompanied with an improvement in inside spread and average trade size, and a worsening of BBO depth.

The paper is organised as follows. Sections 1 and 2 include an introduction and a literature review. Section 3 presents the model and Section 4 its empirical predictions. Section 5 discusses the empirical results. Section 6 presents the conclusions, and Section 7 gives the Appendix to the paper.

2 Literature Review

There is a very limited amount of existing theoretical and empirical literature on the topic of make-take fees. Colliard and Foucault (2012) build a model where agents can be buyers or sellers with equal probability and a discount factor characterises them either as patient or as impatient. Traders choose between trading on a limit order book or on a dealership market. Limit orders expire after one period and if they are not executed on the limit order book, they get executed in the dealership market: the immediate consequence of this assumption is that depth cannot build up on the book and this is crucial to obtaining a stationary equilibrium. However, this feature of the model makes it impossible to generate predictions of the effects of a change in fees on market depth. Within this framework Colliard and Foucault (2012) investigate how investors' order submission strategies and hence market quality and welfare are affected by changes in trading fees. More specifically, they discuss from a theoretical point of view the conclusion reached by Angel et al. (2011) in their work, i.e. that the make-take fee breakdown does not matter as long as the total fee (the sum of make and take fees) is held constant. If for example the rebate (recall this is the absolute value of the make fee) is reduced, then as limit orders become more expensive some traders may decide to submit market orders instead. The increase in market orders increases the execution probability of limit orders, which in turn will prompt traders to post less aggressive limit orders and thus benefit from a price improvement. In this way, though, the reduction of the rebate is offset by the wider spread which is equivalent to an increase in the take fee, with the consequence that the cost of supplying and demanding liquidity remain unchanged. There are three other main comments to this line of argument. The first one is that in real markets there is a

minimum price improvement (tick size) set by exchanges; hence, bid and ask prices may not freely adjust to compensate changes in trading fees. The second one is that there is also no competition among liquidity providing orders since this is a two period model and once an order is submitted in the first period it cannot be undercut in the next period as any order submitted in the second period has zero execution probability (as the second period is the last period of the model) and therefore will never be submitted. The third one is that in real markets a take fee is a certain cost for traders, whereas a make fee will materialise only as long as the limit order is executed. Hence, the breakdown between make and take fees should not be neutral.

The model of Foucault, Kadan, and Kandel (2013) improves on Colliard and Foucault (2012) by including a tick size and shows that in this case the make-take fee breakdown actually matters. In their model, agents monitor the state of the LOB and the level of (make-take) fees affects the intensity with which they do so, and this in turn affects the trading rate. Given a certain level of total fees, there is an optimal breakdown at which the trading rate is maximised, and at which trading platforms consequently maximise their revenue. To maximise the trading rate, platforms use fees to incentivise either makers or takers, depending on who is monitoring the market with lower intensity. As a consequence in this model the make-take fee breakdown matters. Indeed this setup is more realistic than in Colliard and Foucault (2012) even though the underlying assumptions are pretty strong: the book contains at most one limit order; limit orders can only be price improving and buyers and sellers alternate with certainty. It also has a limit, the same as in Colliard and Foucault (2012), in that the equilibrium is stationary and as a consequence predictions can be made for spread but not for depth.

Our model accounts for all of these issues. It does not rely on a stationary equilibrium, being solved by backward induction, and therefore allows to model the dynamic change in the state of the book and therefore draw predictions on market depth. It also includes a tick-size, as well as competition from an alternative trading system that executes at the mid-quote of the primary market bid-ask spread. Our model therefore embeds the effect of competition form a stylised market which works like most of the many alternative trading systems (ATSs) which exist today both in the US and Europe.

3 Model

We model a market for a single security which is traded over three periods, T-2, T-1 and T, and takes value V at period T. The model is sequential in that in each period one agent arrives at the market and trades a single unit of the security. We assume that traders are risk neutral and are characterised by a parameter γ so that their personal valuation of the asset is γV. Furthermore, we let $\gamma \sim U[0,2]$ so that γ takes value 1 in expectation which implies that on average the trader's personal valuation of the asset (γV) equals the fundamental value V at period T. We note that the probability of a trader being a buyer or a seller will depend on the distribution of γ (the greater the value of γ, the more the trader values the asset, and the greater the incentive to hold the asset and hence be a buyer).

Upon arrival at the market the agent can choose between two trading venues, a Limit Order Book (LOB) and an Alternative Trading System (ATS). Our analysis is therefore more realistic than if we only considered a LOB since the ATS, by competing against the LOB, introduces the effects of competition which are ubiquitous in financial markets. This model therefore abandons the classical economic paradigm where there is a single centralised market where the security can be traded in favour of the more realistic paradigm where the same security can be traded on multiple markets simultaneously. In particular, the ATS is characterised by a derivative pricing rule. This means that executions on this trading platform can only take place at the mid-quote of the best bid and the best offer of the LOB. The LOB instead is characterised by a price grid composed of four discrete prices, two prices on the bid side, $\beta_2 < \beta_1$, and two prices on the ask side of the market, $\alpha_1 < \alpha_2$. Discrete prices replicate real world trading platforms which are governed by a minimum price change ω, the tick size. Thus, since the fundamental value of the asset V is taken to be the mid-quote of the best bid and the best ask price, the possible price levels can be written as: $\beta_2 = V - \frac{3}{2}\omega$, $\beta_1 = V - \frac{1}{2}\omega$, $\alpha_1 = V + \frac{1}{2}\omega$ and $\alpha_2 = V + \frac{3}{2}\omega$.[3] Moreover, under normal market conditions there is always someone willing to sell if the price is high enough and willing to buy if the price is low enough, so we assume that there exists a trading crowd available to post a limit

[3] We write each of the four prices as a function of V and ω by solving the simultaneous equations: $(\alpha_1 + \beta_1)/2 = V$ by definition of V and $\omega = \alpha_2 - \alpha_1 = \alpha_1 - \beta_1 = \beta_1 - \beta_2$ by definition of the tick size ω.

sell order at the highest level of the book (the highest ask price α_2) and a limit buy order at the lowest level of the book (the lowest bid price β_2) whenever either of these prices turns empty. The trading crowd is assumed to yield time priority to individual traders and is consistent with Colliard and Foucault (2012) who assume the existence of a dealership market.

Traders can submit two different types of orders and they can do so on either the LOB or the ATS: a market order (M) executes upon submission at the best available price on the opposite side of the market, and a limit order (L) rests on the LOB or the ATS (depending on where it is posted) at the limit price until it is hit by a market order posted by a trader with opposite trading needs. Traders also have the option not to trade (N). Orders are ranked on the LOB according to price and time priority rules, whereas they are ranked on the ATS only according to time priority as on the ATS there is no price grid. To summarise, in each period T-k of the model the trader has five order submission strategies that she can execute on each available price: she could submit a market order to the LOB or ATS ($M^{\text{LOB}}_{\text{T-k}}$ or $M^{\text{ATS}}_{\text{T-k}}$), a limit order to the LOB or ATS ($L^{\text{LOB}}_{\text{T-k}}$ or $L^{\text{ATS}}_{\text{T-k}}$), or she could decide not to trade ($N_{\text{T-k}}$). Consistently with Smart Order Routing Algorithmic Programs we assume that when traders post a market order they hit the best price available on the market. Finally, we assume a make-take pricing structure on both the LOB and the ATS such that traders posting limit orders to the LOB (ATS) gain the make fee MF_{LOB} (MF_{ATS}) on each order executed whereas traders posting market orders to the LOB (ATS) are charged the take fee TF_{LOB} (TF_{ATS}). Notice that for the remainder of this Paper the make fee is a negative fee and therefore a rebate, while the take fee is a positive fee and therefore a straightforward fee.

3.1 Order Submission Strategies

By describing the interaction of agents with the LOB and ATS we compute the set of feasible order submission strategies $\mathbb{S}(\Omega_{\text{T-k}})$ for each of the three periods of the trading game. A strategy is said to be "feasible" if it could be submitted given the state of the limit order book, and is said to be "dominated" given a personal valuation of the asset γ if there exists another feasible strategy that yields a greater payoff.

3.1.1 Period T-2

We assume that both the LOB and the ATS open empty at the start of the trading game and this is consistent with the book being empty at the beginning of each trading day on financial markets. Consequently the state of the book (Ω) at the beginning of period T-2 is equal to $\Omega_{\text{T-2}} = [1001]$, that is there is a limit buy order at β_2 $\left(L^{\beta_2}_{\text{T-2}}\right)$ and limit sell order at α_2 $\left(L^{\alpha_2}_{\text{T-2}}\right)$ in view of the trading crowd which replenishes outside prices.[4] The order submission strategies with their respective expected payoffs available to the trader arriving at the market at T-2 and wishing to position herself on the sell side are as follows:

Case 1 - $\Omega_{\text{T-1}}$=[1001]: market sell order at β_2 $\left(M^{\beta_2}_{\text{T-2}}\right)$ with expected payoff $\beta_2 - \gamma V - \text{TF}_{\text{LOB}}$.[5] We assume the trading crowd immediately replenishes the book by posting a limit buy order at β_2 so the book opens at T-1 as $\Omega_{\text{T-1}} = [1001]$.

Case 2 - $\Omega_{\text{T-1}}$=[1001-S]: limit sell order on the ATS $\left(L^{S}_{\text{T-2}}\right)$ with expected payoff $\left(\frac{\alpha_2+\beta_2}{2} - \gamma V - \text{MF}_{\text{ATS}}\right) \times \Pr\left(M^{B}_{\text{T-1,T}}|\Omega_{\text{T-1}}=[1001\text{-S}]\right)$, where $\Pr\left(M^{B}_{\text{T-1,T}}|\Omega_{\text{T-1}}=[1001\text{-S}]\right)$ is the probability that the limit sell order posted at period T-2 on the ATS is executed by a market buy order on the ATS either at T-1 or at T. In this case the state of the book opens at T-1 as $\Omega_{\text{T-1}}$=[1001-S].

Case 3 - $\Omega_{\text{T-1}}$=[1011]: limit sell order at α_1 $\left(L^{\alpha_1}_{\text{T-2}}\right)$ with expected payoff $(\alpha_1 - \gamma V - \text{MF}_{\text{LOB}}) \times \Pr\left(M^{\alpha_1}_{\text{T-1,T}}|\Omega_{\text{T-1}}=[1011]\right)$, where $\Pr\left(M^{\alpha_1}_{\text{T-1,T}}|\Omega_{\text{T-1}}=[1011]\right)$ is the probability that the limit sell order posted at period T-2 on price α_1 of the LOB is executed by a market buy order either at T-1 or at T. In this case the state of the book opens at T-1 as $\Omega_{\text{T-1}}$=[1011].

Case 4 - $\Omega_{\text{T-1}}$=[1002]: limit sell order at α_2 $\left(L^{\alpha_2}_{\text{T-2}}\right)$ with expected payoff $(\alpha_2 - \gamma V - \text{MF}_{\text{LOB}}) \times \Pr\left(M^{\alpha_2}_{\text{T-1,T}}|\Omega_{\text{T-1}}=[1002]\right)$, where $\Pr\left(M^{\alpha_2}_{\text{T-1,T}}|\Omega_{\text{T-1}}=[1002]\right)$ is the prob-

[4] More generally if the state of the book is $\Omega_{\text{T-k}}$=[$n_1 n_2 n_3 n_4$-N] then the LOB opens at period T-k with $n_1 \in \mathbb{R}$ units on β_2, $n_2 \in \mathbb{R}$ units on β_1, $n_3 \in \mathbb{R}$ units on α_1, $n_4 \in \mathbb{R}$ units on α_2, one unit $N \in \{S, B\}$ (S if it is a sell unit i.e. a limit order to sell and B if it is a buy unit i.e. a limit order to buy) on the ATS, with $\beta_2 < \beta_1 < \alpha_1 < \alpha_2$ the grid of prices on the LOB.

[5] The expected payoff is $\beta_2 - \gamma V - \text{TF}_{\text{LOB}}$ as the trader is paid price β_2 in exchange for the asset which to the trader has a personal valuation of γV and since the trader submitted a market order she needs to pay the take fee TF_{LOB} to the LOB.

ability that the limit sell order posted at period T-2 on price α_2 of the LOB is executed by a market buy order either at T-1 or at T. In this case the state of the book opens at T-1 as $\Omega_{\text{T-1}}$=[1002]. Recall that the trading crowd yields time priority to traders thus an investor positioning herself at α_2 jumps ahead of the limit sell order posted by the trading crowd.

Case 5 - $\Omega_{\text{T-1}}$=[1001]: no trade ($N_{\text{T-2}}$) with expected payoff of zero. The book opens at T-1 as $\Omega_{\text{T-1}}$=[1001].

In summary, the available order submission strategies to the trader at T-2 on the sell side are in order of aggressiveness: $M^{\beta_2}_{\text{T-2}}$, $L^{S}_{\text{T-2}}$, $L^{\alpha_1}_{\text{T-2}}$, $L^{\alpha_2}_{\text{T-2}}$ and $N_{\text{T-2}}$. The order submission strategies on the buy side of the book can be derived by symmetry. We conclude that the complete set of feasible order submission strategies at T-2 (given the book opens as $\Omega_{\text{T-2}}$=[1001]) is
$\mathbb{S}(\Omega_{\text{T-2}}) = \left\{M^{\beta_2}_{\text{T-1}}, L^{S}_{\text{T-1}}, L^{\alpha_1}_{\text{T-1}}, L^{\alpha_2}_{\text{T-1}}, N_{\text{T-1}}, L^{\beta_2}_{\text{T-1}}, L^{\beta_1}_{\text{T-1}}, L^{B}_{\text{T-1}}, M^{\alpha_2}_{\text{T-1}}\right\}$.

3.1.2 Period T-1

The order submission strategies available to the trader entering the market at period T-1 depend on the state of the book which she observes which in turn depends on the strategy undertaken by the previous trader at T-2. Assume the trader at T-2 submitted a sell order and therefore can be reconciled to one of the five cases (Case 1, 2, 3, 4 or 5) listed above. Then the set of feasible strategies to the trader at T-1 can be computed for each case:

Case 1 - $\Omega_{\text{T-1}}$=[1001]: given that at period T-1 the book opens as $\Omega_{\text{T-1}} = [1001]$, the strategies available to the incoming trader are the same (since the book is the same) as those available to the trader at T-2: $\left\{M^{\beta_2}_{\text{T-1}}, L^{S}_{\text{T-1}}, L^{\alpha_1}_{\text{T-1}}, L^{\alpha_2}_{\text{T-1}}\right\}$ on the sell side, $\left\{M^{\alpha_2}_{\text{T-1}}, L^{B}_{\text{T-1}}, L^{\beta_1}_{\text{T-1}}, L^{\beta_2}_{\text{T-1}}\right\}$ on the buy side, and no trade $N_{\text{T-1}}$. The opening state of the book at T depends on the choice of strategy at T-1: on the sell side if $M^{\beta_2}_{\text{T-1}}$ then Ω_{T}=[1001], if $L^{S}_{\text{T-1}}$ then Ω_{T}=[1001-S], if $L^{\alpha_1}_{\text{T-1}}$ then Ω_{T}=[1011], if $L^{\alpha_2}_{\text{T-1}}$ then Ω_{T}=[1002], and if $N_{\text{T-1}}$ then Ω_{T}=[1001]; the buy side follows by symmetry.

Case 2 - $\Omega_{\text{T-1}}$=[1001-S]: under this state of the book the strategies available to the incoming trader are: $\mathbb{S}(\Omega_{\text{T-1}}) = \left\{M^{\beta_2}_{\text{T-1}}, L^{\alpha_1}_{\text{T-1}}, L^{\alpha_2}_{\text{T-1}}, N_{\text{T-1}}, L^{\beta_2}_{\text{T-1}}, L^{\beta_1}_{\text{T-1}}, M^{B}_{\text{T-1}}\right\}$. Notice that in this scenario

the two strategies $L^{S}_{\text{T-1}}$ and $M^{\alpha_2}_{\text{T-1}}$ are not in the set of feasible strategies. The former is a limit sell order posted to the ATS and has zero execution payoff since it has to queue behind the existing sell order on the ATS and therefore will never be executed at the last period T of the model where (as for all periods) a single unit of the asset can be traded. The latter is a market buy order to α_2 and this is simply dominated by the cheaper unit that can be bought on the ATS for $(\alpha_2 + \beta_2)/2$. The opening state of the book at T depends on the choice of strategy at T-1: if $M^{\beta_2}_{\text{T-1}}$ then Ω_{T}=[1001-S], if $L^{\alpha_1}_{\text{T-1}}$ then Ω_{T}=[1011-S], if $L^{\alpha_2}_{\text{T-1}}$ then Ω_{T}=[1002-S], if $N_{\text{T-1}}$ then Ω_{T}=[1001-S], if $L^{\beta_2}_{\text{T-1}}$ then Ω_{T}=[2001-S], if $L^{\beta_1}_{\text{T-1}}$ then Ω_{T}=[1101-S], if $M^{B}_{\text{T-1}}$ then Ω_{T}=[1001].

Case 3 - $\Omega_{\text{T-1}}$=[1011]: under this state of the book the strategies available to the incoming trader are: $\mathbb{S}(\Omega_{\text{T-1}}) = \left\{M^{\beta_2}_{\text{T-1}}, L^{S}_{\text{T-1}}, L^{\alpha_2}_{\text{T-1}}, N_{\text{T-1}}, L^{\beta_2}_{\text{T-1}}, L^{\beta_1}_{\text{T-1}}, L^{B}_{\text{T-1}}, M^{\alpha_1}_{\text{T-1}}\right\}$. Notice that in this scenario the limit orders posted to the ATS are executed at price β_1 since this is the mid-quote of the LOB, and on the buy side a limit order to the ATS $\left(L^{B}_{\text{T-1}}\right)$ executes at the same price as a limit order posted to β_1 $\left(L^{\beta_1}_{\text{T-1}}\right)$ but in view of possibly differing fees on the two markets the expected payoffs of these two strategies may differ. Additionally, $L^{\alpha_1}_{\text{T-1}}$ is not in the set of feasible strategies since it would have to queue behind the existing order on α_1 and therefore has zero expected payoff since it would never be executed at the last period T of the trading game where (as for all periods) a single unit can be traded. The opening state of the book at T depends on the choice of strategy at T-1: if $M^{\beta_2}_{\text{T-1}}$ then Ω_{T}=[1011], if $L^{S}_{\text{T-1}}$ then Ω_{T}=[1011-S], if $L^{\alpha_2}_{\text{T-1}}$ then Ω_{T}=[1012], if $N_{\text{T-1}}$ then Ω_{T}=[1011], if $L^{\beta_2}_{\text{T-1}}$ then Ω_{T}=[2011], if $L^{\beta_1}_{\text{T-1}}$ then Ω_{T}=[1111], if $L^{B}_{\text{T-1}}$ then Ω_{T}=[1011-B], if $M^{\alpha_1}_{\text{T-1}}$ then Ω_{T}=[1001].

Case 4 - $\Omega_{\text{T-1}}$=[1002]: under this state of the book the strategies available to the incoming trader are: $\mathbb{S}(\Omega_{\text{T-1}}) = \left\{M^{\beta_2}_{\text{T-1}}, L^{S}_{\text{T-1}}, L^{\alpha_1}_{\text{T-1}}, N_{\text{T-1}}, L^{\beta_2}_{\text{T-1}}, L^{\beta_1}_{\text{T-1}}, L^{B}_{\text{T-1}}, M^{\alpha_2}_{\text{T-1}}\right\}$. Notice that in this scenario $L^{\alpha_2}_{\text{T-1}}$ has a zero execution probability and is therefore a dominated strategy. The opening state of the book at T depends on the choice of strategy at T-1: if $M^{\beta_2}_{\text{T-1}}$ then Ω_{T}=[1002], if $L^{S}_{\text{T-1}}$ then Ω_{T}=[1002-S], if $L^{\alpha_1}_{\text{T-1}}$ then Ω_{T}=[1012], if $N_{\text{T-1}}$ then Ω_{T}=[1002], if $L^{\beta_2}_{\text{T-1}}$ then Ω_{T}=[2002], if $L^{\beta_1}_{\text{T-1}}$ then Ω_{T}=[1102], if $L^{B}_{\text{T-1}}$ then Ω_{T}=[1002-B], if $M^{\alpha_2}_{\text{T-1}}$ then Ω_{T}=[1001].

Case 5 - Ω_{T-1}=[1001]: under this state of the book, the available strategies are the same as those listed in Case 1 above since the opening book is the same.

These five cases (Case 1-5) list the set of feasible order submission strategies to a trader at T-1 given that the previous trader at T-2 submitted a sell order. In the event a buy order is submitted at T-2, the order submission strategies can be derived by symmetry.

3.1.3 Period T

Period T is the last period of the trading game, and as for periods T-2 and T-1, the set of feasible strategies $\mathbb{S}$ depends on the opening state of the book Ω which in turn depends on the orders submitted at T-1 and T-2. The feasible order submission strategies at T ($\mathbb{S}(\Omega_T)$) are not given for brevity but follow the same methodology that is adopted at the previous periods of the model.

3.2 The Solution to the Model

The model is solved by backward induction starting from period T, the last period of the model. In this period traders do not post limit orders as they would never be executed in a subsequent period (which by definition does not exist as T is the final period of the trading game) and therefore would have a zero execution probability and zero expected payoff, making limit orders in this period dominated by the strategy N_T of no trade. Consequently, in period T there is no uncertainty with regards to the expected payoffs of the feasible order submission strategies and this means we can solve the model.[6]

The model has a two-step solution, which is to be computed for each period of the model, starting from the last. First, we compute the equilibrium strategies as a function of γ, i.e. those strategies that have the highest expected payoff in the set of all feasible strategies. It is clear that such strategies are optimal and thus will be submitted by the trader in equilibrium. Second, we derive the equilibrium submission probabilities, i.e. the probabilities of submitting each equilibrium strategy.

[6]The expected payoff of a market order at T-k is completely known at T-k whereas the expected payoff of a limit order at T-k is not a priori known at T-k since it is a function of the limit order's execution probability which is the submission probability of a market order in one of the following periods T-k+1 or T-k+2.

Once the two-step solution is computed for period T, we can derive the execution probabilities for limit orders submitted at T-1 through the equilibrium submission probabilities of market orders posted at T-2 on the same price level thus eliminating the limit order-induced uncertainty at T-1 which allows to compute the equilibrium strategies at T-1 and the equilibrium probabilities thereof. Finally, we solve for T-2. Notice that the two-step solution to the model has a closed form solution (given the tick size of the LOB and the complete distribution of the traders' evaluation of the asset) and therefore is not attained through repeated numerical simulations of $\gamma \sim U[0,2]$.

Due to the recursive structure of the game and that we disregard orders with a zero execution probability (since these are dominated by N_t), a unique equilibrium always exists and is formally defined as follows: *Let* $\mathbb{S}(\Omega_{T\text{-}k})$ *be the set of feasible order submission strategies under the book* $\Omega_{T\text{-}k}$ *at period T-k. Then the equilibrium is the set of traders' strategies* $\{s^*\}$ *such that at each period the trader maximises his expected payoff* $\mathbb{E}[P(\cdot|\Omega_{T\text{-}k})]$ *over his Bayesian belief of the equilibrium execution probabilities* $\Pr^*(\cdot|\Omega_{T\text{-}k+1})$*. In other words,* $s^* = s^*(\gamma|\Omega_{T\text{-}k}) \in \mathbb{S}(\Omega_{T\text{-}k})$ *is the optimal strategy to the trader* γ *given the book opens as* $\Omega_{T\text{-}k}$ *and out of the set of admissible strategies* $\mathbb{S}(\Omega_{T\text{-}k})$ *if and only if* $\mathbb{E}[P(s^*|\Omega_{T\text{-}k})] \geq \mathbb{E}[P(s|\Omega_{T\text{-}k})]\ \forall\ s \in \mathbb{S}(\Omega_{T\text{-}k})$ *which takes place if and only if* $\gamma \in \Gamma(s^*)$ *where* $\Gamma(s^*) = \Big\{\gamma \in [0,2]$ *such that* $\mathbb{E}[P(s^*|\Omega_{T\text{-}k})] \geq \mathbb{E}[P(s|\Omega_{T\text{-}k})]\ \forall\ s \in \mathbb{S}\Big\}$*. Furthermore, the equilibrium submission probability of the equilibrium strategy* s^* *is given by* $\Pr^*(s^*|\Omega_{T\text{-}k}) = \Pr(\gamma \in \Gamma(s^*)) = \int\limits_{\gamma \in \Gamma(s^*)} \frac{1}{2}\, d\gamma$.

3.2.1 Period T

The two-step backward solution of the model starts by first solving the last period T since Ω_{T} does not allow for any limit orders which would have a zero execution payoff. The state of the book is a function of the strategies submitted at periods T-2 and T-1. Consider Ω_{T}=[1001] and the other states of the book are solved in a similar manner. Under this book, the set of feasible order submission strategies is $\mathbb{S} = \left\{M_{\mathrm{T}}^{\beta_2}, N_{\mathrm{T}}, M_{\mathrm{T}}^{\alpha_2}\right\}$ and their respective expected payoffs are $\mathbb{E}[P(M_{\mathrm{T}}^{\beta_2}|\Omega_{\mathrm{T}}\text{=}[1001])] = \beta_2 - \gamma V - \mathrm{TF}_{\mathrm{LOB}}$, $\mathbb{E}[P(N_{\mathrm{T}}|\Omega_{\mathrm{T}}\text{=}[1001])] = 0$ and $\mathbb{E}[P(M_{\mathrm{T}}^{\alpha_2}|\Omega_{\mathrm{T}}\text{=}[1001])] = \gamma V - \alpha_2 - \mathrm{TF}_{\mathrm{LOB}}$. Thus the equilibrium strategies and the equilibrium execution probabilities are

given by:[7]

$$s^*(\gamma|\Omega_{\mathrm{T}}{=}[1001]) = \begin{cases} M_{\mathrm{T}}^{\beta_2} & \text{if } \gamma \in \left[0, \frac{\beta_2 - \mathrm{TF_{LOB}}}{V}\right) \\ N_{\mathrm{T}} & \text{if } \gamma \in \left[\frac{\beta_2 - \mathrm{TF_{LOB}}}{V}, \frac{\alpha_2 + \mathrm{TF_{LOB}}}{V}\right) \\ M_{\mathrm{T}}^{\alpha_2} & \text{if } \gamma \in \left[\frac{\alpha_2 + \mathrm{TF_{LOB}}}{V}, 2\right] \end{cases} \tag{1}$$

$$\Pr^*(s^*|\Omega_{\mathrm{T}}{=}[1001]) = \begin{cases} \frac{\beta_2 - \mathrm{TF_{LOB}}}{2V} & \text{if } s^* = M_{\mathrm{T}}^{\beta_2} \\ \frac{\alpha_2 - \beta_2 + 2\mathrm{TF_{LOB}}}{2V} & \text{if } s^* = N_{\mathrm{T}} \\ 1 - \frac{\alpha_2 + \mathrm{TF_{LOB}}}{2V} & \text{if } s^* = M_{\mathrm{T}}^{\alpha_2} \end{cases} \tag{2}$$

Having solved the last period of the model, whereby the first step of the two-step solution is given by the equilibrium strategies in (1) and the second step is given by the equilibrium submission probabilities in (2), we proceed to period T-1. Additionally, we note that given the asset value V, the grid prices (β_2, α_2) and the trading fee $\mathrm{TF_{LOB}}$, both the equilibrium strategies and probabilities in (1)-(2) are completely known so that the model does indeed have a closed-form solution.

3.2.2 Period T-1

The two-step solution at period T-1 of the model is computed using the probabilities from period T, in particular the equilibrium execution probabilities of limit orders posted at T-1 are given by the equilibrium submission probabilities of market orders posted at T. The state of the book at T-1 is a function of the order submission strategy posted at T-2, and therefore is one of the seven (since three of the nine cases are identical) cases discussed in Section 3.1.2: $\Omega_{\text{T-1}} = \Big\{$[1001], [1001-S], [1011], [1002], [1001-B], [1101], [2001]$\Big\}$. Consider $\Omega_{\text{T-1}}$=[1001] and the other six possible states of the book at T-1 are solved similarly. The set of feasible order submission strategies under this state of the book is $\mathbb{S}(\Omega_{\text{T-1}}{=}[1001]) = \Big\{M_{\text{T-1}}^{\beta_2}, L_{\text{T-1}}^{S}, L_{\text{T-1}}^{\alpha_1}, L_{\text{T-1}}^{\alpha_2}, N_{\text{T-1}}, L_{\text{T-1}}^{\beta_2}, L_{\text{T-1}}^{\beta_1}, L_{\text{T-1}}^{B}, M_{\text{T-1}}^{\alpha_2}\Big\}$

[7] A seller chooses to sell at β_2 iff $\beta_2 - \gamma V - \mathrm{TF_{LOB}} \geq 0$ iff $\gamma \leq (\beta_2 - \mathrm{TF_{LOB}})/V$ and similarly for a buyer at α_2. $\Pr^*(M_{\mathrm{T}}^{\beta_2}|\Omega_{\mathrm{T}} = [1001]) = \Pr(\gamma \in \Gamma(M_{\mathrm{T}}^{\beta_2})|\Omega_{\mathrm{T}} = [1001]) = \Pr(\gamma \leq (\beta_2 - \mathrm{TF_{LOB}})/V) = (\beta_2 - \mathrm{TF_{LOB}})/2V$ since $\gamma \sim U[0, 2]$ and similarly for a buyer submitting $M_{\mathrm{T}}^{\alpha_2}$ in equilibrium.

and the respective expected payoff of each strategy is: [8]

$$\mathbb{E}[P(s^*|\Omega_{\text{T-1}}=[1001])] = \begin{cases} \beta_2 - \gamma V - \text{TF}_{\text{LOB}} & \text{if } s^* = M^{\beta_2}_{\text{T-1}} \\ \left(\frac{\alpha_2+\beta_2}{2} - \gamma V - \text{MF}_{\text{ATS}}\right) \times \left(1 - \frac{\frac{\alpha_2+\beta_2}{2}+\text{TF}_{\text{ATS}}}{2V}\right) & \text{if } s^* = L^{S}_{\text{T-1}} \\ \left(\alpha_1 - \gamma V - \text{MF}_{\text{LOB}}\right) \times \left(1 - \frac{\alpha_1+\text{TF}_{\text{LOB}}}{2V}\right) & \text{if } s^* = L^{\alpha_1}_{\text{T-1}} \\ \left(\alpha_2 - \gamma V - \text{MF}_{\text{LOB}}\right) \times \left(1 - \frac{\alpha_2+\text{TF}_{\text{LOB}}}{2V}\right) & \text{if } s^* = L^{\alpha_2}_{\text{T-1}} \\ 0 & \text{if } s^* = N_{\text{T-1}} \\ \left(\gamma V - \beta_2 - \text{MF}_{\text{LOB}}\right) \times \left(\frac{\beta_2-\text{TF}_{\text{LOB}}}{2V}\right) & \text{if } s^* = L^{\beta_2}_{\text{T-1}} \\ \left(\gamma V - \beta_1 - \text{MF}_{\text{LOB}}\right) \times \left(\frac{\beta_1-\text{TF}_{\text{LOB}}}{2V}\right) & \text{if } s^* = L^{\beta_1}_{\text{T-1}} \\ \left(\gamma V - \frac{\alpha_2+\beta_2}{2} - \text{MF}_{\text{ATS}}\right) \times \left(\frac{\frac{\alpha_2+\beta_2}{2}-\text{TF}_{\text{LOB}}}{2V}\right) & \text{if } s^* = L^{B}_{\text{T-1}} \\ \gamma V - \alpha_2 - \text{TF}_{\text{LOB}} & \text{if } s^* = M^{\alpha_2}_{\text{T-1}} \end{cases} \tag{3}$$

It therefore follows that the equilibrium strategies $s^*(\gamma|\Omega_{\text{T-1}}=[1001])$ and the equilibrium submission probabilities $\Pr^*(s^*|\Omega_{\text{T-1}}=[1001])$ can be computed using $\Gamma(s^*)$ as for period T in Section 3.2.1 since the execution probabilities of limit orders at T-1 are the submission probabilities of market orders at T which are completely known.

3.2.3 Period T-2

Having derived the equilibrium strategies and probabilities at T-1, we recursively obtain the equilibrium strategies and the equilibrium probabilities for period T-2.

4 Empirical Predictions

The goal of this section is two-fold. First (Sections 4.1 & 4.2), we use the model from Section 3 to understand how traders react to a change in fees. Second (Section 4.3), we corroborate the model from Section 3 and thus its findings by drawing empirical predictions for an event study, namely the change in pricing that took place on the trading centre Toronto Stock Exchange (TSX)

[8] $\mathbb{E}[P(L^{\alpha_1}_{\text{T-1}}|\Omega_{\text{T-1}}=[1001])] = \left(\alpha_1 - \gamma V - \text{MF}_{\text{LOB}}\right) \times \Pr\left(L^{\alpha_1}_{\text{T-1}}|\Omega_{\text{T-1}}=[1001]\right) = \left(\alpha_1 - \gamma V - \text{MF}_{\text{LOB}}\right) \times \Pr\left(M^{\alpha_1}_{\text{T}}|\Omega_{\text{T}}=[1011]\right) = \left(\alpha_1 - \gamma V - \text{MF}_{\text{LOB}}\right) \times \left(1 - \frac{\alpha_1+\text{TF}_{\text{LOB}}}{2V}\right)$ and similarly for the other strategies.

on June 1, 2016, which we will later (Section 5) show match the empirical analysis for the event.

4.1 Increase in Make Fee (Reduction in Rebate)

We start by studying the make fee and in particular, what happens when the make fee (and the make fee alone) is changed. We do this by looking into the effects on traders' strategies and market quality when all fees are kept constant except for the make fee on the limit order book (LOB). Since the numerical fees used in our model are arbitrary, and since we are also interested in the change in pricing on TSX on June 1, 2016, we calibrate our model consistently with the Inter-listed stocks traded on TSX (see Table 1.1) by increasing the make fee on the LOB (MF_{LOB}) by 3bps from -26bps to -23bps whilst keeping constant at 30bps the take fee on the LOB (TF_{LOB}).[9] Moreover, our model consists of a limit order book and an alternative trading system (ATS). As discussed in the introduction, the main Canadian competitor to TSX which makes use of a make-take pricing structure is Chi-X Canada (CXC), thus this trading venue is used as the ATS, whilst TSX corresponds to the LOB. The fees available on the ATS are set consistently with CXC in Table 1.1 to 29bps for the take fee (TF_{ATS}) and -25bps for make fee (MF_{ATS}). We therefore run the model twice, first with the fees in place prior to June 1, 2016 (30bps,-26bps; 29bps,-25bps) and then with the make fee in place post June 1, 2016 (30bps,-23bps; 29bps,-25bps), and we compare the equilibrium submission probabilities from the two periods in Table 1.2.[10] Estimates are not reported for period T (the last period of the model) since in this period limit orders have zero execution probability and therefore are never submitted by traders meaning there is no interaction between market orders and limit orders which is precisely the object of interest.

From Table 1.2 we observe that limit orders to the LOB (L^{LOB}) decrease both in period T-2 (by 0.00229 or 0.68%) and in period T-1 (by 0.03366 or 18.41%), and this happens because following a reduction in the make fee, the incentive to supply liquidity to the LOB via limit orders decreases. In a single market model (that is, a LOB without an ATS), traders switch away from L^{LOB} to market orders (M^{LOB}) and no trade (N) since these are the 2 only other available strategies, so we

[9] Recall that Inter-listed stocks are those stocks that are cross-listed on both the TSX and an American Exchange, and that Non-inter-listed stocks are those stocks which are not Inter-listed.

[10] The notation used (TF_{LOB}, MF_{LOB}; TF_{ATS}, MF_{ATS}) stands for (Take Fee on LOB, Make Fee on LOB; Take Fee on ATS, Make Fee on ATS).

Table 1.2: Increase in the LOB Make Fee (Reduction in the LOB Rebate)

Make Fee Increase (Rebate Reduction)

		Pre-Event	Post-Event	Diff.	% Change
Assumptions	V	1	1		
	ω	0.1	0.1		
	$\mathbf{TF_{LOB}}$	0.0030	0.0030	0	0
	$\mathbf{MF_{LOB}}$	(0.0026)	(0.0023)	0.0003	11.54%
	$\mathbf{TF_{ATS}}$	0.0029	0.0029	0	0
	$\mathbf{MF_{ATS}}$	(0.0025)	(0.0025)	0	0

		Diff.		% Change		Average	
		T-2	T-1	T-2	T-1	Diff.	% Change
Order Submission Strategies	$\mathbf{L^{LOB}}$	(0.00229)	(0.03366)	(0.68%)	(18.41%)	(0.01797)	(9.54%)
	$\mathbf{M^{LOB}}$	(0.00014)	(0.00076)	(0.04%)	(0.12%)	(0.00045)	(0.08%)
	N	0	0.00016	-	1.26%	0.00008	1.26%
	$\mathbf{L^{ATS}}$	0.00243	0.03425	0.95%	21.52%	0.01834	11.24%
	$\mathbf{M^{ATS}}$	0	0	-	-	0	-

		Diff.		% Change		Average	
		T-2	T-1	T-2	T-1	Diff.	% Change
Market Quality Measures	**Spread**	0.00021	0.00352	0.07%	1.13%	0.00187	0.60%
	Volume	(0.00014)	(0.00076)	(0.04%)	(0.12%)	(0.00045)	(0.08%)
	Depth	0.00014	0.00021	0.01%	0.01%	0.00018	0.01%

A detailed explanation of this table is given in Appendix 7.B.

would expect M^{LOB} to increase.[11] However, in a dual market model (that is, a LOB with an ATS - our model) the opposite happens as M^{LOB} fall in both periods (period T-2 by 0.00014 or 0.04% and period T-1 by 0.00076 or 0.12%). Furthermore, we observe that limit orders to the ATS (L^{ATS}) increase in both periods by 0.95% and 21.52%. Combining the fall in M^{LOB} and the rise in L^{ATS} to the fall in L^{LOB}, we understand that traders' activity move away from the LOB (L^{LOB} and M^{LOB}) in favour of the ATS (L^{ATS}). The reason is that if we rank traders' strategies in terms of their aggressiveness, we should consider a L^{LOB} as the order chosen by the most patient trader, followed by L^{ATS} (since the execution price for limit orders to the ATS is worse than the execution price for limit orders to the LOB) and then by M^{LOB} which is the most aggressive order they can select. Clearly when the make fee (which is the rebate traders attain on their limit orders to the limit order book) decreases, the most patient traders (L^{LOB}) switch to their next most profitable strategy which is L^{ATS} explaining the increase in limit orders to the ATS. Moreover, traders anticipate that

[11] Results for the single market model are not reported to economise space, but are available from the author upon request.

the activity is moving to the competing venue (i.e. the ATS) and will consequently submit fewer M^{LOB} given there are fewer L^{LOB} for them to hit, explaining the fall in market orders to the LOB. We therefore conclude that the dual market model differs from the single market model by introducing competition in the form of an ATS, and that liquidity moves from the LOB to the ATS after a change in the make fee on the LOB.

The final effects on market quality are that spread deteriorates and depth improves. In the single market model, the reduction in L^{LOB}, which decreases liquidity supply, and the increase in M^{LOB}, which increases liquidity consumption, lead to two negative effects (namely, the fall in the supply of liquidity and the rise in its consumption) on market quality and thus an overall worsening of market quality.[11] The model with competition (i.e. the dual market model) differs from the former as both L^{LOB} and M^{LOB} fall leading to two opposing effects on market quality with the overall effect being a deterioration of spread (Table 1.2 shows spread increases in both periods by 0.00021 and 0.00352 or 0.07% and 1.13%) and an improvement of depth (shown to increase by Table 1.2 in period T-2 by 0.00014 or 0.01% and in period T-1 by 0.00021 or 0.01%). Spread worsens as the effect of the reduction in L^{LOB} (which affects market quality negatively) on spread overwhelms the effect of the reduction in M^{LOB} (which affects market quality positively) on spread, whereas the opposite happens for depth with the latter improving since the effect of the reduction in M^{LOB} on depth outweighs the effect of the reduction in L^{LOB} on depth.

Table 1.2 is supported by Appendix 7.D Figure 1.2, which shows that the effects on market quality and traders' strategies described in Table 1.2 are qualitatively the same for values of the MF_{LOB} that differ from those used in Table 1.2. Therefore the specific fees used in Table 1.2 generalise to different fees, and the same qualitative results are obtained regardless of the fees chosen. This leads to our first empirical prediction.

Empirical Prediction 1. *In a model where a limit order book (LOB) competes with an alternative trading system (ATS) executing at the mid-quote of the best bid-ask price of the LOB, an increase to the LOB make fee (which is equivalent to a reduction to the LOB rebate) leads to a migration of liquidity from the LOB to the ATS, as well as a worsening of LOB spread and volume, and an improvement in LOB depth.*

4.2 Reduction in Take Fee

We turn our attention to understanding the take fee and in particular, what happens when the take fee is reduced. We do this by looking into the effects on traders' strategies and market quality when all fees are kept constant except for the LOB take fee. Similarly to Section 4.1, we run the model twice, first with the fees in place on TSX and CXC prior to June 1, 2016 (30bps,-26bps; 29bps,-25bps) and then with the take fee in place post June 1, 2016 (27bps,-26bps; 29bps,-25bps), and we compare the equilibrium submission probabilities from the two periods in Table 1.3. We conclude that M^{LOB} and L^{LOB} increase in both periods whilst L^{ATS} falls, so liquidity is drained from the ATS into the LOB. The final effect on market quality is that spread improves and depth deteriorates, leading to a trade off. More generally, under make-take pricing every trading platform faces a trade-off between market quality and volume: on the one hand the exchange wants high volume since this generates revenue, and on the other hand it wants good market quality to attract market participants but the latter are a cost borne in the form of make fees.

Table 1.3: Reduction in the LOB Take Fee

Take Fee Reduction

		Pre-Event	**Post-Event**	**Diff.**	**% Change**
Assumptions	V	1	1		
	ω	0.1	0.1		
	$\mathbf{TF_{LOB}}$	0.0030	0.0027	(0.0003)	(10.00%)
	$\mathbf{MF_{LOB}}$	(0.0026)	(0.0026)	0	0
	$\mathbf{TF_{ATS}}$	0.0029	0.0029	0	0
	$\mathbf{MF_{ATS}}$	(0.0025)	(0.0025)	0	0

		Diff.		**% Change**		**Average**	
		T-2	**T-1**	**T-2**	**T-1**	**Diff.**	**% Change**
Order Submission Strategies	$\mathbf{L^{LOB}}$	0.00154	0.00014	0.46%	0.08%	0.00084	0.27%
	$\mathbf{M^{LOB}}$	0.00116	0.00126	0.29%	0.20%	0.00121	0.24%
	N	0	(0.00017)	-	(1.36%)	(0.00009)	(1.36%)
	$\mathbf{L^{ATS}}$	(0.00270)	(0.00123)	(1.06%)	(0.78%)	(0.00197)	(0.92%)
	$\mathbf{M^{ATS}}$	0	0	-	-	0	-

		Diff.		**% Change**		**Average**	
		T-2	**T-1**	**T-2**	**T-1**	**Diff.**	**% Change**
Market Quality Measures	**Spread**	(0.00004)	(0.00004)	(0.01%)	(0.01%)	(0.00004)	(0.01%)
	Volume	0.00116	0.00126	0.29%	0.20%	0.00121	0.24%
	Depth	(0.00116)	(0.00052)	(0.07%)	(0.03%)	(0.00084)	(0.05%)

A detailed explanation of this table is given in Appendix 7.B.

Similarly to Appendix 7.D Figure 1.2 for Section 4.1, Appendix 7.D Figure 1.3 supports the evidence found in Table 1.3 and shows the results from Table 1.3 are qualitatively the same for different changes in the TF_{LOB}, generalising our findings. This leads to our second empirical prediction.

Empirical Prediction 2. *In a model where a limit order book (LOB) competes with an alternative trading system (ATS) executing at the mid-quote of the best bid-ask price of the LOB, a reduction to the LOB take fee leads to a migration of liquidity from the ATS to the LOB, as well as an improvement in LOB spread and volume, and a worsening of LOB depth.*

4.3 Increase in Make Fee & Reduction in Take Fee

As a final step, we draw an empirical prediction for the change in pricing that took place on the trading centre Toronto Stock Exchange (TSX) on June 1, 2016, and we test that against the results from the empirical analysis in Section 5. If the empirical prediction matches the empirics, then first, this corroborates our model and with it Empirical Predictions 1 and 2, and second, we can use our model to explain the dynamics behind the event and the results from the empirics.

To test for the change in traders' strategies and market quality that occurred on TSX on June 1, 2006, we calibrate our model consistently with TSX and CXC in Table 1.1. In particular, we use inter-listed stocks since the fees on non-inter-listed were left unchanged by TMX. The fees changed from (30bps,-26bps) to (27bps,-23bps) on TSX and from (29bps,-25bps) to (26bps,-22bps) on CXC, using the standard notation (take fee, make fee). We notice that both the take fee and the make fee are changed on both platforms, that is 4 fees are changed simultaneously. We therefore run the model twice, first with the fees in place prior to June 1, 2016 (30bps,-26bps; 29bps,-25bps) and then with the fees in place posterior to June 1, 2016 (27bps,-23bps; 26bps,-22bps), and we compare the equilibrium submission strategies from the two periods in Table 1.4. We conclude that M^{LOB} and L^{LOB} increase in both periods whilst L^{ATS} falls, so liquidity is drained from the ATS into the LOB, and that the final effect on market quality is that spread improves and depth deteriorates. In other words, the change in fees under consideration is qualitatively consistent and in line with Section 4.2 where only the LOB take fee was reduced whilst keeping all other fees constant. In

order to understand why the change in fees on TSX on June 1, 2006 affects traders' strategies and market quality in a way that is qualitatively identical to a single LOB take fee reduction, we need to break down our analysis.

Table 1.4: Increase in LOB/ATS Make Fee & Reduction in LOB/ATS Take Fee

Make Fee Increase & Take Fee Reduction

		Pre-Event	Post-Event	Diff.	% Change
Assumptions	V	1	1		
	ω	0.1	0.1		
	$\mathbf{TF_{LOB}}$	0.0030	0.0027	(0.0003)	(10.00%)
	$\mathbf{MF_{LOB}}$	(0.0026)	(0.0023)	0.0003	11.54%
	$\mathbf{TF_{ATS}}$	0.0029	0.0026	(0.0003)	(10.34%)
	$\mathbf{MF_{ATS}}$	(0.0025)	(0.0022)	0.0003	12.00%

		Diff.		% Change		Average	
		T-2	T-1	T-2	T-1	Diff.	% Change
Order Submission Strategies	$\mathbf{L^{LOB}}$	0.00242	0.00028	0.71%	0.15%	0.00135	0.43%
	$\mathbf{M^{LOB}}$	0.00199	0.00199	0.49%	0.31%	0.00199	0.40%
	N	0	(0.00022)	-	(1.73%)	(0.00011)	(1.73%)
	$\mathbf{L^{ATS}}$	(0.00441)	(0.00204)	(1.73%)	(1.28%)	(0.00323)	(1.51%)
	$\mathbf{M^{ATS}}$	0	0	-	-	0	-

		Diff.		% Change		Average	
		T-2	T-1	T-2	T-1	Diff.	% Change
Market Quality Measures	**Spread**	(0.00004)	(0.00007)	(0.01%)	(0.02%)	(0.00006)	(0.02%)
	Volume	0.00199	0.00199	0.49%	0.31%	0.00199	0.40%
	Depth	(0.00199)	(0.00082)	(0.12%)	(0.05%)	(0.00141)	(0.09%)

A detailed explanation of this table is given in Appendix 7.B.

From Table 1.4, we observe that in absolute value the percentage change in LOB take fee ($\mathrm{TF_{LOB}}$) is smaller than the percentage change in LOB make fee ($\mathrm{MF_{LOB}}$) since $|\%\Delta \mathrm{TF_{LOB}}| = 10.00\% < 11.54\% = |\%\Delta \mathrm{MF_{LOB}}|$ so out of the two mechanisms described in Sections 4.1 & 4.2, the former prevails. Furthermore, the fees on the ATS are also changed and in particular, we have that in absolute value the percentage change in ATS take fee ($\mathrm{TF_{ATS}}$) is smaller than the percentage change in ATS make fee ($\mathrm{MF_{ATS}}$) since $|\%\Delta \mathrm{TF_{ATS}}| = 10.34\% < 12.00\% = |\%\Delta \mathrm{MF_{ATS}}|$ so out of the two changes in fees on the ATS, the ATS make fee prevails. Putting together the fee changes on the LOB with those on the ATS, we observe that the prevailing fee on the LOB is the LOB make fee and the prevailing fee on the ATS is the ATS make fee. We notice that the increase in the ATS make fee has the opposite effect on traders to an increase in the LOB

make fee since the former moves liquidity from the ATS to the LOB while the latter does the opposite, so that two opposing effects are in place. To determine which of these two contrasting effects prevails, we compare the difference in absolute values between the percentage change in LOB make and take fees with the same absolute value percentage difference in ATS make and take fees: $|\%\Delta \mathrm{MF}_{\mathrm{LOB}}| - |\%\Delta \mathrm{TF}_{\mathrm{LOB}}| = 1.54 < 1.66 = |\%\Delta \mathrm{MF}_{\mathrm{ATS}}| - |\%\Delta \mathrm{TF}_{\mathrm{ATS}}|$. We therefore conclude that the ATS make fee is the dominant fee change and this explains why the overall effect we observe in Table 1.4 is a movement of liquidity from the ATS to the LOB leading to results on traders' strategies and market quality that are in line with Empirical Prediction 2 from Section 4.2 where only the LOB take fee is reduced.

This leads to our third and final empirical prediction.

Empirical Prediction 3. *According to our model, the change in make-take pricing that took place on Toronto Stock Exchange on June 1, 2006, should lead to an improvement in inside spread and a worsening of BBO depth, as well as an increase in executed volume. This result is in line with a single reduction in the limit order book take fee, and this is due to the fact that the change in fees on the alternative trading system overwhelms that on the limit order book.*

5 Empirical Analysis

In the remainder of the paper we focus on the second stage of the Make-Take Reduction Program (discussed in greater detail in Appendix 7.C) that took place on the Toronto Stock Exchange in 2016. The maker-taker fees were changed as per Table 1.1 on the inter-listed stocks and were left unchanged on the non-inter-listed stocks which we can therefore use, as explained in Section 5.4, as the control group. In particular, we assess empirically the effect this change in fees has on market quality, measured as inside spread and BBO depth, and executed volume. We then check whether the empirics match Empirical Prediction 3, where the model is calibrated consistently with the event under scrutiny, and if this is the case, this corroborates our model and with it, Empirical Predictions 1 & 2, which examine the impact on market quality and traders' strategies of a change in the make-take fees. Moreover, if the empirics match Empirical Prediction 3, we can use the

model to explain the dynamics behind the changes in market quality we observe in the empirics.

5.1 Sample Selection

We focus on June 1, 2016, and in particular, we take an event window of four months, two before and two after the event, so our sample period starts on April 1, 2016 and ends on July 31, 2016 for a total of 84 trading days. Notice there are only 84 trading days as May 23, 2016 and July 1, 2016 are public holidays in Canada. Over the period ranging from April 1, 2016 to July 31, 2016, there are three abnormal days that should be adjusted for before proceeding with our analysis. These are the Memorial Day (May 30), Brexit (June 24) and Independence Day (July 4). May 30 and July 4 are two federal holidays in the United States (but not in Canada), and the reduced market activity in the U.S. is likely to have affected inter-listed stocks differently from non-inter-listed stocks. June 24 is the day when the results on Brexit were released and this caused a spike in volatility which again is likely to have influenced inter-listed stocks differently from non-inter-listed stocks, since the former have more of their exposure internationally than the latter which are predominantly Canadian-focused. These three days are structurally different for exogenous reasons from the other trading days and therefore we exclude them from the regression analysis (Section 5.3).

The selection of our sample starts by downloading from the TSX website all 1,487 listed companies, and then separating the 180 inter-listed companies from the 1,307 that are non-inter-listed. We then excluded all securities, such as ETFs and preferred equities, that are not ordinary stocks, and out of the latter we only kept those that are continuously traded over the sample period, leaving 853 stocks. Furthermore, we removed securities that became inter-listed over the event window, as well as securities that had less than 20 trades per day on average, and out of the remaining stocks we picked for each category (inter- or non-inter- listed) the 50 stocks that had the largest market capitalisation. The final sample is found in Appendix 7.E Table 1.9. Notice that all stocks are priced above $ 1.00, and thus are traded under the make-take pricing in Table 1.1. Having selected the sample, we download from Thompson Reuters Tick History time and sales TAQ high frequency data down to the microsecond for the 100 stocks in our sample over the 4 months that constitute the sample period. This dataset consists of 2 billion observations with information both on price

and volume of all executed trades and on the best ask and bid prices and sizes of all quotes posted during the continuous trading day.

5.2 Descriptive Statistics

We start by computing six daily metric (defined in Appendix 7.E Table 1.8) that summarise intraday trading and that will allow us to gauge market quality over time. Indeed, it is these metrics that we use in the next section to study the effect on market quality of the change in fees that took place on June 1, 2016 on TSX. The measures of market quality we are interested in are Spread, Quote Size, Trading Activity and Volatility, and the daily metrics used to evaluate them are Time Wighted Percentage Spread (TWPS) and Time Wighted Percentage Cum-fee Spread (TWPCS) for Spread, Time Weighted Depth (TWD) for Depth, Executed Volume and Average Trade Size (AVTS) for Trading Activity, and Realised Volatility (RV) for Volatility. The advantage of using TWPS over Quoted Spread is that the former is more informative since it accounts for both price (via the mid-quote) and time (via the weights $w_{i,t,j}$ which measure the time the quote is on the book for before being updated). The TWPCS is an improvement on TWPS since it factors in the take fee which must be paid by liquidity takers.

The measures of market quality are averaged over stocks and time in Appendix 7.E Table 1.10, and for each metric we compute four summary statistics, namely the mean, standard deviation, minimum and maximum, for each category of stock (inter-listed or non-inter-listed) and for each time period (before or after June 2016). Furthermore, we average each metric over stocks and plot it against time for both inter-listed and non-inter-listed securities in Appendix 7.D Figure 1.4.

5.3 Regression Analysis

On June 1, 2016, the Toronto Stock Exchange (TSX) altered the maker-taker fees it offers to investors that trade on its platform (Table 1.1), and we measure the impact this event has on market quality by means of a difference-in-difference event study estimation consisting of the following

regression for stock $i \in \{1, 2, ..., 100\}$ and time (measured in trading days) $t \in \{1, 2, ..., 81\}$:[12]

$$\text{Market Quality}_{it} = \alpha_0 + \alpha_1 * \text{Event}_t + \alpha_2 * \text{Treatment}_i + \alpha_3 * \text{Interaction}_{it} + \alpha_4 \text{VIX}_{it} + \upsilon_{it} \quad (4)$$

Notice we include the CBOE volatility index since inter-listed securities on TSX are cross-listed on a U.S. trading venue and thus are more exposed to U.S.-based volatility than non-inter-listed securities. The difference-in-difference event study approach we adopt does account for world-wide volatility and other shocks that affect the two groups in the same way, but inter-listed securities will react differently from non-inter-listed securities to a U.S. based shock, and this is why we include the VIX index as one of the independent variables. Moving on to the dependent variable, the following three metrics are used to gauge market quality (MQ): the time weighted percentage cum-fee spread (TWPCS), the time weighted depth (TWD) and the average trade size (AVTS). We use the TWPCS since this measure of spread is directly comparable with the measure of spread computed by our model. In our model, traders have to pay the take fee to consume liquidity and thus the take fee affects the trader's expected payoff, which in turn affects the trader's strategy, thus making the resulting spread net of take fee.[13] The TWD matches the depth computed by the model, and the AVTS is to be used for comparison against the Volume measured by the model, for in the model each trader can submit one order per period, and normalising our empirical measure for Volume to a single trade gives the AVTS. Finally, under difference-in-difference estimation with no additional exogenous shocks on the event date, we may omit α_1 which only increases the R^2 by 0.0003 and α_2 which is co-linear with the other independent variables, while we focus our analysis on α_3 which measures the effect of the change in fees on market quality controlling for non-inter-listed

[12] Market Quality$_{it}$ is the daily realisation of one of the metrics discussed in Section 5.2;
$\alpha_i \in \mathbb{R} \; \forall \; i$;
Event$_t$ is 1 if trading day t corresponds to a date equal to or posterior to June 1, 2016 ($t \in \{42, 43, ..., 81\}$) and 0 otherwise ($t \in \{1, 2, ..., 41\}$);
Treatment$_i$ is 1 if i is an inter-listed stock ($i \in \{1, 2, ..., 50\}$) and 0 otherwise ($i \in \{51, 52, ..., 100\}$);
Interaction$_{it}$ is the product of Event$_t$ and Treatment$_i$;
VIX$_{it}$ is the closing value of the CBOE volatility index;
υ_{it} is the error component.
Furthermore, there are 84 trading days in the four months under scrutiny but we remove the three abnormal days (Memorial Day, Brexit and Independence Day) leaving 81 trading days.

[13] The time weighted percentage spread yields qualitatively consistent results.

securities. We are thus left with the following specification:

$$\text{Market Quality}_{it} = \beta_0 + \beta_1 * \text{Interaction}_{it} + \beta_2 * \text{VIX}_{it} + \upsilon_{it} \tag{5}$$

Equation (5) is estimated using the classical panel data regression estimator with fixed effects (as we make no assumptions about the unobserved cross-sectional heterogeneity and the ensuing correlation across stocks), cluster robust standard errors (to account for the unobserved time-dependant heterogeneity and the ensuing correlation within stocks across trading days), and Newey-West standard errors (to deal with the serial correlation and heteroskedasticity of υ_{it}). The estimates of our estimator are reported in Table 1.5. We find that Interaction_{it} (which is the product of Event_t and Treatment_i therefore it is 1 for inter-listed stocks traded after June 1, 2016) for the TWPCS is negative (-0.0085) and since its p-value is 0.0000 to four significant figures, our estimate is statistically significant. In summary, the interaction term for the TWPCS is negative and significant, implying that the TWPCS falls as a result of the change in fees even controlling for non-inter-listed stocks which were not affected by the change in fees. Moreover, the VIX index has a coefficient of 0.0003 and is significant at the 1% significance level, whilst the constant of our linear regression is 0.0940 (4 s.f.) and is statistically significant at the 1% significance level. Turning our attention to the TWD, we observe that the interaction term is negative but not significant at the 5% significance level indicating the fall in TWD is not statistically significant when controlling for non-inter-listed stocks. Finally, the AVTS has β_1 equal to 5.5998 and this is significant at the 1% significance level since the p-value is 0.0000, thus the change in fees that took place on TSX on June 1, 2016 led to an increase in the average trade size after having controlled for the increase in AVTS of non-inter-listed stocks.

Empirical Result 1. *To summarise our empirical analysis of the change in make-take fees that took place on Toronto Stock Exchange on June 1, 2016, after controlling for non-inter-listed stocks, we find that on inter-listed stocks the time weighted percentage cum-fee spread falls and the fall is statistically significant, the time weighted depth falls and the fall is not statistically significant, and the average trade size increases and the increase is statistically significant.*

Table 1.5: Panel Data Regression

$$\textbf{Market Quality}_{it} = \beta_0 + \beta_1 * \textbf{Interaction}_{it} + \beta_2 * \textbf{VIX}_{it} + \upsilon_{it}$$

Market Quality	Independent Variable	Coef.	Std. Err.	p-value
Time Weighted Percentage Cum-fee Spread	Interaction	-0.0085	0.0006	0.0000***
	VIX	0.0003	0.0001	0.0030***
	Constant	0.0940	0.0013	0.0000***
Time Weighted Depth	Interaction	-0.9740	0.5405	0.0720*
	VIX	-0.1794	0.0818	0.0280**
	Constant	30.37	1.2278	0.0000***
Average Trade Size	Interaction	5.5998	0.7172	0.0000***
	VIX	-0.0884	0.1085	0.4160
	Constant	156.38	1.6292	0.0000***

* if p-value < 0.1; ** if p-value < 0.05; *** if p-value < 0.01

Empirical Prediction 3 predicts an improvement of inside spread, a worsening of BBO depth, and an increase in executed volume. Comparing Empirical Result 1 with Empirical Prediction 3, we conclude that the two match. This corroborates the model and its findings. Moreover, we can use the model to explain the dynamics behind what we observe empirically. In particular, the make fee on TSX (which corresponds to MF_{LOB} in our model) is increased and using Empirical Prediction 1 it follows that this leads to liquidity being drained from the TSX, but the take fee (TF_{LOB}) too is changed and its reduction lures liquidity into the TSX according to Empirical Prediction 2, so the increase in make fee and reduction in take fee lead to two contrasting effects on liquidity. Since the change in make fee is greater in absolute value than the change in take fee, we expect the former to dominate the latter. However, the make fee (MF_{ATS} in the model) and take fee (TF_{ATS}) on the competing platform (Chi-X Canada) were also altered and this will affect the TSX. The increase in ATS make fee has the opposite effect to an increase in the LOB make fee and therefore moves liquidity into the TSX away from CXC, whilst the reduction in ATS take fee (which has the opposite effect to a reduction in LOB take fee) drains liquidity away from the TSX and into CXC. These are two opposing effects on liquidity, and we expect the effect driven by the ATS make fee to dominate that from the ATS take fee since it has a greater percentage change in absolute value. Thus the change in fees on CXC moves liquidity into TSX, whilst the change in fees on

TSX moves liquidity away from the TSX. To gauge the final effect on liquidity, we need to compare the change in fees on TSX against those on CXC. The difference in absolute values between the percentage change in CXC make and take fees is greater than that for TSX, meaning the change in the change in fees on CXC is greater than on TSX, and therefore the former dominates the latter. The resulting effect is that there is an injection of liquidity (measured as market orders and limit orders, that is all trading activity) into the TSX, and the final effect on market quality is in line with Empirical Prediction 2: inside spread and average trade size improve whilst BBO depth worsens. But the latter is what we observe empirically, so the model indeed explains the dynamics behind what is empirically observed.

Empirical Result 2. *Comparing our findings in Empirical Result 1 with those in Empirical Prediction 3, we conclude that the empirics match the theory and this has two implications. First, this corroborates the Model and therefore Empirical Predictions 1 and 2. Second, we can use the Model (through Empirical Predictions 1 and 2) to explain the dynamics behind the changes in market quality that we observe empirically, namely that inside spread and average trade size improve whilst BBO depth worsens as a result of liquidity moving away from Chi-X Canada and into the Toronto Stock Exchange.*

5.4 Robustness of the Control Group

We are left to show that non-inter-listed stocks are a valid control group which really does isolate the event under scrutiny. In order to do this, we need to show that prior to the event the control group (non-inter-listed stocks - which we abbreviate to "NIL") tracks the treatment group (inter-listed stocks - which we abbreviate to "IL"), also known in the literature as the parallel trend assumption, and that posterior to the event the control group is not affected by the event itself, or in other words that there is no leakage from the treatment to the control group. Having a valid control is necessary since it is practically impossible to completely eliminate all of the bias and outside influence that could alter the results of the experiment, and control groups make it possible to focus on the variables of interest. It is for this reason that we focus on the dataset at hand, namely the change in fees that took place on TSX in 2016. Indeed, as we argue below this dataset

satisfies the parallel trend assumption before the event and is leakage-free after the event.

Firstly, let us show that before the event the parallel trend assumption holds true. We want to show that prior to the event the control tracks the treatment and we do this by breaking up the pre-event period into sub-intervals. In particular, we take as many sub-intervals as possible, that is one for each available date, and we compare the average market quality of the two groups over each sub-interval. We show the two groups are not statistically different over each sub-interval using backward induction on interval k where $k \in [1, 41]$ since our sampling window comprises 41 trading days before the event date June 1, 2016 if we exclude abnormal days. As part of our inductive proof, we also need to extend Equation (5) so to regress market quality on the interval-specific interaction term ($\text{Interaction}_{it}^{k}$) which equals 1 for inter-listed securities traded on or after day k and 0 otherwise:

$$\text{Market Quality}_{it} = \beta_0 + \left[\sum_{k=1}^{41} \beta_1^k * \text{Interaction}_{it}^{k}\right] + \beta_2 * \text{VIX}_{it} + \upsilon_{it} \tag{6}$$

$$H_0 : \beta_1^k = 0 \; \forall \; k \in [1, 41] \tag{7}$$

We exploit Equation (6) by testing the null hypothesis (7) and a failure to reject the null $H_0 : \beta_1^k = 0$ for a specific interval k under the dual assumption that it holds true for all future intervals and that there are no additional macroeconomic stock-specific exogenous shocks, implies we cannot reject $\mathbb{E}[\text{MQ}_{it}|i = \text{IL}, k \leq t \leq 41] = \mathbb{E}[\text{MQ}_{it}|i = \text{NIL}, k \leq t \leq 41]$, that is we cannot reject that market quality on interval k (MQ_{it}^{k}) is not statistically different between inter- and non-inter-listed securities, showing the control indeed matches and is similar to the treatment group. The inductive proof may be broken down into the base case and the inductive case:

(i) **Base Case** ($k = 41$): we test $H_0 : \beta_1^{41} = 0$. Failure to reject the null implies we may not reject $\mathbb{E}[\text{MQ}_{it}|i = \text{IL}, 41 \leq t \leq 41] = \mathbb{E}[\text{MQ}_{it}|i = \text{NIL}, 41 \leq t \leq 41]$, or in other words $\mathbb{E}[\text{MQ}_{it}|i = \text{IL}, t = 41] = \mathbb{E}[\text{MQ}_{it}|i = \text{NIL}, t = 41]$, since there are no future intervals to the last interval and since we assume there are no additional macroeconomic stock-specific exogenous shocks.

(ii) **Continued Inductive Case** (assume control matches treatment $\forall$ intervals s such that

$k \leq s \leq 41$ and show this holds true for interval $k-1$): we test $H_0 : \beta_1^{k-1} = 0$. Failure to reject the null implies $\mathbb{E}[\text{MQ}_{it}|i = \text{IL}, k-1 \leq t \leq 41] = \mathbb{E}[\text{MQ}_{it}|i = \text{NIL}, k-1 \leq t \leq 41]$ since by the continued inductive assumption $\beta_1^s = 0 \ \forall \ k \leq s \leq 41$ and since we assume there are no additional macroeconomic stock-specific exogenous shocks:

(A) $\mathbb{E}[\text{MQ}_{it}|i = \text{IL}, k-1 \leq t \leq 41] = \beta_0 + \beta_1^{k-1} + \beta_1^k + .. + \beta_1^{41} + \beta_2 * \mathbb{E}[\text{VIX}_{it}|i = \text{IL}, k-1 \leq t \leq 41]$

(B) $\mathbb{E}[\text{MQ}_{it}|i = \text{NIL}, k-1 \leq t \leq 41] = \beta_0 + 0 + 0 + ... + 0 + \beta_2 * \mathbb{E}[\text{VIX}_{it}|i = \text{NIL}, k-1 \leq t \leq 41]$

(A) - (B) $\implies \beta_1^{k-1} = \mathbb{E}[\text{MQ}_{it}|i = \text{IL}, k-1 \leq t \leq 41] - \mathbb{E}[\text{MQ}_{it}|i = \text{NIL}, k-1 \leq t \leq 41]$ by our continued inductive assumption and since we assume there are no additional shocks

$\implies \beta_1^{k-1} = 0 \iff \mathbb{E}[\text{MQ}_{it}|i = \text{IL}, k-1 \leq t \leq 41] = \mathbb{E}[\text{MQ}_{it}|i = \text{NIL}, k-1 \leq t \leq 41]$ QED

Thus our inductive proof relies on testing Equation (7) which is reported in Table 1.6. We observe that all p-values except for a select few are not significant completing the proof. We turn our attention to the validity of the control posterior to the event.

Now (secondly posterior to firstly), let us show that after the event the dataset is leakage-free. There are two ways for the event to affect the control group: this can happen through a direct channel or indirectly. The direct effect is not possible since fees are only changed on inter-listed stocks and they are left unchanged on non-inter-listed stocks; and the indirect effect, also known as leakage, is absent since brokers can only choose the platform they trade on and have no input over the stocks they trade since these are chosen by the client/investor. Therefore the broker cannot move from inter-listed stocks to non-inter-listed stocks as a result of the change in fees, what she may do is to pick a different trading venue to TSX which is fine since this does not affect the control.

Putting the above together, the control is valid after the event and this, in conjunction with the parallel trend assumption, shows the validity of the control group across the sampling window. It is worth noticing that having a valid control is somewhat difficult to attain. Indeed, it is rare for an exchange to change its trading fees on some stocks but not on others as is the case in our dataset. What usually happens is that trading fees are changed on one platform (e.g. TSX) but not on others (e.g. CXC) but this constitutes a major problem for anyone attempting to study the fee changes. For it is not possible to use TSX as the treatment group and CXC as the control

Table 1.6: Robustness of Control Group – Parallel Trend Assumption

$$\mathbf{MQ}_{it} = \beta_0 + \left[\sum_{k=1}^{41} \beta_1^k * \mathbf{Interaction}_{it}^k\right] + \beta_2 * \mathbf{VIX}_{it} + \upsilon_{it}$$

$\mathbf{Inter}_{it}^k$	TWPCS	TWD	AVTS	$\mathbf{Inter}_{it}^k$	TWPCS	TWD	AVTS
β_1^1	co-linear	co-linear	co-linear	β_1^{22}	0.561	0.819	0.020**
β_1^2	0.797	0.250	0.340	β_1^{23}	0.267	0.814	0.255
β_1^3	0.851	0.111	0.963	β_1^{24}	0.059*	0.257	0.513
β_1^4	0.958	0.268	0.647	β_1^{25}	0.698	0.367	0.785
β_1^5	0.914	0.455	0.899	β_1^{26}	0.989	0.699	0.085*
β_1^6	0.975	0.369	0.291	β_1^{27}	0.918	0.611	0.239
β_1^7	0.605	0.060*	0.569	β_1^{28}	0.237	0.394	0.28
β_1^8	0.917	0.503	0.340	β_1^{29}	0.987	0.582	0.364
β_1^9	0.612	0.374	0.463	β_1^{30}	0.645	0.624	0.694
β_1^{10}	0.568	0.072*	0.658	β_1^{31}	0.872	0.949	0.863
β_1^{11}	0.995	0.376	0.478	β_1^{32}	0.356	0.201	0.13
β_1^{12}	0.616	0.144	0.872	β_1^{33}	0.760	0.683	0.474
β_1^{13}	0.722	0.029**	0.022**	β_1^{34}	0.567	0.077*	0.133
β_1^{14}	0.224	0.614	0.131	β_1^{35}	0.236	0.259	0.762
β_1^{15}	0.373	0.140	0.250	β_1^{36}	0.037**	0.060*	0.370
β_1^{16}	0.937	0.583	0.339	β_1^{37}	0.679	0.963	0.222
β_1^{17}	0.336	0.266	0.574	β_1^{38}	0.802	0.858	0.674
β_1^{18}	0.552	0.921	0.680	β_1^{39}	0.968	0.875	0.111
β_1^{19}	0.578	0.761	0.372	β_1^{40}	0.976	0.365	0.272
β_1^{20}	0.844	0.701	0.241	β_1^{41}	0.665	0.487	0.328
β_1^{21}	0.950	0.471	0.854				

* if p-value < 0.1; ** if p-value < 0.05; *** if p-value < 0.01

This table reports the p-values of the null hypothesis specified in Equation (7) for the three measures of market quality (MQ) used throughout Section 5: the time weighted percentage cum-fee spread (TWPCS), the time weighted depth (TWD), and the average trade size (AVTS). Consistently with the notation used, β_1^1 refers to the coefficient of the interaction term on April 1, 2016, β_1^{21} to April 29, 2016, β_1^{22} to May 2, 2016, and β_1^{41} to May 31, 2016. Coefficients are estimated by the same estimator used in Section 5.3 Equation (5).

group since brokers will move from TSX to CXC (or vice-versa depending on the change in fees) invalidating the control group which must be kept separate from the treatment group. In this instance, it is leakage from the treatment group TSX to the control group CXC that invalidates the control group and thus the resulting analysis.

6 Conclusions

This paper develops a theoretical model with a closed-form solution that proxies a modern electronic trading platform. The model consists of a limit order book (LOB) and a competing trading venue (ATS) which executes at the mid-quote of the BBO on the LOB. Trading takes place over three periods and in each period all traders arrive in probability, and they may submit an order which may be a sell/buy market/limit order on the LOB/ATS. The model is solved by backward induction using a two-step solution: first, we compute the traders' optimal submission strategy as a function of their personal valuation of the traded asset; second, we compute the submission probability of the traders' optimal submission strategy. The model is used to study the impact a change in fees has on trading activity and market quality (Empirical Predictions 1-3), and to explain the dynamics behind the changes in market quality (Empirical Result 1) that we observed empirically on the Toronto Stock Exchange (TSX) on June 1, 2016 when the make-take pricing scheme was altered (Empirical Result 2).

The model's findings are summarised in three empirical predictions. First, we find that following an increase in the LOB make fee, the incentive to supply liquidity to the LOB via limit orders decreases so market makers move to the ATS as a limit order to the ATS is their next most profitable strategy, leading to a widening of inside spread and a rise in BBO depth on the LOB (which now measures depth at the outside quotes). Since market takers follow market makers, volume falls and there is an overall movement of liquidity from the LOB to the ATS. Second, the model shows that opposite dynamics take place when the LOB take fee is reduced. Finally, we consider a simultaneous change in both the LOB make and take fees, and the ATS make and take fees. Since four fees are changed, there are four different effects at work. The final effect on market quality is in line with a reduction in the LOB take fee. The reason for this is that by looking at

the absolute value of the percentage change, we observe that the fee changes on the ATS dominate those on the LOB. Within the ATS, the make fee dominates the take fee. So the dominating effect is that of an increase in the ATS make fee which has the opposite effect of an increase in the LOB make fee.

The results of the empirical analysis can be summarised in two findings. First, after controlling for non-inter-listed stocks which were not affected by the fee change as shown in Section 5.4, we find that following the June 1, 2016 event inter-listed stocks traded on TSX experienced a fall in time weighted percentage cum-fee spread, an increase in average trade size, and a not significant reduction in time weighted depth. The regression analysis uses diff-in-diff estimation, where the estimator used is the classical panel data regression estimator with fixed effects, cluster robust and Newey-West standard errors. The second finding from our empirical analysis compares the theoretical results from our model with the empirical results and we observe that the two match, corroborating our model. Under the change in fees that took place on TSX on June 1, 2016, our model predicts a narrowing of spread and increase in volume as traders move to TSX away from the competing venues, and this is exactly what we observe empirically.

7 Appendix

7.A Legend of Table 1.1

This table illustrates how the major Canadian trading centres, namely Toronto Stock Exchange (TSX - https://www.tsx.com/), Chi-X Canada (CXC - http://business.nasdaq.com/chi-x-canada), TSX Alpha (TSXA - https://www.tsx.com/trading/tsx-alpha-exchange) and Omega ATS (Omega - https://omegaats.com/), changed their maker-taker pricing on June 1, 2016. Since our analysis only concerns high-priced stocks, that is stocks whose price is $\geq$ \$ 1.00, we report only those changes in fees that concern high-priced stocks. All fees are given in Canadian Dollars.

Consider TSX. We distinguish between Inter-listed stocks, i.e. stocks that are cross-listed on both TSX and an American Exchange, and Non-inter-listed stocks, i.e. all those stocks which are not Inter-listed. We also distinguish between the take fee, which is charged on orders that remove

(take) liquidity from the market (hence the term "taker" or "take fee"), and a make fee, which is assessed on orders that provide (make) liquidity to the market (hence the term "maker" or "make fee"). A negative fee is paid from the Exchange to the broker and therefore is a rebate - negative fees are written in brackets. So on TSX, the take fee on Inter-listed stocks prior to June 1, 2016 is $ 0.0030 for each traded share, and the make fee is ($ 0.0026) per traded share (or equivalently the rebate is $ 0.0026 per traded share). Comparing for Inter-listed stocks across the event, we notice that both make and take fees are changed by 3 basis points (bps), thus keeping the Exchange Fee (namely, the sum of the make and the take fees) constant at 4 bps. Fees on Non-inter-listed stocks are left unchanged.

A similar analysis follows for CXC. TSXA makes no distinction between stocks and on June 1, 2016 switched from classical (positive take fee and negative make fee) to inverted (negative take fee and positive make fee) make-take pricing. OMG adopts classical make-take pricing throughout and distinguishes between trades that have an Executed Order Value (EOV), measured as the number of shares traded multiplied by the value of each share, smaller than $ 100 million, trades with an EOV $\in$ [$ 100MM, $ 200MM), and trades with an EOV $\geq$ $ 200MM.

7.B Legend of Table 1.2

This table reports the assumptions needed to run our model (developed in Section 3) twice: first, for the fees in place on Toronto Stock Exchange (which represents the LOB in our model) and Chi-X Canada (which represents the ATS in our model) prior to June 1, 2016 (which we call Pre-Event in the table), and then, for the same fees except that the limit order book make fee (MF_{LOB}) is increased by 3bps from ($ 0.0026) per traded share to ($ 0.0023) per traded share (column Post-Event in the table). Table 1.1 reports the make-take fees adopted by TSX (Toronto Stock Exchange) and CXC (Chi-X Canada) prior to June 1, 2016. Using the notation (TF_{LOB}, MF_{LOB}; TF_{ATS}, MF_{ATS}), we run the model first for (30bps,-26bps; 29bps,-25bps) and then for (30bps,-23bps; 29bps,-25bps). Comparing the two sets of fees, we observe that the MF_{LOB} is the only fee that is changed, and therefore this table examines the effect of a change in the LOB make fee on Order Submission Strategies and Market Quality. The 5th column in the top third of the

table (Diff.) reports the difference in fees (Post-Event - Pre-Event) and the 6th column (% Change) reports the percentage change (11.54%) measured as [Post-Event - Pre-Event] / Pre-Event. For both numerical simulations, the value V of the traded security is \$ 1 and the tick size ω is 0.1.

Consider the second third of the table, i.e. the one describing the results for the Order Submission Strategies. The 3rd and 4th columns (under Diff.) report the difference in the equilibrium submission probabilities between the model with the parameters specified in Post-Event and the model with the parameters specified in Pre-Event (so Post-Event - Pre-Event). The 3rd column details the difference in period T-2 (the first period of the model), and the 4th in period T-1 (the second period of the model). Notice estimates are not reported for period T (the last period of the model) since in this period limit orders have zero execution probability and therefore are never submitted by traders meaning there is no interaction between market orders and limit orders which is what we are interested in. The 2nd column reports the order submission strategies available to traders, so for example the difference in the equilibrium submission probabilities of LOB limit orders (L^{LOB}) at period T-2 is -0.00229 and that at period T-1 is -0.03366. N^{LOB} stands for limit order book no trade. The 5th and 6th columns report the same estimates as in columns 3 and 4, but instead of taking the difference between post and pre, they take the percentage difference ([Post-Event - Pre-Event] / Pre-Event). Columns 7 and 8 report respectively the average between periods T-2 and T-1 for the difference (average of columns 3 and 4) and for the percentage change (average of columns 5 and 6).

Moving to the bottom third of the table, the Market Quality Measures for the LOB are reported in the same format as the Order Submission Strategies. So for example, the difference in Spread at T-2 is 0.00021, the % difference in Volume at T-1 is -0.12%, and the average difference in Depth is 0.00018 whilst the average % difference in Depth is 0.01%.

7.C The Toronto Stock Exchange

The Toronto Stock Exchange (TSX) was established on October 25, 1861 and is owned by the TMX Group, which is the ninth largest stock exchange group worldwide in terms of market

capitalisation as of January 2018.[14] Prior to October 1, 2005, securities traded on TSX were subject to a value-based pricing, whereby liquidity takers were charged a fee of 2bps, while liquidity makers were not. The current make-take pricing structure, by which liquidity providers are offered a rebate (the absolute value of a negative make fee) and liquidity takers are charged a fee (the take fee), was introduced over two stages between 2005 and 2006. On October 1, 2005, TSX switched to the make-take pricing only for those stocks that were also cross-listed on NASDAQ or AMEX, and kept the old value-based pricing mechanism for the remaining stocks. Stocks cross-listed on NASDAQ and AMEX were offered a rebate of $ 27.5bps (i.e. a make fee of ($ 27.5bps)) per traded share and charged a take fee of $ 40.0bps per traded share, whilst for the remaining stocks the fee was reduced from 2bps to 1.82bps. It was only on July 1, 2006 that the program was extended to all shares traded on TSX. Over the last decade, make-take fees have come under the scrutiny of the regulatory authorities and the question as to whether they should be regulated has risen. It is for this reason that the TMX opted to remove both fees on securities trading below $ 1.00, and two years later it launched the Make-Take Reduction Program on securities trading above $ 1.00. In order to alleviate the burden on traders that comes with new fees, TMX opted to cut the fees over a period that could take from two to three years.

The first stage of the program came into force on June 1, 2015. The program involved all securities trading above $ 1.00 and the fees were cut from (35bps, -31bps) to (30bps, -26bps) for inter-listed stocks and from (35bps, -31bps) to (23bps, -19bps) for the remaining stocks, so that the extent of the reduction differed depending on the stock traded. The list of stocks considered to be inter-listed for billing purposes is updated by TMX every four weeks, and it refers to stocks that are listed on both TSX and a U.S. exchange. TMX performed a similar change in fees on the two other platforms it owns: TSX Venture (TSXV), dedicated to emerging companies, and TSX Alpha (TSXA), which is an alternative trading system.

The second stage of the program was implemented on June 1, 2016, exactly one year after the first reduction. A summary of the fee changes is presented in Table 1.1. TMX changed its fees on both TSX and TSXA: on the one hand, TSX left non-inter-listed securities unchanged whilst

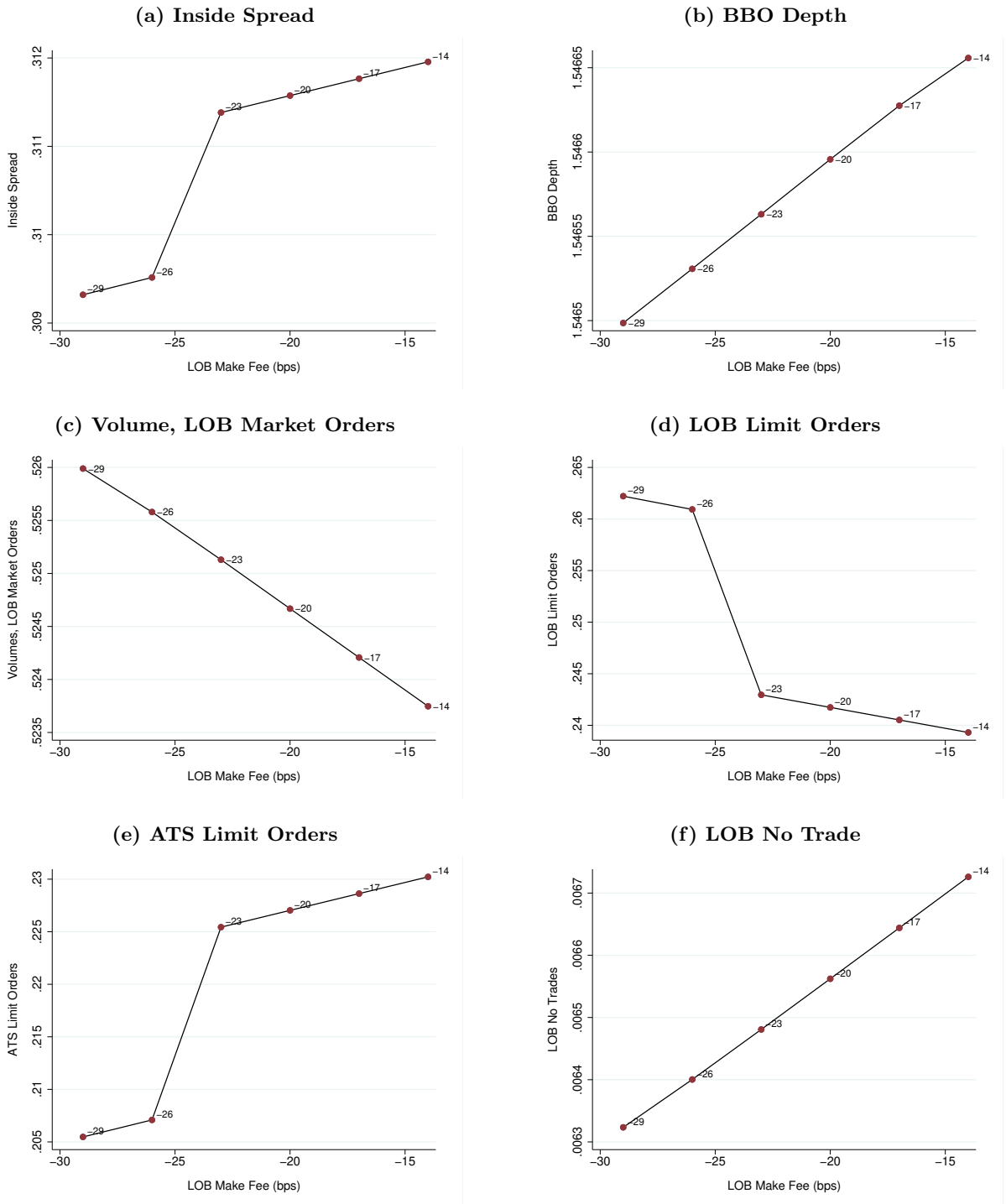
[14] https://www.tsx.com/

cutting both fees on inter-listed securities by 3bps thus keeping the exchange fee (defined as the sum of the take fee and the make fee) constant; and on the other hand, TSXA moved from the standard make-take pricing to an inverted one, where the rebate is given to liquidity takers instead (i.e. the take fee is now negative, and the make fee is now positive). Moreover, TMX invited other Canadian trading venues to follow its lead, in an attempt to increase the effectiveness of the program, and both Chi-X Canada (CXC) and Omega ATS (OMG) got involved. CXC, in line with TSX, cut both fees by 3bps, and OMG, which implements an inverted make-take fee, expanded both fees non-uniformly.

The third and final stage of the program, which consisted in cutting the fees on non-inter-listed stocks, never took place. TMX was discouraged from altering its fees on non-inter-listed stocks by an amendment proposal published on April 7, 2016 by the Canadian Securities Administrators (CSA), which, if approved, would set a lower take fee cap for high-priced non-inter-listed equities at $ 17bps per share. Since the TMX program was planning on cutting the make-take fee on non-inter-listed stocks by 3bps, as it did for inter-listed stocks, this would have brought the take fee down to $ 20bps per share which is still higher than the proposed cap of $ 17bps, necessitating further reductions and with it further disruptions to market participants. It is for this reason that TMX delayed the third stage of the program until the discussed proposed amendments were finalised.

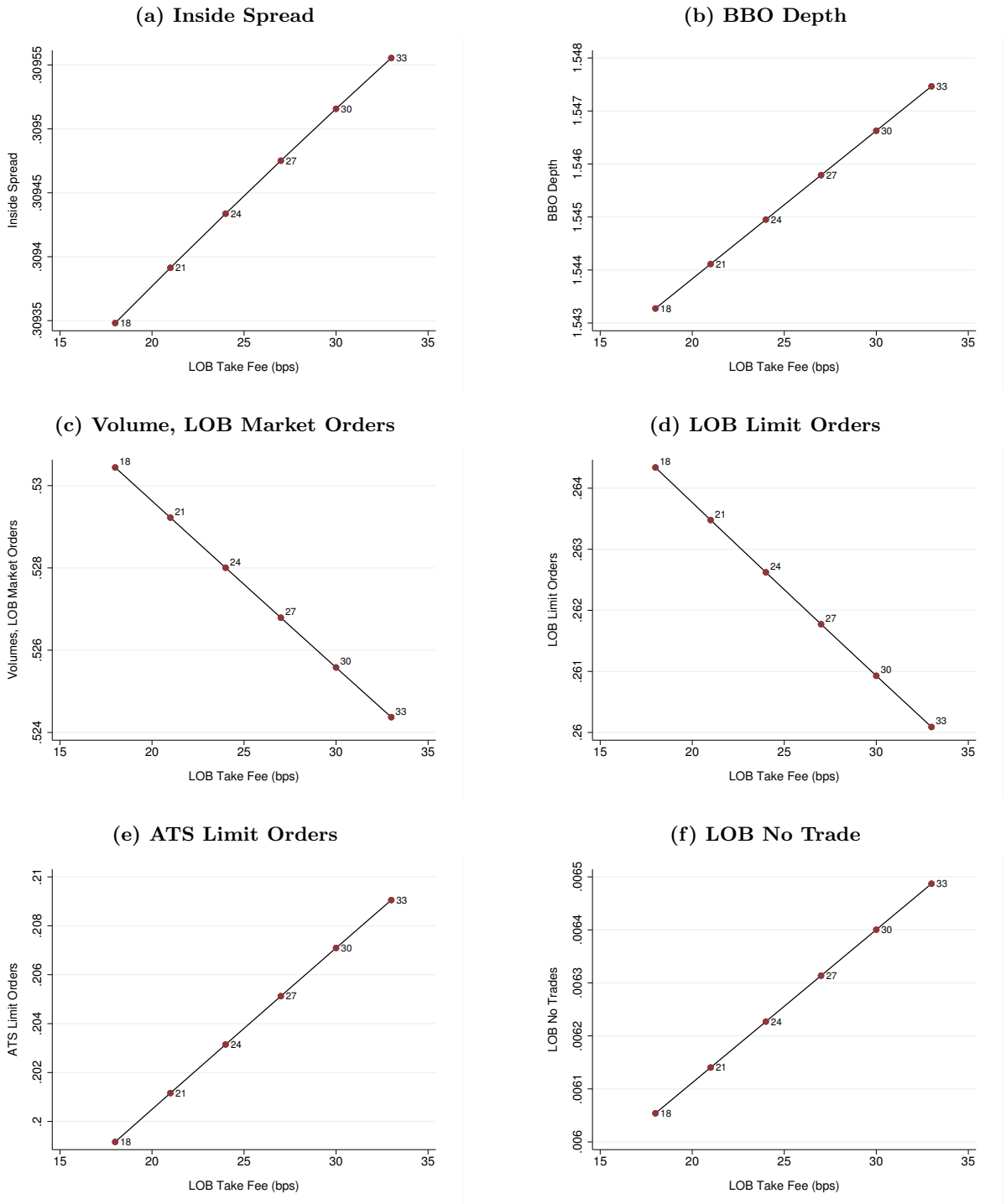
7.D Appendix Figures

Figure 1.2: Generalising the Increase in the LOB Make Fee



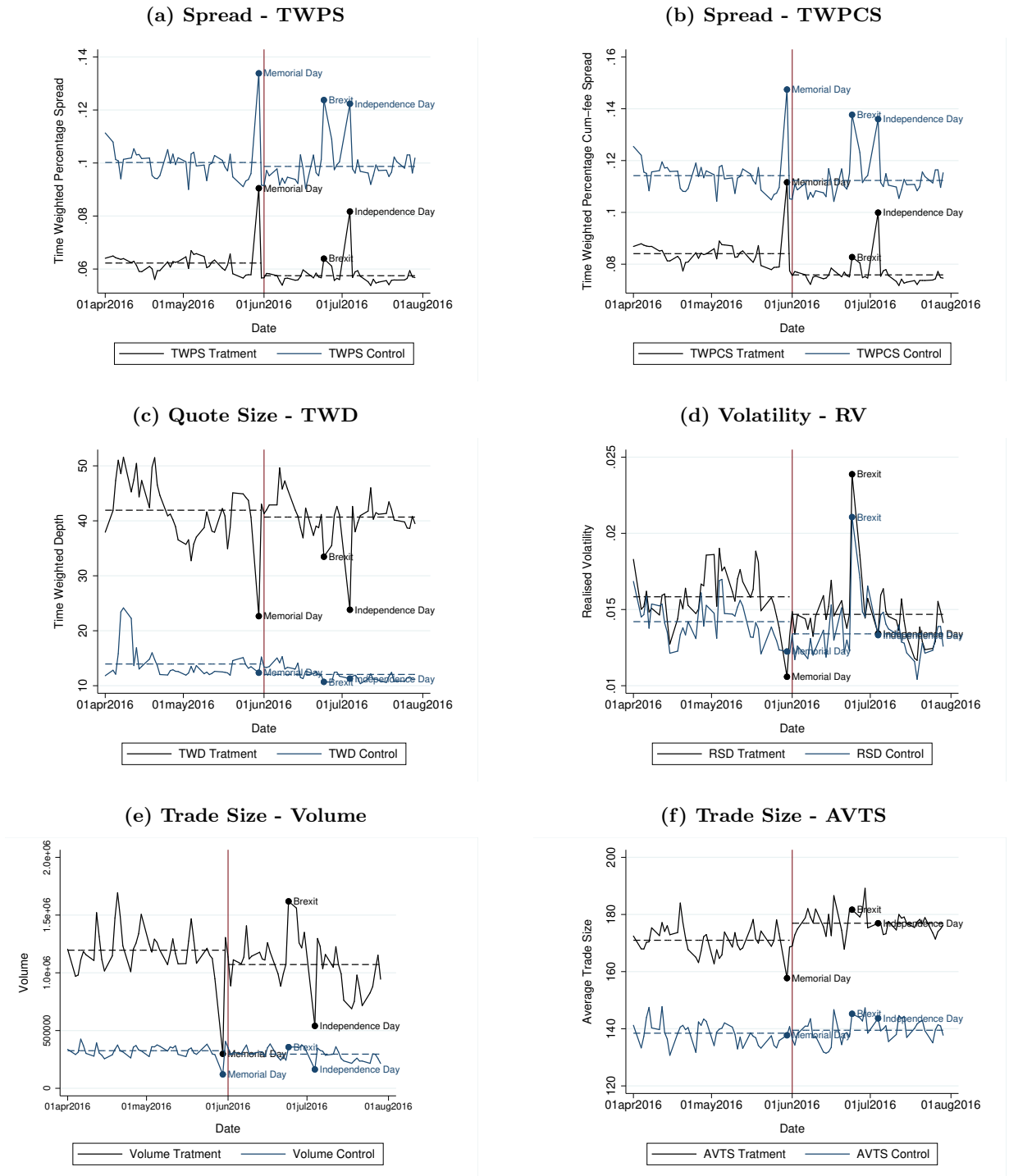
This figure exhibits the market quality metric / trader strategy specified in the subplot title for different values of the limit order book (LOB) make fee (shown on the x-axis in basis points). The plotted measures of market quality and traders' strategies are computed by our model keeping the remaining fees, namely the LOB take fee, the ATS take fee and the ATS make fee, constant at 30bps, 29bps and -25bps respectively.

Figure 1.3: Generalising the Reduction in the LOB Take Fee



This figure exhibits the market quality metric / trader strategy specified in the subplot title for different values of the limit order book (LOB) take fee (shown on the x-axis in basis points). The plotted measures of market quality and traders' strategies are computed by our model keeping the remaining fees, namely the LOB make fee, the ATS take fee and the ATS make fee, constant at -26bps, 29bps and -25bps respectively.

Figure 1.4: Market Quality Metrics – Qualitative Descriptive Statistics



This figure exhibits each market quality metric over the sampling period (April 1, 2016 - July 31, 2016) as a time series plot of the daily observations averaged over inter-listed stocks (the treatment group) as a solid black line and the daily observations averaged over non-inter-listed stocks (the control group) as a solid blue line, together with the pre- and post- event averages drawn as dashed horizontal lines. The three abnormal days (Memorial Day, Brexit, Independence Day) and the event itself (June 1, 2016) are highlighted.

7.E Appendix Tables

Table 1.7: Daily Executed Volume for Canadian Equities

Daily Volume		**Non-Inter-listed**		**Inter-listed**		**Non-Inter-listed + Inter-listed**			
		Pre-Event	Post-Event	Pre-Event	Post-Event	Pre-Event	Post-Event	Pre + Post	% Tot
OMG	Sum	156,333,159	138,237,500	211,123,446	240,812,767	367,456,605	379,050,267	746,506,872	4%
	Mean	53,242	45,187	86,394	83,990	69,818	64,589	67,203	
	Median	26,125	28,909	53,249	53,589	39,687	41,249	40,468	
TSXA	Sum	259,963,169	200,383,500	423,443,821	447,911,266	683,406,990	648,294,766	1,331,701,756	7%
	Mean	75,141	58,891	122,626	129,129	98,883	94,010	96,447	
	Median	34,220	32,191	70,163	76,823	52,191	54,507	53,349	
CXC	Sum	587,251,261	520,971,253	1,245,381,935	1,240,756,919	1,832,633,196	1,761,728,172	3,594,361,368	20%
	Mean	165,157	150,054	357,464	357,762	261,311	253,908	257,609	
	Median	64,339	57,091	183,574	186,467	123,957	121,779	122,868	
TSX	Sum	2,373,977,111	1,962,321,298	4,099,173,943	3,731,247,413	6,473,151,054	5,693,568,711	12,166,719,765	68%
	Mean	560,219	430,528	1,176,391	1,073,033	868,305	751,780	810,043	
	Median	245,001	221,177	644,117	583,223	444,559	402,200	423,380	
Total	Sum	3,377,524,700	2,821,913,551	5,979,123,145	5,660,728,365	9,356,647,845	8,482,641,916	17,839,289,761	100%
	Mean	170,752	136,932	348,575	328,783	259,663	232,857	246,260	
	Median	73,937	67,874	190,221	180,020	132,079	123,947	128,013	

This table reports the total (Sum), average (Mean) and median (Median) daily executed volume between April 1, 2016 and July 31, 2016 over stock category (Non-Inter-listed, Inter-listed or the sum of the two) and time period (Pre- or Post- Event) for each of the four major Canadian trading centres, namely Omega ATS (OMG), TSX Alpha (TSXA), Chi-X Canada (CXC) and Toronto Stock Exchange (TSX). The final column gives the weight of each trading centre as a percentage of the total (Sum) cumulative daily executed volume.

Table 1.8: Market Quality Metrics – Definition

Market Quality	**Metric**	**Formula**
Spread	Time Weighted Percentage Spread	$\frac{1}{\bar{w}_{i,t}} \sum_{j=1}^{J} w_{i,t,j} * \left(\frac{p^{\alpha}_{i,t,j} - p^{\beta}_{i,t,j}}{\frac{1}{2}(p^{\alpha}_{i,t,j} + p^{\beta}_{i,t,j})} \right)$
	Time Weighted Percentage Cum-fee Spread	$\frac{1}{\bar{w}_{i,t}} \sum_{j=1}^{J} w_{i,t,j} * \left(\frac{(p^{\alpha}_{i,t,j} + \text{TF}) - (p^{\beta}_{i,t,j} - \text{TF})}{\frac{1}{2}(p^{\alpha}_{i,t,j} + p^{\beta}_{i,t,j})} \right)$
Quote Size	Time Weighted Depth	$\frac{1}{\bar{w}_{i,t}} \sum_{j=1}^{J} w_{i,t,j} * (q^{\alpha}_{i,t,j} + q^{\beta}_{i,t,j})$
Volatility	Realised Volatility	$\sqrt{\sum_{j=1}^{J} \log^2 \left(\frac{\frac{1}{2}(p^{\alpha}_{i,t,j} + p^{\beta}_{i,t,j})}{\frac{1}{2}(p^{\alpha}_{i,t,j-2} + p^{\beta}_{i,t,j-2})} \right)}$
Trading Activity	Executed Volume	$\sum_{j=1}^{J} q_{i,t,j}$
	Average Trade Size	$\frac{1}{J} \sum_{j=1}^{J} q_{i,t,j}$

This table reports the formulae used to compute the daily market quality metrics, where $(p/q)^{\alpha/\beta}_{i,t,j}$ = Best Ask/Bid Price/Size at intra-day observation (quote) j on stock i traded on day t; $(p/q)_{i,t,j}$ = trading price/quantity at intra-day observation (trade) j on stock i traded on day t; $w_{i,t,j}$ = time difference between observation j and observation $j+1$ for stock i traded on day t and $\bar{w}_{i,t} = \sum_{j=1}^{N} w_{i,t,j}$; TF = take fee on the limit order book; and J is the number of observations in the trading day under consideration.

Table 1.9: Sample

IL	Stock	RIC	Mkt Cap (CAD)
1	Royal Bank of Canada	RY	1.03E+11
2	Toronto-Dominion Bank	TD	9.53E+10
3	Bank of Nova Scotia	BNS	6.51E+10
4	Canadian National Railway Co.	CNR	5.76E+10
5	BCE Inc.	BCE	4.74E+10
6	Bank of Montreal	BMO	4.70E+10
7	Suncor Energy Inc.	SU	4.61E+10
8	Valeant Pharmaceuticals International Inc.	VRX	4.46E+10
9	Brookfield Asset Management Inc.	BAM	4.02E+10
10	Thomson Reuters Corp.	TRI	3.93E+10
11	Enbridge Inc.	ENB	3.88E+10
12	Manulife Financial Corp.	MFC	3.70E+10
13	Imperial Oil Ltd.	IMO	3.52E+10
14	Canadian Imperial Bank Of Commerce	CM	3.49E+10
15	TransCanada Corp.	TRP	3.19E+10
16	Canadian Natural Resources Ltd.	CNQ	2.90E+10
17	Rogers Communications Inc.	RCI	2.50E+10
18	Canadian Pacific Railway Ltd.	CP	2.45E+10
19	Sun Life Financial Inc.	SLF	2.42E+10
20	TELUS Corp.	T	2.24E+10
21	Restaurant Brands International Inc.	QSR	2.20E+10
22	Brookfield Property Partners L.P.	BPY	2.10E+10
23	Magna International Inc.	MG	2.03E+10
24	Potash Corporation of Saskatchewan Inc.	POT	1.90E+10
25	CGI Group Inc.	GIB	1.71E+10
26	Agrium Inc.	AGU	1.68E+10
27	Cenovus Energy Inc.	CVE	1.42E+10
28	Barrick Gold Corp.	ABX	1.40E+10
29	Goldcorp Inc.	G	1.30E+10
30	Shaw Communications Inc.	SJR	1.15E+10
31	Pembina Pipeline Corp.	PPL	1.08E+10
32	Brookfield Infrastructure Partners L.P.	BIP	1.05E+10
33	Franco-Nevada Corp.	FNV	1.00E+10
34	Brookfield Renewable Partners L.P.	BEP	9.32E+09
35	Gildan Activewear Inc.	GIL	9.08E+09
36	Agnico Eagle Mines Ltd.	AEM	8.75E+09
37	Open Text Corp.	OTC	8.01E+09
38	Crescent Point Energy Corp.	CPG	6.94E+09
39	Silver Wheaton Corp.	SLW	6.55E+09
40	Cameco Corp.	CCO	6.41E+09
41	Turquoise Hill Resources Ltd.	TRQ	5.60E+09
42	BlackBerry Ltd.	BB	5.53E+09
43	Encana Corp.	ECA	5.08E+09
44	CAE Inc.	CAE	4.05E+09
45	Vermilion Energy Inc.	VET	3.87E+09
46	Methanex Corp.	MX	3.41E+09
47	Ritchie Bros. Auctioneers Inc.	RBA	3.38E+09
48	Stantec Inc.	STN	3.06E+09
49	Domtar Corp.	UFS	2.94E+09
50	Teck Resources Ltd.	TCK	2.77E+09

NIL	Stock	RIC	Mkt Cap (CAD)
1	Alimentation Couche-Tard Inc.	ATD	3.41E+10
2	Great-West Lifeco Inc.	GWO	3.33E+10
3	Loblaw Companies Ltd.	L	2.66E+10
4	Power Financial Corp.	PWF	2.18E+10
5	Fairfax Financial Holdings Ltd.	FFH	1.65E+10
6	George Weston Ltd.	WN	1.33E+10
7	Husky Energy Inc.	HSE	1.33E+10
8	Hydro One Ltd.	H	1.32E+10
9	Power Corporation Of Canada	POW	1.29E+10
10	Saputo Inc.	SAP	1.28E+10
11	National Bank of Canada	NA	1.28E+10
12	Intact Financial Corp.	IFC	1.12E+10
13	Constellation Software Inc.	CSU	1.09E+10
14	Fortis Inc.	FTS	1.06E+10
15	Metro Inc.	MRU	9.49E+09
16	Dollarama Inc.	DOL	9.28E+09
17	ONEX Corp.	OCX	9.21E+09
18	Canadian Utilities Ltd.	CU	8.58E+09
19	IGM Financial Inc.	IGM	8.29E+09
20	CI Financial Corp.	CIX	8.08E+09
21	CCL Industries Inc.	CCL	7.55E+09
22	Riocan Real Estate Investment Trust	REI	7.44E+09
23	Inter Pipeline Ltd.	IPL	7.07E+09
24	Empire Company Ltd.	EMP	6.92E+09
25	Keyera Corp.	KEY	6.30E+09
26	Emera Inc.	EMA	6.15E+09
27	Element Financial Corp.	EFN	5.84E+09
28	SNC-Lavalin Group Inc.	SNC	5.83E+09
29	ARC Resources Ltd.	ARX	5.65E+09
30	Tourmaline Oil Corp.	TOU	5.33E+09
31	H&R Real Estate Investment Trust	HR	5.29E+09
32	Choice Properties Real Estate Investment Trust	CHP	4.65E+09
33	Smart Real Estate Investment Trust	SRU	4.57E+09
34	PrairieSky Royalty Ltd.	PSK	4.47E+09
35	AltaGas Ltd.	ALA	4.39E+09
36	Peyto Exploration & Development Corp.	PEY	4.25E+09
37	Quebecor Inc.	QBR	4.21E+09
38	Industrial Alliance Insur. and Fin. Services Inc.	IAG	4.17E+09
39	Atco Ltd.	ACO	4.10E+09
40	First Capital Realty Inc.	FCR	4.09E+09
41	Linamar Corp.	LNR	3.95E+09
42	WSP Global Inc.	WSP	3.94E+09
43	West Fraser Timber Co. Ltd.	WFT	3.69E+09
44	Seven Generations Energy Ltd.	VII	3.49E+09
45	Canadian Apartment Prop. Real Est. Invest. Trust	CAR	3.48E+09
46	Jean Coutu Group (PJC) Inc.	PJC	3.40E+09
47	DH Corp.	DH	3.34E+09
48	Stella-Jones Inc.	SJ	3.18E+09
49	Northland Power Inc.	NPI	3.09E+09
50	Maple Leaf Foods Inc.	MFI	3.08E+09

The sample used in the empirical part of the paper consists of 50 Inter-listed (IL) stocks, i.e. stocks that are cross-listed on both TSX and an American Exchange, and 50 Non-Inter-listed (NIL) stocks, i.e. stocks which are not Inter-listed. The market capitalisation (Mkt Cap) is reported in Canadian Dollars (CAD) and for each security.

Table 1.10: Market Quality Metrics – Quantitative Descriptive Statistics

Descriptive Statistics		**Non-Inter-listed**		**Inter-listed**	
		Pre-Event	**Post-Event**	**Pre-Event**	**Post-Event**
Time Weighted Percetage Spread	**Mean**	10.02%	9.87%	6.23%	5.75%
	Std. Dev.	7.11%	7.32%	4.70%	3.94%
	Min.	2.27%	2.39%	1.29%	1.34%
	Max.	63.25%	86.29%	37.51%	31.19%
Time Weighted Percentage Cum-fee Spread	**Mean**	11.42%	11.23%	8.41%	7.59%
	Std. Dev.	6.98%	7.16%	6.73%	5.46%
	Min.	3.28%	3.03%	2.04%	2.02%
	Max.	63.34%	86.38%	49.29%	41.80%
Time Weighted Depth	**Mean**	14	12	42	41
	Std. Dev.	24	7	74	66
	Min.	3	2	4	4
	Max.	613	96	742	457
Realised Volatility	**Mean**	1.42%	1.34%	1.58%	1.47%
	Std. Dev.	0.87%	0.72%	1.05%	0.81%
	Min.	0.19%	0.35%	0.38%	0.40%
	Max.	9.70%	7.86%	10.66%	8.18%
Executed Volume	**Mean**	325,602	295,684	1,195,895	1,071,624
	Std. Dev.	303,897	266,890	1,435,226	1,165,213
	Min.	6,536	6,927	2,856	8,191
	Max.	4,071,046	2,539,660	16,300,000	8,796,482
Average Trade Size	**Mean**	138	139	171	177
	Std. Dev.	42	29	88	87
	Min.	62	53	93	97
	Max.	927	367	778	632

This table reports four (mean, standard deviation, minimum and maximum) summary statistics averaged over stock category (Non-Inter-listed or Inter-listed) and time period (Pre- or Post- Event) for each of the six market quality metrics defined in Appendix 7.E Table 1.8.

Chapter 2

Assessing the Attributes and Implications of Mini Flash Crashes

Abstract

In this paper we perform an empirical study of all Mini Flash Crashes (MFCs) that took place on U.S. equities between January 3, 2006 and February 3, 2011. These are sharp price oscillations that quickly recover to the original price. The phenomenon of MFCs has received little attention by researchers and yet it has deep repercussions not just on traders that are caught trading on the wrong side of the book during the MFC, but also on path dependent options and brokers trading at the VWAP. The mandate of this paper is to show that MFCs have a statistically significant impact on capital markets by investigating the attributes and implications of MFCs, and that therefore this phenomenon should be comprehensively addressed by regulators.

1 Introduction

As we approach the third decade of the XXI century, financial markets are changing faster than ever before. The ever changing nature of financial markets is driven partly by advances in technology and partly by the new regulations that are necessary to keep up to speed with the evolving technology. Many argue that technology has improved markets, for example by improving the speed at which transactions take place. However, it is also true that markets have become somewhat less stable and more susceptible to spikes in volatility that can very quickly wipe out significant sums of money, the prime example being what is known as the May 2010 Flash Crash, or simply The Flash Crash. On May 6, 2010, the Dow Jones Industrial Average lost almost 9% of its value leading to a trillion-dollar stock market crash in the U.S. This, however, was no ordinary crash as the DJIA recovered within approximately 36 minutes, and it is precisely this type of phenomenon that we study in this paper, that is crashes that recover quickly to their starting price. In particular, while "major" flash crashes, such as the May 2010 flash crash or the 2013 fake-news-induced crash or the British pound flash crash, have been widely investigated, we focus on "mini" flash crashes.

We define a mini flash crash (MFC) as an event that meets the following criteria: the price change has to be larger than 0.8% in less than 1.5 seconds for at least 10 consecutive trades before recovering to within 0.1% of the starting price. This definition is consistent with that used by Nanex, one of the first firms to study the concept of a mini flash crash, and we notice that a MFC may be an upward spike (also known as a "rally") or a downward spike (also known as a "crash"). We include in our sample all mini flash crashes that characterised the U.S. equity markets between January 3, 2006 and February 3, 2011.

This is an empirical paper which studies mini flash crashes (MFCs) empirically. We are not concerned with what causes MFCs from a theoretical point of view, neither are we interested in the theoretical implications of this phenomena. Papers such as Li, Wang, and Ye (2017) deal with these issues. Instead, we pursue two studies: firstly, we investigate the "empirical" drivers of MFCs; and secondly, we investigate the "empirical" implications of MFCs. The first study involves analysing the impact a stock's risk and return have on the frequency, amplitude and distribution

of MFCs where comparisons are drawn between daily and intraday distributions. Since correlation does not imply causation, we do not look at the "empirical" causes of MFCs but rather at what correlates with MFCs, namely their attributes. The second study involves investigating the impact MFCs have on market quality such that, so armed, we can decide whether the marketplace should be left to its own devices, as was preached in the lead up to the Global Financial Crisis, and therefore we should consider MFCs as the inevitable consequence of letting price be determined by the intersection of the supply and demand for liquidity, or whether the phenomenon known as a mini flash crash is in fact pernicious and detrimental to the wellbeing of financial intermediaries. We pursue this issue by using market quality as a proxy for the "empirical" welfare of the social planner.

The findings of this paper are summarised as follows. With regards to the first study relating to the attributes of MFCs, the findings of our analysis are three-fold. First, addressing the risk-return impact on the frequency of MFCs we find that both risk and return affect the frequency of MFCs since stocks with negative returns and higher volatility exhibit more MFCs. We also find that the intraday frequency of MFCs is U-shaped and has gradual increments with the exception of the first half hour of trading. Second, addressing the risk-return impact on the amplitude of MFCs we find that only risk affects the amplitude of MFCs since stocks with higher volatility exhibit greater MFCs. Third, addressing the risk-return impact on the distribution of MFCs we find that only risk affects the distribution of MFCs since days with higher volatility exhibit greater likelihood of MFCs. We also find that MFCs are persistent across days. With regards to the second study relating to the implications of MFCs on market quality, the findings of our analysis are again three-fold. First, we find that inside spread increases both during the MFC itself and over the 10 minutes around the MFC so that MFCs lead to a widening of inside spread. Second, we find that BBO depth increases during the MFC itself but is decreasing over the 10 minutes around the MFC so that MFCs lead to a thinning of BBO depth. Third, we find that both average trade size and intraday volatility increase during the MFC itself while they remain unchanged over the 10 minutes around the MFC so that MFCs have no permanent effect on average trade size and intraday volatility.

In summary, this manuscript provides insights into all mini flash crashes (MFCs) that took

place on US equities in the 5 year period that spans January 3, 2006 to February 3, 2011. A MFC is a sudden change in the stock price that quickly recovers to its original price. More formally, a MFC arises when the stock price changes by at least 0.8% in less than 1.5 seconds for at least 10 consecutive trades and recovers to within 0.1% of the starting price over this time horizon. These small (but recurrent - there is more than three a day on average in our sample) financial earthquakes have far-reaching repercussions on financial markets and their participants: they affect equity investors with the average MFC costing up to $ 2 million, they affect brokers trading at the VWAP who fail to take advantage of the MFC, and they affect option traders caught on the wrong side of the book. But above all, they signal a failure of the price in its primary function which is that to reflect the fundamental value of the underlying stock, since companies are made of people and capital, and neither can change so much so quickly only to recover to its starting value within fractions of a second. To study MFCs this paper looks at two sets of metrics: it studies how MFCs correlate with daily stock returns and volatility, and it investigates market quality, measured as inside spread and BBO depth, over the course of the MFC. It is found that MFCs are predominantly correlated with daily volatility and that they have a negative, and at times lasting, effect on market quality. The above captures our contribution to the Literature, and it is the above that leads us to believe that mini flash crashes ought to be further regulated since they often times slip under circuit breakers which are designed by Regulators for the bigger flash crashes. Furthermore, we also postulate that there may exist a link between the Regulation National Market System (Reg NMS), High Frequency Traders (HFTs) and MFCs. It is indeed possible that Reg NMS, in an attempt to enhance competition by pulling together trading venues, freed up HFTs, which in turn come at the expense of MFCs.

The paper is organised as follows. Sections 1 and 2 include an introduction and a literature review. Section 3 discusses the relevance, definition and dataset of MFCs. Section 4 analyses the attributes of MFCs in relation to the stock's risk and return, while Section 5 investigates the implication of MFCs on market quality. Section 6 presents the conclusions, and Section 7 gives the Appendix to the paper.

2 Literature Review

A survey of both the empirical and theoretical papers that relate to our study is given in Appendix 7.A. On the empirical side, we are interested in papers that describe the phenomenon of a mini flash crash, and on the theoretical side, we are interested in papers that give empirical predictions of the effects risk and return have on mini flash crashes as well as the effects mini flash crashes have on the market quality of the limit order book. However, neither is possible due to the limited research conducted in this area. Instead, we give an overview of those empirical papers that investigate fields connected to MFCs, as well as theoretical papers that give alternative empirical predictions on MFCs.

3 MFC Overview – Relevance, Definition and Dataset

3.1 The Relevance of a Mini Flash crash

Mini flash crashes (MFCs) despite taking place over a very short period of time can have deep repercussions on financial markets and their participants. A MFC is a sharp price movement which then recovers to the original price and if we believe that the primary responsibility of a price is to reflect the fundamental value of the stock, we immediately conclude that over those 1.5 seconds the price is failing in its primary duty since it no longer reflects the true value of the company, which cannot possibly jump for 1.5 seconds before recovering to its starting position. Besides the assets, capital and employees of the company are the same at the start of the MFC as they are at the end.

Therefore, MFCs represent a failure of the market and quantifying said failure is a rather challenging task. Nonetheless, we compute an upper-bound for the loss borne by all traders that are adversely affected by a MFC by measuring the loss from selling during a crash and buying during a rally. The loss over a single trade is given by multiplying the number of shares traded by the change in price, measured as the difference between the price at which the trade takes place and the price prior to the start of the MFC. We then add up all trades that take place during the MFC. Over our sample which excludes the financial crisis, an upper-bound for the average loss borne by all adversely affected traders during a single MFC is $ 2,344,326. However, the final number is

much greater when we include the impact MFCs have on options and on brokers trading at the volume-weighted average price (VWAP).

The greatest source of loss borne by brokers is the VWAP, and the problem arises when the broker's client demands the VWAP even when the broker was not able to take advantage of the MFC, as they were not actively trading during the MFC. Monitoring the limit order book is a cost for the broker as is trading at high frequency, and no matter how often the broker trades, there are going to be periods of inactivity during which a MFC could take place. Not taking advantage of a MFC is a daily struggle for brokers and financial intermediaries as the VWAP at which they trade is worse than what they can relay to their client, thus producing a monetary loss.

Furthermore, MFCs affect all path dependent options, such as Asian or Russian options, as well as options that expire during the MFC. Since investors leverage on options, changes in their price will have a significant impact on the investor's portfolio, and if we factor in that option prices are more volatile than stock prices in percentage terms, the final impact of MFCs on options will be significant. To put options into perspective, the Securities Industry and Financial Markets Association reported that the total amount of debt outstanding at the end of 2017 was more than $ 40.7 trillion, and according to a study from the Bespoke Investment Group, the total capitalisation of the Russell 3000 Index, which includes more than 98% of the market's current capitalisation, has recently reached $ 30 trillion. Since options are traded on all these (currency, bond, commodities, equities) markets as well as on unlisted OTC markets, some market analysts place the size of the derivatives market at more than 10 times that of the total world gross domestic product ($1 quadrillion). There are researchers that challenge these estimates, such as the Bank for International Settlements which found from its most recent data that the total notional amounts outstanding for contracts in the derivatives market is an estimated $ 542.4 trillion. However, by any calculation, the derivatives market is sizeable and significant in the overall picture of worldwide investments.

To conclude, an upper-bound for the loss incurred by market participants trading during a mini flash crash is $ 2,344,326 in our sample where we exclude the financial crisis, but the total loss is much greater if we account for options derived from the stock that crashed/rallied and for the losses sustained by brokers trading at the VWAP. Over the entire sample that spans from January

3, 2006 to February 3, 2011, up to $ 8,500,525,684 could be lost from trading during a MFC, and that number excludes the financial crisis, options and VWAP-trading.

3.2 Competing Definitions of a Mini Flash Crash

We define a mini flash crash as a phenomenon that meets the following criteria: the underlying price of the security fluctuates by more than 0.8% in less than 1.5 seconds for at least 10 consecutive trades before recovering to within 0.1% of the starting price. This definition is taken by Nanex, the firm that first studied this phenomenon. However, we may question the validity of this definition and whether a different definition would yield different results to the ones we put forward, thus compromising our analysis. We show our work is robust to the definition used by considering two competing definitions of a MFC: first, we consider using a different fixed threshold, and second, we consider using a volatility-dependent threshold.

Let us start by assessing whether the frequency and severity of MFCs is sensitive to our fixed threshold, namely 0.8%. To this end, we exploit the fact that returns are approximately normally distributed over short time horizons and we calculate the annualised volatility a stock must have for mini flash crashes on that stock to come from the centre of the distribution. If MFCs really did come from the centre of the distribution, then a small change in the chosen threshold would significantly alter the number of mini flash crashes included in our sample and thus the ensuing analysis. Setting the threshold to one standard deviation, and since the standard deviation of our distribution is the square root of the annualised variance divided by the number of trading days (252) and the number of seconds in a trading day (23,400) then multiplied by the maximum length of time over which a MFC takes place (1.5 seconds), we find that the stock requires an annualised volatility of approximately 1,600% for mini flash crashes on that stock to come from the centre of the distribution. This number is much greater than the volatility of some of the most volatile U.S. stocks, so we can safely conclude that the threshold is in the tail of the distribution and our analysis is not significantly affected by the chosen threshold. As a reference point, the average annualised volatility of the S&P 500 over the last 10 years is just under 20% which is 80 times smaller than the required volatility.

Now let us continue by assessing whether the frequency and severity of MFCs is sensitive to adopting a volatility-dependent threshold in place of our fixed threshold. One example of such a volatility-dependent threshold is 0.8% + k * VIX, where k is a constant. We note that under a fixed threshold definition (of e.g. 0.8%), we do not include mini flash crashes that are just below the threshold on low volatility days (LVDs) but we do include mini flash crashes that are just above the threshold on high volatility days (HVDs). Mini flash crashes that are just below the threshold on LVDs have a large impact on market quality since our measures of market quality are a function of, and thus embed, the stock-specific daily volatility (this is a well-known empirical fact e.g. Budish, Cramton, and Shim (2015)), so the overall effect on market quality would only be bigger if we included these mini flash crashes, making the analysis based on a fixed threshold more conservative than one based on a volatility-dependent threshold. Similarly, mini flash crashes that are just above the threshold on HVDs have a small impact on market quality, since the effect of a mini flash crash on market quality is net of volatility, and thus including these mini flash crashes dampens the overall effect on market quality, once again making the analysis based on a fixed threshold conservative as stronger results would ensue if we excluded these mini flash crashes. To conclude, the exclusion of "large" mini flash crashes on LVDs and the inclusion of "small" mini flash crashes on HVDs make our analysis more conservative than if we adopted a volatility-dependent threshold.

3.3 Dataset, Sample Selection and Descriptive Statistics

The dataset was originally inspired by Nanex, a firm that offers streaming market data services and real time analysis for financial markets. It consists of all mini flash crashes that according to Nanex took place in the U.S. market between January 3, 2006 and February 3, 2011. The three main American stock exchanges involved are the New York Stock Exchange (NYSE), the National Association of Securities Dealers Automated Quotations (NASDAQ) and the Archipelago Exchange (NYSE ARCA). We therefore use as a reference the list of MFCs identified by Nanex for the years 2006-2011, and the data was manually downloaded from the Nanex website, where they appear

by stock.[1] We then select the 100 stocks with the most MFCs having filtered our data to drop illiquid stocks, stocks that went bust before 2011 and stocks that were listed after 2006. Finally, for each mini flash crash we download from Thomson Reuters Tick History, time and sales (quotes and trades) high frequency data (down to microsecond) for the day of the mini flash crash, a day before and a day after. We then checked the data downloaded from Nanex by scanning our sample for MFCs. To this point, we ensure 4 conditions are satisfied: at least 10 consecutive trades take place with at least 10 price movements; the price changes by at least 0.8%; the time interval for the mini flash crash and recovery is at most 1.5 seconds; and the recovery price is within 0.1% of the starting price. We found some inconsistencies with the Nanex dataset and therefore we rectified our dataset. Hence the dataset from Nanex was only used as a starting point but is then verified and checked using TRTH intraday data.

Before proceeding with the analysis, we discard MFCs that take place within 10 minutes of each other, equities that are not ordinary stocks or that are not continuously traded over the sample period, and we also remove the Global Financial Crisis (GFC), a period which is structurally different from what precedes and follows the crisis. To identify the crisis we use the VIX index. Data on VIX is made available by the Chicago Board Options Exchange (CBOE) and is presented in Appendix 7.E Figure 2.7 with the GFC highlighted. Thus to account for the financial crisis, we omit transactions that take place between August 29, 2008 (1 month before September 29, 2008, which is the date known as the GFC) and December 29, 2008 (3 months after September 29, 2008). Finally, keeping the 100 stocks with the most MFCs leaves 3,626 mini flash crashes in our sample. The stocks included in our sample are tabulated in Tables 2.1-2.2 which report the first and second respectively 50 stocks in alphabetical order together with some descriptive statistics.

As further descriptive statistics: a time-series plot of the number of mini flash crashes in the selected sample over the period of interest is provided in Figure 2.1; tabulated frequencies of the number of mini flash crashes are reported in Table 2.3; and a breakdown of mini flash crashes by stock exchange is given in Figure 2.2.

[1] http://www.nanex.net/FlashCrashEquities/FlashCrashAnalysis_Equities.html

Table 2.1: Sample I/II

	Stock	RIC	Close (USD)	Volume	Mkt Cap (USD)	NoC
1	Aluminium Corp. of China Limited	ACH	16.50	166,057	2,035,020,000	20
2	Aecom	ACM	22.90	707,280	1,948,270,000	23
3	Aflac Inc.	AFL	48.49	1,681,329	23,890,950,000	54
4	Assured Guaranty Ltd.	AGO	25.24	268,790	1,722,000,000	20
5	American International Group Inc	AIG	1125.21	452,300	168,922,100,000	211
6	AK Steel Holding Corporation	AKS	7.58	2,487,167	796,100,000	29
7	Anadarko Petroleoum Corporation	APC	44.55	4,629,914	21,187,100,000	16
8	Bank of America Corp	BAC	45.38	12,606,514	175,181,300,000	94
9	Banco Bradesco SA	BBD	4.99	7,461,157	13,316,610,000	20
10	BB&T Coproration	BBT	43.07	1,511,886	23,096,380,000	18
11	Boardwalk Pipeline Partners, LP	BWP	18.76	682,380	1,293,450,000	24
12	Blackstone Group LP	BX	29.38	14,724,495	8,227,440,000	54
13	Beazer Homes USA, Inc	BZH	327.53	344,690	2,508,590,000	141
14	Morgan Stanley China A Share Fund	CAF	20.40	88,364	263,090,000	42
15	Chicago Bridge & Iron Company N.V.	CBI	24.23	790,405	2,079,890,000	20
16	Cheasapeake Energy Corporation	CHK	27.32	14,841,695	10,719,820,000	17
17	Cliffs Natural Resources Inc	CLF	22.36	2,766,514	1,774,490,000	45
18	Chipotle Mexican Grill	CMG	44.60	968,500	346,670,000	15
19	CNH Global NV	CNH	17.08	61,705	2,414,070,000	53
20	Capital One Financial Corp	COF	80.87	2,732,071	20,206,430,000	35
21	CVR Energy	CVI	22.08	787,948	1,653,240,000	31
22	Chevron Corporation	CVX	57.74	8,645,976	129,400,800,000	19
23	Concho Resources Inc	CXO	12.83	502,881	935,250,000	18
24	Deutsche Bank	DB	74.34	198,914	50,757,880,000	12
25	Diana Shipping Inc	DSX	14.35	220,290	636,000,000	16
26	Morgan Stanley Emerging Markets	EDD	21.32	139,935	150,360,000	26
27	New Oriental Education & Tech Grp	EDU	6.08	2,720,365	736,600,000	15
28	E-House (China) Holdings Limited	EJ	16.23	1,549,550	1,449,630,000	22
29	Wells Fargo Global Dividend Opp.	EOD	20.20	133,936	870,100,000	39
30	Eversource Energy	ES	18.28	814,529	2,347,760,000	47
31	Eaton Vance	ETJ	19.60	244,864	1,300,100,000	19
32	Energy transfer Partners LP	ETP	6.39	183,400	668,050,000	52
33	Freeport-McMoRan inc	FCX	25.61	4,372,029	9,204,080,000	26
34	Fortress Investment Group	FIG	28.78	3,525,800	2,770,070,000	28
35	Greenberier Companies Inc	GBX	26.64	295,471	430,320,000	34
36	Genesco Inc	GCO	35.20	187,395	840,960,000	33
37	The Gabelli Dividend and Income Trust	GDV	16.73	204,605	8,820,000	18
38	Greif Inc	GEF	30.24	83,352	688,070,000	15
39	Greenhill & Co., Inc.	GHL	52.55	175,733	1,423,970,000	35
40	Goldman Sachs Group Inc	GS	130.30	2,909,100	57,071,120,000	36
41	Genesee & Wyoming Inc	GWR	22.59	193,833	856,830,000	24
42	Halliburton Company	HAL	30.61	15,335,838	30,790,910,000	37
43	The India Fund, Inc.	IFN	33.36	304,086	1,196,770,000	34
44	Interoil Corporation	IOC	19.55	182,600	552,200,000	8
45	JA Solar Holdings Co., Limited	JASO	30.63	882,090	474,670,000	26
46	JPMorgan Chase & Co.	JPM	38.00	11,322,476	128,940,400,000	35
47	KKR Financial Holdings LLC	KFN	22.67	397,681	1,796,360,000	48
48	Kayne Anderson Energy	KYE	21.86	143,843	671,700,000	28
49	Kayne Anderson MLP Investment Co.	KYN	25.23	134,971	29,960,000	33
50	Lazard Ltd.	LAZ	29.12	388,705	976,870,000	16

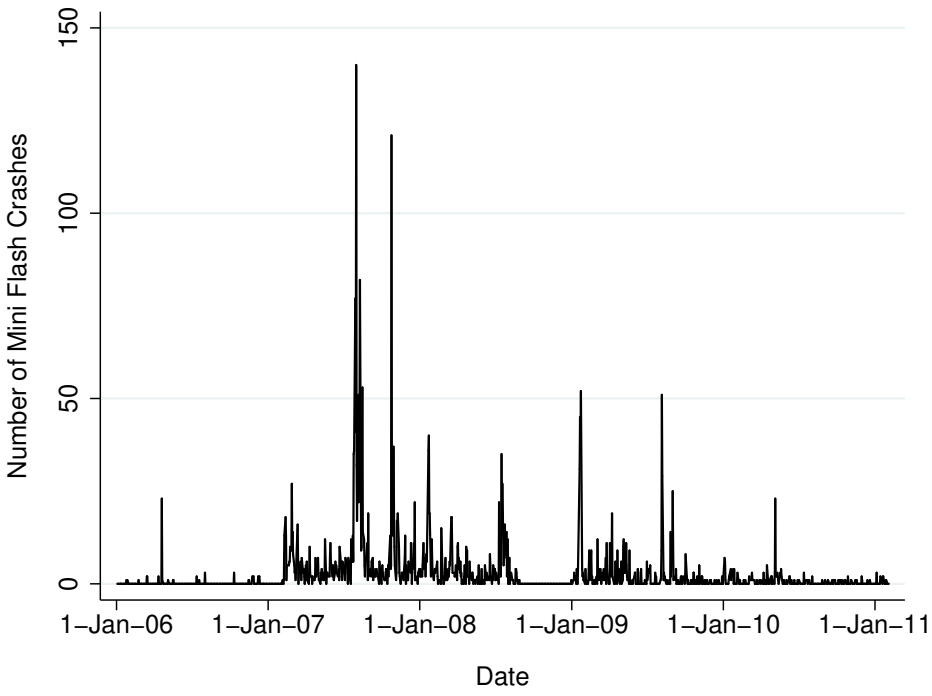
This table reports the first 50 stocks in our sample in alphabetical order together with four descriptive statistics for each stock: the average closing price in USD and volume over November 2005; the market capitalisation in USD on November 1, 2005; and the number of times the stock had a mini flash crash over the sample period (NoC).

Table 2.2: Sample II/II

	Stock	RIC	Close (USD)	Volume	Mkt Cap (USD)	NoC
51	Life Time Fitness, Inc.	LTM	38.79	156,833	1,295,780,000	40
52	Macy's, Inc.	M	33.17	5,300,533	16,085,600,000	16
53	Mastercard Inc	MA	4.62	47,179,562	3,450,830,000	24
54	MBIA, Inc.	MBI	61.86	1,168,081	7,707,220,000	73
55	Moody's Corporation	MCO	56.76	1,146,667	15,499,630,000	12
56	Marcus Corp	MCS	23.98	70,538	465,680,000	29
57	MGM Resorts international	MGM	39.04	2,247,995	10,557,060,000	41
58	Monsanto Company	MON	34.86	3,943,124	17,064,840,000	19
59	Mosaic Co	MOS	13.28	911,976	5,160,850,000	17
60	Merck & Co., Inc.	MRK	29.85	10,971,657	62,362,700,000	31
61	Morgan Stanley	MS	45.64	6,469,048	57,162,140,000	26
62	MGIC Investment Corp	MTG	63.95	752,905	5,414,820,000	37
63	MTL	MTL	19.25	186,100	4,166,870,000	40
64	Allianzgi NFJ Dividend int & Prm Stgy Fd	NFJ	21.80	341,095	366,730,000	23
65	Northstar Realty Finance Corp	NRF	15.18	22,329	199,470,000	63
66	NuStar GP Holdings, LLC	NSH	21.49	610,773	935,000,000	28
67	Novo Nordisk	NVO	5.36	297,286	15,655,460,000	15
68	Plains All American Pipeline, L.P.	PAA	20.23	255,210	2,827,980,000	53
69	PNC Financial Services Group Inc.	PNC	62.85	1,292,548	17,604,220,000	32
70	Potash Corporation of Saskatchewan	POT	8.85	7,031,871	8,683,390,000	18
71	Par Pharmaceutical Companies, Inc.	PRX	25.81	528,976	863,520,000	111
72	Prudential Public Limited Companies, Inc.	PUK	18.10	48,819	20,848,240,000	22
73	Transocean LTD	RIG	64.82	4,912,119	18,475,790,000	12
74	Rambus Inc	RMBS	14.37	1,778,752	1,280,610,000	45
75	Redwood Trust, Inc.	RWT	44.38	343,167	1,131,480,000	59
76	Sears Holding Corp	SHLD	81.95	4,026,852	19,496,450,000	18
77	ProShares UltraShort Financials	SKF	1100.74	1,168	15,430,000	17
78	SL Green Realty Corp	SLG	70.33	177,048	2,829,230,000	43
79	SanDisk Corporation	SNDK	56.62	14,816,695	10,921,690,000	14
80	Seaspan Corporation	SSW	19.90	170,829	555,110,000	22
81	State Street Corp	STT	56.97	1,616,971	18,011,820,000	47
82	SolarWinds Inc	SWI	14.29	1,060,141	882,560,000	18
83	Systemax Inc.	SYX	6.43	32,871	239,660,000	23
84	Target Corporation	TGT	55.93	5,846,590	48,618,820,000	12
85	Terra Nitrogen Company	TNH	24.35	78,376	453,110,000	18
86	Trina Solar Limited	TSL	9.55	1,291,580	430,240,000	37
87	Total System Services, Inc.	TSS	21.64	107,290	4,160,980,000	15
88	UIL Holdings Corporation	UIL	29.26	68,895	710,490,000	13
89	U.S. Bancorp	USB	29.96	4,290,557	53,327,870,000	60
90	Universal Corp	UVV	38.54	153,286	945,030,000	15
91	VMware Inc	VMW	67.83	5,982,617	3,831,120,000	42
92	Vanda Pharmaceuticals	VNDA	10.36	206,159	209,510,000	112
93	Verizon Communications Inc	VZ	28.17	12,268,167	86,654,190,000	13
94	WellCare Health Plans, Inc	WCG	37.41	423,767	1,247,840,000	190
95	Wells Fargo & Co	WFC	30.99	9,612,619	100,150,600,000	38
96	Willbros Group Inc	WG	15.54	151,590	323,060,000	16
97	Wipro Limited	WIT	2.92	961,452	14,449,070,000	52
98	Waste Management, Inc.	WM	30.25	1,872,114	16,174,310,000	59
99	Aqua America Inc	WTR	20.91	664,448	3,194,310,000	39
100	Yingli Green Energy Holdings Co Ltd	YGE	130.54	156,705	1,332,690,000	26

This table reports the second 50 stocks in our sample in alphabetical order together with four descriptive statistics for each stock: the average closing price in USD and volume over November 2005; the market capitalisation in USD on November 1, 2005; and the number of times the stock had a mini flash crash over the sample period (NoC).

Figure 2.1: Mini Flash Crashes over Sampling Period



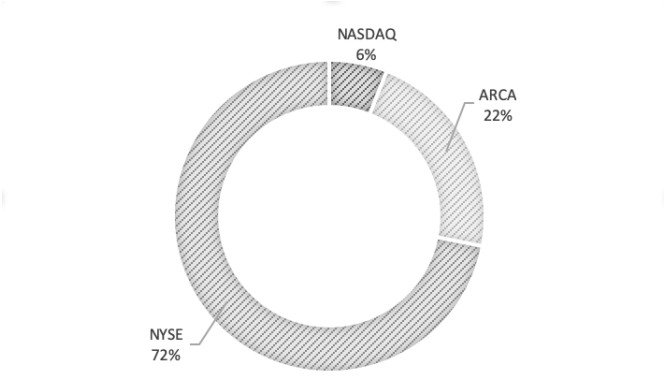
This figure exhibits the daily occurrence of mini flash crashes that took place on stocks in our sample (listed in Tables 2.1-2.2) over the sampling period (January 3, 2006 - February 3, 2011). The flat line at 0 between August 29, 2008 - December 29, 2008 is due to the fact that the we removed the Global Financial Crisis from our sample.

Table 2.3: Mini Flash Crashes by Tabulated Frequency

MFC	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	22
Frequency	196	110	72	53	34	29	19	8	9	13	10	8	7	5	1	4	2	3	7	3
MFC	23	25	27	29	31	32	35	36	37	40	41	45	48	51	52	53	77	82	121	140
Frequency	2	1	2	2	1	1	3	3	1	1	1	1	1	2	1	1	1	1	1	1

This table reports the occurrence of the number of mini flash crashes, for example the second entry asserts that there are 110 days with 2 mini flash crashes happening on that day across all stocks.

Figure 2.2: Mini Flash Crashes by Stock Exchange



This figure exhibits the share of each stock exchange where the security was traded when the mini flash crash took place: thus 72% of the MFCs in our sample took place on NYSE.

4 MFC Attributes – Risk and Return

This section studies the attributes of mini flash crashes, and in particular the impact of a stock's risk and return on the frequency, amplitude and distribution of its MFCs where comparisons are drawn between daily and intraday distributions.

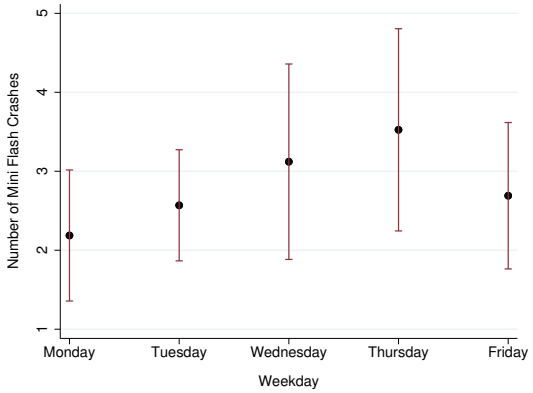
4.1 Frequency of MFCs

Prior to investigating the impact of a stock's risk and return, we start by analysing the daily and intraday distributions of the frequency of MFCs. The number of mini flash crashes is broken down by the day of the week in Figure 2.3a. The latter displays the average number of MFCs per weekday together with a 95% confidence interval. We observe that there are fewer MFCs per weekday at the beginning of the week on Monday, and that the number gradually increases on Tuesday and Wednesday to peak on Thursday, before falling back on Friday. However, since all confidence intervals overlap, no day of the week is statistically different from the others. Turning our attention to the intraday distribution in Figure 2.3b which gives a half hourly breakdown of the frequency of MFCs, we observe that the most prolific period is the first half hour of trading as investors trade frenetically to adjust their positions after the markets were closed between 16:00 and 9:30 local time. We notice that the average number of MFCs per time period gradually falls until 12 noon when it starts rising again thus producing a U-shape. It is a well-know empirical fact

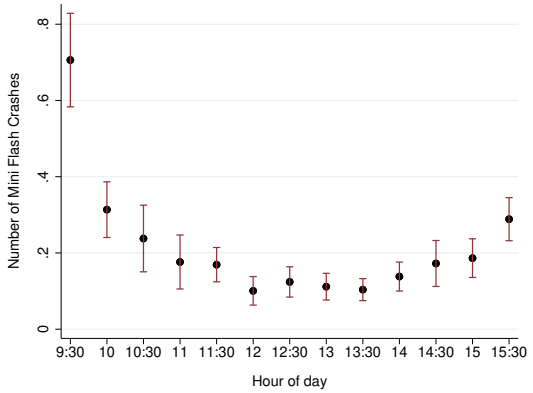
that intraday volatility is also U-shaped thus we conclude that the frequency of MFCs is positively correlated with intraday volatility. Moreover, we observe that all adjacent 95% confidence intervals overlap except for first one (trading between 9:30 and 10:00) which has (0.7) more than double the number of MFCs than the second busiest period (0.3 mini flash crashes on average between 10:00 and 10:30) indicating that MFCs cluster in the first half hour of continuous trading.

Figure 2.3: Frequency of MFCs – Daily and Intraday Distributions

(a) Daily Frequency. This figure shows the weekday breakdown of the frequency of mini flash crashes that took place on stocks in our sample. Thus on average there are 2.14 mini flash crashes on a Monday. The red line gives a 95% confidence interval.



(b) Intraday Frequency. This figure shows the half-hourly breakdown of the frequency of mini flash crashes that took place on stocks in our sample. Thus on average there are 0.7 mini flash crashes between 9:30 and 10:00. The red line gives a 95% confidence interval.



This figure gives a breakdown of the frequency of mini flash crashes by the day of the week (daily) and the half hour of the day (intraday) when the MFC took place.

In order to address the impact of a stock's risk and return on the frequency of MFCs, we run the dynamic linear model with the following specification for stock $i \in \{1, 2, ..., 100\}$ and time (measured in trading days) $t \in \{2, 3, ..., 1282\}$:[2]

$$\text{NMFC}_{it} = \beta_0 + \beta_1 * \text{NMFC}_{i,t-1} + \beta_2 * \pi_{it} + \beta_3 * \sigma_{it} + \epsilon_{it} \tag{1}$$

Notice daily volatility is used as a proxy for risk and is measured by the daily Realised Volatility

[2] $\text{NMFC}_{it} \in \{0 \cup \mathbb{N}\}\ \forall\ i, t$ is the daily Number of Mini Flash Crashes i.e. the number of times stock i had a mini flash crash on day t;
$\beta_i \in \mathbb{R}\ \forall\ i, t$ and $|\beta_1| < 1$;
$\pi_{it}\ \forall\ i, t$ is the daily return of stock i on day t;
$\sigma_{it}\ \forall\ i, t$ is the daily volatility of stock i on day t; and
$\epsilon_{it}\ \forall\ i, t$ is the error/shock of stock i on day t.

metric discussed in Appendix 7.F Table 2.8. Furthermore, we compute the optimal two-step generalised method of moments (GMM) estimator of Equation (1) and estimates are reported in Table 2.4. The econometrics behind our estimates are discussed in Appendix 7.B. From Table 2.4, we find that the daily stock-specific return, π_{it}, is partly negatively correlated with the number of MFCs if we hold the lagged daily number of MFCs and the daily stock volatility fixed since $\beta_2 = -71.1\%$ and its p-value is significant at the 5% significance level. This implies that the number of MFCs is greater on days with negative returns all else being equal. Turning our attention to the daily stock-specific volatility, σ_{it}, we deduce the latter is partly positively correlated with the number of MFCs since β_3 is positive and statistically significant. Therefore all else being equal the number of MFCs is greater on days with higher low-frequency daily volatility, and recall from Section 2.2 that this result is non-trivial since the MFC-induced high-frequency volatility does not necessarily have a significant impact on daily volatility.

Table 2.4: Frequency of MFCs – Risk and Return

$$\mathbf{NMFC_{it} = \beta_0 + \beta_1 * NMFC_{i,t-1} + \beta_2 * \pi_{it} + \beta_3 * \sigma_{it} + \epsilon_{it}}$$

Frequency of MFCs	**Variable**	**Beta**	**Coef.**	**Std.Err..**	**p-value**
Lagged Daily Number of MFCs	$\text{NMFC}_{i,t-1}$	β_1	0.259	0.003	0.000***
Daily Stock Return	π_{it}	β_2	-0.711	0.046	0.000***
Daily Stock Volatility	σ_{it}	β_3	2.718	0.048	0.000***
Constant	1	β_0	-0.090	0.003	0.000***

* if p-value < 0.1; ** if p-value < 0.05; *** if p-value < 0.01

MFC Attribute 1. *To summarise our findings on the frequency of mini flash crashes, we conclude that both risk (measured as daily stock volatility) and return (measured as daily stock return) are respectively positively and negatively correlated with the frequency of mini flash crashes. Therefore, we expect to observe a greater number of mini flash crashes on days with negative returns and high volatility. Furthermore, comparing the daily and intraday distributions of the frequency of mini flash crashes, we find that while the day of the week does not have a statistically significant impact on MFCs, the hour of day does: MFCs are U-shaped over a day of trading and cluster in the first half hour of trading.*

4.2 Amplitude of MFCs

Having studied the frequency at which mini flash crashes occur, we turn our attention to their amplitude. The amplitude of mini flash crash k is defined as:

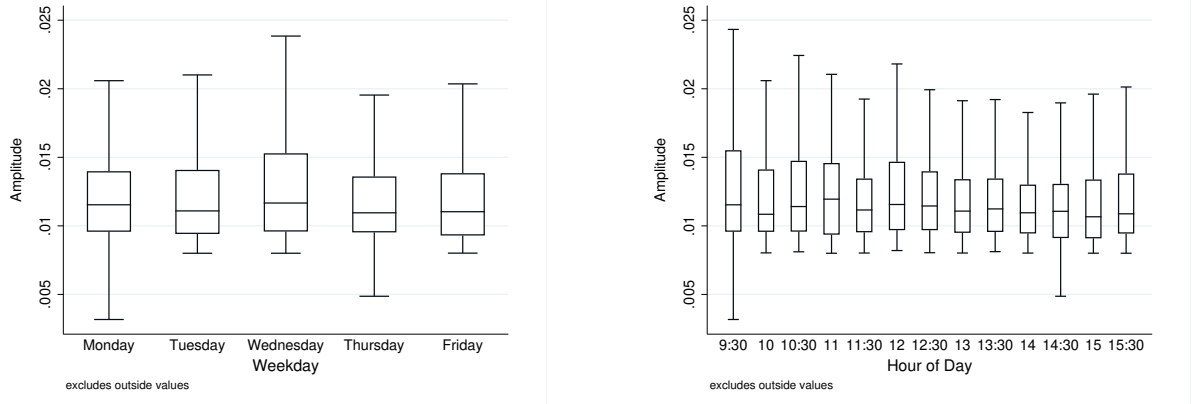
$$\text{Amplitude}_k = \left| \frac{\text{Extremum_Price}_k - \text{Start_Price}_k}{\text{Start_Price}_k} \right| \tag{2}$$

for $k \in \{1, 2, ..., 3626\}$ and where Start_Price_k is the price of the stock at the start of mini flash crash k and Extremum_Price_k is the price of the stock that is furthest away in absolute terms from Start_Price_k.

Figure 2.4: Amplitude of MFCs – Daily and Intraday Distributions

(a) Daily Amplitude. This figure shows the weekday breakdown of the amplitude of mini flash crashes that took place on stocks in our sample and is given in the form of a box-and-whicker plot. Thus the first box-and-whisker plot is for all mini flash crashes that took place on a Monday.

(b) Intraday Amplitude. This figure shows the half-hourly breakdown of the amplitude of mini flash crashes that took place on stocks in our sample and is given in the form of a box-and-whicker plot. Thus the first box-and-whisker plot is for all mini flash crashes that took place between 9:30 and 10:00.



This figure gives a breakdown of the amplitude of mini flash crashes by the day of the week (daily) and the half hour of the day (intraday) when the MFC took place.

Prior to investigating the impact of a stock's risk and return, we start by analysing the daily and intraday distributions of the amplitude of MFCs. Figure 2.4a gives a breakdown of the amplitude of mini flash crashes by the day of the week in the form of a box and whisker plot. Recall that the ends of the box are the upper and lower quartiles, so the box spans the interquartile range, that the median is marked by a vertical line inside the box, and that the whiskers are the two lines outside the box that extend to the highest and lowest observations. We observe that the median is

similar across all days of the week, and so is the range (defined as the highest observation minus the lowest observation) so we conclude that just as for the frequency of MFCs, their amplitude does not vary significantly across weekdays. Turning our attention to the intraday distribution in Figure 2.4b which gives a half hourly breakdown of the amplitude of MFCs, we observe that the range of MFCs is significantly higher in the first hour of trading compared to the rest of the day. This implies that between 9:30 and 10:00 we observe the widest range of MFCs, and combining this result with that on the frequency of MFCs, we conclude that not only do MFCs cluster in the first hour of trading but that this is also when they have the widest range of amplitudes across the whole trading day.

In order to address the impact of a stock's risk and return on the amplitude of MFCs, we run the classical cross-sectional linear regression model with the following specification for mini flash crash $k \in \{1, 2, ..., 3626\}$:[3]

$$\begin{aligned} \text{Amplitude}_k &= \beta_0 + \beta_1 * \text{NMFC}_k + \beta_2 * \pi_k + \beta_3 * \sigma_k + \epsilon_k \quad (3) \\ &= x_k'\beta + \epsilon_k \quad (4) \end{aligned}$$

Notice the data on (Amplitude_k, x_k') is assumed to be i.i.d. with $\text{E}[x_k\epsilon_k] = 0 \ \forall \ k$ and both the 4×4 matrices $\text{E}[x_k x_k']$ and $\text{E}[\epsilon_k^2 x_k x_k']$ are assumed to exist and be non-singular. Furthermore, since the dataset is large we can exploit the asymptotic distribution of the OLS estimator to conduct hypothesis tests on the coefficients in Equation (3) which together with their estimates are reported in Table 2.5.[4] The daily stock return has a coefficient of $\beta_2 = -0.3\%$ which is not significant at any reasonable significance level so we may conclude that the partial correlation between the amplitude of MFCs and their daily stock return which holds fixed the daily number of MFCs and the daily stock volatility is not statistically significant. Turning our attention to the daily stock volatility we infer that its coefficient is $\beta_3 = 0.5\%$ and that this is significant at the 5%

[3] NMFC_k, π_k, and σ_k are the daily number of MFCs, the daily stock return, and the daily stock volatility respectively of stock i and day t when mini flash crash k occurred;
$\beta = (\beta_0 \ \beta_1 \ \beta_2 \ \beta_3)' \in \mathbb{R}^4$; and
$x_k = (1 \ \text{NMFC}_k \ \pi_k \ \sigma_k)' \in \mathbb{R}^4$.

[4] The distribution of the OLS estimator is given by $\sqrt{3626}(\hat{\beta}_{OLS} - \beta) \xrightarrow{D} \text{N}(0, \text{avar}(\hat{\beta}_{OLS}))$ where $\hat{\beta}_{OLS}$ is the OLS estimator for the population parameter β and $\text{avar}(\hat{\beta}_{OLS})$ is the asymptotic variance of $\hat{\beta}_{OLS}$.

significance level. Therefore the amplitude of MFCs is partly positively correlated with the daily stock volatility indicating that the amplitude of MFCs is greater on days with higher short-term low-frequency daily volatility. In summary, the amplitude of MFCs is positively correlated with risk but is uncorrelated with returns, which adds to the results from Section 4.1 where we found that the frequency of MFCs is negatively correlated with returns and positively correlated with risk. We note that risk has a stronger impact on the frequency of MFCs than their amplitude since against the latter risk is not significant at the 1% significance level.

Table 2.5: Amplitude of MFCs – Risk and Return

$$\textbf{Amplitude}_{\boldsymbol{k}} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1} * \textbf{MFC}_{\boldsymbol{k}} + \boldsymbol{\beta_2} * \boldsymbol{\pi_k} + \boldsymbol{\beta_3} * \boldsymbol{\sigma_k} + \boldsymbol{\epsilon_k}$$

Amplitude of MFCs	**Variable**	**Beta**	**Coef.**	**Std.Err.**.	**p-value**
Daily Number of MFCs	MFC_k	β_1	-6.69E-05	2.83E-05	0.018**
Daily Stock Return	π_k	β_2	-0.003	0.002	0.155
Daily Stock Volatility	σ_k	β_3	0.005	0.002	0.013**
Constant	1	β_0	1.35E-02	2.98E-04	0.000***

* if p-value < 0.1; ** if p-value < 0.05; *** if p-value < 0.01

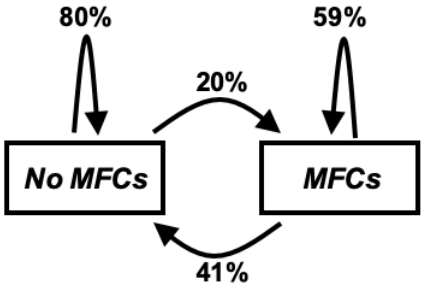
MFC Attribute 2. *To summarise our findings on the amplitude of mini flash crashes, we conclude that only risk (measured as daily stock volatility) and not return (measured as daily stock return) is correlated with the amplitude of mini flash crashes. In other words, we expect to observe greater mini flash crashes on days with high volatility. Finally, comparing the daily and intraday distributions of the amplitude of mini flash crashes, we find that while the day of the week does not have a statistically significant impact on MFCs, the hour of day does: MFCs in the first half hour of trading have a greater range of amplitudes (both big and small) when compared to the rest of the trading day.*

4.3 Distribution of MFCs

The final attribute to be investigated is the impact of a stock's risk and return on the distribution of MFCs. Prior to this we study the persistence of MFCs across trading days by analysing the conditional probability of observing one or more MFCs on any given day conditional on the previous day. The probabilities under examination are displayed in Figure 2.5, from which we conclude that

the persistence between days without any MFCs is of 80% and the persistence between days with at least one MFC is approximately 60%. In other words, given there are no MFCs today, we expect to see no MFCs tomorrow 8 times out of 10, while if there was one or more MFC today we expect a MFC tomorrow 60% of the time.

Figure 2.5: Distribution of MFCs – Daily Persistence



This figure exhibits the conditional probabilities for the two states "No MFCs" and "MFCs" where the former state indicates a day with no mini flash crashes across all stocks and the latter a day with at least one mini flash crash across all stocks. Probabilities are identified by arrows as is custom when dealing with Markov Probabilities and are defined as follows: $\Pr[TNMFC_t = 0 | TNMFC_{t-1} = 0] = 80\%$, $\Pr[TNMFC_t \geq 1 | TNMFC_{t-1} = 0] = 20\%$, $\Pr[TNMFC_t \geq 1 | TNMFC_{t-1} \geq 1] = 59\%$ *and* $\Pr[TNMFC_t = 0 | TNMFC_{t-1} \geq 1] = 41\%$, *where* $TNMFC_t$ *is the Total Number of MFCs across all stocks on day t.*

In order to address the impact of a stock's risk and return on the distribution of MFCs, we run the logistic regression with the following specification for time (measured in trading days) $t \in \{2, 3, ..., 1282\}$:[5]

$$\text{logit}(\mathbb{E}[\text{BMFC}_t]) = \beta_0 + \beta_1 * \pi_{t-1} + \beta_2 * \sigma_{t-1} + \beta_3 * \nu_{t-1} + \epsilon_{it} \tag{5}$$

The coefficients in Equation (5) are estimated by iteratively re-weighted least squares (IRLS) and are reported in Table 2.6. The daily return is found to have a p-value of 30.7% and therefore does not have a statistically significant impact on the probability of observing at least one MFC across

[5] BMFC_t is a Bernoulli random variable that equals 1 if $\text{TNMFC}_t \geq 1$ and 0 otherwise;
$\beta_i \in \mathbb{R} \; \forall \; i$;
$\pi_t \; \forall \; t$ is the daily return on day t (averaged across all stocks);
$\sigma_t \; \forall \; t$ is the daily volatility on day t (averaged across all stocks);
$\nu_t \; \forall \; t$ is the daily normalised volume on day t (averaged across all stocks) - we normalise by the daily executed volume of each stock on January 3, 2006 (the first day in the sample) so to make volume comparable across stocks; and $\epsilon_t \sim \text{Logistic}(0, 1)$ is the error/shock on day t assumed to be distributed according to the standard logistic distribution.

all stocks on any given day. Turning our attention to the daily volatility we find that its coefficient $\beta_3 = 4.641$ is both positive and significant at the 5% significance level thus implying that daily volatility is partly positively correlated with the probability of observing a MFC and consequently days with higher volatility have a greater likelihood of featuring a MFC holding volume and return fixed. Finally, to emphasise the role played by the logistic function we compute the fitted values which give the predicted unconditional probability of observing one or more MFCs on any given day:

$$\widehat{\Pr}[\text{TNMFC}_t \geq 1] = \frac{1}{1 + e^{-(-0.282 + 0.002 * \nu_{t-1} + 2.667 * \pi_{t-1} + 4.641 * \sigma_{t-1})}} \tag{6}$$

Table 2.6: Distribution of MFCs – Risk and Return

$$\boldsymbol{logit}(\mathbb{E}[\mathbf{BMFC}_t]) = \boldsymbol{\beta_0 + \beta_1 * \nu_t + \beta_2 * \pi_t + \beta_3 * \sigma_t + \epsilon_t}$$

Distribution of MFCs	**Variable**	**Beta**	**Coef.**	**Std.Err.**.	**p-value**
Daily Volume	ν_t	β_1	0.002	0.003	0.494
Daily Return	π_t	β_2	2.667	2.613	0.307
Daily Volatility	σ_t	β_3	4.641	2.145	0.031**
Constant	1	β_0	-0.282	0.109	0.010***

* if p-value < 0.1; ** if p-value < 0.05; *** if p-value < 0.01

MFC Attribute 3. *To summarise our findings on the distribution of mini flash crashes, we conclude that only risk (measured as daily volatility) and not return (measured as daily return) is correlated with the distribution of mini flash crashes. In other words, we expect a greater likelihood of observing a mini flash crash on days with high volatility. Finally, a study of the conditional distribution of mini flash crashes shows there is persistence across days.*

5 MFC Implications – Market Quality

This section studies the implications of MFCs, namely their impact on market quality.

5.1 Methodology

In order to investigate the implications of MFCs on market quality we perform two separate studies. First, we test market quality during the mini flash crash itself against the 10-minute window

around the MFC by means of a 99% confidence intervals (Test 1 – MFC against 10-minute window); and second, we test market quality before the mini flash crash against the same metric after the MFC by means of a Welch test (Test 2 – pre-MFC against post-MFC). We start by breaking down the trading day in 10-minute time intervals around each mini flash crash (and ignore trades that take place outside that 10-minute window), specifically we take 5 minutes before the mini flash crash and 5 minutes after the mini flash crash. We then divide each 10-minute time interval in 20 mini time intervals (MTIs), each 30 seconds long. Finally, we average market quality over each mini time interval and across mini flash crashes. The measures of market quality used include time weighted percentage spread (TWPS), time weighted depth (TWD), average trade size (AVTS), and realised volatility (RV) to gauge inside spread, BBO depth, trading activity, and intraday volatility respectively as defined in Appendix 7.D Table 2.8. The ensuing analysis does not differentiate MFCs between rallies and crashes since the results are qualitatively equivalent.

The first test (MFC against 10-minute window) determines whether our measures of market quality are statistically different around a mini flash crash. In order to test for changes around the mini flash crash, or indeed changes away from the mini flash crash, we compute the 99% confidence interval around the mean and we do so for each measure of market quality. We use 99% confidence intervals (as opposed to 95% confidence intervals) as they will only detect extreme changes in our metrics.

The second test (pre-MFC against post-MFC) establishes whether our measures of market quality are affected by the MFC. This is done by comparing the metric before the MFC against

the same metric after the MFC in the Welch t-test defined as follows:[6]

$$\begin{aligned} B &= \{B_{10}, B_9, B_8, B_7, B_6, B_5, B_4, B_3, B_2, B_1\} \\ A &= \{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}\} \\ H_0 &: d \coloneqq \mu_b - \mu_a = 0 \\ H_1 &: d \neq 0 \\ T &\coloneqq \frac{(\bar{X}_b - \bar{X}_a) - d}{s.e.} \sim t_{df} \end{aligned} \tag{7}$$

This test assumes the two populations have unequal variances and are normally distributed, and that the two samples are independent with the degrees of freedom discussed in Appendix 7.C. Assessing the validity of the null hypothesis in the Welch t-test yields the result of the second test we make to investigate the implications of MFCs on market quality.

5.2 Market Quality

We study the implications of mini flash crashes on market quality by means of two tests on four measures of market quality. The metrics used are time weighted percentage spread (TWPS), time weighted depth (TWD), average trade size (AVTS), and realised volatility (RV); and the two tests involve testing each metric during the MFC against the 10-minute window, and testing the metric before the MFC against the same metric after the MFC. Results for the first and second tests are presented in Figure 2.6 and Table 2.7 respectively, with the former figure tabulated in Appendix 7.F Table 2.10. Furthermore, we breakdown our analysis by the stock exchange where the MFC took place and since both tests show that the results do not vary significantly by market, we report these results in Appendix 7.E Figures 2.9-2.10 and Appendix 7.F Table 2.9 for Tests 1

[6] $B_i = \frac{1}{J}\left[\sum_{j=1}^{J} \text{Metric}_j\right]$ where $\{1, ..., J\}$ is the set of all observations that take place in the *i-th* mini time interval prior to the MFC for $i \in \{1, ..., 10\}$. Sample B therefore comprises the 10 MTIs prior to (before) the MFC and each element B_i in sample B is the metric under consideration averaged over all observations in that MTI.

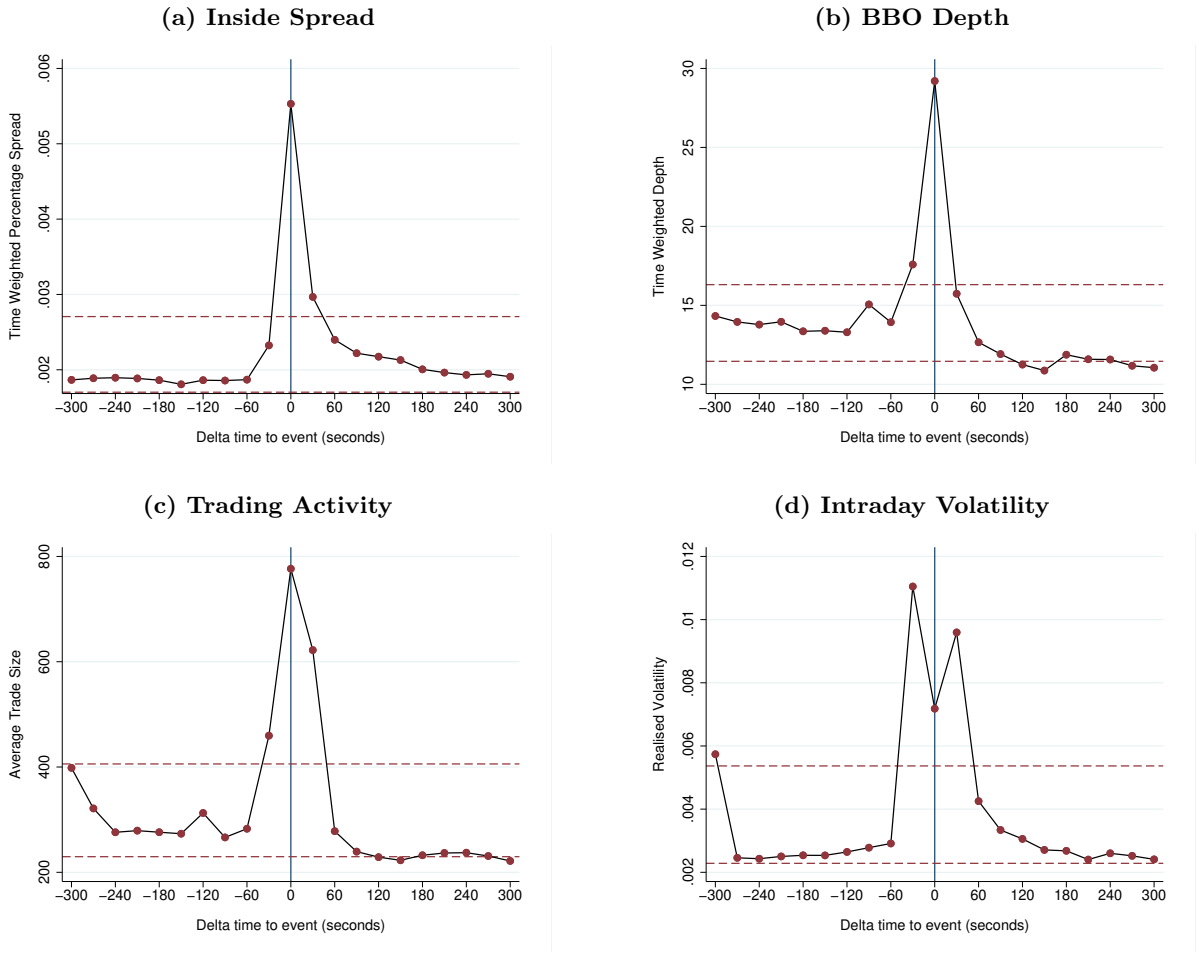
$A_i = \frac{1}{J}\left[\sum_{j=1}^{J} \text{Metric}_j\right]$ where $\{1, ..., J\}$ is the set of all observations that take place in the *i-th* mini time interval posterior to the MFC for $i \in \{1, ..., 10\}$. Sample A therefore comprises the 10 MTIs posterior to (after) the MFC and each element A_i in sample A is the metric under consideration averaged over all observations in that MTI.

H_0, μ_b, μ_a are the null hypothesis, and the population means of samples B and A respectively.

$\bar{X}_{b/a}$, $S_{b/a}$, $N_{b/a}$ are the mean, standard deviation, and size respectively of sample B/A; $s.e. \coloneqq \sqrt{\frac{S_b}{N_b} + \frac{S_a}{N_a}}$; and t_{df} is the Student's t-distribution with *df* degrees of freedom.

and 2 respectively.

Figure 2.6: Test 1 – MFC against 10-minute window



This figure exhibits each market quality metric as a time series plot over the 10-minute window around the MFC (highlighted by the vertical blue line) together with a 99% confidence interval (dashed red line). The x-axis is reported as number of seconds to the event (the mini flash crash): -300 is 5 minutes prior to the MFC and 300 is 5 minutes posterior to the MFC, for a total of a 10-minute window.

Out of the four market quality metrics we start by investigating inside spread through the time weighted percentage spread (TWPS). Figure 2.6a shows that the TWPS is outside the 99% confidence interval and therefore significant at the 1% significance level during the mini flash crash (which corresponds to mini time interval (MTI) E) so we conclude that during the MFC the TWPS is statistically higher than its 10-minute average around the MFC at 99% significance. Turning our attention to Test 2 in Table 2.7 we find that the p-value of the Welch t-test on TWPS is 4.1% and therefore significant at the 5% significance level. We may then reject the null hypothesis that

Table 2.7: Test 2 – pre-MFC against post-MFC

The Welch t-Test – $H_0 : d := \mu_b - \mu_a = 0$ VS $H_1 : d \neq 0$

	Group	Observations	Mean	Std.Err.	99% Conf.	Int.
Inside Spread	Before	10	0.0019	0.0000	0.0018	0.0021
	After	10	0.0022	0.0001	0.0018	0.0025
	Combined	20	0.0020	0.0001	0.0019	0.0022
	Difference	20	-0.0003	0.0001	-0.0006	0.0001
	t-statistic	-2.26	d.f.	13	Pr(|T|>t)	0.041**
	Group	**Observations**	**Mean**	**Std.Err.**	**99% Conf.**	**Int.**
BBO Depth	Before	10	14.26	0.405	12.95	15.58
	After	10	11.97	0.449	10.51	13.43
	Combined	20	13.12	0.395	11.99	14.25
	Difference	20	2.29	0.605	0.57	4.02
	t-statistic	3.80	d.f.	20	Pr(|T|>t)	0.001***
	Group	**Observations**	**Mean**	**Std.Err.**	**99% Conf.**	**Int.**
Trading Activity	Before	10	315	20	248	381
	After	10	275	39	149	401
	Combined	20	295	22	232	357
	Difference	20	40	44	-90	169
	t-statistic	0.90	d.f.	15	Pr(|T|>t)	0.382
	Group	**Observations**	**Mean**	**Std.Err.**	**99% Conf.**	**Int.**
Intraday Volatility	Before	10	0.0038	0.0009	0.0009	0.0066
	After	10	0.0036	0.0007	0.0013	0.0058
	Combined	20	0.0037	0.0005	0.0021	0.0052
	Difference	20	0.0002	0.0011	-0.0030	0.0034
	t-statistic	0.18	d.f.	19	Pr(|T|>t)	0.571

* if p-value < 0.1; ** if p-value < 0.05; *** if p-value < 0.01

This table reports the Welch t-Test of each market quality metric as a test between the 5 minutes that lead up to the mini flash crash (aka the event) and the 5 minutes that immediately follow the mini flash crash, therefore testing for a difference between the pre-MFC average (μ_b) and the post-MFC average (μ_a).

the 5-minute average before the MFC (μ_b) is not statistically different from the 5-minute average after the MFC (μ_a) and since the t-statistic is negative we conclude that MFCs lead to an increase and consequent worsening of inside spread. This leads to our first implication of MFCs on market quality:

MFC Implication 1. *Mini flash crashes lead to a widening of inside spread both during the MFC itself and over the 10-minute window around the MFC. This can be explained by liquidity being drained from all levels of the limit order book during the frenetic trading that takes place during a MFC.*

The second measure of market quality we investigate is best bid-offer depth through time weighted depth (TWD). Figure 2.6b shows that the TWD is outside the 99% confidence interval during the mini flash crash so we conclude that during the MFC the TWD is statistically higher than its 10-minute average. Turning our attention to Test 2 in Table 2.7 we find that the Welch t-test on TWD has a p-value of 0.1% and a positive t-statistic. We therefore reject that the two samples have the same population mean in favour of the hypothesis that the sample before the mini flash crash has a greater population mean than the one after the mini flash crash, thus establishing that MFCs lead to a decrease and consequent worsening of BBO depth. This leads to our second implication of MFCs on market quality:

MFC Implication 2. *Mini flash crashes lead to a deepening of BBO depth during the MFC and a thinning of BBO depth over the 10-minute window around the MFC. This can be explained as being a consequence of the wider spread that accompanies MFCs: BBO depth initially increases as the BBO moves to outer levels of the limit order book with deeper depth, but this is not sustainable as the time horizon widens so that over the 10-minute window as liquidity is drained from the LOB by the MFC with it comes a fall of BBO depth.*

The third measure of market quality we investigate is trading activity through average trade size (AVTS). We prefer this metric to daily executed volume since the latter is a cumulative metric which places great importance on the length of the MTI, whereas AVTS does not since it divides the number of traded shares by the number of trades. Figure 2.6c shows that the AVTS is outside

the 99% confidence interval during the mini flash crash so we conclude that during the MFC the AVTS is statistically higher than its 10-minute average. Turning our attention to Test 2 in Table 2.7 we find that the p-value of the Welch t-test on AVTS is not significant at the 5% significance level, and therefore we cannot reject the null hypothesis that observations prior to the MFC are not statistically different from those posterior to the MFC. We conclude that MFCs lead to no statistically significant impact on trading activity. This leads to our third implication of MFCs on market quality:

MFC Implication 3. *Mini flash crashes lead to a thickening of trading activity as measured by average trade size during the MFC and to no statistically significant change over the 10-minute window around the MFC. This can be explained by traders hectically trying to liquidate their positions during the MFC while this is no longer necessary once the MFC is over and normal market conditions are restored.*

The fourth and final measure of market quality we investigate is intraday volatility through realised volatility (RV). Figure 2.6d shows that the RV is outside the 99% confidence interval and therefore significant at the 1% significance level during the mini flash crash so we conclude that during the MFC the RV is statistically higher than its 10-minute average around the MFC at 99% significance. Turning our attention to Test 2 in Table 2.7 we find that the p-value of the Welch t-test on RV is not significant at the 5% significance level, and therefore we cannot reject the null hypothesis that observations prior to the MFC are not statistically different from those posterior to the MFC. We conclude that MFCs lead to no statistically significant impact on intraday volatility. This leads to our fourth and final implication of MFCs on market quality:

MFC Implication 4. *Mini flash crashes lead to a spike in intraday volatility during the MFC and to no statistically significant impact over the 10-minute window around the MFC. This can be explained by the very definition of a mini flash crash which we define as a surge of intraday volatility that quickly recovers to its original level.*

6 Conclusions

This paper studies all mini flash crashes (MFCs) that characterised U.S. financial markets between January 3, 2006 and February 3, 2011. The dataset is initially taken from Nanex, a firm that offers streaming market data services and real time analysis for financial markets, but is then independently checked by downloading high frequency data from Thomson Reuters. The definition used for a "mini flash crash" is that formulated by Nanex i.e. that the stock's price must change by at least 0.8% in less than 1.5 seconds for at least 10 consecutive trades and recover to within 0.1% of the starting price, and this definition is robust both to changes in the fixed threshold and to a volatility-dependent threshold. Finally, this is an empirical paper that aims at studying MFCs empirically. In particular, this manuscript is concerned with the "empirical" drivers and the "empirical" implications of MFCs, or in other words with what drives MFCs and with the implications of MFCs on financial markets. Since when studying the former we can only establish correlation which does not imply causation, we limit our analysis to the attributes of MFCs as well as their implications on the quality of financial markets. Topics for future research are discussed in Appendix 7.D.

The attributes of mini flash crashes can be summarised into three findings. First, in relation to the frequency of mini flash crashes, we find that both risk (measured as daily stock volatility) and return (measured as daily stock return) are respectively positively and negatively correlated with the frequency of mini flash crashes. Therefore, we expect to observe a greater number of mini flash crashes on days with negative returns and high volatility. Furthermore, comparing the daily and intraday distributions of the frequency of mini flash crashes, we find that while the day of the week does not have a statistically significant impact on MFCs, the hour of day does: MFCs are U-shaped over a day of trading and cluster in the first half hour of trading. Second, in relation to the amplitude of mini flash crashes, we find that only risk and not return is correlated with the amplitude of mini flash crashes. In other words, we expect to observe greater mini flash crashes on days with high volatility. Moreover, comparing the daily and intraday distributions of the amplitude of mini flash crashes, we find that while the day of the week does not have a statistically significant impact on MFCs, the hour of day does: MFCs in the first half hour of trading have a

greater range of amplitudes (both big and small) when compared to the rest of the trading day. Third, in relation to the distribution of mini flash crashes, we find that only risk is correlated with the distribution of mini flash crashes. In other words, we expect a greater likelihood of observing a mini flash crash on days with high volatility. Additionally, a study of the conditional distribution of mini flash crashes shows there is persistence across days.

The implications of mini flash crashes can be summarised into four findings. First, mini flash crashes lead to a widening of inside spread both during the MFC itself and over the 10-minute window around the MFC. This can be explained by liquidity being drained from all levels of the limit order book during the frenetic trading that takes place during a MFC. Second, mini flash crashes lead to a deepening of BBO depth during the MFC and a thinning of BBO depth over the 10-minute window around the MFC. This can be explained as being a consequence of the wider spread that accompanies MFCs: BBO depth initially increases as the BBO moves to outer levels of the limit order book with deeper depth, but this is not sustainable as the time horizon widens so that over the 10-minute window as liquidity is drained from the LOB by the MFC with it comes a fall of BBO depth. Third, mini flash crashes lead to a thickening of trading activity as measured by average trade size during the MFC and to no statistically significant change over the 10-minute window around the MFC. This can be explained by traders hectically trying to liquidate their positions during the MFC while this is no longer necessary once the MFC is over and normal market conditions are restored. Fourth, mini flash crashes lead to a spike in intraday volatility during the MFC and to no statistically significant impact over the 10-minute window around the MFC. This can be explained by the very definition of a mini flash crash which we define as a surge of intraday volatility that quickly recovers to its original level.

7 Appendix

7.A Literature Review

7.A.1 Empirical Literature Review

Mini flash crashes have received little attention by researchers, but there are however several streams in the empirical literature which are relevant to the analysis conducted in this study. First, our research relates to empirical papers that investigate the causes of intraday market crashes starting from the empirical evidence (e.g. the May 6, 2010 flash crash). Second, it is connected with trading curbs known as circuit breakers. Third, it is part of the growing literature on High Frequency Trading (HFT) as trading at extremely high speed has been acknowledged to be one of the causes of the increased frequency of intraday flash crashes. Fourth, it builds on papers that investigate intraday volatility given the close connection between MFCs and short-term volatility. Finally, due to the high degree of fragmentation that followed the implementation of Regulation National Market System (Reg NMS) our research also adds to the literature on the impact of market fragmentation on the quality of financial markets.

We begin with papers that look into the causes of mini flash crashes starting from the empirical evidence. The biggest crash ever recorded and the one that has drawn by far the most interest from researchers and academics is the May 6, 2010 flash crash, also known as the Flash Crash. Following the Flash Crash and ever since then, a lot of attention has been paid to these phenomena and while the causes of such an earthquake in financial markets is still disputed, the first reaction to the event was to blame high frequency traders (HFTs) for triggering it. Recently fingers have been pointed to a single trader, Navinder Singh Sarao. On April 21, 2015 the U.S. Department of Justice laid "22 criminal counts, including fraud and market manipulation" against Mr Sarao. Among the charges included was the use of a spoofing algorithm. Spoofing is a way to fool other market participants and to manipulate the price of the security. Before the Flash Crash, Mr Sarao placed orders in E-mini contracts (electronically traded futures contracts which track the S&P 500) betting on that the market would fall. The orders had a total value of $200 million and were replaced or amended 19,000 times before they were cancelled. Even though Mr Sarao's spoofing algorithm probably was

an important trigger, it is hard to give him full responsibility for the crash of a trillion $ market. It is reasonable to believe that a number of concomitant factors contributed to the crash. A number of empirical works were produced to explain the Flash Crash and to understand whether high frequency trading (HFT) was the cause of the crash and the discussion was broadened to the effect of HFT on market fragility. Kirilenko, Kyle, Samadi, and Tuzun (2017) study the May 6 U.S. flash crash by using audit-trail transaction level data for all regular transaction in the June 2010 E-mini S&P 500 futures contract during three days before the crash and conclude that even though HFTs did not trigger the Flash Crash, their responses to the unusually large selling pressure coming from fundamental traders on that day exacerbated market volatility. Chakravarty, Jain, Upson, and Wood (2012) investigate inter-market sweep orders (ISOs). These are orders that take liquidity by sweeping the entire book of several different markets. They found an increase in ISOs around the time of the Flash Crash and conclude that by aggressively draining liquidity from the bid side of the book HFTs may have contributed to the crash. McInish, Upson, and Wood (2014) too find that before the Flash Crash there was an unusually high level of ISOs activity by informed traders (i.e. HFTs). Easley, López de Prado, and O'Hara (2012) find that the order flow as captured by the VPIN metric became increasingly toxic immediately before the crash, thus exposing market makers to adverse selection on the day of the crash. They conclude that informed traders were responsible for the Flash Crash. Conversely, Menkveld and Yueshen (2016) conclude that an interruption in arbitrage activity (i.e. HFT) across markets originated the Flash Crash showing that evidence on what caused the Flash Crash is inconclusive and indeed that researchers are still unsure as to what sparked a trillion-dollar stock market crash in approximately 36 minutes.

Furthermore, our research relates to papers on circuit beakers. Trading venues that work like limit order books usually operate a price monitoring functionality, called circuit breaker, that automatically tracks trade prices and halts execution in the event of price tolerance levels being breached (Goldstein and Kavajecz (2004) and Goldstein (2015)). Circuit breakers aim at temporarily suspending trading under extreme market conditions to prevent excessive intraday price fluctuations. Morris (1990) and Subrahmanyam (1994) provide theoretical arguments in favour of coordinated circuit breakers in fragmented markets. In a similar vein, Chakrabarty, Corwin, and Panayides

(2011) show empirically that during NYSE trading halts off-NYSE executed volume significantly increases and contributes to the price discovery process on NYSE, which is however, an argument against coordinated circuit breakers. More recently, Cui and Gozluklu (2016) filter out a set of 54 intraday crashes at the firm level, involving only S&P 500 or Russell 1000 stocks, that triggered the new single stock circuit breaker (SSCB) introduced by the SEC for the U.S. market after the May 6 flash crash. They find that the extreme price changes that triggered the SSCB at the firm level are accompanied by a massive surge in volume, spread and short-term volatility, which gradually revert back to normal. They also find that speculative strategies of arbitrageurs such as momentum trading cause cross-sectional spillovers to related stocks. Clapham, Gomber, and Panz (2017) analyse volatility interruptions on Deutsche Bourse and find that, differently from the prevailing theoretical rationale, executed volume on alternative venues significantly decreases during circuit breakers on the main market and do not find any volatility spillover.

Moreover, our paper contributes to the burgeoning academic literature on high frequency trading (for a survey see Jones (2013), Biais and Foucault (2014), and O'Hara (2015)). This literature focuses both on the empirical study of the effect of high frequency trading on market quality (this literature is vast, including Hendershott, Jones, and Menkveld (2011), Hasbrouck and Saar (2013), Brogaard et al. (2014), Foucault, Kozhan, and Tham (2016), and Menkveld and Zoican (2014)), and on the theoretical discussion of the effects of HFT on welfare of market participants (see Biais, Foucault, and Moinas (2015)).

Additionally, the analysis in this paper adds to our understanding of intraday volatility given the close connection MFCs have with short-term high-frequency volatility. Bellia, Pelizzon, Subrahmanyam, Uno, and Yuferova (2017) and Boussetta, Lescourret, and Moinas (2017) recently showed that the first half hour of the continuous auction at the NYSE Euronext Paris exchange is characterised by the highest level of intraday volatility. It is also a well-known fact that empirical intraday volatility is U-shaped, therefore if mini flash crashes were positively correlated with volatility we would expect a similar U-shape intraday pattern.

Finally, our work is related to the literature that focuses on the debate over the fragmentation and consolidation of trading venues. As mentioned above, the level of fragmentation among both

U.S. and European trading venues has grown rapidly over the last decade since the introduction of Reg NMS and MiFID Level 1 in the U.S. and E.U. respectively. O'Hara and Ye (2011) as well as Gomber, Sagade, Theissen, Weber, and Westheide (2017) provide comprehensive overviews of the existing literature in this field. However, despite the most recent literature, such as Battalio (1997), Boehmer and Boehmer (2003), Foucault and Menkveld (2008), Hengelbrock and Theissen (2009), Bennett and Wei (2006), and O'Hara and Ye (2011), positively evaluates the effects of fragmentation on market quality and transaction costs, market fragmentation could interfere with the effectiveness of circuit breakers due to possible volume migration and volatility spillover if circuit breakers are not coordinated across all trading venues on which an instrument is traded.

7.A.2 Theoretical Literature Review

On the theoretical side too, mini flash crashes have received little attention. Additional to the lack of empirical predictions on the effect of a stock's risk and return on MFCs and in turn the effect of MFCs on the quality of financial markets, few models have been put forward to simulate and reproduce the behaviour of MFCs. We describe three different models which generate empirical predictions for MFCs and although either these predictions are not the ones we test empirically in the data or the models only indirectly relate to MFCs, they are the closest empirical predictions to the ones needed that are currently available in the literature.

Our first model, and the only one that was specifically designed to explain the behaviour of MFCs, is the model put forward by Li et al. (2017). This model extends Budish et al. (2015) along two dimensions: it allows non-HFTs (non- high frequency traders) to provide as well as demand liquidity; and it considers discrete price improvements to reflect the tick size (minimum price variation) imposed by the U.S. Security and Exchange Commission's Regulation National Market Systems Rule 612. The model makes three empirical predictions: first, non-HFTs are more likely than HFTs to supply liquidity at price levels that cross the midpoint, and these limit orders are more likely to be taken by HFTs; second, a mini flash crash is more likely to occur for stocks with smaller tick sizes and higher adverse selection risk; third, an upward (downward) mini flash crash is more likely to follow an initial price jump in the same direction.

The second model is that by Dugast (2018) who show theoretically that the effects of a liquidity shock is to dry up liquidity by reducing depth at the Best Bid Offer (BBO) and widening inside spread. Moreover, due to the sudden reaction of traders it will also be associated with an increase in executed volume. The intuition is that, conditional - for example - on a negative news, a limit seller will cancel her orders and opt for new market sell orders (volume) to take advantage of the current stale ask prices. Furthermore, Biais, Hombert, and Weill (2014) analyse optimal trading and equilibrium pricing under preference uncertainty, and consider an aggregate liquidity shock that affects traders' preferences. Following an aggregate positive liquidity shock, all investors switch to a high-valuation type, and since data collection and aggregation is time consuming, it will take time for some traders to revert to low-valuation and therefore after a liquidity shock we will observe first an increase in imbalance and then an increase in volume. So both Dugast (2018) and Biais et al. (2014) predict a deterioration in market quality and an increase in volume following a liquidity shock which can be used as a proxy for a mini flash crash.

Finally, mini flash crashes generally originate from an aggregate liquidity shock that generates an imbalance between the demand and the supply of liquidity. For example, if following a good piece of news that stimulates an increase in the demand for a security, liquidity supply does not adequately accommodate, the result is a sudden imbalance which may dry up the book and deteriorate market quality. Dugast (2018) models a limit order book in the spirit of Duffie, Gârleanu, and Pedersen (2005) and focuses on an equilibrium in which the market prior to the news arrival is in a steady state. He shows that when an unscheduled piece of news arrives, order flow becomes imbalanced and the gradual reaction of liquidity demanders and liquidity suppliers generates an increase in volume as well as a reduction in market depth which is consistent with the empirical prediction made by Dugast (2018) and Biais et al. (2014). Furthermore, this means that the arrival of unexpected news (volatility shock) will generate both book imbalances and sudden changes in prices (MFCs), predicting that MFCs are positively correlated with short term volatility, which is consistent with the third empirical prediction in Li et al. (2017). Notice that the relation between high frequency volatility and MFCs makes no assumptions about causality. Indeed higher news arrival could be associated with higher volatility, or it may generate an asynchronous reaction in liquidity demanders

and liquidity suppliers, hence inducing order imbalances and sudden price changes (MFCs). In other words, more MFCs may lead to greater high frequency volatility, and at the same time greater high frequency volatility may lead to more mini flash crashes. Interestingly, whilst greater low frequency volatility may lead to more MFCs, we can confidently say that the converse does not hold. There are 3.03 mini flash crashes per day on average in our sample, and each MFC takes place over at most 1.5 seconds, thus mini flash crashes occur for at most 5 seconds out of the 23,400 seconds present in a trading day. It is therefore difficult to argue that mini flash crashes impact a low frequency volatility such as daily volatility.

7.B The Optimal Two-step Generalised Method of Moments Estimator

The model under consideration is

$$\begin{aligned} \text{NMFC}_{it} &= \beta_0 + \beta_1 * \text{NMFC}_{i,t-1} + \beta_2 * \pi_{it} + \beta_3 * \sigma_{it} + \epsilon_{it} \\ &= \beta_0 + \beta_1 * \text{NMFC}_{i,t-1} + \beta' * x_{it} + \epsilon_{it} \\ &= \delta' * X_{it} + \epsilon_{it} \end{aligned} \tag{8}$$

where $\beta = (\beta_2, \beta_3)' \in \mathbb{R}^2$; $x_{it} = (\pi_{it}, \sigma_{it})' \in \mathbb{R}^2$; $\delta = (\beta_0, \beta_1, \beta')' \in \mathbb{R}^4$; $X_{it} = (1, \text{NMFC}_{i,t-1}, x'_{it})' \in \mathbb{R}^4$; and $\epsilon_{it} = \eta_i + \upsilon_{it}$ is the error/shock for stock i on day t which we breakdown into the time-invariant stock-specific error component and the time-dependant stock-specific error component.

We make the following assumptions which constitute our set of orthogonality restrictions (moment conditions):

- Slope Parameter Homogeneity: the parameters in β are common to all $i \in \{0, 1, 2, 3\}$.
- Error Components: $E[\eta_i] = E[\upsilon_{it}] = E[\eta_i \upsilon_{it}] = 0 \ \forall \ i, t$.
- Serially Uncorrelated Shocks: $E[\upsilon_{it}\upsilon_{is}] = 0$ for $s \neq t$ and $\forall \ i$.
- Homoskedasticity: $E[\eta_i^2] = \sigma_\eta^2$ and $E[\upsilon_{it}^2] = \sigma_\upsilon^2 \ \forall \ i$ since our dataset is over a long enough period to cover business cycles.
- Predetermined Initial Conditions: $E[\text{NMFC}_{i1}\upsilon_{it}] = 0$ for $t \in \{2, ..., T\}$ and $\forall \ i$.

- Predetermined Independent Variable: $E[X_{it}\upsilon_{is}] = 0$ for $t \leq s$ and $\forall\ i$. The assumption that $\text{NMFC}_{i,t-1}$ is predetermined follows naturally from assuming that υ_{it} are serially uncorrelated shocks, provided the initial conditions $\text{NMFC}_{i,1}$ and the contemporaneous regressors x_{it} are also uncorrelated with subsequent υ_{it} shocks, thus making X_{it} predetermined.

- Fixed/Random Effects: We test the validity of random effects for x_{it} using the Difference Sargan Test. This test procedure can be used to test nested hypotheses, where only a subset of the moment conditions that are valid under the null hypothesis remain valid under a less restrictive alternative hypothesis. We therefore test:

$$\text{H}_0 : E[Z_{it}\epsilon_{it}^{+}] = 0\ \forall\ i,t \text{ against } \text{H}_1 : E[Z_{it}^{+}\epsilon_{it}^{+}] = 0\ \forall\ i,t \tag{9}$$

  where $E[Z_{it}^{+}\epsilon_{it}^{+}] = 0\ \forall\ i,t$ is the above set of moment conditions and $E[Z_{it}\epsilon_{it}^{+}] = 0\ \forall\ i,t$ is $\{E[Z_{it}^{+}\epsilon_{it}^{+}] = 0\ \forall\ i,t\} \cup \{E[x_{it}\eta_i] = 0\ \forall\ i,t\}$. Note Z_{it}^{+} is a strict subset of Z_{it}. We estimate the coefficients by GMM both under the null and under the alternative, and construct the Sargan(1958) statistics S under H_0 and S^{+} under H_1. If the null is valid, neither S nor S^{+} should reject, and the difference between them should not be too large. If the null if invalid but the alternative is valid, then S should reject while S^{+} should not reject, thus the difference between them should be larger. And if both the null and the alternative are invalid, then both S and S^{+} should reject, again in this case the difference between them should be larger than in the case where the null is valid. The difference S - S^{+} is therefore informative about the validity of the null hypothesis, and formally $DS = S - S^{+} \overset{a}{\sim} \chi^2(L - L^{+})$ under H_0, where L is the number of columns in Z_{it} and L^{+} is the (strictly smaller) number of columns in Z_{it}^{+}. Large values of the DS statistic reject the null hypothesis. Our p-value is 0.003 which is smaller than the 5% significance level, thus we reject H_0 and the hypothesis that the individual effects are uncorrelated. We therefore assume fixed effects instead.

In summary, we have the following model for $i \in \{1, 2, ..., 100\}$ and $t \in \{2, 3, ..., 1282\}$:

$$\begin{aligned} \text{NMFC}_{it} &= \delta' * X_{it} + \epsilon_{it} \\ s.t. \quad & E[Z_{it}^{+} \epsilon_{it}^{+}] = 0 \end{aligned} \tag{10}$$

where $E[Z_{it}^{+} s\epsilon_{it}^{+}] = 0$ is the whole set of moment conditions, summarised as follows. Error component, Serially uncorrelated shocks and Predetermined Initial conditions imply $(T-2)(T-1)/2$ linear moments conditions, namely $E[MFC_{i,t-s}\Delta\upsilon_{it}] = 0$ for $t \in \{3, 4, ..., T\}$ and $2 \leq s < t$, and a further $T-3$ quadratic moment conditions, which we now write as $E[\Delta\epsilon_{is}\epsilon_{iT}] = 0$ for $s \in \{3, 4, ..., T-1\}$. Under the homoskedasticity assumption, there are a further $T-3$ linear moment conditions $E[\text{NMFC}_{i,t-2}\Delta\upsilon_{i,t-1} - \text{NMFC}_{i,t-1}\Delta\upsilon_{it}] = 0$ for $t \in \{4, ..., T\}$. Finally, the assumption X_{it} is predetermined implies the $(T-2)T/2$ moment conditions $E[X_{i,t-s}\Delta\upsilon_{it}] = 0$ for $t \in \{3, 4, ..., T\}$ and $1 \leq s < t$.

We can now compute the optimal (two-step) GMM estimator as

$$\hat{\delta}_{GMM} = (X_{it}' Z_{it}^{+} W_N Z_{it}^{+'} X_{it})^{-1} X_{it}' Z^{+} W_N Z^{+'} \text{NMFC}_{it} \overset{a}{\sim} N(\delta, avar(\hat{\delta}_{GMM})) \tag{11}$$

where W_N is chosen so to minimise $avar(\hat{\delta}_{GMM})$. This is reported in Table 2.4.

7.C Degrees of Freedom used in Welch's t-test

The degrees of freedom used in the Welch t-test in Section 5.1 are computed using the following formula:

$$df = \begin{cases} N_b + N_a - 2 & \text{if Pooled Standard Error} \\ \Delta & \text{otherwise} \end{cases} \tag{12}$$

where the standard error is pooled if the 2 populations have the same standard deviation and a sufficient condition for this to hold when dealing with sample standard deviations is to ensure

$\frac{1}{2} < \frac{S^b}{S^a} < 2$. Furthermore,

$$\Delta = \begin{cases} \frac{\left(S_b^2/N_b + S_a^2/N_a\right)^2}{\left(S_b^2/N_b\right)^2/(N_b-1)+(S_a^2/N_a)^2/(N_a-1)} & \text{Satterthwaite's formula} \\ -2 + \frac{\left(S_b^2/N_b + S_a^2/N_a\right)^2}{\left(S_b^2/N_b\right)^2/(N_b+1)+(S_a^2/N_a)^2/(N_a+1)} & \text{Welch's formula} \\ \min(N_b - 1, N_a - 1) & \text{rough approximation} \end{cases} \tag{13}$$

Since the standard error in our dataset is not pooled and since Welch's formula (1947) is a refinement of Satterthwaite's formula (1946), we use Welch's formula to compute the degrees of freedom.[7,8]

7.D Topics for Future Research

The scope of this paper is to give an overview of mini flash crashes so that regulators can determine whether this phenomenon should be regulated or if the market should instead be left to its own devices. In light of the negative implications MFCs have on inside spread and BBO depth our policy recommendation is to mitigate the full force of these phenomena through regulation. Several possible resolutions come to mind and these are beyond the scope of this paper. As one example, Li et al. (2017) consider increasing the tick size to soften the impact of MFCs, however they find this would have the negative knock-on effect of also decreasing liquidity in normal times which is not desirable especially from the exchange's perspective. As an alternative solution, they propose to use circuit breakers as a way of containing MFCs, and circuit breakers are already being used by some stock exchanges. We must also point out that the empirical predictions of Li et al. (2017) have not been tested empirically so further work is still needed before we can advocate for the use of circuit breakers. Other papers exist on this issue. Furthermore, we may also argue that instead of trying to mitigate the effects of mini flash crashes once they take place, we should address what started this phenomenon, which hardly existed as little as just over a decade ago, so that we can pre-empty the threat of MFCs and make sure they never occur in the first place. This can be clearly seen in Appendix 7.E Figure 2.8 which gives a count of the daily occurrence of MFCs between January 3, 2006 and February 3, 2011 (our sampling period) with crashes depicted

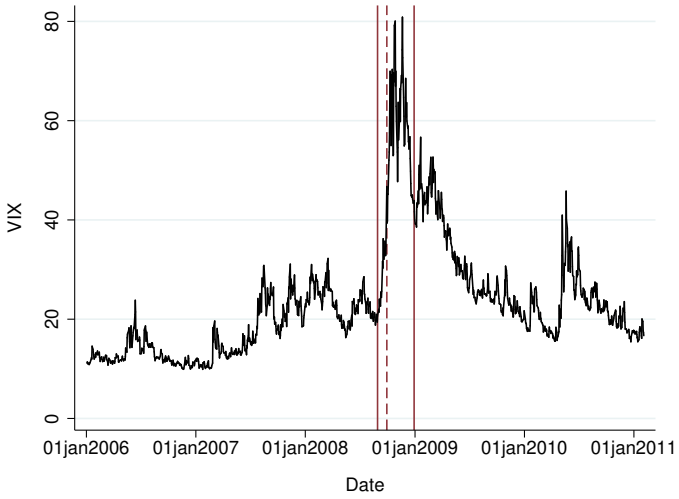
[7] Welch (1947)
[8] Satterthwaite (1946)

in red and rallies in blue. To highlight this point, the lower chart uses a logarithmic scale, and it is difficult not to notice that something must have changed in financial markets in late 2006 - early 2007, a time when the number of MFCs increased exponentially as one can see from the structural break in the mean of the logarithmic chart. Interestingly, on August 29, 2005 Regulation National Market System (Reg NMS) was introduced by the Security and Exchange Commission (SEC). This is a regulatory reform comprising a set of rules that was designed to strengthen the national market system for equity securities by enhancing both competition among individual markets and competition among individual orders.[9] The latter was to be achieved by linking together all trading platforms whether stock exchanges or alternative trading systems (ATS) thus allowing securities to be traded freely, on different exchanges, simultaneously. Connecting Regulation National Market System and mini flash crashes, it is indeed possible that the two are linked and that it is this very regulatory change that is responsible for the surge in MFCs we observe in Appendix 7.E Figure 2.8. This link however necessitates further investigation so we can only speculate that Reg NMS created the conditions for the exponential increase in high frequency trading (HFT) we observe in those same years (late 2006 - early 2007), and that HFT in turn led to the surge in MFCs. Indeed, Angel et al. (2011) show in Figures 2.13-2.16 of their paper that shortly after Reg NMS was introduced HFT grew exponentially, supporting our theory that Reg NMS and HFT are connected. Moreover, many of the papers published on the May 6, 2010 flash crash, such as Kirilenko et al. (2017), as well as other papers, such as Menkveld (2016) and Boehmer, Fong, and Wu (2012), find that extremely fast trading comes at the cost of (mini) flash crashes, which supports our theory that HFT and MFCs are connected. Although the evidence here gathered is not conclusive, there is ground to believe that Regulation National Market System freed up high frequency trading which comes at the cost of mini flash crashes. This remains a burgeoning topic of research and one that is scope for further research.

[9]For a more detailed discussion of Regulation National Market System one should revert to: SEC, Concept Release on Regulation NMS, Release No. 34-51808, File No. S7-10-04.

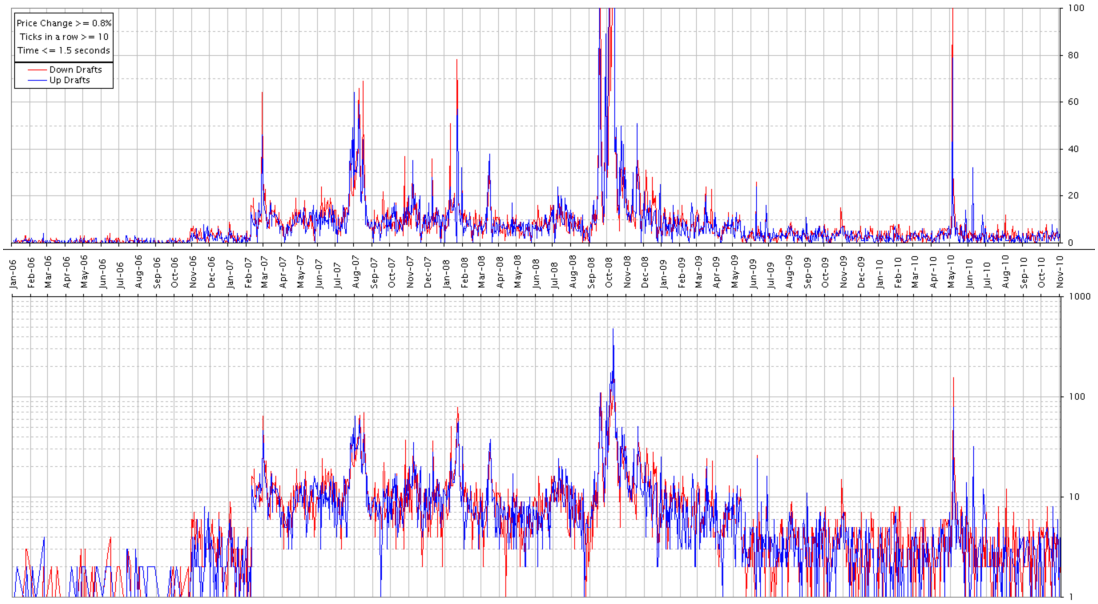
7.E Appendix Figures

Figure 2.7: The Great Financial Crisis



This figure exhibits the end of day (close) Chicago Board Options Exchange (CBOE) Volatility Index (VIX) over the period January 3, 2006 - February 3, 2011. The 2 red lines indicate the period which is to be removed as affected by the Global Financial Crisis, namely August 29, 2008 - December 29, 2008, and the dashed red line is the GFC itself i.e. September 29, 2008.

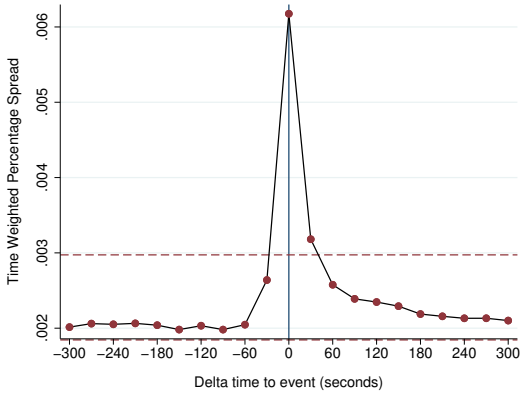
Figure 2.8: Regulation NMS, High Frequency Trading, and Mini Flash Crashes



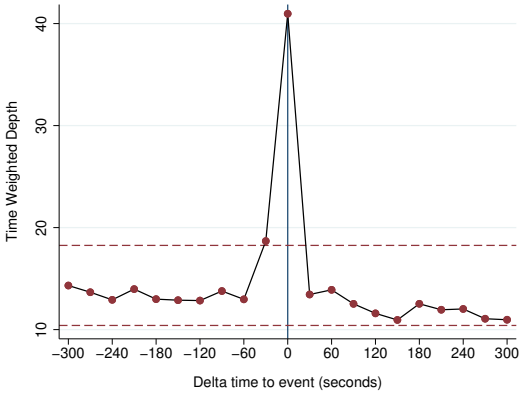
This figure exhibits the daily occurrence of mini flash crashes in the U.S. over January 3, 2006 and February 3, 2011. Furthermore, crashes are reported in red and rallies in blue, while the lower chart uses a logarithmic scale on the upper chart. Source: http://www.nanex.net/FlashCrashEquities/FlashCrashAnalysis_Equities.html

Figure 2.9: Test 1 by Stock Exchange I/II

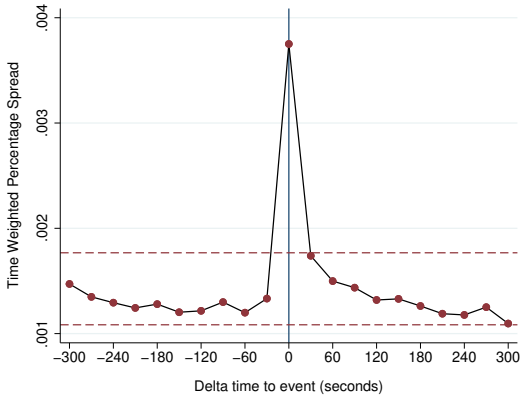
(a) NYSE - Inside Spread



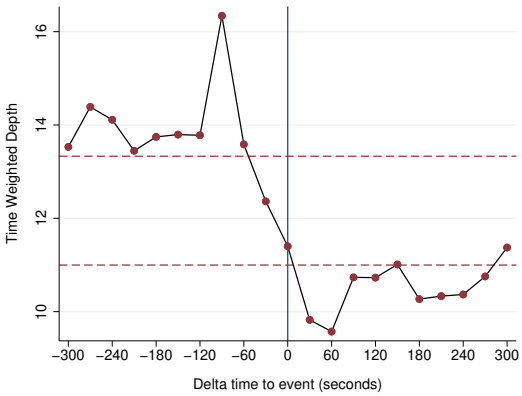
(b) NYSE - BBO Depth



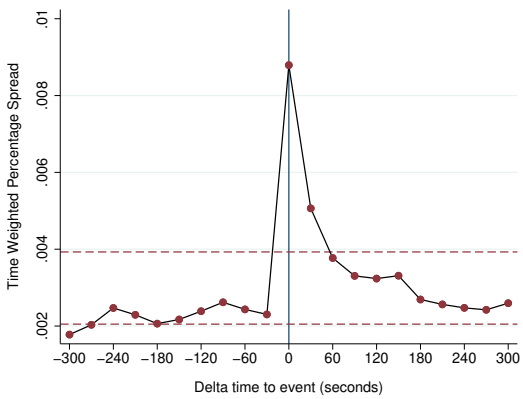
(c) ARCA - Inside Spread



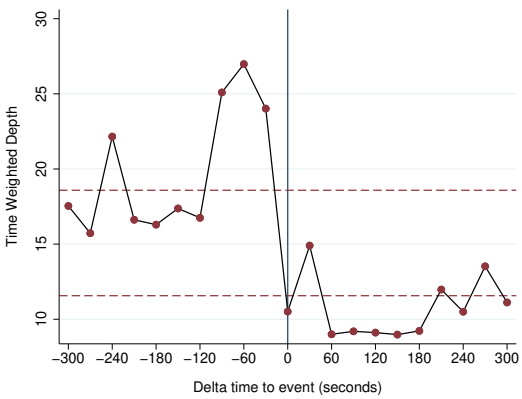
(d) ARCA - BBO Depth



(e) NASDAQ - Inside Spread



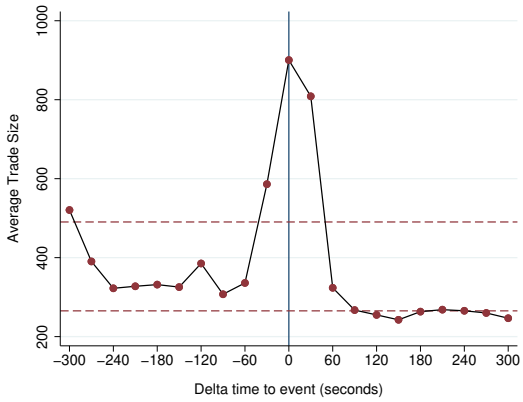
(f) NASDAQ - BBO Depth



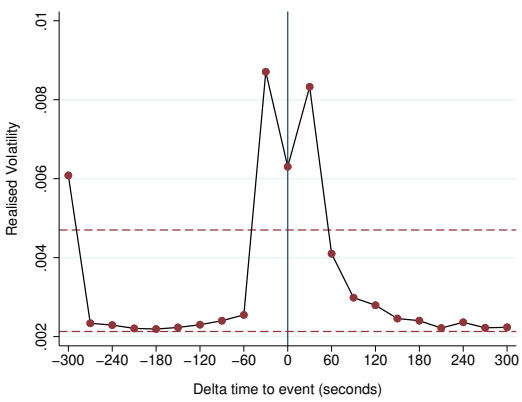
This figure exhibits each market quality metric by the stock exchange where the MFC took place as a time series plot over the 10-minute window around the MFC (highlighted by the vertical blue line) together with a 99% confidence interval (dashed red line). The x-axis is reported as number of seconds to the event (the mini flash crash).

Figure 2.10: Test 1 by Stock Exchange II/II

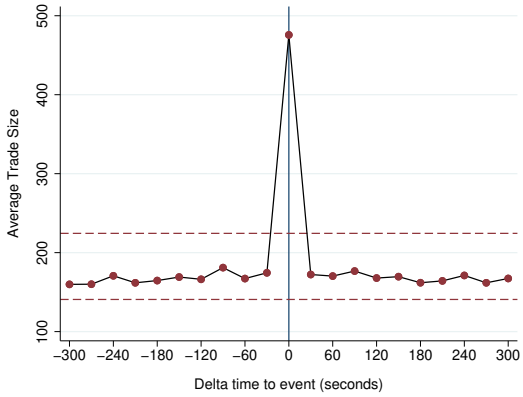
(a) NYSE - Trading Activity



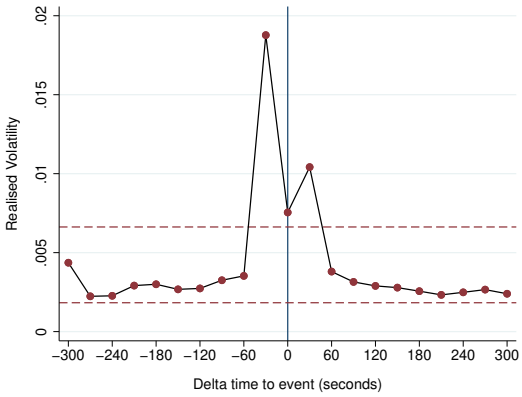
(b) NYSE - Intraday Volatility



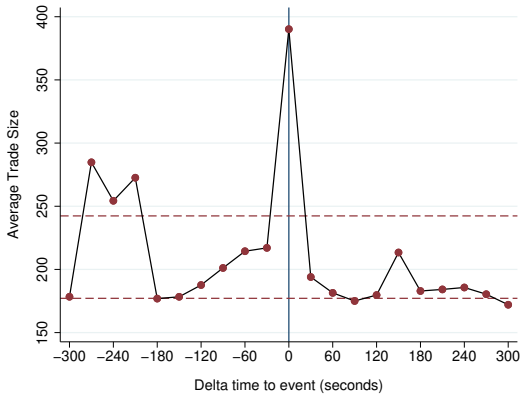
(c) ARCA - Trading Activity



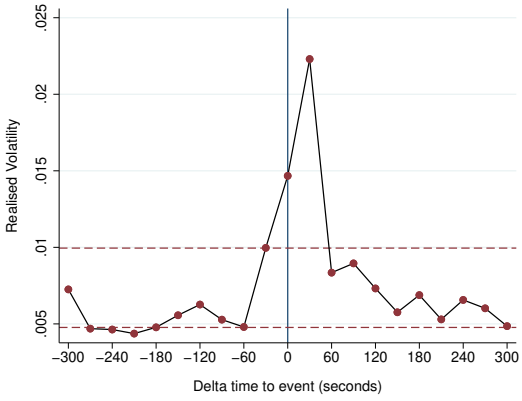
(d) ARCA - Intraday Volatility



(e) NASDAQ - Trading Activity



(f) NASDAQ - Intraday Volatility



This figure exhibits each market quality metric by the stock exchange where the MFC took place as a time series plot over the 10-minute window around the MFC (highlighted by the vertical blue line) together with a 99% confidence interval (dashed red line). The x-axis is reported as number of seconds to the event (the mini flash crash).

7.F Appendix Tables

Table 2.8: Market Quality Metrics – Definition

Market Quality	**Metric**	**Formula**
Inside Spread	Size Weighted % Spread	$\frac{1}{J}\sum_{j=1}^{J}\frac{p^{\alpha}_{i,MTI,j}*q^{\alpha}_{i,MTI,j}-p^{\beta}_{i,MTI,j}*q^{\beta}_{i,MTI,j}}{\frac{1}{2}(p^{\alpha}_{i,MTI,j}+p^{\beta}_{i,MTI,j})}$
	Time Weighted % Spread	$\frac{1}{\bar{w}_{i,MTI}}\sum_{j=1}^{J}w_{i,MTI,j}*\left(\frac{p^{\alpha}_{i,MTI,j}-p^{\beta}_{i,MTI,j}}{\frac{1}{2}(p^{\alpha}_{i,MTI,j}+p^{\beta}_{i,MTI,j})}\right)$
	Size Weighted % Effec. Spread	$\frac{1}{\bar{\omega}_{i,MTI}}\sum_{j=1}^{J}\omega_{i,MTI,j}*\frac{2*I^{TD}_{i,MTI,j}*(p_{i,MTI,j}-\frac{1}{2}(p^{\alpha}_{i,MTI,j}+p^{\beta}_{i,MTI,j}))}{\frac{1}{2}(p^{\alpha}_{i,MTI,j}+p^{\beta}_{i,MTI,j})}$
BBO Depth	Time Weighted Depth	$\frac{1}{\bar{w}_{i,MTI}}\sum_{j=1}^{J}w_{i,MTI,j}*(q^{\alpha}_{i,MTI,j}+q^{\beta}_{i,MTI,j})$
	Time Weigthed Best Ask Size	$\frac{1}{\bar{w}_{i,MTI}}\sum_{j=1}^{J}w_{i,MTI,j}*q^{\alpha}_{i,MTI,j}$
	Time Weighted Best Bid Size	$\frac{1}{\bar{w}_{i,MTI}}\sum_{j=1}^{J}w_{i,MTI,j}*q^{\beta}_{i,MTI,j}$
Trading Activity	Executed Volume	$\sum_{j=1}^{J}q_{i,MTI,j}$
	Average Trade Size	$\frac{1}{J}\sum_{j=1}^{J}q_{i,MTI,j}$
Intraday Volatility	Realised Volatility	$\sqrt{\sum_{j=1}^{J}\log^2\left(\frac{\frac{1}{2}(p^{\alpha}_{i,MTI,j}+p^{\beta}_{i,MTI,j})}{\frac{1}{2}(p^{\alpha}_{i,MTI,j-2}+p^{\beta}_{i,MTI,j-2})}\right)}$
	Volatility Price Range	$\log\left(\frac{\max_{j\in\{1,..J\}}p_{i,MTI,j}}{\min_{j\in\{1,..J\}}p_{i,MTI,j}}\right)$

This table reports the formulae used to compute the daily market quality metrics, where $(p/q)^{\alpha/\beta}_{i,MTI,j}$ *= Best Ask/Bid Price/Size at intra-day observation (quote)* j *on stock* i *traded in mini time interval* MTI*;* $(p/q)_{i,MTI,j}$ *= trading price/quantity at intra-day observation (trade)* j *on stock* i *traded in mini time interval* MTI*;* $w_{i,MTI,j}$ *= time difference between observation* j *and observation* $j+1$ *for stock* i *traded in mini time interval* MTI *and* $\bar{w}_{i,MTI}=\sum_{j=1}^{N}w_{i,MTI,j}$*;* $\omega_{i,MTI,j}$ *= number of traded shares at intra-day observation* j *on stock* i *in mini time interval* MTI *(may be at the best ask or best bid depending on trade direction) and* $\bar{\omega}_{i,MTI}=\sum_{j=1}^{N}\omega_{i,MTI,j}$*; and* $I^{TD}_{i,MTI,j}$ *= 1 if a market order to buy is submitted and -1 if a market order to sell is submitted; and* J *is the number of observations in the mini time interval under consideration.*

Table 2.9: Test 2 by Stock Exchange

The Welch t-Test – $\boldsymbol{H_0: d=0}$ **VS** $\boldsymbol{H_1: d\neq 0}$

Market Quality	**NYSE**	**ARCA**	**NASDAQ**
Inside Spread	-2.14**	-0.63	-3.30***
BBO Depth	2.81**	9.46***	6.07***
Trading Activity	1.01	-0.31	2.38**
Intraday Volatility	0.13	0.58	-1.45

** if p-value < 0.05; *** if p-value < 0.01

This table reports the Welch t-Test of each market quality metric by the stock exchange where the MFC took place as a test between the pre-MFC average (μ_b) *and the post-MFC average* (μ_a) *summarised as* $d := \mu_b - \mu_a$*.*

Table 2.10: Tabulation of Figure 2.6

MTI	**Inside Spread**		**BBO Depth**		**Trading Activity**		**Intraday Volatility**	
	Estimate	**99% Conf. Int.**	**Estimate**	**99% Conf. Int.**	**Estimate**	**99% Conf. Int.**	**Estimate**	**99% Conf. Int.**
B10	0.0019	(0.0017,0.0027)	14.33	(11.46,16.31)	398	(230, 406)	0.0057***	(0.0023,0.0054)
B9	0.0019	(0.0017,0.0027)	13.95	(11.46,16.31)	321	(230, 406)	0.0025	(0.0023,0.0054)
B8	0.0019	(0.0017,0.0027)	13.78	(11.46,16.31)	276	(230, 406)	0.0024	(0.0023,0.0054)
B7	0.0019	(0.0017,0.0027)	13.96	(11.46,16.31)	279	(230, 406)	0.0025	(0.0023,0.0054)
B6	0.0019	(0.0017,0.0027)	13.36	(11.46,16.31)	276	(230, 406)	0.0025	(0.0023,0.0054)
B5	0.0018	(0.0017,0.0027)	13.39	(11.46,16.31)	273	(230, 406)	0.0025	(0.0023,0.0054)
B4	0.0019	(0.0017,0.0027)	13.30	(11.46,16.31)	313	(230, 406)	0.0026	(0.0023,0.0054)
B3	0.0019	(0.0017,0.0027)	15.06	(11.46,16.31)	266	(230, 406)	0.0028	(0.0023,0.0054)
B2	0.0019	(0.0017,0.0027)	13.93	(11.46,16.31)	283	(230, 406)	0.0029	(0.0023,0.0054)
B1	0.0023	(0.0017,0.0027)	17.59***	(11.46,16.31)	460***	(230, 406)	0.0110***	(0.0023,0.0054)
E	0.0055***	(0.0017,0.0027)	29.20***	(11.46,16.31)	777***	(230, 406)	0.0072***	(0.0023,0.0054)
A1	0.0030***	(0.0017,0.0027)	15.74	(11.46,16.31)	622***	(230, 406)	0.0096***	(0.0023,0.0054)
A2	0.0024	(0.0017,0.0027)	12.66	(11.46,16.31)	278	(230, 406)	0.0043	(0.0023,0.0054)
A3	0.0022	(0.0017,0.0027)	11.92	(11.46,16.31)	239	(230, 406)	0.0033	(0.0023,0.0054)
A4	0.0022	(0.0017,0.0027)	11.25***	(11.46,16.31)	229***	(230, 406)	0.0031	(0.0023,0.0054)
A5	0.0021	(0.0017,0.0027)	10.87***	(11.46,16.31)	223***	(230, 406)	0.0027	(0.0023,0.0054)
A6	0.0020	(0.0017,0.0027)	11.88	(11.46,16.31)	232	(230, 406)	0.0027	(0.0023,0.0054)
A7	0.0020	(0.0017,0.0027)	11.59	(11.46,16.31)	237	(230, 406)	0.0024	(0.0023,0.0054)
A8	0.0019	(0.0017,0.0027)	11.57	(11.46,16.31)	237	(230, 406)	0.0026	(0.0023,0.0054)
A9	0.0019	(0.0017,0.0027)	11.18***	(11.46,16.31)	231	(230, 406)	0.0025	(0.0023,0.0054)
A10	0.0019	(0.0017,0.0027)	11.05***	(11.46,16.31)	222***	(230, 406)	0.0024	(0.0023,0.0054)

* if p-value < 0.1; ** if p-value < 0.05; *** if p-value < 0.01

This table reports the tabulated values of Figure 2.6.

Chapter 3

Leadership of the Price Discovery Process at Outer Levels of the Book

Abstract

In this paper we study the leadership of the price discovery process and we show that outer levels of the limit order book (LOB) need not exhibit the same properties as the inner levels (i.e. levels closer to the BBO). This is relevant as many studies in the literature focus exclusively on the BBO and tacitly assume that all other levels of the LOB exhibit the same price leadership. Our contribution to the literature is two-fold. Theoretically, we build on the model by Hasbrouck (1995) to derive a metric which can be used to measure price discovery leadership. Empirically, we provide one specific case study (BATS Chi-X between October 1, 2012 and March 28, 2013) which exhibits a switch in the directional leadership of the price discovery process across the LOB.

1 Introduction

In ancient times the modern financial system as we know it today did not exist. There was no Teìckhos Hòdos (Wall Street) among the ancient streets of Greece, nor was there an exchange where ancient Romans could convert their treasures in the different currencies of the countries they had conquered. In fact it was not until May 17, 1792 that twenty-four brokers met beneath a buttonwood tree in lower Manhattan to sign the Buttonwood Agreement thus giving birth to one of the first stock exchanges, known today as the New York Stock Exchange. For many years stock exchanges were a physical location with a floor where buyers and sellers would meet and negotiate, setting prices. Then, in the late 20th century, financial markets reached a major milestone as traders started to transact remotely in what became known as electronic trading. The digital age revolutionised financial markets as they turned from relatively simple exchanges where people had to shout at each other in other to make deals, to a complex system of computers where trades take place at the speed of light. Evolving regulations, such as Regulation NMS in the U.S. and MiFID in Europe, also contributed to this change, lowering barriers to entry and thus contributing to the proliferation of multiple trading venues, in turn reducing the dominance of incumbent exchanges. Market fragmentation is a burgeoning topic of research and one that features heavily in the agenda of most regulators.

Fragmentation means that we can no longer rely on the classical economic paradigm of one or more securities that trade in a single centralised market. Instead, a new framework is needed, one which accounts for the dispersion of trading in a security to multiple sites. The exploitation of arbitrage opportunities keeps prices in the different markets from diverging without bound from each other, which econometrically translates to the prices being co-integrated with one or more common stochastic factors. Hasbrouck (1995) puts forward a model that exploits this property. Furthermore, given the same security is now traded on multiple platforms and that each platform has its own price, logically the same asset may have different prices. Each price is determined, and therefore discovered, by news being gathered and interpreted in that particular market, so we may be interested in knowing which market produces the price discovery. In this framework, we assume there exists an implicit efficient price and that this should be common to all markets; and that

the sources of variation in this efficient price are attributed to the different markets the security is traded on, thus providing a way to measure a market's contribution to price discovery. We define price leadership as the ability of a trading venue to impound the fundamental information about the value of an instrument into its prices ahead of its competitors.

Understanding price discovery and which market is at the forefront in this process is important for several applications. For example Ibikunle (2015) explains the successful breakthrough of BATS Chi-X Europe in the European market by showing that the high-tech entrant holds leadership in the price discovery process over the LSE and that this was key to gain market share. Price discovery also explains why traders are able to exploit some trading venues in their inter-market arbitrage, or why traders choose national exchanges instead of foreign (Lieberman, Ben-Zion, and Hauser (1999)) or regional (Garbade and Silber (1979)) exchanges. From the broker's perspective, price discovery signals which market provides the informationally best quote, and from the regulator's perspective, price discovery informs which type of investor is responsible for the inclusion of fundamental information in the current price. Indeed Brogaard et al. (2014) shows that on NASDAQ between 2008 and 2009 the market orders of high frequency traders (HFTs) play a positive role in the price discovery process since they are posted in the direction of permanent price changes and in the opposite direction of transitory pricing errors, while the opposite is true about HFTs' limit orders.

This paper contributes to the price discovery literature by showing that in order to fully understand price leadership one cannot simply examine the best bid-offer. The reason we cannot exclusively focus on the BBO is twofold: first, less than 50% of liquidity is posted to the BBO (evidence is provided in Section 5 of this manuscript as well as in other research papers) making outer levels of the LOB just as important as the BBO; and second, outer levels of the LOB may exhibit a switch in the directional leadership of the price discovery process so that leadership at the BBO does not imply leadership on outer levels. The combination of these two points, namely the importance of outer levels and their potential dissimilarity to the BBO, leads to the understanding that price discovery is not exclusively driven by and related to the BBO, and in this lies our

contribution to the literature.[10]

This paper is organised as follows. Sections 1 and 2 include an introduction and a literature review. Section 3 extends the model developed by Hasbrouck (1995) and Baillie, Booth, Tse, and Zabotina (2002). Section 4 derives from the model a metric to measure leadership of the price discovery process and Section 5 provides the empirical evidence namely a case study which uses the aforementioned metric to measure price leadership and show this may differ between outer and inner levels of the LOB. Section 6 presents the conclusions, and Section 7 gives the Appendix to the paper.

2 Literature Review

Price discovery is a timely research topic since it asks a very basic question: under the assumption that prices are efficient, by which we mean that they include all available information, how do prices adjust as new information becomes available? This issue can then be expanded into how quickly they adjust and what drives their adjustment. This subject became of even greater relevance with the advent of electronic trading and with it of machines that are capable of trading within milliseconds or even microseconds of the outbreak of new information. Indeed, it is now believed that to gauge the immediate response of a political event to the economy, one should look at financial markets which react within seconds of the news and in any case much faster than the real economy. For example, many hold that as the Brexit saga drags on, the best indicator of how events in Westminster are perceived by the rest of the country is to use the value of the pound against other major currencies such as the euro or the dollar. It will therefore be of no surprise that the existing literature on price discovery is extensive and the survey here provided is by no means exhaustive.

[10] A level of the Limit Order Book (LOB) is a price level at which there exists at least one limit order. Consider the following example. Suppose the LOB is consists of 4 limit orders to sell at $ 104, 1 limit order to sell at $ 102, 2 limit order to sell at $ 101, 1 limit order to buy at $ 99, and one limit order to buy at $ 97. Then $ 104 is the third level on the sell side of the LOB, $ 102 is the second level on the sell side of the LOB, $ 101 is the first level on the sell side of the LOB (also known as the "best offer" price), $ 99 is the first level on the buy side of the LOB (also known as "best bid" price), and $ 97 is the second level on the buy side of the LOB. Notice, that the mid-quote of the Best Bid Offer (BBO) is $ 100, that the LOB is not symmetric between the buy and sell sides, and that there is a "gap" on both the sell side (namely $ 103) and on the buy side (namely $ 98). In real case scenarios, while gaps are unlikely near the mid-quote of the BBO, they may appear as one moves away from the BBO.

A financial asset usually trades on multiple markets, and since the asset is exactly the same on all markets it is safe to assume that inter-market arbitrage ensures the price differences between markets do not diverge without bound. In econometrics terms, this is equivalent to saying that prices are co-integrated variables, which means that the price series share one or more common stochastic factors. In most applications there is only one common factor which we refer to as the implicit efficient price, and it is this price that driven by news constitutes the source of permanent movements in the prices of all markets. There are several models that assume prices have one common stochastic factor but the most commonly used in the study of the price discovery process are undoubtedly those developed by Hasbrouck (1995) and Gonzalo and Granger (1995). Baillie et al. (2002) explores the relationship between these two common factor models, while many extensions and empirical applications exist in the literature. We provide a brief overview of the aforementioned papers.

Hasbrouck (1995) develops an econometric approach based on an implicit unobservable efficient price common to all markets. The paper then defines price discovery in terms of the variance of each market's innovation to the common implicit efficient price and builds a metric to measure each market's contribution to this variance. Said metric is the information share associated with a particular market and is taken to be the proportional contribution of that market's innovation to the innovation in the common factor. As an empirical example, this measure of price discovery is applied to quotes for the thirty stocks in the Dow Jones Index and it is concluded that the New York Stock Exchange leads the price discovery with a median information share of 92.7%.

Gonzalo and Granger (1995) believe that in order to study co-integration in large systems, it is required to reduce their dimensionality, and to this end they suggest using an $I(1)$ common factor model. They then put forward a different approach to Hasbrouck's (1995) information share which relies on the decomposition of the vector error correction representation of the co-integrated time series, namely the observed price, into transitory and permanent shock components. While the former component is completely ignored in this study, a new way of estimating the latter is provided, and it relies on identifying the $I(1)$ common factors by imposing that they be linear combinations of the original variables. In other words, Gonzalo and Granger (1995) use the common

trend representation of Stock and Watson (1988) to break the observed prices into a common factor component and a transitory component, and define the market's price discovery as the market's contribution to the common factor which they measure as a linear combination of the observed prices. In the price discovery context, under this model only permanent shocks result in disequilibria which occur because markets process news at different rates. The price discovery measure is dubbed component share and several empirical examples are given in the paper to illustrate the procedure.

The relationship between the Hasbrouck (1995) and Gonzalo and Granger (1995) common factor models is examined in Baillie et al. (2002) with a view that the two models complement each other since they measure and therefore define price discovery in slightly different ways. Baillie et al. (2002) show that the two models are closely related and provide similar results if the residuals are uncorrelated between markets, but they also show the same cannot be said if there is substantial correlation and in this instance the results typically differ. They give several analytic examples to illustrate this as well as real world examples of Nasdaq market makers. Baillie et al. (2002) do not conclude that either model provides a better measure of price discovery since this depends on whether price discovery is considered to be solely an error correction phenomenon, or it should also consider the correlations between the markets' innovations. However they believe Hasbrouck (1995) has greater economic appeal due to its dependence on the relative innovation variances.

Moreover, Putniņš (2013) argues that both measures of price discovery, namely the information and component shares, correctly allocate price discovery leadership only if the price series have similar levels of trading noise. Whenever there are differences in the noise levels, the two metrics measure a combination of speed of impounding information and relative avoidance of noise rather than price discovery, although this kind of problem is more likely to arise when the price series are for different asset classes. In any case, Putniņš (2013) solves this issue by showing that the information leadership metric derived by Yan and Zivot (2010) correctly assigns contributions to price discovery even when the noise levels differ. The information leadership share cleverly combines the information and component shares to remove the relative noise levels and obtain a clear measure of relative information leadership.

Finally, there are numerous studies that build on what we have just discussed. For instance,

Taylor (2011) and Frijns, Indriawan, and Tourani-Rad (2015) show that price discovery varies intraday in relation to macro-economic news announcements, and Ibikunle (2015) gives evidence that BATS Chi-X Europe, a high-tech entrant, was able to enter the British market and gain market share at the expense of the established national exchange (the London Stock Exchange) thanks to its superior price discovery. Moreover, Harris, McInish, Shoesmith, and Wood (1995) investigate the distribution of price discovery between the New York Stock Exchange and regional stock exchanges, and while Bacidore and Sofianos (2002) show that the home market usually leads the price discovery process, this result has not been consistent across the literature as different study settings reach different conclusions. For example, Lieberman et al. (1999) show that for Israeli stocks cross-listed on the New York Stock Exchange the home market is informationally dominant. On the other hand, according to Kadapakkam, Misra, and Tse (2003) cross-listed Indian stocks on the London Stock Exchange receive equal contribution in terms of price discovery from the home and foreign market.

3 Model

We aim is to show that leadership of the price discovery process is not determined by the BBO and that it may change with the level of the LOB. In order to pursue this we need a model from which we can build a metric to measure price discovery leadership. As discussed in Section 2, it is reasonable to assume that the prices of the same asset traded on two different trading platforms are first-order co-integrated with the common stochastic trend referred to as the implicit efficient price or the common efficient price:[11]

$$\{p_t\}_{t=1}^{\infty} = \{(p_{1,t}, p_{2,t})\}_{t=1}^{\infty} \sim \text{CI}(1,0) \text{ with co-integrating vector } \beta = (1,-1)' \tag{14}$$

The model (14) is further justified in Appendix 7.A.1, and is also extensively used in the literature as argued in Section 2 and by Campbell and Shiller (1979) who show that if the security price in one market is a rational forecast of that in another market then the two prices will be co-integrated. The

[11] $p_{i,t}$ is the price of the asset at time t on trading platform i where $i \in \{1,2\}$. The model can be extended to more than two platforms if needed.

key intuition to fully take advantage of this model is to exploit the vector moving average representation (Section 3.1) to solve the model and understand which component of the price is permanently impounded into the common efficient price, and the vector error correction representation (Section 3.2) to estimate the parameters in the model. This technique was developed by Hasbrouck (1995), Baillie et al. (2002), Yan and Zivot (2007), and Ibikunle (2015). Our contribution lies in assembling the above analysis in a coherent and cogent manner.

3.1 Vector Moving Average

The model in (14) implies that the first difference of price is second-order stationary and therefore that it can be expressed as an infinite-order vector moving average (VMA): [12,13]

$$\Delta p_t = \tau + \sum_{k=0}^{\infty} \Theta_k \eta_{t-k} \tag{15}$$

$$= \tau + \Theta(L)\eta_t \tag{16}$$

Exploiting the Beveridge-Nelson Representation Theorem, as shown in Appendix 7.A.2, we can summarise the solution to the VMA representation of the model in (16) as:

$$p_t = \widehat{p_0} + (\mu + \tau t) + \Theta(1) \sum_{s=1}^{t} \eta_s + \Theta^*(L)\eta_t \tag{17}$$

The solution in (17) breaks down the stochastic process $\{p_t\}_{t=1}^{\infty}$ into four terms: the first two are deterministic and the second two are stochastic. The first term, $\widehat{p_0}$, is a vector of adjusted initial values and reflects non-stochastic differences between the price variables. Therefore if we

[12]Equation (15):

$\{p_t\}_{t=1}^{\infty}$ is a bivariate stochastic process with $p_t = (p_{1,t}, p_{2,t}) \in \mathbb{R}^2$ of the quoted prices $p_{i,t}$ on platform i where the prices could be taken from any level of the LOB;

$\mathbb{E}_t[p_t] = \mu + \tau t \in \mathbb{R}^2$ is the deterministic trend of p_t since higher order terms $O(t^2)$ are not possible due to the stationarity of the first-order price-difference Δp_t (note $\mu, \tau \in \mathbb{R}^2$);

$\{\Theta_k\}_{k=0}^{\infty}$ is a sequence of 2 x 2 scalar matrices whose elements are Cesaro summable and $\Theta_0 \neq I_2$ and is invertible;

$\{\eta_t\}_{t=1}^{\infty}$ is a bivariate second-order stationary stochastic process with $\eta_t = (\eta_{1,t}, \eta_{2,t}) \in \mathbb{R}^2$, $\mathbb{E}[\eta_t] = 0$, $\text{Cov}[\eta_t] = \Sigma < \infty$, and η_1, η_2 are serially uncorrelated (but may be cross-correlated).

Restrictions need to be placed on Θ_0 as this matrix contains the initial impacts of the shocks on Δp_t and consequently defines the contemporaneous correlation structure of Δp_t.

Finally, note the RHS is a generic second-order stationary process.

[13]Equation (16):

$\Theta(L) = \sum_{k=0}^{\infty} \Theta_k L^k \in \mathbb{R}^2$ x $\mathbb{R}^2$ is a matrix polynomial in the lag operator L.

compare quotes between two different trading venues it matters not that one market has a wider spread than the other since this factor is captured and dealt with by p_0 which makes up $\widehat{p_0}$. This term allows us to focus solely on the speed at which quotes update without having to worry about the exact position of the quotes we are comparing. We can therefore focus on the relative position of quotes without having to worry about their absolute position. Since no two platforms have the exact same spread, this term is crucial for our analysis to be consistent and for us to be able to compare two exchanges whose levels do not match one for one. The second term, $\mu + \tau t$ is the non-stochastic trend of our price series. The third term, $\Theta(1)\sum_{s=1}^{t}\eta_s$, is impounded into all prices posterior to time t since $p_{t+\delta} = \widehat{p_0} + (\mu + \tau(t+\delta)) + \left[\sum_{s=1}^{t}\Theta(1)\eta_s + \sum_{s=t+1}^{t+\delta}\Theta(1)\eta_s\right] + \Theta^*(L)\eta_{t+\delta}$ for $\delta > 0$. It follows that $\Theta(1)\sum_{s=1}^{t}\eta_s$ is the permanent component of the price p_t (abbreviated as $\text{pc}(p_t)$) or in other words the common stochastic trend at time t. This is consistent with $\{p_t\}_{t=1}^{\infty} \sim \text{I(1)}$ since I(1) processes have a unique stochastic trend identified by a random walk (note that $\sum_{s=1}^{t}\eta_s$ is the sum of predetermined white noise shocks and therefore a random walk). Moreover, we infer that $\Theta(1)\eta_t = \sum_{s=1}^{t}\Theta(1)\eta_s - \sum_{s=1}^{t-1}\Theta(1)\eta_s = \text{pc}(p_t) - \text{pc}(p_{t-1}) = \text{pc}(\Delta p_t)$, that is $\Theta(1)\eta_t$ is the component of the price change that is permanently impounded into future (to time t) prices. In summary, $\sum_{s=1}^{t}\Theta(1)\eta_s = \text{pc}(p_t)$ is the component of the contemporaneous price p_t that is permanently impounded into the common efficient price and it cumulates the individual contributions made by predetermined prices, namely $\Theta(1)\eta_t = \text{pc}(\Delta p_t)$, where each contribution is the component of the contemporaneous innovation Δp_t that is permanently impounded into the common efficient price. The fourth and final term, $\Theta^*(L)\eta_t$, differs from the first two terms (which are non-stochastic) and the second (which is a function of a random walk) as it only comprises the contemporaneous shock η_t. It follows that this term has no bearing on prices posterior to time t since it wears off with the passage of time and therefore that it constitutes the temporary component of the price p_t. We have consequently broken down the price p_t into four terms: an adjusted vector of initial values, a non-stochastic trend, a permanent component which lasts over time and a temporary component which vanishes over time.

Furthermore, we show in Appendix 7.A.3 that:

$$\Theta(1) = \frac{\partial}{\partial \eta_t} \lim_{k \to \infty} \mathbb{E}_t[p_{t+k}] \tag{18}$$

which implies that $\Theta(1)$ is the long-run impact matrix of a contemporaneous shock on price (so intuitively it can be thought of as aggregating all short-run dynamic multipliers Θ_k) and $\Theta(1)\eta_t$ is the rational forecast of the long-run price $\lim_{k\to\infty} \mathbb{E}_t[p_{t+k}]$.

3.2 Vector Error Correction

The contribution of the vector moving average representation in (16) to the overall analysis is that it allows to solve the model and interpret the stochastic process that is captured by the model. In particular, the VMA representation shows that $\Theta(1)\eta_t$ constitutes the permanent component of the price change taking place at time t that is permanently impounded into the common efficient price. However, the VMA representation does not favour the estimation of the parameters in the model since the current and lagged shocks are not observable. The key intuition behind this piece of theoretical work lies in the realisation that the model in (14) has several parameterisations and that we can link these to extract the most out of the model. In particular, while we employ the VMA for intuition, we take advantage of the vector error correction (VEC) for estimation purposes. The parameters of interest are $\Theta(1)$ and $\Sigma = \text{Cov}(\eta_t)$ since these capture the object of interest in the study of the price discovery process, namely $\Theta(1)\eta_t$.

Exploiting the Simplified Granger Representation Theorem, as shown in Appendix 7.A.4, we can summarise the solution to the VEC representation of the model in (14) as:

$$\Delta p_t = -\alpha(\beta' p_{t-1} + \widehat{\mu} + \widehat{\tau} t) - \sum_{k=1}^{K} \Gamma_k \Delta p_{t-k} + \eta_t \tag{19}$$

The VEC representation of the model in (19) has the following interpretation. At equilibrium, by definition we have $0 = \Delta p_t = \eta_t \; \forall \, t$ and this applied to the VEC in (19) implies $\beta' p_{t-1} + \widehat{\mu} + \widehat{\tau} t = 0$ which captures the dynamics of the price series p_t at equilibrium. Note the series is said to be trend stationary around the deterministic trend $\widehat{\mu} + \widehat{\tau} t$ and that the co-integrating vector β determines

how the individual markets $p_{1,t}$ and $p_{2,t}$ relate to the trend. Off equilibrium, $\beta' p_{t-1} + \widehat{\mu} + \widehat{\tau} t \neq 0$ and $-\alpha$ determines how the system p_t converges or corrects back to equilibrium (hence why α is referred to as the error correction vector). Additionally, $\sum_{k=1}^{K} \Gamma_k \Delta p_{t-k}$ accounts for innovations (Δp_{t-k}) that took place in the near past and still have an effect on the current innovation (Δp_t). Since $\Delta p_t \to 0$ as $t \to$ equilibrium, past innovations Δp_{t-k} are only present off-equilibrium and can be viewed as depicting short-run off-equilibrium dynamics induced by market imperfections.

The parameters of the VEC $(\alpha, \widehat{\tau}, \{\Gamma_k\}_{k=1}^{K}, \widehat{\widehat{\mu}}, \Sigma)$ in (19) are estimated by the maximum likelihood estimator developed by Johansen (1988, 1995). Recall from Section 3.1 that out of the parameters of the VMA representation $(\widehat{p_0}, \mu, \tau, \Theta(1), \{\Theta_k^*\}_{k=0}^{\infty}, \Sigma)$ in (17) we are interested in $(\Theta(1), \Sigma)$ for the study of the price discovery process thus after estimation of the VEC we are missing $\Theta(1)$ since Σ is the same for both representations. This parameter can be computed, as shown in Appendix 7.A.5, from:

$$g(\theta, \alpha) = 0 \tag{20}$$

Therefore, we exploit $g(\theta, \alpha) = 0$ to compute the parameters of the VMA $(\Theta(1), \Sigma)$ from the estimated parameters of the VEC (α, Σ). The function $g(\theta, \alpha)$ links the two representations of the model in (14), and in particular it links the permanent/long-run component of the VMA ($\Theta(1)$) to the equilibrium component of the VEC (α).

In summary, the VEC representation of the model has two purposes: first to estimate the parameters of the VEC itself; and second to estimate the parameters of the VMA exploiting $g(\theta, \alpha)$ in (20) and that the error is the same in both representations. Furthermore, we found that the rows of $\Theta(1)$ are identical and this implies that the efficient price of the first market is the same as the efficient price of the second market, or in other words that the two markets have a common efficient price. This is to be expected since the system is co-integrated CI(1,0). For this reason, going forwards we shift our attention from $\Theta(1)\eta_t$, that is the component of the price change permanently impounded into each market's efficient price, to $\theta\eta_t$, that is the component of the price change permanently impounded into either market's efficient price ($\text{pc}(\Delta p_t)|_{\text{Mkt}}$).[14]

[14] Appendix 7.A.6 gives a breakdown.

4 Price Discovery Leadership

This section joins together the model in Section 3 with the empirical evidence in Section 5 by deriving from the model a metric to measure price discovery leadership which can then be used in empirical applications. Let us start by defining the topic at hand. Price discovery is the dynamic process by which a market incorporates new fundamental information about an asset's value into the asset's price. A market is then said to lead the price discovery process over an another market if it incorporates a larger amount of the new information more quickly. In particular, market i leads the price discovery over market j if and only if i makes a greater contribution than j to the discovery of price. Furthermore, since prices are co-integrated $(1, 0)$ we may write $p_{k,t} = p_{k,t-1} + \Delta p_{k,t} = p_{t-1} + \Delta p_{k,t}$ which implies that informationally comparing prices $(p_{k,t})$ across markets is equivalent to comparing innovations $(\Delta p_{k,t})$ since the lagged term is the same for both markets.[15] Moreover, out of the innovation $\Delta p_{k,t}$ only its permanent component $(\text{pc}(\Delta p_{k,t}))$ is of interest informationally as its temporary component $(\text{tc}(\Delta p_{k,t}))$ eventually wears off and therefore does not impact the fundamental value of the asset (namely the common efficient price). Additionally, in Section 3.2 we showed the two rows of $\text{pc}(\Delta p_{k,t}) = \Theta(1)\eta_t = (\theta\eta_t, \theta\eta_t)'$ are identical so we may focus on either row $\theta\eta_t = \text{pc}(\Delta p_{k,t}|_{\text{Mkt}})$. In summary, market i leads the price discovery process over market j if and only if market i makes a greater contribution than market j to the component of the price change that is permanently impounded into the common efficient price.

Let ι be the object of interest, namely the component of the price change that is permanently impounded in the common efficient price, that is $\iota := \theta\eta_t$. Since the market that makes the greatest contribution to ι is said to lead the price discovery process, we need a metric $S_k(\iota)$ that measures the contribution $C_k(\iota)$ each market k makes to ι. It follows that the market with the greatest $S_k(\iota)$ leads the price discovery process. For a bivariate price series p_t, both $S_k(\iota)$ and $C_k(\iota)$ will vary depending on whether the correlation ρ of the two shocks $\eta_{1,t}$, $\eta_{2,t}$ is zero. In the case the two shocks are uncorrelated so that $\rho = \text{Corr}(\eta_{1,t}, \eta_{2,t}) = 0$ then the contribution each market makes

[15] $p_{k,t} \in \mathbb{R}$ is the price on market k at time t so that $\{p_t = (p_{1,t}, p_{2,t})\}_{t=1}^{\infty}$ is the bivariate stochastic process under consideration.

to $\iota = \theta\eta_t = \theta_1\eta_{1,t} + \theta_2\eta_{2,t}$ is simply:[16]

$$C_k^u = \theta_k\eta_{k,t} \tag{21}$$

Let us define the magnitude of each contribution as the variance of said contribution:[17]

$$M_k^u \coloneqq \text{Var}[C_k^u] = \text{Var}[\theta_k\eta_{k,t}] = \theta_k^2\sigma_k^2 \tag{22}$$

Finally, we define the metric to measure price discovery leadership as the normalised magnitude of the contribution each market makes to ι:[18]

$$S_k^u \coloneqq \frac{M_k^u}{M_1^u + M_2^u} = \frac{\theta_k^2\sigma_k^2}{\theta_1^2\sigma_1^2 + \theta_2^2\sigma_2^2} \tag{23}$$

The metric in (23) is known in the literature as the Information Share (IS) of market k for two uncorrelated markets, and the market with the greatest IS leads the price discovery process. Note the metric has been normalised so that the two information shares add up to unity ($S_1^u + S_2^u = 1$) which implies that the two markets are together considered to be fully responsible for the common efficient price of the asset. This is clearly a limitation as any liquid stock is traded on more than two trading centres all of which contribute to the price discovery. In other words, we may only use the IS here constructed as a relative measure of price discovery between the two markets under study, and should not be interpreted as an absolute measure of the total information that is permanently impounded in prices. Furthermore, we observe that while an asymptotic standard error for the information share can be computed from the estimated parameters, this is cumbersome to do in

[16]Equation (21):
$C_k^u \coloneqq$ the contribution market k makes to ι in the case when the shocks $\eta_{1,t}$ and $\eta_{2,t}$ are uncorrelated (hence the superscript u). Notice therefore that C_k^u is short-hand notation for $C_k^u(\iota)$.

[17]Equation(22):
$M_k^u \coloneqq$ the magnitude of the contribution market k makes to ι in the case when the shocks $\eta_{1,t}$ and $\eta_{2,t}$ are uncorrelated (hence the superscript u). Notice therefore that M_k^u is short-hand notation for $M_k^u(\iota)$;

$$\text{Cov}[\eta_t] = \Sigma \coloneqq \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

We take the variance as we measure the magnitude as the squared distance from the mean which is a measure of variation or squared volatility. There are, of course, other ways to measure the contribution's magnitude.

[18]Equation (23):
$S_k^u \coloneqq$ the normalised magnitude of the contribution market k makes to ι in the case when the shocks $\eta_{1,t}$ and $\eta_{2,t}$ are uncorrelated (hence the superscript u). Notice therefore that S_k^u is short-hand notation for $S_k^u(\iota)$.

practice so we carry out statistical inference by considering the cross-sectional estimates of the information share. Finally, we should also stress that our model does not attempt to determine neither which market has the "best" prices or quotes intended as the narrowest spread nor which market "moves first" in response to new information becoming available but rather that it seeks to measure which market "informationally moves first" in the process of price adjustment to the outbreak of news. Indeed, it is possible for a market to have the wider spreads and the slower adjustment of prices, and yet have the higher information share. This happens whenever the innovations in that market drive those in its competing markets making that market informationally dominant even if it comes at the expense of increased spreads and sluggish price adjustments.

Let us consider the case when the two markets are correlated, or in other words when $\rho = \text{Corr}(\eta_{1,t}, \eta_{2,t}) \neq 0$. Under this case, there exist many ways to compute each market's contribution to ι. We follow the literature developed by Hasbrouck (1995) but offer a different reasoning for the methodology implmented as instead of simply stating the metric used, we derive it from the model. This is carried out in Appendix 7.B which shows that the contribution to ι reduces to:

$$\begin{cases} C_i^{c,i} &= \theta_i \eta_{i,t} + \rho \theta_j \eta_{j,t} \\ C_j^{c,i} &= \sqrt{1-\rho^2}\theta_j \eta_{j,t} \end{cases} \tag{24}$$

Consistently with the case of uncorrelated shocks, we define the magnitude of each contribution under correlated shocks where market i affects market j but not vice-versa as the variance of the contribution:[19]

$$\begin{cases} M_i^{c,i} &:= \text{Var}[C_i^{c,i}] = (\theta_i \sigma_i + \rho \theta_j \sigma_j)^2 \\ M_j^{c,i} &:= \text{Var}[C_j^{c,i}] = (1-\rho^2)(\theta_j \sigma_j)^2 \end{cases} \tag{25}$$

Consequently, we define the metric to measure price discovery leadership (namely the Information

[19] Equation(25):
$M_k^{c,i} :=$ the magnitude of the contribution market $k \in \{i, j\} = \{1, 2\}$ makes to ι in the case when the shocks $\eta_{1,t}$ and $\eta_{2,t}$ are correlated (hence the superscript c) and under the scenario when shock $\eta_{i,t}$ affects shock $\eta_{j,t}$ but the converse does not hold (hence the superscript i). Notice therefore that $M_k^{c,i}$ is short-hand notation for $M_k^{c,i}(\iota)$.

Share) as the normalised magnitude of the contribution each market makes to ι:[20]

$$\begin{cases} S_i^{c,i} & := \quad \frac{M_i^{c,i}}{M_1^{c,i}+M_2^{c,i}} = \frac{(\theta_i\sigma_i+\rho\theta_j\sigma_j)^2}{(\theta_i\sigma_i+\rho\theta_j\sigma_j)^2+(1-\rho^2)(\theta_j\sigma_j)^2} \\ S_j^{c,i} & := \quad \frac{M_j^{c,i}}{M_1^{c,i}+M_2^{c,i}} = \frac{(1-\rho^2)(\theta_j\sigma_j)^2}{(\theta_i\sigma_i+\rho\theta_j\sigma_j)^2+(1-\rho^2)(\theta_j\sigma_j)^2} \end{cases} \tag{26}$$

Finally, having computed the Information Share in (26) assuming that $\eta_{j,t}$ does not affect shock $\eta_{i,t}$, we can also compute $(S_i^{c,j}, S_j^{c,j})$ under the assumption that $\eta_{i,t}$ does not affect shock $\eta_{j,t}$. We then define the Information Share for two correlated markets, that is two markets with correlated shocks, as the average of the information shares computed under the assumption that one market does not affect the other:

$$\begin{cases} S_i^c & := \quad \frac{S_i^{c,i}+S_i^{c,j}}{2} = \frac{1}{2}\left[\frac{(\theta_i\sigma_i+\rho\theta_j\sigma_j)^2}{(\theta_i\sigma_i+\rho\theta_j\sigma_j)^2+(1-\rho^2)(\theta_j\sigma_j)^2} + \frac{(1-\rho^2)(\theta_i\sigma_i)^2}{(1-\rho^2)(\theta_i\sigma_i)^2+(\theta_j\sigma_j+\rho\theta_i\sigma_i)^2}\right] \\ S_j^c & := \quad \frac{S_j^{c,i}+S_j^{c,j}}{2} = \frac{1}{2}\left[\frac{(1-\rho^2)(\theta_j\sigma_j)^2}{(\theta_i\sigma_i+\rho\theta_j\sigma_j)^2+(1-\rho^2)(\theta_j\sigma_j)^2} + \frac{(\rho\theta_i\sigma_i+\theta_j\sigma_j)^2}{(\rho\theta_i\sigma_i+\theta_j\sigma_j)^2+(1-\rho^2)(\theta_i\sigma_i)^2}\right] \end{cases} \tag{27}$$

The definition of Information Share for correlated markets in (27) takes the average of the two "pure" extreme scenarios where only one market affects the other so that it "averages out" to a scenario where both markets affect each other which is the likely outcome in any empirical application. While this definition is not the only one present in the literature, it is both empirically tractable and intuitive from a financial viewpoint. We are only left to write (27) as a function of the parameters estimated by the VEC representation of the model in (19) since the VMA representation of the model in (16) cannot be estimated in its current format as explained in Section 3.2. To do this we first divide the numerator and denominator of (27) by θ_j since this writes S_k^c ($k \in \{i,j\}$) as a function of $\frac{\theta_i}{\theta_j}$, and subsequently we substitute θ for α in (27) by exploiting the function $g(\theta,\alpha)$ in (20) which yields:

$$\begin{cases} S_i^c = \frac{1}{2}\left[\frac{(\alpha_i^\perp\sigma_i+\rho\alpha_j^\perp\sigma_j)^2}{(\alpha_i^\perp\sigma_i+\rho\alpha_j^\perp\sigma_j)^2+(1-\rho^2)(\alpha_j^\perp\sigma_j)^2} + \frac{(1-\rho^2)(\alpha_i^\perp\sigma_i)^2}{(1-\rho^2)(\alpha_i^\perp\sigma_i)^2+(\alpha_j^\perp\sigma_j+\rho\alpha_i^\perp\sigma_i)^2}\right] \\ S_j^c = \frac{1}{2}\left[\frac{(1-\rho^2)(\alpha_j^\perp\sigma_j)^2}{(\alpha_i^\perp\sigma_i+\rho\alpha_j^\perp\sigma_j)^2+(1-\rho^2)(\alpha_j^\perp\sigma_j)^2} + \frac{(\rho\alpha_i^\perp\sigma_i+\alpha_j^\perp\sigma_j)^2}{(\rho\alpha_i^\perp\sigma_i+\alpha_j^\perp\sigma_j)^2+(1-\rho^2)(\alpha_i^\perp\sigma_i)^2}\right] \end{cases} \tag{28}$$

[20] Equation(26):
$S_k^{c,i} :=$ the normalised magnitude of the contribution market $k \in \{i,j\} = \{1,2\}$ makes to ι in the case when the shocks $\eta_{1,t}$ and $\eta_{2,t}$ are correlated (hence the superscript c) and under the scenario when shock $\eta_{i,t}$ affects shock $\eta_{j,t}$ but the converse does not hold (hence the superscript i). Notice therefore that $S_k^{c,i}$ is short-hand notation for $S_k^{c,i}(\iota)$.

The Information Share (S_k^c) as defined in (28) can be computed from the estimated parameters (α, Σ) of the VEC representation of the model. Thus we may now use this definition of S_k^c in empirical applications such as the one in Section 5. Recall that the Information Share may only be used to compare markets since it does not measure the total information that is permanently impounded in prices but rather the information relative to the other platform. It is also worth noting that as expected the Information Share for correlated markets in (28) simplifies to that for uncorrelated markets in (23) if $\rho = 0$ (that is if the two markets are uncorrelated).

5 Empirical Evidence

This paper contributes to the price discovery literature by showing that leadership of the price discovery process needs not be homogeneous across the limit order book, or in other words that outer levels of LOB (intended as deep levels that are distant from the BBO) and inner levels (intended as levels that are close to the BBO such as levels 2 and 3 although the exact boundary between inner and outer levels is discretionary and depends on the context) may exhibit a switch in the directional leadership of the price discovery process. As a consequence price discovery cannot be limited to the BBO. We show that a switch in the directional leadership is possible by providing one example in the form of a case study. After discussing the dataset (Section 5.1) and descriptive statistics thereof (Section 5.2) we compute (Section 5.3) the information share (developed in Section 4) of the two markets under investigation to measure price discovery leadership under the model (discussed in Section 3) of a co-integrated price series. By showing that the market with the greatest information share changes as we sweep the LOB, we have provided evidence that it is indeed possible for the directional leadership to switch as we move away from the BBO.

5.1 Dataset

The dataset used records all transactions that take place on BATS Chi-X Europe between October 1, 2012 and March 28, 2013. In other words our dataset records every transaction down to the microsecond that is placed over those six months on any of the six books operated by BATS Chi-X Europe, which comprise a lit, dark and hidden order book for both CXE and BXE (a short

overview of BATS Chi-X Europe can be found in Appendix 7.C).[21] The transactions recorded include all market orders as well as limit orders, cancellations and hidden orders that take place during the continuous trading session which opens at 8:00 UK local time and closes at 16:30 UK local time. Out of the dataset at hand we focus on the two lit order books of CXE and BXE, and in particular we take the first 25 levels of both the ask and bid side of each book for a total of 50 quotes on each book to be recorded at each transaction. Given the size of the dataset we restrict our analysis to the FTSE100 with the exception of stocks that exited the index or moved to a different stock exchange over the sampling period and we also exclude stocks that are not class A or that were subject to a buyout leaving 80 stocks and approximately 1 terabyte of trading. The sample can be found in Appendix 7.E Tables 3.3-3.4.

The nature of the dataset requires to build the lit limit order book of CXE and that of BXE, and this is instrumental to this paper.[22] Indeed, to show a switch in the directional leadership of price discovery we need to compute the information share at each individual level of the book so that quotes for each level are required. Most datasets, even at intraday frequency, do not offer this feature as they limit themselves to providing the best bid-offer (that is the first level of the LOB) without any data on the other levels of the LOB.

In the subsequent analysis we focus on the limit orders posted to the lit order book as they are both necessary and sufficient for the computation of the information share, and we restrict ourselves to the first 25 levels of the book since this suffices to show the price discovery leadership changes as we swipe the LOB. Furthermore, we only briefly discuss the two other books, namely the dark and hidden order books, and orders that are not limit orders, such as market orders and cancellations, as part of the descriptive statistics of the dataset. A more detailed study of the hidden and dark order books is beyond the scope of this paper but constitutes an interesting topic of research and one that is scope for further research. We restrict ourselves to defining cancellations and hidden orders since this will be required in the next section: a cancellation is any order that cancels some

[21] Strictly speaking the hidden order book is part of the lit order book but the two books are treated as if they were distinct in order to emphasise their difference.

[22] The first step to building the lit limit order book of either CXE or BXE lies in de-coding each of the 12-character base-36 strings that identify an order submitted to the platform. The second step is to break our analysis by date, stock and book-side (ask or bid). Third, we match orders by visibility (lit, dark or hidden) and record orders as they are submitted to the platform thus gradually building the book.

of the shares placed on either side of the book by a preceding order (and therefore does not require the cancellation of the entire order - a point worth noticing); while a hidden order is a special type of limit order that conceals both the price and quantity to the other market participants.[23]

5.2 Descriptive statistics

We are primarily interested in descriptive statistics on the 25 levels of the LOB that we will use in the computation of the information share. Further descriptive statistics on the BBO (i.e. the first level of the book), on the BBO by price and market tercile, and on the dark and hidden order books are given in Appendix 7.D. To compute all descriptive statistics we divide the trading day into ten 51-minute time intervals starting from 8:00 (when the continuous trading session opens) and ending at 16:30 (when the continuous trading session closes). For each descriptive statistic over each time interval we first collapse observations in the same time interval to a single estimate, and we then take the median observation over stocks. How we collapse observations in the same time interval depends on the descriptive statistic: the number of limit orders, market orders and cancellations is counted; the executed volume and intraday volatility are summed (the latter is the squared root of summed squared returns); and the inside spread, BBO depth, average trade size, size of limit orders and size of market orders are averaged. Finally, we take the average of the two sides (ask or bid) of the book, and perform the above analysis for CXE and BXE (the two platforms run by BATS Chi-X) which we keep separate. Since both platforms exhibit the same statistical properties we limit ourselves to providing results for CXE in the remainder of this section.

Given our interest in the quotes of the first 25 levels of the book we start by analysing the number and average size of those limit orders that are responsible for the quotes of interest. We then compute the distribution of these limit orders measured as the product of the number and average size of the limit orders divided by the total number of shares posted to the LOB, and

[23] A hidden order (HO) is a special type of limit order with the following properties: a HO yields priority to visible orders thus priority follows a price-visibility-time rule; the HO's size must exceed the Large Order Threshold (LOT) set by BATS so that only large orders can be hidden orders; a HO only executes orders that exceed the Minimum Execution Size (MES) and this is an anti-gaming feature known as "pinging" designed to protect hidden orders from small orders which aim at discovering their presence. In summary, hidden orders are invisible limit orders, not included in the calculation of the BBO, which interact with both displayed orders and other hidden orders on the book, and have been designed to both reduce the cost of trading and unlock latent liquidity.

this metric helps determine where the bulk of the liquidity is at. Finally we give the cancellation-to-limit order ratio which divides the number of cancellations by the number of limit orders, to gauge the stability of the limit orders that are posted. To ease exposition we divide the 25 levels into 3 groups: we keep the BBO separate since this is the level researchers usually look at when studying the price discovery process, so that the BBO constitutes the first group; the second group consists of the inner levels 2 to 10 inclusive where levels are summed for the number, distribution and cancellation of limit orders while they are averaged for the size of limit orders; and the third group includes the remaining levels that is outer levels 11 to 25 and the descriptive statistics are grouped over these levels as per levels 2 to 10. A qualitative representation of these descriptive statistics is given in Figure 3.1 and a quantitative representation can be found in Table 3.1.

Table 3.1: Quantitative Descriptive Statistics by Book Level

The two-Sample Paired t-Test
$\boldsymbol{H_0 : \mathbb{E}[\text{Diff}_i] = 0}$ **VS** $\boldsymbol{H_1 : \mathbb{E}[\text{Diff}_i] \neq 0}$ $\boldsymbol{i \in \{1, 2\}}$

The Levels of the LOB	**Number of Limit Orders**			**Size of Limit Orders**		
	Mean	**Std.Err.**	**t-statistic**	**Mean**	**Std.Err.**	**t-statistic**
BBO	762	261	-	470	21	-
Levels #2-10	760	173	-	527	28	-
Levels #11-25	57	27	-	1026	140	-
Diff$_1$	2	35	0.07	-57	10	-5.98***
Diff$_2$	703	53	13.18***	-499	40	-12.36***
The Levels of the LOB	**Distribution of Limit Orders**			**Cancellation of Limit Orders**		
	Mean	**Std.Err.**	**t-statistic**	**Mean**	**Std.Err.**	**t-statistic**
BBO	45%	5%	-	0.90	0.04	-
Levels #2-10	47%	2%	-	1.31	0.12	-
Levels #11-25	8%	5%	-	1.68	0.14	-
Diff$_1$	-2%	2%	-1.39	-0.41	0.03	-15.18***
Diff$_2$	39%	2%	22.11***	-0.37	0.02	-16.69***

* if p-value < 0.1; ** if p-value < 0.05; *** if p-value < 0.01

This table reports the mean and standard error of the three groups (BBO, L2-10 and L11-25) plotted in Figure 3.1 for each of the four descriptive statistics (number, size, distribution and cancellation of limit orders) on the levels of the LOB of CXE. Additionally, this table also reports a two-sample paired t-test on the differences between the first and second group (Diff$_1$) and the second and third group (Diff$_2$) for each descriptive statistic.

Figure 3.1 displays the intraday dynamics of the number, average size, distribution and cancellation of limit orders over the trading day and by group level (BBO, L2-L10, and L11-25) on CXE. These results are tabulated in Table 3.1 where the mean and standard error of each group

Figure 3.1: Qualitative Descriptive Statistics by Book Level



This figure exhibits four descriptive statistics (number, size, distribution and cancellation of limit orders) on the first 25 levels of the LOB of CXE over the trading day. Levels are arranged into three groups: the BBO, levels 2-10 (also referred to as the inner levels), and levels 11-25 (also referred to as the outer levels). Descriptive statistics are summed over levels except for the size of limit orders which is averaged (i.e. the number of limit orders over levels 2-10 is the sum of the number of limit orders on each of these nine levels).

is given. The table also reports the mean, standard error and t-statistic of the difference between the first two groups, namely the BBO and Levels #2-10, under the name Diff_1, and the mean, standard error and t-statistic of the difference between the second (Levels #2-10) and third (Levels #11-25) group under the name Diff_2. The two differences Diff_i are the output of a two-sample paired t-test (which is a one-sample t-test on the differences of the paired observations) between groups 1 and 2 for Diff_1 and groups 2 and 3 for Diff_2. We therefore conclude from Table 3.1 that the number of limit orders is significantly higher on the BBO and levels 2 to 10 than on the levels 11-25 (since Diff_2 has a t-statistic of 13.18 which is significant to the 1% significance level) so the number of limit orders decreases as we move away from the BBO and inner levels (L2-10) to outer levels (L11-25). However the opposite is true about the size of limit orders since both Diff_1 and Diff_2 are negative and significant (-5.98 and -12.36 respectively) so limit orders posted to the BBO have smaller average size than those posted to inner levels (L2-10) which in turn are thinner than those posted to outer levels (L11-25). We conclude that as one moves away from the BBO limit orders become fewer but larger. If we combine size and frequency (and divide by the total number of shares posted to the LOB at that time of the day) we get the distribution of limit orders. We find that on average 45% of liquidity (i.e. all shares posted by limit orders) are posted to the BBO which is less than half, with a similar amount on inner levels, and some 10% on the outer levels. Therefore reducing the analysis of the price discovery process to the BBO misses the whole picture (55% of liquidity that is) if, as we show in the next section, a directional switch in the leadership of the price discovery is possible. Finally, the number of cancellations per number of limit orders is higher on outer levels than on inner levels which is higher than on the BBO (Diff_1 and Diff_2 are -15.18 and -16.69 respectively and both are statistically significant). This is likely due to the fact that as traders who post to the BBO are undercut, they subsequently cancel their stale orders on outer levels which are unlikely to be executed and place themselves on the BBO or closer to it. Moreover, the number of cancellations per limit order posted seems to be increasing over the trading day from Figure 3.1d, and this is possibly due to traders becoming more impatient as the day draws to a close. There are on average 0.9 cancellations for each limit order posted to the BBO compared against 1.7 on the outer levels L11-25 meaning that there are on average almost 2

cancellations (modifications) for each limit order posted to levels 11 to 25.

5.3 Empirical Evidence

In order to show that the leadership of the price discovery process between two trading centres can change over the levels of the LOB, we need to provide one example where this happens. Consider the two trading venues run by BATS Chi-X Europe, namely CXE and BXE, and out of the three order books each trading venue operates consider the lit order book. That is we aim at comparing the lit order books of BATS Chi-X Europe CXE and BATS Chi-X Europe BXE between October 1, 2012 and March 28, 2013. For each market we take the first 25 quotes at each order submission on both sides of the book giving a total of 100 quotes (2 markets (CXE or BXE) times 2 sides of the book (ask or bid) times 25 quotes). Since the dataset is at microsecond frequency we sample the limit order book at one-second frequency which suffices for the study of the price discovery process. We therefore have two price series (one for CXE and one for BXE) for each side s and each level l of the book. Assuming the model discussed in Section 3 and summarised in Equation (14) holds we can compute the information share for each statistic process $\{p_t^{s,l}\}_{t=1}^{T}$ using the formula derived in Section 4 and outlined in Equation (28).[24]

The information share is a relative measure of price informativeness so that we are only interested in the relative value of one market's information share compared against that of the other market and we have no interest in the absolute value of one individual market's information share. Therefore for side s and level l we define the ratio R_l^s of the two markets' information share as follows:[25]

$$R_l^s := \frac{S_{\text{CXE}}^{s,l}}{S_{\text{BXE}}^{s,l}} \tag{29}$$

It is apparent that whenever $R_l^s > 1$ then CXE on that side s and level l of the book has a greater

[24] $\{p_t^{s,l}\}_{t=1}^{T}$ is a bivariate stochastic process where:
$p_t^{s,l} = (p_{\text{CXE},t}^{s,l}, p_{\text{BXE},t}^{s,l}) \in \mathbb{R}^2$ is the price series on side $s \in \{\text{Ask}, \text{Bid}\}$ and level $l \in \{1, .., 25\}$ of the book;
$p_{k,t}^{s,l} \in \mathbb{R}$ is the quote at time t on market $k \in \{\text{CXE}, \text{BXE}\}$ side $s \in \{\text{Ask}, \text{Bid}\}$ level $l \in \{1, ..., 25\}$ of the book;
$T = 3,825,000$ is the number of seconds of trading over the sampling window (125 trading days x 8.5 daily hours of trading x 3,600 seconds).

[25] $S_k^{s,l}$ is the information share defined for correlated markets in Equation (28) where $k \in \{\text{CXE}, \text{BXE}\}$ is the market, $s \in \{\text{Ask}, \text{Bid}\}$ is the side of the book and $l \in \{1, ..., 25\}$ is the level of the book. Notice that since we are using the information share for correlated markets, $S_k^{s,l}$ is shorthand for $S_k^{c,s,l}$ where the superscript c is that used in Equation (28).

information share than BXE and therefore leads the price discovery process, while the opposite holds true whenever $R_l^s < 1$ that is BXE leads the price discovery process over CXE. Estimates of the information share ratio R_l^s are reported in Table 3.2.

Table 3.2: Information Share Ratio

R_l^s	**Level #1**	**Level #2**	**Level #3**	**Level #4**	**Level #5**
Ask	27.99***	32.78***	30.84***	34.65***	17.73***
Bid	12.69***	17.92***	21.43***	18.61***	9.52***
R_l^s	**Level #6**	**Level #7**	**Level #8**	**Level #9**	**Level #10**
Ask	10.56***	3.47***	1.94*	2.62	2.33
Bid	9.07***	6.30***	4.08***	1.63	0.91
R_l^s	**Level #11**	**Level #12**	**Level #13**	**Level #14**	**Level #15**
Ask	1.28	0.55	0.28*	0.23***	0.20***
Bid	0.41**	0.30**	0.31***	0.27***	0.27***
R_l^s	**Level #16**	**Level #17**	**Level #18**	**Level #19**	**Level #20**
Ask	0.11***	0.11***	0.17***	0.26***	0.13***
Bid	0.19***	0.21***	0.19***	0.13***	0.16***
R_l^s	**Level #21**	**Level #22**	**Level #23**	**Level #24**	**Level #25**
Ask	0.12***	0.14***	0.08***	0.07***	0.16***
Bid	0.25***	0.20***	0.12***	0.09***	0.10***

*This table reports the ratio R_l^s of the information shares of the two markets CXE and BXE on side s and level l of the limit order book. As an example the first (last) entry $R_1^{Ask} = 27.99$ ($R_{25}^{Bid} = 0.10$) conveys that CXE (BXE) is 27.99 (10) times more informative than BXE (CXE) on level 1 (25) of the ask (bid) side. The asterisks reflect the statistical significance of each estimate of R_l^s under a Wilcoxon t-Test on $H_0 : R_l^s = 1$ VS $H_1 : R_l^s \neq 1$. In particular, *** indicates 1% significance; ** indicates 5% significance; and * indicates 10% significance. As an example the first entry R_1^{Ask} is significant at the 1% significance level.*

Table 3.2 reports the estimated value of R_l^s for each side $s \in \{\text{Ask}, \text{Bid}\}$ and each level $l \in \{1, ...25\}$ of the limit order book so, for example, $R_1^{\text{Ask}} = 27.99$ implies that on level 1 of the ask side of the book the prices on CXE are almost 30 times more informative than those on BXE. To determine whether these estimates are statistically greater than one we could compute the asymptotic standard error and distribution of the test statistic R_l^s but since this is cumbersome to do we prefer to carry out statistical inference by means of the Wilcoxon t-Test on the null hypothesis that the quotes on CXE have the same distribution as the quotes on BXE or in other words on $H_0 : R_l^s = 1$ VS $H_1 : R_l^s \neq 1$.[26]

[26]The Wilcoxon t-Test, or Wilcoxon Signed-rank t-Test, is a non-parametric test of the null hypothesis that it is

The p-values of the Wilcoxon t-Test are computed for each side $s \in \{\text{Ask}, \text{Bid}\}$ and each level $l \in \{1, ...25\}$ of the limit order book and are reported in Appendix 7.E Table 3.5. The p-values are also summarised in Table 3.2 by means of asterisks (*** indicates 1% significance; ** indicates 5% significance; and * indicates 10% significance) so, for example, R_1^{Ask} is significantly greater than one at the 1% significance level and this implies that CXE leads the price discovery process over BXE at 1% significance. By sweeping through the levels of the LOB in Table 3.2 we conclude that CXE leads BXE on both sides of the book up to level 8 and that BXE leads CXE on both sides after level 13. We have therefore shown that a switch in the directional leadership of the price discovery process is indeed possible.

Finally we have assumed that prices on CXE and BXE can be modelled by a co-integrated CI$(1, 0)$ stochastic process with co-integrating vector $\beta = (1, -1)'$. We need to show this model fits our dataset for the above analysis to be justified, and we do so by testing the two assumptions required for the model to hold namely that the individual quoted prices are integrated of order 1 and that their difference is second-order stationary. For the former assumption we use the augmented Dickey-Fuller test of the null hypothesis that the price series has a unit root, or equivalently that it follows a random walk, against the alternative that the price series was generated by a stationary process (Dickey and Fuller (1979)). And for the latter assumption we employ the Johansen trace test of the null hypothesis that there are no more than r co-integrating equations, where r is the rank of the co-integrating vector, and iterate over r until we reach a value that is not rejected (Johansen (1995)). Neither of these tests when implemented by stock and quote level finds evidence against the assumptions of the model.

equally likely that a randomly selected value from one sample will be less than or greater than a randomly selected value from a second sample, and therefore can be used to establish whether two paired samples were selected from populations having the same distribution. In particular, the Wilcoxon t-Test tests the equality of matched pairs of observations by either using the median of their differences or by using the Wilcoxon matched-pairs (Snedecor and Cochran (1989) give a summary of both techniques). We prefer the former approach as we do not wish to assume that the distribution of R_l^s is symmetric about 1.

6 Conclusions

This paper's contribution to the literature on price discovery is twofold. First, it shows by means of a case study that a switch in the directional leadership of the price discovery process is possible over the levels of the LOB. In other words, levels close to the BBO (known as inner levels) may exhibit a different price discovery leadership from levels that are further away from the BBO (known as outer levels). This implies that the price discovery process cannot be limited to the BBO under the tacit assumption that whichever market leads the price discovery process at the BBO will also lead on all other levels of the LOB. A methodology to determine which trading venue as a whole leads the price discovery process, that is constructing a metric that captures all price levels of the limit order book, is interesting scope for future research. The empirical evidence provided to support that a switch in the directional leadership of the price discovery process is possible, consists of all trading activity that takes place on the first 25 levels (of both the ask and bid side) of the lit order books of the two trading venues CXE and BXE, run by BATS Chi-X Europe, over October 1, 2012 and March 28, 2013 and we find that CXE leads BXE over the first 8 levels while BXE leads CXE after level 13 so leadership indeed switches between levels 9 and 12. Additionally, less than 50% of liquidity is posted to the BBO on BATS Chi-X Europe and this corroborates that price discovery cannot be limited to the BBO.

The second contribution of this manuscript to the literature lies in providing a complete derivation of a market's information share, namely the metric used to measure a market's contribution to the component of the price change that is permanently impounded into the common efficient price and that is therefore used to measure a market's contribution to the discovery of price. This derivation stems from assuming that while individual market prices have a unit root, a linear combination thereof is second-order stationary, or in other words that the prices are co-integrated. The model has a vector moving average representation that is used for intuition and a vector error correction representation that is used for estimation, and linking the two representations of the model is the key to deriving the information share. While the latter theoretical contribution does not put forward something that did not already exist in the literature (the model was first developed by Hasbrouck (1995)), previous studies did not present this topic in its entirety.

7 Appendix

7.A Model

7.A.1 Motivating the Model

Mainstream economic theory such as ordinary least squares regression requires that the time series be second-order stationary (also known as covariance-stationary or simply stationary) that is the first two moments be finite and time-invariant. However this assumption may not hold true and in such a statistical environment a different approach needs to be implemented since conventional estimators have non-standard asymptotic distributions and different rates of convergence. One plausible question is to ask whether the time series instead of being second-order stationary in its levels is second-order stationary in its differences. Indeed many economic time series are first-difference stationary meaning that the level of the time series is not stationary itself but its first-difference is. Such first-difference stationary processes are also known as integrated processes of order 1, or I(1) processes. The most common example of an I(1) process is the simple random walk that is a stochastic variable $\{z_t\}_{t=1}^{\infty}$ that can be written as $z_t = z_{t-1} + \epsilon_t$ where ϵ_t is white noise.[27] More generally, a statistic process whose *d-th* difference is stationary is an integrated process of order d, or an I(d) process (thus making a second-order stationary process I(0)).

Consider dealing with not one but two I(1) time series: $\{y_{1,t}\}_{t=1}^{\infty}$ and $\{y_{2,t}\}_{t=1}^{\infty}$. We understand the dynamics of each individual process since it is I(1) so we are interested in how the two series relate to each other or in other words what can be said about the linear combination $\lambda_1 y_{1,t} + \lambda_2 y_{2,t}$ for $\lambda_1, \lambda_2 \in \mathbb{R}$. Granger (1981) pointed out that the equation $\lambda_1 y_{1,t} + \lambda_2 y_{2,t} = e_t$ must "balance" in the order of integration that is that the process on each side of the equal sign must have the same order of integration. This gives three possibilities for e_t: $\{e_t\}_{t=1}^{\infty} \sim$ I(1), I(0), or not integrated at all. In the latter scenario, there is little that can be done and we must try a different linear combination of $y_{1,t}$ and $y_{2,t}$; in the first case ($\{e_t\}_{t=1}^{\infty} \sim$ I(1)), Phillips and Durlauf (1986) derive

[27] A white noise process is a sequence of independent and identically distributed random variables with zero mean and a finite variance σ^2;

We readily establish that z_t is an I(1) process since although its expected value $\mathbb{E}[z_t] = 0$ is independent of time, its variance $\text{Var}[z_t] = t \times \sigma^2$ is not time invariant and therefore $\{z_t\}_{t=1}^{\infty}$ is not second-order stationary itself but its first-difference is stationary since $\{\Delta z_t\}_{t=1}^{\infty} = \{z_t - z_{t-1}\}_{t=1}^{\infty} = \{\epsilon_t\}_{t=1}^{\infty} \sim$ I(0).

the asymptotic theory for the OLS estimator though it has not been widely used in applied work; and the case $\{e_t\}_{t=1}^{\infty} \sim \text{I}(0)$ is by far the most interesting so much so that it has a name in its own right and is called co-integration. Two variables are said to be co-integrated with order of co-integration (1,0) that is $\{(y_{1,t}, y_{2,t})\}_{t=1}^{\infty} \sim \text{CI}(1,0)$ if and only if $\{y_{1,t}\}_{t=1}^{\infty}, \{y_{2,t}\}_{t=1}^{\infty} \sim \text{I}(1)$ & $\exists\ (\lambda_1, \lambda_2) \in \mathbb{R}^2$ s.t. $\{\lambda_1 y_{1,t} + \lambda_2 y_{2,t}\}_{t=1}^{\infty} \sim \text{I}(0)$. The vector $\lambda = (\lambda_1, \lambda_2)$ is called the co-integrating vector. In Section 3 we pursue co-integration analysis since it provides a framework for estimation, inference and interpretation when the variables of interest are not second-order stationary but a linear combination of them is. Notice that the simple regression of $\Delta y_{1,t}$ on $\Delta y_{2,t}$ (since both terms are stationary) is mis-specified since it omits the variable λy_t (shown in Section 3.2).

7.A.2 Vector Moving Average - Beveridge-Nelson Representation Theorem

Consider the Beveridge-Nelson Representation Theorem (Beveridge and Nelson (1981)): *Let* $\{x_t\}_{t=1}^{\infty}$ *be a bivariate linear stochastic process such that* $x_t = \sum_{i=0}^{\infty} \Phi_i \epsilon_{t-i}$ *with* $\{\epsilon_t\}_{t=1}^{\infty}$ *a zero-mean second order stationary process whose elements* $\epsilon_{1,t}$, $\epsilon_{2,t}$ *are serially uncorrelated, and* $\{\Phi_i\}_{i=0}^{\infty}$ *a sequence of 2 x 2 scalar matrices whose elements are summable in the sense that* $\sum_{i=0}^{\infty} i|\phi_i| < \infty$. *Then:*

$$\sum_{s=1}^{t} x_s = \Phi(1) \sum_{s=1}^{t} \epsilon_s + \xi_t - \xi_0 \tag{30}$$

where $\Phi(1) = \sum_{i=0}^{\infty} \Phi_i$, $\xi_t = \sum_{i=0}^{\infty} \Lambda_i \epsilon_{t-i}$, $\Lambda_i = -\sum_{j=1}^{\infty} \Phi_{i+j}$, *and* $\sum_{i=0}^{\infty} |\Lambda_i| < \infty$.

The stochastic process $\{\Delta p_t\}_{t=1}^{\infty}$ from the VMA representation in (15) does not satisfy in its current format the assumptions of the Beveridge-Nelson Representation Theorem (BNRT) since it may have a non-zero mean τ. We therefore apply the BNRT to $\left\{\widetilde{\Delta p_t}\right\}_{t=1}^{\infty} = \{\Delta p_t - \tau\}_{t=1}^{\infty}$ which is a bivariate linear stochastic process with the properties required by the BNRT since Cesaro convergence is a stronger requirement than that needed for the theorem. Applying the BNRT to

$\left\{\widetilde{\Delta p_t}\right\}_{t=1}^{\infty}$ yields:[28,29,30]

$$\implies \quad \sum_{s=1}^{t} \widetilde{\Delta p_s} = \Theta(1) \sum_{s=1}^{t} \eta_s + \zeta_t - \zeta_0 \tag{31}$$

$$\implies \quad \sum_{s=1}^{t} [\Delta p_t - \tau] = \Theta(1) \sum_{s=1}^{t} \eta_s + \sum_{k=0}^{\infty} \Theta_k^* \eta_{t-k} - \sum_{k=0}^{\infty} \Theta_k^* \eta_{-k} \tag{32}$$

$$\implies \quad p_t - p_0 - \tau t = \Theta(1) \sum_{s=1}^{t} \eta_s + \Theta^*(L)\eta_t - \Theta_0^* \eta_0 \tag{33}$$

$$\implies \quad p_t = \widehat{p_0} + (\mu + \tau t) + \Theta(1) \sum_{s=1}^{t} \eta_s + \Theta^*(L)\eta_t \tag{34}$$

Finally, we summarise the solution to the VMA representation of the model in (16) as:

$$\begin{cases} p_t = \widehat{p_0} + (\mu + \tau t) + \Theta(1) \sum_{s=1}^{t} \eta_s + \Theta^*(L)\eta_t \text{ s.t.} \\ \widehat{p_0} = p_0 - \mu - \Theta_0^* \eta_0, \\ \mathbb{E}_t[p_t] = \mu + \tau t, \\ \Theta(1) = \sum_{k=0}^{\infty} \Theta_k, \\ \Theta^*(L) = \sum_{k=0}^{\infty} \Theta_k^* L^k, \\ \Theta_k^* = -\sum_{j=1}^{\infty} \Theta_{k+j}, \\ \sum_{k=0}^{\infty} |\Theta_k^*| < \infty. \end{cases} \tag{17}$$

[28]Equation (31):
$\Theta(L) = \sum_{k=0}^{\infty} \Theta_k L^k \in \mathbb{R}^2$ x $\mathbb{R}^2$ is a matrix polynomial in the lag operator L and therefore $\Theta(1) = \sum_{k=0}^{\infty} \Theta_k$;
$\{\zeta_t\}_{t=1}^{\infty}$ is a bivariate stochastic process such that $\zeta_t = \sum_{k=0}^{\infty} \Theta_k^* \eta_{t-k}$ with $\{\Theta_k^*\}_{k=0}^{\infty}$ a sequence of 2 x 2 scalar matrices.

[29]Equation (33):
$\Theta^*(L) = \sum_{k=0}^{\infty} \Theta_k^* L^k \in \mathbb{R}^2$ x $\mathbb{R}^2$ is a matrix polynomial in the lag operator L;
$\{\eta_t\}_{t=1}^{\infty}$ is the bivariate second-order stationary stochastic process, with $\eta_0 \coloneqq 1 \in \mathbb{R}^2$ and $\eta_t \coloneqq 0 \in \mathbb{R}^2$ if $t < 0$, from Equation (15) since $\{\Delta p_t\}_{t=1}^{\infty}$ and $\left\{\widetilde{\Delta p_t}\right\}_{t=1}^{\infty}$ have the same error.

[30]Equation (34):
$\widehat{p_0} \coloneqq p_0 - \mu - \Theta_0^* \eta_0 \in \mathbb{R}^2$ is a vector of adjusted initial values.

7.A.3 Vector Moving Average - Proof of Equation (18)

Let $\Theta_i(1)$ be the *i-th* column of $\Theta(1)$. Then:[31]

$$\begin{aligned}
\Theta_i(1) &= \left[\sum_{k=0}^{\infty}\Theta_k\right]_i && (35)\\
&= \sum_{k=0}^{\infty}\Theta_{i,k} && (36)\\
&= \lim_{j\to\infty}\sum_{k=0}^{j}\Theta_{i,k} && (37)\\
&= \lim_{j\to\infty}\sum_{k=0}^{j}\mathbb{E}_t\left[\frac{\partial}{\partial\eta_{i,t}}\left\{(\Theta_{1,k},\Theta_{2,k})(\eta_{1,t},\eta_{2,t})'\right\}\right] && (38)\\
&= \lim_{j\to\infty}\sum_{k=0}^{j}\mathbb{E}_t\left[\frac{\partial}{\partial\eta_{i,t}}\Theta_k\eta_t\right] && (39)\\
&= \lim_{j\to\infty}\sum_{k=0}^{j}\mathbb{E}_t\left[\frac{\partial}{\partial\eta_{i,t}}\sum_{l=0}^{\infty}\Theta_l\eta_{t+k-l}\right] && (40)\\
&= \lim_{j\to\infty}\sum_{k=0}^{j}\mathbb{E}_t\left[\frac{\partial}{\partial\eta_{i,t}}\Delta p_{t+k}\right] && (41)\\
&= \frac{\partial}{\partial\eta_{i,t}}\lim_{j\to\infty}\mathbb{E}_t\left[\sum_{k=0}^{j}\Delta p_{t+k}\right] && (42)\\
&= \frac{\partial}{\partial\eta_{i,t}}\lim_{j\to\infty}\mathbb{E}_t\left[p_{t+j}\right] && (43)
\end{aligned}$$

Therefore

$$\Theta(1) \coloneqq (\Theta_1(1),\Theta_2(1)) = (\frac{\partial}{\partial\eta_{1,t}}\lim_{k\to\infty}\mathbb{E}_t[p_{t+k}], \frac{\partial}{\partial\eta_{2,t}}\lim_{k\to\infty}\mathbb{E}_t[p_{t+k}]) =: \frac{\partial}{\partial\eta_t}\lim_{k\to\infty}\mathbb{E}_t[p_{t+k}] \qquad (44)$$

which completes the proof.

[31] Equation (35) by definition of $\Theta_i(1)$;
Equation (36) by $\Theta_{i,k} \coloneqq$ *i-th* column of Θ_k;
Equation (38) by definition of partial differentiation;
Equation (40) by assumption that η_t is serially uncorrelated;
Equation (41) by VMA representation in (15);
Equation (42) by Theorem 7.17 of Walter Rudin's Principle of Mathematical Analysis (3rd edition) which shows that limit and derivative can be exchanged since $\{\Delta p_t\}_{t=1}^{\infty}$ converges uniformly;
Equation (43) by summation of the telescopic sum and noting p_{t-1} is known at time t so it can be taken out of the expectation operator and is then suppressed by the partial differentiation operator.

7.A.4 Vector Error Correction - Simplified Granger Representation Theorem

Consider the Simplified Granger Representation Theorem (the complete theorem is given in Engle and Granger (1987)): *Let* $\{x_t\}_{t=1}^{\infty}$ *be a bivariate purely non-deterministic stochastic process such that* $\{x_t\}_{t=1}^{\infty} \sim \text{CI}(1,0)$ *with co-integrating vector* $\alpha \in \mathbb{R}^2$. *Then the following vector error correction representation exists:*

$$A(L)(1-L)x_t = -\gamma z_{t-1} + d(L)\epsilon_t \tag{45}$$

where $A(L)$ *is a matrix polynomial in the lag operator* L *with* $A(0) = I_2$, $\gamma \in \mathbb{R}^2$, $z_t = \alpha' x_t$ *is second-order stationary,* $d(L)$ *is a scalar polynomial in the lag operator* L *with* $d(1)$ *finite, and* $\{\epsilon_t\}_{t=1}^{\infty}$ *is a bivariate second-order stationary process such that* $\Delta x_t = \Omega(L)\epsilon_t$ *for* $\Omega(L) = \sum_{k=0}^{\infty} \Omega_k L^k \in \mathbb{R}^2 \; x \; \mathbb{R}^2$ *a matrix polynomial in the lag operator* L. *Furthermore, if* $A(L)$ *can be written as a finite matrix polynomial then* $d(L) = 1$.

The stochastic process $\{p_t\}_{t=1}^{\infty}$ from the model in (14) does not satisfy in its current format the assumptions of the Simplified Granger Representation Theorem (SGRT) since it may have a non-zero deterministic trend of the form $\mu + \tau t$ (higher order terms are not possible due to the stationarity of the first-order price-difference Δp_t). We therefore apply the SGRT to $\{\widetilde{p}_t\}_{t=1}^{\infty} = \{p_t - \mathbb{E}_t[p_t]\}_{t=1}^{\infty} = \{p_t - (\mu + \tau t)\}_{t=1}^{\infty}$ which is both purely non-deterministic and CI(1,0) as required by the SGRT. Additionally, we assume that a finite vector autoregressive representation for the left-hand side of (45) exists or in other words that the infinite matrix polynomial can be approximated by one with a finite number of terms K. Applying the SGRT to $\{\widetilde{p}_t\}_{t=1}^{\infty}$ and assuming a finite

autoregression yields:[32,33]

$$\implies \quad \Gamma(L)(1-L)\widetilde{p_t} = -\alpha p^*_{t-1} + d(L)\eta_t \tag{46}$$

$$\implies \quad \left[\sum_{k=0}^{K} \Gamma_k L^k\right] \Delta\widetilde{p_t} = -\alpha\left(\beta'\widetilde{p_{t-1}}\right) + \eta_t \tag{47}$$

$$\implies \quad \left[I_2 + \sum_{k=1}^{K} \Gamma_k L^k\right] \Delta\left(p_t - \mu - \tau t\right) = -\alpha\beta'\left(p_{t-1} - \mu - \tau(t-1)\right) + \eta_t \tag{48}$$

$$\implies \quad \Delta p_t = -\alpha(\beta' p_{t-1} + \widehat{\mu} + \widehat{\tau} t) - \sum_{k=1}^{K} \Gamma_k \Delta p_{t-k} + \eta_t \tag{49}$$

Finally, we summarise the VEC representation from the model in (14) as:

$$\begin{cases} \Delta p_t \;=\; -\alpha(\beta' p_{t-1} + \widehat{\mu} + \widehat{\tau} t) - \sum_{k=1}^{K} \Gamma_k \Delta p_{t-k} + \eta_t \text{ s.t.} \\ \qquad\quad \widehat{\mu}, \widehat{\tau} \in \mathbb{R},\ \alpha \in \mathbb{R}^2,\ \Gamma_k \in \mathbb{R}^2 \text{ x } \mathbb{R}^2 \ \forall\ k, \\ \qquad\quad \beta = (1,-1)',\ \eta_t \text{ is the error in VMA (16).} \end{cases} \tag{19}$$

7.A.5 Vector Error Correction - Proof of Equation (20)

The parameter $\Theta(1)$ can be calculated following Johansen (1991) who shows that:[34]

$$\Theta(1) \;=\; \beta_\perp \Upsilon \alpha'_\perp \text{ s.t.} \tag{50}$$

$$\Upsilon \;=\; \left(\alpha'_\perp \left(I_2 - \sum_{k=1}^{K} \Gamma_k\right) \beta_\perp\right)^{-1} \tag{51}$$

---

[32]Equation (46):
$\Gamma(L) = \sum_{k=0}^{K} \Gamma_k L^k \in \mathbb{R}^2$ x $\mathbb{R}^2$ is a matrix polynomial in the lag operator L with $\{\Gamma_k\}_{k=0}^{K}$ a sequence of 2 x 2 scalar matrices satisfying $\Gamma(0) = \Gamma_0 = I_2$;
$\alpha \in \mathbb{R}^2$;
$\{p^*_t\}_{t=1}^{\infty}$ is a bivariate second-order stationary stochastic process with $p^*_t = \beta'\widetilde{p_t}$ where $\beta = (1,-1)'$ is the co-integrating vector from the model in (14);
$d(L) = 1 \in \mathbb{R}$ since $\Gamma(L)$ can be written as a finite matrix polynomial (this is the additional assumption we make when we say "$A(L)$ can be written as a finite matrix polynomial");
$\{\eta_t\}_{t=1}^{\infty}$ is the bivariate second-order stationary stochastic process with $\eta_t = (\eta_{1,t}, \eta_{2,t}) \in \mathbb{R}^2$, $\mathbb{E}[\eta_t] = 0$, $\text{Cov}[\eta_t] = \Sigma < \infty$, and η_1, η_2 serially uncorrelated (but may be cross-correlated), from Eq. (17). Note that η_t is the error to $\widetilde{p_t}$.
[33]Equation (57):
$\widehat{\mu} = \beta'(\tau - \mu) + \widehat{\widehat{\mu}} \in \mathbb{R}$; $\widehat{\tau} = -\beta'\tau \in \mathbb{R}$ as $\widehat{\tau} \in \mathbb{R}$; $\widehat{\widehat{\mu}}$ is such that $-\alpha\widehat{\widehat{\mu}} = \sum_{k=0}^{K} \Gamma_k \tau \in \mathbb{R}^2$ as $\tau \in \mathbb{R}^2$.
[34]Equation (50):
$\beta_\perp = (1,1)' \in \mathbb{R}^2$ is the orthogonal vector to β;
$\alpha_\perp = (\alpha_1^\perp, \alpha_2^\perp)' \in \mathbb{R}^2$ is the orthogonal vector to α.

Subsequently, Baillie et al. (2002) demonstrate that the calculation in (50)-(51) simplifies under the model in (14), and this is because under a bivariate price series, Υ in (51) reduces to a scalar which is therefore commutative and when substituted back into (50) yields:

$$\Theta(1) = \beta_{\perp}\Upsilon\alpha'_{\perp} = \Upsilon\beta_{\perp}\alpha'_{\perp} = \Upsilon \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} \alpha_1^{\perp} & \alpha_2^{\perp} \end{pmatrix} = \Upsilon \begin{pmatrix} \alpha_1^{\perp} & \alpha_2^{\perp} \\ \alpha_1^{\perp} & \alpha_2^{\perp} \end{pmatrix} \tag{52}$$

We may simplify the computation of $\Theta(1)$ further by noticing from (52) that the rows of $\Theta(1)$ must be identical since Υ is a scalar and therefore we may write:

$$\begin{pmatrix} \theta_1 & \theta_2 \\ \theta_1 & \theta_2 \end{pmatrix} = \Theta(1) = \Upsilon \begin{pmatrix} \alpha_1^{\perp} & \alpha_2^{\perp} \\ \alpha_1^{\perp} & \alpha_2^{\perp} \end{pmatrix} \tag{53}$$

Equating the left-hand side and right-hand side of (53) by matrix entry and then dividing the resulting equations gives:[35]

$$\begin{cases} \theta_1 = \Upsilon\alpha_1^{\perp} \\ \theta_2 = \Upsilon\alpha_2^{\perp} \end{cases} \implies \frac{\theta_1}{\theta_2} = \frac{\alpha_1^{\perp}}{\alpha_2^{\perp}} \implies g(\theta, \alpha) = 0 \tag{20}$$

The function $g(\theta, \alpha)$ in (20) suffices for our purposes of the price discovery process since we only deal with two markets (otherwise Υ is not a scalar), and we can exploit $g(\theta, \alpha)$ to compute θ (and therefore $\Theta(1) = (\theta, \theta)'$) from α in place of the more general computation in (50)-(51).

7.A.6 Vector Error Correction - Common Efficient Price

First, let us adopt the following notation:

$$\begin{cases} \theta_{\text{Mkt}_1} & \coloneqq & (\theta_1, \theta_2) \in \mathbb{R} \text{ x } \mathbb{R} \text{ is the first row of } \Theta(1); \\ \theta_{\text{Mkt}_2} & \coloneqq & (\theta_3, \theta_4) \in \mathbb{R} \text{ x } \mathbb{R} \text{ is the second row of } \Theta(1) \end{cases} \tag{54}$$

[35] Equation (20):
$\theta \coloneqq (\theta_1, \theta_2) \in \mathbb{R}$ x $\mathbb{R}$ is either row of $\Theta(1)$;
$g(\theta, \alpha) \coloneqq h(\theta', \alpha_{\perp}) = \frac{\theta_1}{\theta_2} - \frac{\alpha_1^{\perp}}{\alpha_2^{\perp}}$ with $h\colon \mathbb{R}^2$ x $\mathbb{R}^2 \to \mathbb{R}$.

Second, we can write $\Theta(1)$ as:

$$\begin{aligned}
\Theta(1) &= \begin{pmatrix} \theta_{\text{Mkt}_1} \\ \theta_{\text{Mkt}_2} \end{pmatrix} = \begin{pmatrix} \theta_1 & \theta_2 \\ \theta_3 & \theta_4 \end{pmatrix} \in \mathbb{R}^2 \text{ x } \mathbb{R}^2 \text{ is the matrix polynomial defined in (31)} \\
&= \begin{pmatrix} \theta_1 & \theta_2 \\ \theta_1 & \theta_2 \end{pmatrix} \text{ since rows of } \Theta(1) \text{ are identical by (52)} \\
&= \begin{pmatrix} \theta \\ \theta \end{pmatrix} \text{ where } \theta \text{ is the 2 x 1 vector defined in (20)}
\end{aligned}$$

Third, and finally, by taking advantage of $\text{pc}(\Delta p_t)|_{\text{Mkt}_1} = \text{pc}(\Delta p_t)|_{\text{Mkt}_2}$, we define $\text{pc}(\Delta p_t)|_{\text{Mkt}}$ as follows:

$$\begin{pmatrix} \text{pc}(\Delta p_t)|_{\text{Mkt}_1} \\ \text{pc}(\Delta p_t)|_{\text{Mkt}_2} \end{pmatrix} \coloneqq \text{pc}(\Delta p_t) = \Theta(1)\eta_t = \begin{pmatrix} \theta \\ \theta \end{pmatrix} \eta_t = \begin{pmatrix} \theta\eta_t \\ \theta\eta_t \end{pmatrix} \tag{55}$$

$$\implies \quad \text{pc}(\Delta p_t)|_{\text{Mkt}_i} = \theta\eta_t \text{ for } i \in \{1, 2\} \tag{56}$$

$$\implies \quad \text{pc}(\Delta p_t)|_{\text{Mkt}} \coloneqq \text{pc}(\Delta p_t)|_{\text{Mkt}_i} = \theta\eta_t \tag{57}$$

7.B Price Discovery Leadership

Let us start by assuming that shock $\eta_{j,t}$ does not affect shock $\eta_{i,t}$. Under this scenario, the opposite, i.e. shock $\eta_{i,t}$ affects $\eta_{j,t}$, must hold since otherwise the two shocks would be uncorrelated which is not the case under consideration. Then the contribution to ι reduces to:[36]

$$\begin{cases} C_i^{c,i} &= \theta_i\eta_{i,t} + \rho\theta_j\eta_{j,t} \\ C_j^{c,i} &= f(\rho)\theta_j\eta_{j,t} \end{cases} \tag{58}$$

[36] Equation (58):
$C_k^{c,i} \coloneqq$ the contribution market $k \in \{i, j\} = \{1, 2\}$ makes to ι in the case when the shocks $\eta_{1,t}$ and $\eta_{2,t}$ are correlated (hence the superscript c) and under the scenario when shock $\eta_{i,t}$ affects shock $\eta_{j,t}$ but the converse does not hold (hence the superscript i). Notice therefore that $C_k^{c,i}$ is short-hand notation for $C_k^{c,i}(\iota)$. Furthermore, shock k is short for $\eta_{k,t}$;
$f(\rho) \coloneqq$ any decreasing function of ρ with $f\colon \mathbb{R} \to \mathbb{R}$.

The reason we can break down $C_i^{c,i}$ as in Equation (58) is that the shock in market i affects ι both directly (i.e. through $\theta_i\eta_{i,t}$) and indirectly via the shock in market j (i.e. through $\theta_j\eta_{j,t}$) where the strength of the latter effect is determined by the degree of correlation between the two markets (i.e ρ). Combining the direct effect $\theta_i\eta_{i,t}$ and the indirect effect $\rho\theta_j\eta_{j,t}$ gives the result for $C_i^{c,i}$ in Equation (58). For the contribution market j makes to ι ($C_j^{c,i}$) there is no indirect effect through market i since under the current scenario shock j does not affect shock i, and therefore only the direct effect of $\eta_{j,t}$ on ι (i.e. $\theta_j\eta_{j,t}$) remains. However, market j is only partly (i.e. $f(\rho)$) responsible for the direct effect of $\eta_{j,t}$ on ι (i.e. through $f(\rho)\theta_j\eta_{j,t}$) given market i is responsible for the other part (i.e. through $\rho\theta_j\eta_{j,t}$). Thus $C_j^{c,i}$ can be broken down as in Equation (58) as the fraction $f(\rho)\theta_j\eta_{j,t}$ of the effect shock j has on ι. Furthermore, let us assume that the shock in market j can be written as $\eta_{j,t} = \alpha + \beta + \gamma \in \mathbb{R}$ where the first term α is due to market j, the second term β is due to market i, and the third term γ is due to exogenous factors. Since α and β are orthogonal, and γ is non-stochastic we may write $\text{Var}[\eta_{j,t}] = \text{Var}[\alpha] + \text{Var}[\beta]$. Additionally, the component of $\eta_{j,t}$ due to market j is $f(\rho)\eta_{j,t}$ and the component of $\eta_{j,t}$ due to market i is $\rho\eta_{j,t}$ which we can use to substitute for α and β giving $1 = f(\rho)^2 + \rho^2$ or after re-arranging $f(\rho) = \sqrt{1-\rho^2}$. Substituting for $f(\rho)$ in Equation (58) yields:

$$\begin{cases} C_i^{c,i} &= \theta_i\eta_{i,t} + \rho\theta_j\eta_{j,t} \\ C_j^{c,i} &= \sqrt{1-\rho^2}\theta_j\eta_{j,t} \end{cases} \tag{24}$$

7.C BATS Chi-X Europe

BATS Global Markets was founded in 2005 in the U.S. by David Cummings, a computer programmer, in the belief that the ongoing consolidation within the industry led by NASDAQ and NYSE was leading to higher trading fees. His aim was for BATS to apply lower fees and to become "a neutral, private, broker-dealer owned, semi-profitable utility". Other brokerage firms, hedge funds and institutional investors gradually became involved with the company with no party owning more than 20%. In 2008 BATS entered the European equity markets by establishing a Multilateral Trading Facility (MTF), named BATS Europe, to compete on a pan-european basis against the incumbent exchanges and in February 2011 it bought Chi-X Europe, its main compet-

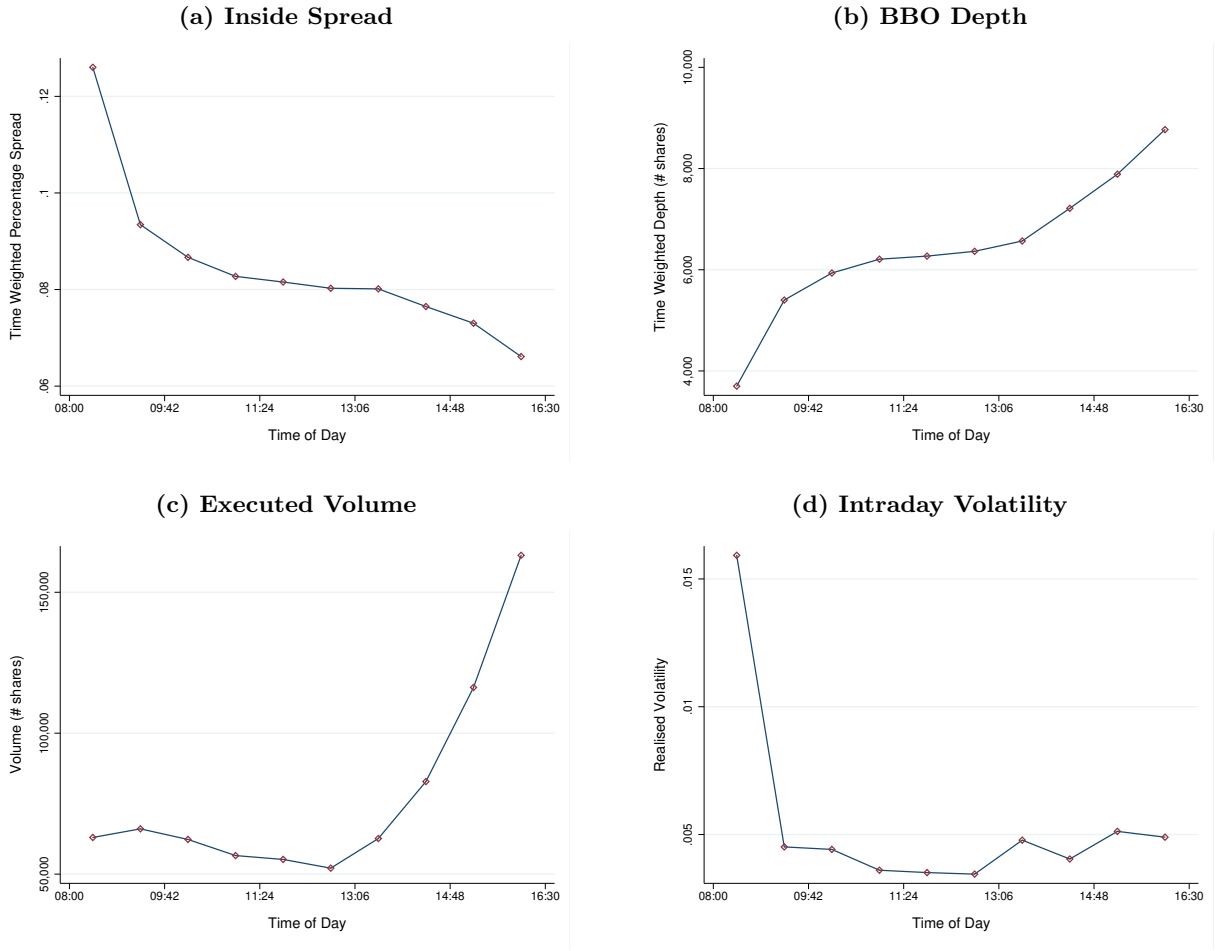
itor and largest pan-european MTF. As a result of this acquisition the new entity, named BATS Chi-X Europe, became the largest MTF in Europe by market share and traded volume offering trading in more than 3,600 securities across some 25 indices and 15 markets and it also offers trading in Exchange-Traded Funds, Equipment Trust Certificates and International Depository Receipts. Noticeably, in 2013 BATS Chi-X Europe became the first MTF to obtain full exchange status as it gained the status of Recognised Investment Exchange by the Financial Conduct Authority. Currently BATS Chi-X Europe operates six order books, namely CXE and BXE lit order books, CXE and BXE dark order books, and CXE and BXE hidden order books, that altogether account for approximately 25% of daily traded volume in European equities and alternate with the LSE for the title of Europe's largest share trading platform. This achievement is even more remarkable when one considers that previous entry attempts into the European market have largely failed.

7.D Additional Descriptive Statistics

Additionally to the descriptive statistics on the first 25 levels of the book described in the main body of the paper, we present three supplementary sets of descriptive statistics. These are not directly related to the study of the price discovery process so we only give a brief and qualitative overview. Once again we only report results for CXE since BXE is statistically equivalent. Figure 3.2 exhibits four descriptive statistics on the BBO of the LOB. In particular, the inside spread (Figure 3.2a) measured as the time weighted percentage spread appears to fall (and therefore improve) over the trading day. The BBO depth (Figure 3.2b) displays the opposite trend to that observed on the inside spread since the time weighted depth increases over time once again showing an improvement as the day draws to a close. Thus market quality measured as inside spread and BBO depth seems to improve over time. Executed volume (Figure 3.2c) appears to increase significantly towards the end of the trading day, while intraday volatility (Figure 3.2d), measured as the realised volatility over each 51-minute time interval, seems to be left-skewed U-shaped that is U-shaped if we exclude the first time intervals which has an exceptionally large volatility.

If we break up each descriptive statistic in Figure 3.2 by the price or the market capitalisation of the 80 stocks in our sample we gain insight into how the stock price and market capitalisation affect

Figure 3.2: The Best Bid-Offer of the Limit Order Book



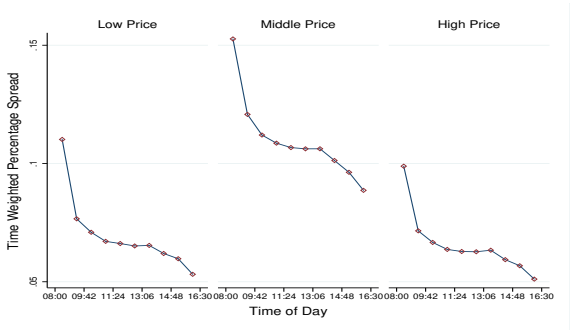
This figure exhibits four descriptive statistics (inside spread, BBO depth, executed volume and intraday volatility) for the Best Bid-Offer of the Limit Order Book on CXE over the trading day.

trading at the BBO. We group stocks into terciles first by the closing price on September 3, 2012 and second by the closing market capitalisation on the same day. Close stock prices on September 3, 2012 range from £0.12 (Vedanta Resources plc) to £74.30 (Randgold Resources plc) and the two threshold prices are the 33rd percentile at £5.53 (RSA Insurance Group plc) and the 66th percentile at £13.80 (Tullow Oil plc). Correspondingly, close market capitalisations on September 3, 2012 range from £2.1 billion (Melrose Industries plc) to £101.0 billion (HSBC Holdings plc) and the two threshold market capitalisations are the 33rd percentile at £4.8 billion (British Land Company plc) and the 66th percentile at £11.1 billion (Antofagasta plc). The four descriptive statistics for the BBO of the LOB by tercile are reported in Figure 3.3. We observe that the BBO depth (Figure 3.3c) and the executed volume (Figure 3.3e) seem to be decreasing in the stock price. Although this result is only qualitative it does give a first indication of how stock prices are correlated with the trading at the BBO. A further more rigorous analysis is beyond the scope of this manuscript. Turning our attention to the market capitalisation, it appears that inside spread (Figure 3.3b) and intraday volatility (Figure 3.3h) are decreasing in the market cap, while BBO depth (Figure 3.3d) and executed volume (Figure 3.3f) are increasing.

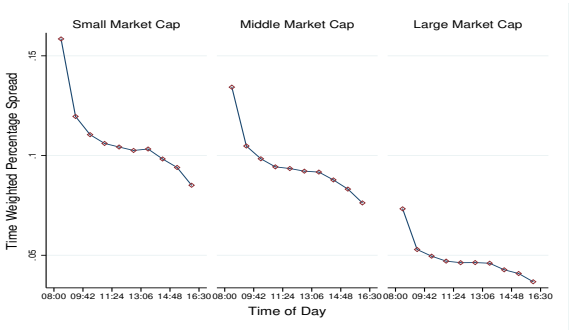
The final set of descriptive statistics in Figure 3.4 aims at comparing the three order books run by BATS Chi-X Europe on CXE, namely the lit, dark and hidden order books. As for the descriptive statistics above on the BBO in Figure 3.2 and on the BBO by tercile in Figure 3.3 we only provide a qualitative analysis since this is a side topic (albeit an interesting one) to our study of the price discovery process. We notice from Figure 3.4a that the executed volume is U-shaped over the trading day on both the dark and hidden order books since the first, middle and last observations are outside the 95% confidence interval (dotted horizontal lines). Moreover, the volume on the lit book is right-skewed U-shaped since only the right-hand side of the curve lies outside the confidence interval. Figure 3.4b illustrates the average trade size for the three books and concludes that the average trade size is greatest on the dark book followed by the lit and finally by the hidden book. This result is statistically significant since the confidence intervals (dotted horizontal lines) of the three different books do not overlap and neither does any of the curves cross one of the other confidence intervals.

Figure 3.3: The Best Bid-Offer of the Limit Order Book by Tercile

(a) Inside Spread by Stock Price



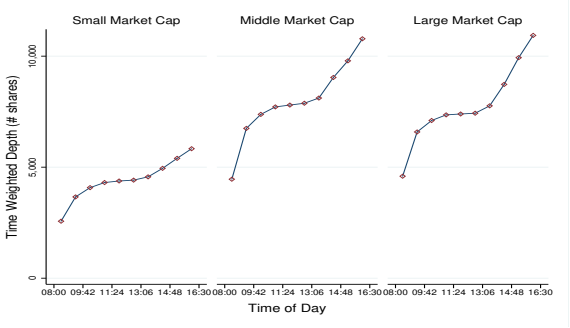
(b) Inside Spread by Market Cap



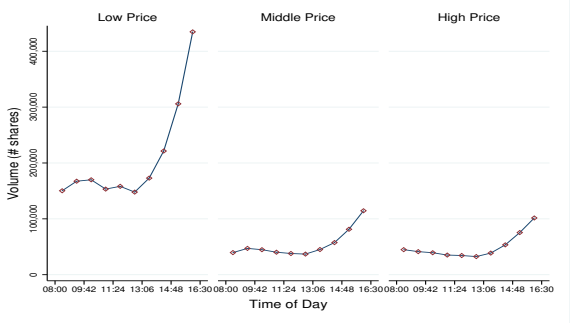
(c) BBO Depth by Stock Price



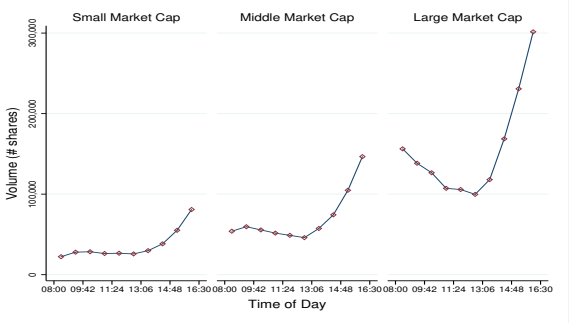
(d) BBO Depth by Market Cap



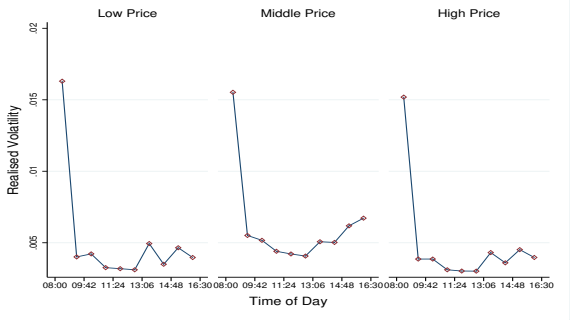
(e) Executed Volume by Stock Price



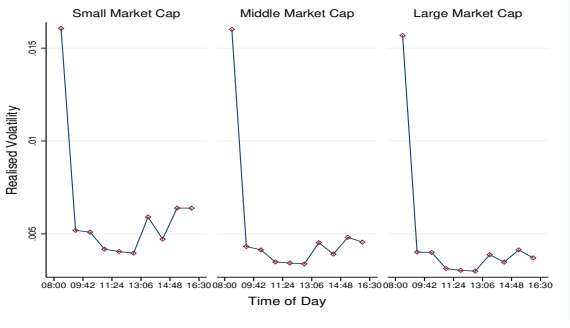
(f) Executed Volume by Market Cap



(g) Intraday Volatility by Stock Price



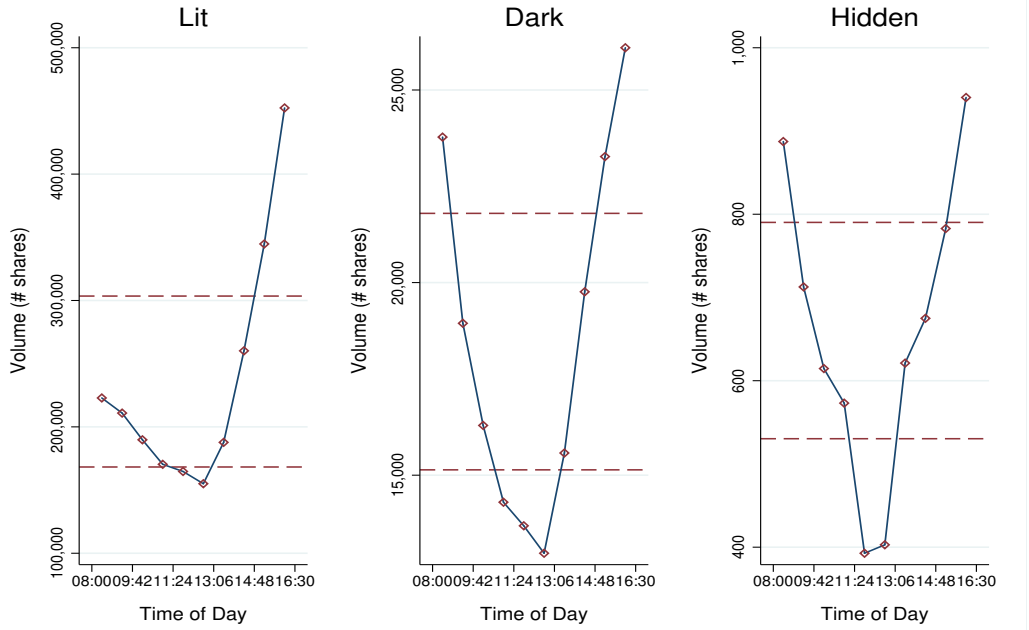
(h) Intraday Volatility by Market Cap



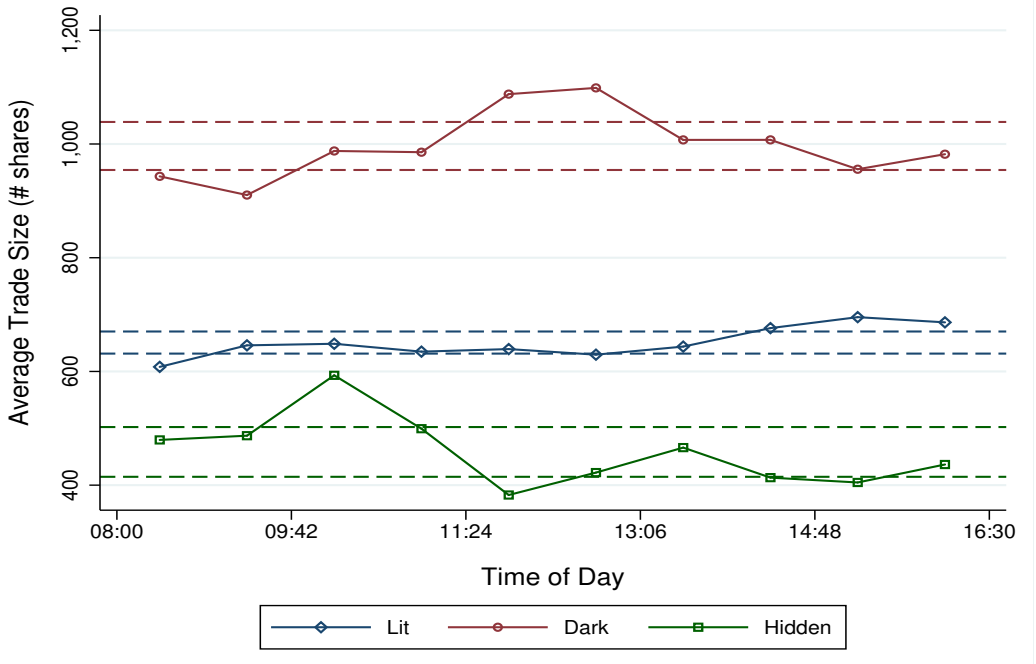
This figure exhibits the four descriptive statistics on the BBO of the LOB in Figure 3.2 while grouping the stocks in our sample into terciles first by the closing stock price and second by the closing market cap on September 3, 2012.

Figure 3.4: The Best Bid-Offer of the LOB by Market

(a) Executed Volume



(b) Average Trade Size



This figure exhibits two descriptive statistics (executed volume and average trade size) with their respective 95% confidence intervals (dotted horizontal lines) for the BBO of the three order books run by BATS Chi-X Europe on CXE over the trading day.

7.E Appendix Tables

Table 3.3: Sample I/II

	Stock	RIC	Close	Volume	Mkt Cap
1	Anglo American plc	AAL	1776.18	2,835,660	24.7
2	Associated British Foods plc	ABF	1326.79	348,711	10.5
3	Admiral Group plc	ADML	1143.56	610,334	3.1
4	AMEC Foster Wheeler plc	AMFW	1114.00	697,314	4.4
5	Aggreko plc	AGGK	2343.02	486,555	6.3
6	Antofagasta plc	ANTO	1125.71	1,237,865	11.1
7	Aviva plc	AV	327.20	4,379,688	9.5
8	AstraZeneca plc	AZN	2946.50	1,386,998	36.7
9	BAE Systems plc	BAES	323.42	4,389,046	10.5
10	Babcock International Group plc	BAB	937.89	692,199	3.4
11	Barclays plc	BARC	184.30	12,860,353	22.6
12	British American Tobacco plc	BATS	3333.50	1,027,012	64.8
13	BG Group plc	BG	1299.50	2,112,067	34.5
14	British Land Company plc	BLND	541.50	1,350,622	4.8
15	BHP Billiton plc	BLT	1648.00	1,212,095	3.5
16	Bunzl plc	BNZL	1126.55	361,171	3.7
17	BP plc	BP	443.41	13,343,190	84.4
18	Burberry Group plc	BRBY	1353.70	955,057	6.0
19	BT Group plc	BT	219.63	16,225,221	17.3
20	Canival plc	CCL	2156.08	365,667	3.9
21	Centrica plc	CNA	328.90	3,295,163	17.1
22	Compass Group plc	CPG	710.26	2,273,724	13.2
23	Capita plc	CPI	720.50	639,646	4.7
24	Croda International plc	CRDA	2375.00	129,152	3.2
25	Diageo plc	DGE	1738.40	1,679,653	43.6
26	Experian plc	EXPN	1014.48	1,233,177	10.2
27	Fresnillo plc	FRES	1625.76	890,833	11.7
28	G4S plc	GFS	254.01	3,537,785	3.6
29	GlaxoSmithKline plc	GSK	1436.03	6,953,818	71.4
30	Hammerson plc	HMSO	459.16	1,197,403	3.3
31	HSBC Holdings plc	HSBA	551.60	9,272,263	101.0
32	International Consolid. Airlines plc	ICAG	142.31	4,079,750	2.6
33	InterContinental Hotels Group plc	IHG	1612.82	480,516	4.7
34	IMI plc	IMI	858.55	717,761	2.8
35	Imperial Tobacco Group plc	IMB	2487.00	1,043,432	23.7
36	Intertek Group plc	ITRK	2811.53	284,834	4.5
37	ITV plc	ITV	83.90	5,704,373	3.3
38	Johnson Matthey plc	JMAT	2412.59	293,368	4.9
39	Kingfisher plc	KGF	274.24	7,334,956	6.5
40	Land Securities group plc	LAND	791.05	890,381	6.2

This table reports the first 40 stocks in our sample in alphabetical order and for each stock we compute three summary statistics on September 3, 2012: the closing price in GBX, the executed volume as number of traded shares, and the market capitalisation in billion GBP.

Table 3.4: Sample II/II

	Stock	RIC	Close	Volume	Mkt Cap
41	Legal & General Group plc	LGEN	128.78	5,400,270	7.6
42	Lloyds Banking Group plc	LLOY	33.43	69,432,512	23.5
43	Meggitt plc	MGGT	396.65	564,124	3.1
44	Marks and Spencer Group plc	MKS	357.40	3,577,015	5.7
45	Melrose Industries plc	MRO	44.31	9,530,744	2.1
46	WM Morrison Supermarkets plc	MRW	278.20	5,454,152	6.7
47	National Grid plc	NG	687.50	2,528,249	23.7
48	NEXT plc	NXT	3609.88	433,233	6.0
49	Petrofac lmd	PFC	1503.09	762,261	5.2
50	Prudential plc	PRU	785.47	1,831,614	20.1
51	Pearson plc	PSON	1207.05	670,316	9.9
52	Royal Bank of Scotland Group plc	RBS	226.50	7,257,280	13.6
53	Royal Dutch Shell plc Class A	RDSa	2209.50	1,398,884	82.1
54	Rexam plc	REX	427.00	948,889	2.5
55	Rio Tinto plc	RIO	2796.49	5,146,944	39.5
56	Rolls-Royce Holding plc	RR	835.50	2,961,338	15.6
57	Randgold Resources plc	RRS	7430.00	1,336,773	7.2
58	RSA Insurance Group plc	RSA	553.04	1,053,817	5.7
59	J Sainsbury plc	SBRY	328.25	2,074,872	6.2
60	Schroders plc	SDR	1420.02	118,916	3.2
61	Shire plc	SHP	1922.00	559,930	17.5
62	Standard Life Aberdeen plc	SL	266.80	2,696,899	6.7
63	The Sage Group plc	SGE	303.50	4,687,160	3.8
64	Smith Group plc	SMIN	1042.97	405,926	4.1
65	Smith & Nephew plc	SN	667.75	950,610	6.0
66	Serco Group plc	SRP	567.65	826,477	2.8
67	Standard Chartered plc	STAN	1398.83	1,693,107	33.5
68	Severn Trent plc	SVT	1736.07	340,207	4.1
69	Tate & Lyle plc	TATE	660.03	584,239	3.1
70	Tullow Oil plc	TLW	1380.16	2,135,010	12.5
71	Tesco plc	TSCO	338.01	6,450,096	27.2
72	Unilever plc	ULVR	2272.59	829,577	29.2
73	Unite Utilities Group plc	UU	712.76	1,442,403	4.9
74	Vedanta Resources plc	VED	11.62	182,200	10.7
75	Vodafone Group plc	VOD	183.20	29,912,183	90.1
76	Weir Group plc	WEIR	1637.19	501,323	3.5
77	John Wood Group plc	WG	830.50	568,295	3.1
78	Ferguson plc	WOS	2602.00	655,861	7.7
79	WPP plc	WPP	817.03	2,636,210	10.3
80	Whitbread plc	WTB	2127.55	265,694	3.8

This table reports the second 40 stocks in our sample in alphabetical order and for each stock we compute three summary statistics on September 3, 2012: the closing price in GBX, the executed volume as number of traded shares, and the market capitalisation in billion GBP.

Table 3.5: p-Values of Information Share Ratio

p-value	**Level #1**	**Level #2**	**Level #3**	**Level #4**	**Level #5**
Ask	7.37E-13	1.34E-14	1.34E-14	9.58E-19	1.52E-16
Bid	6.54E-06	6.67E-07	1.05E-13	1.05E-13	4.69E-12
p-value	**Level #6**	**Level #7**	**Level #8**	**Level #9**	**Level #10**
Ask	7.37E-13	1.30E-04	5.03E-02	0.3857	0.1284
Bid	7.08E-10	1.35E-08	1.03E-06	0.2780	0.9999
p-value	**Level #11**	**Level #12**	**Level #13**	**Level #14**	**Level #15**
Ask	0.8284	0.2780	5.03E-02	8.84E-03	4.47E-04
Bid	0.0165	0.0165	1.02E-03	7.40E-05	1.02E-03
p-value	**Level #16**	**Level #17**	**Level #18**	**Level #19**	**Level #20**
Ask	1.03E-06	2.77E-05	9.81E-06	2.77E-05	3.28E-06
Bid	3.04E-07	1.03E-06	2.77E-05	5.24E-09	7.40E-05
p-value	**Level #21**	**Level #22**	**Level #23**	**Level #24**	**Level #25**
Ask	3.04E-07	2.67E-04	1.43E-07	2.42E-07	1.12E-04
Bid	2.19E-03	4.47E-04	9.81E-06	2.73E-06	2.18E-06

This table reports the p-values of the Wilcoxon t-Test on $H_0 : R_l^s = 1$ *VS* $H_1 : R_l^s \neq 1$ *where* R_l^s *is the information share ratio of level* $l \in \{1, ..., 25\}$ *on side* $s \in \{Ask, Bid\}$ *of the limit order book.*

References

Angel, James J, Lawrence E Harris, and Chester S Spatt, 2011, Equity trading in the 21st century, *The Quarterly Journal of Finance* 1, 1–53.

Bacidore, Jeffrey M, and George Sofianos, 2002, Liquidity provision and specialist trading in NYSE-listed non-US stocks, *Journal of Financial Economics* 63, 133–158.

Baillie, Richard T, Geoffrey G Booth, Yiuman Tse, and Tatyana Zabotina, 2002, Price discovery and common factor models, *Journal of Financial Markets* 5, 309–321.

Battalio, Robert H, 1997, Third market broker-dealers: cost competitors or cream skimmers?, *The Journal of Finance* 52, 341–352.

Bellia, Mario, Loriana Pelizzon, Marti G Subrahmanyam, Jun Uno, and Darya Yuferova, 2017, Coming early to the party, *Working-paper at SSRN: https://ssrn.com/abstract=3038699* .

Bennett, Paul, and Li Wei, 2006, Market structure, fragmentation, and market quality, *Journal of Financial Markets* 9, 49–78.

Beveridge, Stephen, and Charles R Nelson, 1981, A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle', *Journal of Monetary Economics* 7, 151–174.

Biais, Bruno, and Thierry Foucault, 2014, HFT and market quality, *Bankers, Markets & Investors* 128, 5–19.

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

Biais, Bruno, Pierre Hillion, and Chester Spatt, 1995, An empirical analysis of the limit order book and the order flow in the paris bourse, *The Journal of Finance* 50, 1655–1689.

Biais, Bruno, Pierre Hillion, and Chester Spatt, 1999, Price discovery and learning during the preopening period in the paris bourse, *Journal of Political Economy* 107, 1218–1248.

Biais, Bruno, Johan Hombert, and Pierre-Olivier Weill, 2014, Equilibrium pricing and trading volume under preference uncertainty, *The Review of Economic Studies* 81, 1401–1437.

Boehmer, Beatrice, and Ekkehart Boehmer, 2003, Trading your neighbor's ETFs: competition or fragmentation?, *Journal of Banking & Finance* 27, 1667–1703.

Boehmer, Ekkehart, Kingsley Y L Fong, and Juan J Wu, 2012, Algorithmic trading and changes in firms' equity capital, *Working-paper available at SSRN: https://ssrn.com/abstract=2050856* .

Boussetta, Selma, Laurence Lescourret, and Sophie Moinas, 2017, The role of pre-opening mechanisms in fragmented markets, *Working-paper at SSRN: https://ssrn.com/abstract=2939502* .

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

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

Campbell, John Y, and Robert J Shiller, 1979, Interpreting cointegrated models, *Journal of Economic Dynamics and Control* 12, 505–522.

Cardella, Laura, Jia Hao, and Ivalina Kalcheva, 2012, Competition in make-take fees in the us equity market, Technical report, Working paper Texas Tech University http://ssrn. com/abstract= 2149302.

Chakrabarty, Bidisha, Shane A Corwin, and Marios A Panayides, 2011, When a halt is not a halt: an analysis of off-NYSE trading during NYSE market closures, *Journal of Financial Intermediation* 20, 361–386.

Chakravarty, Sugato, Pankaj Jain, James Upson, and Robert Wood, 2012, Clean sweep: informed trading through intermarket sweep orders, *Journal of Financial and Quantitative Analysis* 47, 415–435.

Clapham, Benjamin, Peter Gomber, and Sven Panz, 2017, Coordination of circuit breakers? Volume migration and volatility spillover in fragmented markets, *Working-paper available at SSRN: https://ssrn.com/abstract=2906719* .

Colliard, Jean-Edouard, and Thierry Foucault, 2012, Trading fees and efficiency in limit order markets, *The Review of Financial Studies* 25, 3389–3421.

Cui, Bei, and Arie E Gozluklu, 2016, Intraday rallies and crashes: spillovers of trading halts, *International Journal of Finance & Economics* 21, 472–501.

Dickey, David A, and Wayne A Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association* 74, 427–431.

Duffie, Darrell, Nicolae Gârleanu, and Lasse H Pedersen, 2005, Over-the-counter markets, *Econometrica* 73, 1815–1847.

Dugast, Jérôme, 2018, Unscheduled news and market dynamics, *The Journal of Finance* 73, 2537–2586.

Easley, David, Marcos M López de Prado, and Maureen O'Hara, 2012, Flow toxicity and liquidity in a high-frequency world, *The Review of Financial Studies* 25, 1457–1493.

Engle, Robert F, and Clive W J Granger, 1987, Co-integration and error correction: representation, estimation, and testing, *Econometrica* 1, 251–276.

Foucault, Thierry, Ohad Kadan, and Eugene Kandel, 2013, Liquidity cycles and make/take fees in electronic markets, *The Journal of Finance* 68, 299–341.

Foucault, Thierry, Roman Kozhan, and Wing W Tham, 2016, Toxic arbitrage, *The Review of Financial Studies* 30, 1053–1094.

Foucault, Thierry, and Albert J Menkveld, 2008, Competition for order flow and smart order routing systems, *The Journal of Finance* 63, 119–158.

Frijns, Bart, Ivan Indriawan, and Alireza Tourani-Rad, 2015, Macroeconomic news announcements and price discovery: evidence from Canadian–US cross-listed firms, *Journal of Empirical Finance* 32, 35–48.

Garbade, Kenneth D, and William L Silber, 1979, Dominant and satellite markets: a study of dually-traded securities, *The Review of Economics and Statistics* 1, 455–460.

Goldstein, Michael A, 2015, Circuit breakers, trading collars, and volatility transmission across markets: evidence from NYSE Rule 80A, *Financial Review* 50, 459–79.

Goldstein, Michael A, and Kenneth A Kavajecz, 2004, Trading strategies during circuit breakers and extreme market movements, *Journal of Financial Markets* 7, 301–333.

Gomber, Peter, Satchit Sagade, Erik Theissen, Moritz C Weber, and Christian Westheide, 2017, Competition between equity markets: a review of the consolidation versus fragmentation debate, *Journal of Economic Surveys* 31, 792–814.

Gonzalo, Jesus, and Clive W J Granger, 1995, Estimation of common long-memory components in cointegrated systems, *Journal of Business and Economic Statistics* 13, 27–35.

Granger, Clive W J, 1981, Some properties of time series data and their use in econometric model specification, *Journal of Econometrics* 16, 121–130.

Harris, Frederick H. D., Thomas H. McInish, Gary L. Shoesmith, and Robert A. Wood, 1995, Cointegration, error correction, and price discovery on informationally linked security markets, *Journal of Financial and Quantitative Analysis* 30, 563–579.

Hasbrouck, Joel, 1995, One security, many markets: determining the contributions to price discovery, *The Journal of Finance* 50, 1175–1199.

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

Hendershott, Terrence, Charles M Jones, and Albert J Menkveld, 2011, Does algorithmic trading improve liquidity?, *The Journal of Finance* 66, 1–33.

Hengelbrock, Jördis, and Erik Theissen, 2009, Fourteen at one blow: the market entry of Turquoise, *Working-paper available at SSRN: https://ssrn.com/abstract=1743589* .

Ibikunle, Gbenga, 2015, Competition for order flow and price discovery: the curious case of high-tech entrants, *Working-paper available at SSRN: https://ssrn.com/abstract=2440691* .

Johansen, Søren, 1988, Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control* 12, 231–254.

Johansen, Søren, 1991, Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica* 1, 1551–1580.

Johansen, Søren, 1995, Identifying restrictions of linear equations with applications to simultaneous equations and cointegration, *Journal of Econometrics* 69, 111–132.

Jones, Charles M, 2013, What do we know about high-frequency trading?, *Columbia Business School Research Paper No. 13-11* .

Kadapakkam, Palani-Rajan, Lalatendu Misra, and Yiuman Tse, 2003, International price discovery for emerging market stocks: evidence from Indian GDRs, *Review of Quantitative Finance and Accounting* 21, 179–199.

Kirilenko, Andrei, Albert S Kyle, Mehrdad Samadi, and Tugkan Tuzun, 2017, The flash crash: high-frequency trading in an electronic market, *The Journal of Finance* 72, 967–998.

Li, Sida, Xin Wang, and Mao Ye, 2017, Who supplies liquidity, and when?, *Working-paper available at SSRN: https://ssrn.com/abstract=3231166* .

Lieberman, Offer, Uri Ben-Zion, and Shmuel Hauser, 1999, A characterization of the price behavior of international dual stocks: an error-correction approach, *Journal of International Money and Finance* 18, 289–304.

Lutat, Marco, 2010, The effect of maker-taker pricing on market liquidity in electronic trading systems – empirical evidence from European equity trading, *Working-paper available at SSRN: https://ssrn.com/abstract=1752843* .

Malinova, Katya, and Andreas Park, 2015, Subsidizing liquidity: the impact of make/take fees on market quality, *The Journal of Finance* 70, 509–536.

McInish, Tom, James Upson, and Robert A Wood, 2014, The flash crash: trading aggressiveness, liquidity supply, and the impact of intermarket sweep orders, *Financial Review* 49, 481–509.

Menkveld, Albert J, 2016, The economics of high-frequency trading: taking stock, *Annual Review of Financial Economics* 8, 1–24.

Menkveld, Albert J, and Bart Z Yueshen, 2016, The flash crash: a cautionary tale about highly fragmented markets, *Working-paper available at SSRN: https://ssrn.com/abstract=2243520* .

Menkveld, Albert J, and Marius A Zoican, 2014, Need for speed? Exchange latency and market quality, *Journal of Financial Economics* 14, 71–100.

Morris, Charles S, 1990, Coordinating circuit breakers in stock and futures markets, *Economic Review* 75, 35–48.

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

O'Hara, Maureen, and Mao Ye, 2011, Is market fragmentation harming market quality?, *Journal of Financial Economics* 100, 459–474.

Phillips, Peter C B, and Steven N Durlauf, 1986, Multiple time series regression with integrated processes, *The Review of Economic Studies* 53, 473–495.

Putniņš, Tālis J, 2013, What do price discovery metrics really measure?, *Journal of Empirical Finance* 23, 68–83.

Satterthwaite, Franklin E, 1946, An approximate distribution of estimates of variance components, *Biometrics Bulletin* 2, 110–114.

Snedecor, George W, and William G Cochran, 1989, Statistical methods, *Iowa State University Press* 8, 138–142.

Stock, James H, and Mark W Watson, 1988, Testing for common trends, *Journal of the American Statistical Association* 83, 1097–1107.

Subrahmanyam, Avanidhar, 1994, Circuit breakers and market volatility: a theoretical perspective, *The Journal of Finance* 49, 237–254.

Taylor, Nicholas, 2011, Time-varying price discovery in fragmented markets, *Applied Financial Economics* 21, 717–734.

Welch, Bernard L, 1947, The generalization of Student's problem when several different population variances are involved, *Biometrika* 34, 28–35.

Yan, Bingcheng, and Eric Zivot, 2007, The dynamics of price discovery, *Working-paper available at SSRN: https://ssrn.com/abstract=249501* .

Yan, Bingcheng, and Eric Zivot, 2010, A structural analysis of price discovery measures, *Journal of Financial Markets* 13, 1–19.