High Frequency Trading, Market Fragmentation and Market Quality: An EU Market Evidence

by

Shahadat Hossain

Supervisor:

Prof. Marco Pagano

Program Coordinator:

Prof. Maria Gabriella Graziano

A thesis submitted in fulfilment of the requirements of the degree of
Doctor of Philosophy in Economics
(XXX cycle)
to the

Department of Economics and Statistics
University of Naples Federico II
June 2018

High Frequency Trading, Market Fragmentation and Market Quality: An EU Market Evidence

By

Shahadat Hossain

Abstract

This Ph.D. thesis joins the debate regarding the social benefit of HFT with the aim of contributing to HFT research originally. My research design integrates both HFT and market fragmentation and extends analyses within and across markets. I use the European equity markets as a laboratory as Europe has been confronting the issue of HFT influx and market fragmentation since the adoption of the Market in Financial Instruments and Directives (MiFID) in November 2007 by the European Parliament. I employ an extremely large dataset with the highest granularity, which gives the most recent and longest coverage of data in HFT research to date. I mainly examine the effects of HFT and market fragmentation on market liquidity within and across European markets.

In chapter 1, I review the literature and develop my arguments that rationalise the studies presented in this thesis.

In chapter 2, I examine the impact of HFT and market fragmentation on market liquidity within a market by applying three alternative estimations: OLS, IV-GMM and simultaneous equations model. I document that HFT improves liquidity, but market fragmentation appears detrimental to liquidly. I show that the interaction between HFT and market fragmentation has significant impact on market environment. It seems that in the absence of HFT, a fragmented market would be more detrimental to liquidity.

In chapter 3, I extend the analysis of the previous chapter to incorporate all fragmented markets, and present a novel approach to creating full view HFT image from HFT activities across markets. I primarily examine how HFT and fragmentation affect market liquidity across markets by using the simultaneous equations model. The results show that HFT improves liquidity across markets, whereas market fragmentation harms liquidity in the primary exchange but improves in alternative exchanges. I also provide evidence that HFT activities are linked across markets, and HFTs provide liquidity when spreads are wider. It seems that HFTs concentrate in the primary exchange during periods of high market volatility.

To

my sisters and brothers;

Fatima, Noorhan and Khadiza;

&

the departed souls of my parents.

(May Allah place them in Jannah)

Acknowledgements

My Ph.D. study has made me indebted to a number of kind and generous souls who have been so supportive throughout my studies without which it would have been so difficult if not impossible to complete this Program.

First and foremost, I would like to thank my supervisor, Marco Pagano, who has been at the core in providing me his unconditional support throughout these years. His invaluable guidance, highly experienced insights into my needs to undertake the research that I have conducted, and his excellent academic networks have been of immense benefit to me. His extraordinary patience in correcting my academic flaws has been unparalleled. I also gratefully acknowledge the scholarship that I received to finance part of my Ph.D. study from his FINLAB project, and his support for bringing my family in Italy while I was struggling to keep face with bureaucratic requirements.

I would like to express my deepest gratitude to my program coordinator, Maria Gabriella Graziano, who has been so supportive and kind in organizing every single request that I placed to her during my study at the Department of Economics and statistics of the University of Naples Federico II. I would also like to thank Tullio Jappelli , Marco Pagnozzi for their encouragements. I thank Antonio Acconcia, Luigi Benfratello, Giovanni Populo, Annalisa Scognamilio and Tommaso Oliviero for their valuable suggestions.

I would like to take this opportunity to thank my fellow Ph.D. students for the quality time that we spent together. In particular, I thank Claudio, Aniello, Annamaria, Luca , Muhammed, Amadou, and Taslim for their accompanies and supports in several occasions.

I am thankful to Stefania, Micol, Mario and Enrico who have been helping me from the date I enrolled in this program.

Back to my home, I express my heartfelt gratitude to all my teachers and colleagues for their spontaneous supports. I would like to thank Bro. Tanvir and Bro. Hasmat who have indebted me for the rest of my life. I am grateful to Bro. Habib and Bro. Aziz, two members of my extended family, who took pains to visit us in Italy.

The last but not the least, no thanks is enough for Fatima (my better half), Noorhan and Khadiza (my angels) whom I deprived the most during this period.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

For the last two decades, the advent of sophisticated computing technology has been changing the financial market structure unprecedentedly. Machines are gradually occupying the places for which formerly human interaction was necessary. The rise of machines' intelligence has enabled the human civilization to do things at ease, speed, and economy, and at the same time, raised the concern of welfare damaging inequitable competition between man and machines. Among many, High Frequency Trading (hereafter referred to as HFT) is one of the instances of such competition, and there has been debate since 2009 about its effect (Menkveld, 2016). Recent financial market regulations in the US and EU have opened the avenue for order flow fragmentation in equity markets and augmented the process of creating access for more machines amid regulators' concern for promoting greater transparency and competition. The rise of machines has also given rise to the proliferation of machine-friendly trading venues and their high speed connecting channels. At present, society is facing a real dilemma, mainly to support or hinder the proliferation of HFT, and a wrong decision might come at a high social cost. I join this debate with the aim to provide empirical evidence. This Ph.D. thesis takes an intra market microstructure approach motivated from the recent HFT literature and attempts to address the issue holistically which is often overlooked in the heated debate. I develop my arguments in the subsequent paragraphs.

The arrival of HFT coincided with the entry of new markets, and subsequently, strong fragmentation of order flow. The existing market microstructure literature has considered HFT and market fragmentation as two separate strands so far and their interaction has been ignored mostly. As academics, practitioners and regulators strive on this issue, new realities of the high frequency world become more visible. Menkveld (2014, 2016) argues "Electronic trading, new venues, and HFTs[1] are intimately related. There is arguably a symbiotic relationship between new electronic venues and HFTs. These new venues need HFTs to insert aggressively priced bid and ask quotes, and HFTs need the new venues to satisfy their requirements in terms of automation, speed, and low fees." When market fragmentation is overlooked in HFT research or vice versa, at least in the existing equity market structures of EU and US, one risks missing the complexity of the problem and it could lead misspecification.

In a highly fragmented marketplace, the potential counterparties for HFT market-makers have a large selection of trading venues on which they can trade. To interact with this order flow, HFT market makers must be present on all these trading venues (The Netherlands Authority for the Financial Markets, 2016). This cross-market HFT market making makes the order books linked and so, too, order flows and price behaviour. Studying HFT in a traditional empirical setup which generally focused on a market in isolation often misses the fact that individual markets are tied together, and incorporating market links in the analysis is crucial for determining how well overall market functions in equilibrium. O'Hara (2015) argues that traditionally employed empirical methods might no longer be appropriate to tackle this cross-market HFT complexity. She suggests "Theoretically modelling such interrelation is daunting, so empirical analyses focusing on the predictive power of market variables both within and across markets can be a good place to start."

Besides, the literature examining the effect of HFT on market quality in a fragmented market setting is comparatively new and seems to be in its infancy. Since HFT got its momentum in the 2000s, a vibrant and large literature has developed in this area mostly

[1] High Frequency Traders (HFTs)

focused on a single venue, notably NYSE or NASDAQ in US markets. HFT research on EU equity markets data started to increase in recent years, but no comprehensive study to date has examined the effect of HFT on the European market environment. This thesis aims to shift the traditional HFT research focus to an intra market microstructure approach that is capable of answering the interrelated questions in market microstructures. My research design integrates both HFT and market fragmentation and extends analyses within and across markets. I use the European equity markets as a laboratory as Europe has been confronting the issue of market fragmentation and HFT influx since the adoption of the regulation Market in Financial Instruments and Directives (MiFID) in November 2007 by the European Parliament.

Research outline

The primary goal of this thesis is to examine the effects of HFT and market fragmentation on market quality across European markets. In doing so, it also sheds light on the drivers of HFT and market fragmentations. It concentrate on the liquidity dimension of market quality and use the liquidity, HFT and market fragmentation measures that are commonly employed in empirical market microstructure literature. It uses the millisecond time-stamped trades and quotes data stemming from the Thomson Reuters Tick History (TRTH) for the selected LSE listed stocks traded across four rivalry exchanges in Europe: the London Stock Exchange (LSE), BATS Chi-X Europe (CHIX), BATS Europe (BATS) and Turquoise (TURQ). This Ph.D. thesis specifically examines the following research questions within and across markets:

- How do HFT and market fragmentation, and their interaction impact liquidity?
- Does the impact of HFT and market fragmentation on liquidity change over time?
- Does the impact of HFT and market fragmentation on liquidity change across cross sections?
- What drives HFT and market fragmentation?

Structure of the thesis

The remainder of the thesis is organized as follows. Section 1.2 reviews the literature relevant to studies presented in chapters 2 and 3, and section 1.3 presents a brief overview of the European equity market structure in which the studies of this thesis are designed. Chapter 2 examines the impact of HFT and market fragmentation on market liquidity in a traditional HFT research setting focusing on a single market (LSE). Unlike the existing literature, my research design incorporates both HFT and market fragmentation. First, it provides details on data preparation, relevant measures and research strategies. Second, it analyses the relation between market liquidity, HFT and market fragmentation using three alternative estimation methods—OLS, IV-GMM and simultaneous equations model, and discusses the results. Chapter 3 examines the impact of HFT and market fragmentation on liquidity in a cross-market setting by integrating all major lit trading platforms in Europe that facilitate trading on LSE listed stocks. It provides details on data, measures and models specification. It analyses the relation between market liquidity, HFT and market fragmentation across four markets applying the simultaneous equations model, and discusses the results.

1.2 Literature Review

This thesis is relevant to two main strands of literature in market microstructure, HFT and market fragmentation. The HFT literature is relatively young and growing with respect to the research in order flow fragmentation. The main contribution of this thesis is to focus on a new issue, i.e. the interaction of HFT and order flow fragmentation within and across markets and covers three related aspects (i) HFT (ii) order flow fragmentation (iii) interaction of HFT and order flow fragmentation and their impacts on market liquidity.

Before moving to the literature discussion, it is useful to clarify some terminology which is used extensively in HFT research. Many papers (Gomber, Arndt, Lutat and Uhle, 2011; The Netherlands Authority for the Financial Markets, 2010; Aldridge, 2013) attempt to distinguish among the prevailing terminologies applied to modern electronic exchanges providing definitions. HFT and Algorithmic Trading, hereafter referred to as AT, are the mostly referred two also in this thesis. Often, it creates confusion when they are used interchangeably without mentioning the context, as both are dominant in literature but have different roles to play in the market. To be on the safe side, I take the definitions which are well founded in academic literature and use them within the context interchangeably to refer a HFT firm.

Practically, HFT and AT share features like low-latency in order routing and execution, use of algorithms, less human intervention, direct and high speed market access but differ in the use of algorithms, inventory position taking, proprietary/agency trading, frequency of quote update, market strategy etc. Gomber et al. (2011) provide an excellent coverage on the concepts of HFT and AT. The distinct characteristics of HFT are high-speed quote updates, reflected in huge number of orders and rapid order cancellation, proprietary trading, generation of profit from market making, very short holding period with zero net inventory, use of low-latency technology like colocation, sponsored market access and direct data feeds. On the contrary, agent trading, trading to minimize price impacts, relatively long holding period possibly days/week/months characterize AT. Therefore, HFT is considered a subset of AT.

1.2.1 Theoretical predictions

The literature has identified the possible mechanisms through which HFT and fragmentation and also their interaction might affect liquidity. Biais and Foucault (2014) suggest the economic channels through which HFT could impact market liquidity. HFT can be imagined as a low-latency based modern trading and perform the same functions that are expected from a traditional market maker but differently. HFT firms develop dynamic algorithms based on real-time feed and perform market making automatically at millisecond speed, which essentially does not change the economic role of a market maker. Modern technology enables HFT to (i) get fastest access to the market (ii) acquire and processes information in almost real time (iii) watch and routes orders across markets. Thus, as a market maker, HFT should be able to provide liquidity at a lower cost. But the concern is that slow traders are at an information disadvantage relative to HFT, which might create adverse selection cost.

Foucault, Pagano and Roell (2013) and Cantillon and Yin (2011) explain the possible effects of order flow fragmentation. As per market fragmentation concerns, there are two opposite views, one holding that fragmentation improves liquidity, and another that it harms liquidity. The main arguments against fragmentation or in favour of exchange floor monopoly are scale economics and the existence of network externalities in trading. In a concentrated market, it is easier to find a counter party, which minimizes search costs. The price impact of transactions also tends to be smaller in markets that attract higher trading volume. The dominant argument in favour of order flow fragmentation is that it places competitive pressure on the transaction fees charged by the exchanges. It also forces exchanges to install cutting-edge trading technology and intensifies competition among market makers.

Menkveld (2014) has explained the possible effect of interaction between HFT and electronic market fragmentation, and concludes that HFT may benefit or hurt market quality through adverse selection on price quotes, a technology arms race, or high-risk trading strategy. This thesis particularly contributes to this research topic, by examining the net impact of HFT on liquidity in the fragmented market.

Models describing the effect of HFT on market quality are not enormous, limited to a few aspects of HFT and market fragmentation, and generate different predictions depending on their assumptions and focuses. Goettler, Parlour and Rajan (2009) predict when market makers like HFT are more informed about fundamental value it can improve liquidity. In Biais, Foucault and Moinas (2011), HFT generates adverse selection problem which harms market liquidity. Some other models also predict detrimental liquidity effect of HFT through front-running in Li (2014), winner's curse in Han, Khapko and Kyle (2014) and a wasteful arms race for speed in Budish, Cramton and Shim (2015). In contrast, Jovanovic and Menkveld (2015) and Ait-Sahalia and Saglam (2017) predict both positive and negative liquidity effects of HFT depending on the market environments in which HFT works.

1.2.2 Empirical evidence

The empirical HFT literature has been growing on several strands. I concentrate here the ones which are relevant to this thesis.

Hendershott et al. (2011) and Hasbrouck and Saar (2013) are considered to be the seminal papers in empirical HFT literature to propose HFT identification measures based on limit order book updates. Hendershott et al. (2011) is the first to examine causality between AT and market quality, where they introduce an algorithmic trading proxy based on electronic message rate (order book update). They target a market structure change—the introduction of the auto quote in 2003 in NYSE—which made the exchange more accessible to algorithmic traders. To address the suspected endogeneity between algorithmic trading and market quality they use a dummy variable instrument based on the event date. Using monthly observations for NYSE listed 943 common stocks for the period 2001–2005, they report positive impact of AT on liquidity, for large stocks in particular. Hasbrouck and Saar (2013) propose a different low-latency measure, 'strategic-run' based on a particular trading executing behaviour of HFT and use the same to assess the impact of low-latency trading on market quality. Their findings suggest that low-latency activity improves liquidity. Boehmer, Fong and Wu (2015) study the effect of AT on market quality using data on 42 equity markets for the period 2001–2011. They use colocation as an instrument variable to tackle the endogeneity and provide evidence that AT improves liquidity.

In recent years, the speed competition forced exchanges to provide cutting-edge technology, and which eventually benefitted HFTs to update their quotes more rapidly. Many papers examine the impact of speed on market quality. Riordan and Storkenmaier (2012) study the system upgrade of Deutsche Boerse with the 8.0 release of Xetra on April 23, 2007 which reduced system latency from 50ms to 10ms. Frino, Mollica and Webb (2014) examine the impact of allowing traders to co-locate near exchange server on the liquidity of future contracts traded on the Australian Securities Exchange. Murray, Pham and Singh (2016) investigate the role of latency in market quality in the Australian Securities Exchange following the introduction of the Integrated Trading Platform and ASXTrade which reduced latency from 70 ms to 30 ms. Frino, Mollica, Monaco and Palumbo (2017) examine the impact of AT on market liquidity following the implementation of proximity hosting service by Borsa Italiana. Hendershott and Moulton (2011) study the impact of introducing hybrid market in New York Stock Exchange in 2006 which reduced the execution time of market order from 10 seconds to less than one second. Brogaard, Hagströmer, Nordén and Riordan (2015) exploit an optional colocation upgrade at NASDAQ OMX Stockholm to assess how speed affects market liquidity. All of these papers on external HFT shocks other than Hendershott and Moulton (2011) provide evidence that reduction in system latency or enhancing speed improves liquidity, and the converse is shown in Hendershott and Moulton (2011).

The literature examining the effect of HFT on market quality in a fragmented market setting is comparatively new. Brogaard, Hendershott and Riordan (2014b) study the interaction of HFT and fragmentation to help understand the role HFT have in enhancing or harming market quality via market integration/fragmentation based on Canadian market data. Their preliminary results show that HFTs play a key role in tying market together, but are inconclusive regarding the impact on liquidity. Aitken, Cumming and Zhan (2014) use the joint modelling of AT and market quality in their research. They examine the relation between market quality, AT, market fragmentation and market manipulation in U.S. equity market following the fragmentation regulation RegNMS using a simultaneous equations model. They find that fragmentation of the lit market order flow and the ensuing increase in competition, especially from HFT/AT and alternative trading systems, impact market liquidity positively. This thesis mainly contributes in this research area, by adopting a novel approach with a rich panel dataset.

Several papers provide evidence regarding HFT market making within and across markets. Hendershott and Riordan (2013) examine the role of algorithmic traders in liquidity supply and demand in the 30 Deutscher Aktien Index stocks on the Deutsche

Boerse in January 2008 and report an AT participation rate of 52% and 64% in marketable limit order and nonmarketable limit order volume respectively. They find that algorithm traders take liquidity when spreads are narrow and provide liquidity when spreads are wide, and when spreads are narrow algorithmic traders are less likely to submit new orders and less likely to cancel their orders. Algorithmic traders cluster their trades together and initiate trade quickly when quoted spreads are small. Carrion (2013) also reports similar findings using a sample of NASDAQ trades and quotes that directly identifies HFT participation. He provides evidence that HFTs supply liquidity when it is scarce and consume liquidity when it is plentiful. Jarnecic and Snape (2014) compare the liquidity supply by HFT with the remainder of participants in the order book on LSE data for April 2009–June 2009. The evidence is consitence with the view that HFT improves liquidity. Brogaard, Hendershott, Hunt and Ysusi (2014) use the technology upgrades that lower the latency of London Stock Exchange, following which the level of HFT increased, to examine a claim that execution cost could be increasing because of HFT. They use data for the period 2007—2011, and provide no clear evidence that HFT impact institutional execution costs. Menkveld (2013) studies the strategies of a large HFT firm that started trading after CHIX entered as a new venue for European equities. His evidence suggests that HFTs supply liquidity across markets.

Several recent papers examine the impact of market fragmenatation on market quality. O'Hara and Ye (2011) is an original study which documents the causality between fragmentation and market quality on a dataset of 262 U.S. stocks over six months in 2008. They find that more fragmented stocks are associated with lower transaction cost and fastest execution speed. Unlike O'Hara and Ye (2011), some recent papers, mostly focused on order flow fragmentation, analyse the cross market liquidity. Upson and Van Ness (2017) study the volume fragmentation, cross market competition of AT, and their impact on liquidity using NYSE data only for the first quarter of 2012. They find that volume fragmentation has a positive effect on the best depth level across markets but venue competition and excess AT activities harm market liquidity. Degryse, De Jong and Kervel (2015) is the first to address the issue of fragmentation in European equity markets by introducing cross market measures of depth and spreads similar to Foucault and Menkveld (2008), but it does not consider any aspect of HFT. They study 51 Dutch stocks across European venues for the initial post-MiFID period (November 2007–December 2009) and provide evidence that lit fragmentation improves liquidity. Gresse (2017) also assesses the impact of both lit and dark fragmentation in European markets in a cross market setting. She uses data on the LSE and Euronext's blue-chip stocks for the period October 2007–November 2009 and establishes control for AT in regressions. Her findings suggest that lit fragmentation improves liquidity across markets.

Competition for order flow is at the core of exchange competition. Foucault and Menkveld (2008) study the rivalry between Euronext and the LSE in the Dutch stock market. They test hypothesis about the effect of market fragmentation and provide evidence that fragmentation of order flow can enhance liquidity. He, Jarnecic and Liu (2015) examine the market share drivers of CHIX in an international context. Their evidence shows that alternative venue's market share is negatively related to trading fees and latencies and positively related to liquidity relative to primary exchanges. They also show that trading tend to concentrate on the primary exchanges during market stress and tick constraint in primary exchanges moves order flow to alternative exchanges. Riordan, Storkenmaier and Wagener (2011) study the market quality of FTSE 100 constituents traded on the LSE, and three MTFs for April–May 2010 to examine the impact of alternative trading venues on market quality. They provide evidence that alternative trading venues contribute positively to market quality and exchange competitions benefits investors.

1.3 Market Background

Two of the most recent striking changes in global equity market design (those have been proliferating trading venues) are the adoption of 'Regulation National Market System (RegNMS)' in the US in 2005, and the enactment of MiFID in Europe in 2007, following the development in the US. After the enactment of MiFID, traditional stock exchanges in Europe lost monopoly on trading which existed until the beginning of 2007. The types of trading venues defined by MiFID include regulated markets (RMs), multilateral trading facilities (MTFs), and systematic internaliser (SIs) [2].

In broad terms, RMs and MTFs operate in a similar fashion, providing an electronic platform for users to transact orders multilaterally. These trading venues generally match orders on a non-discretionary basis according to pre-defined rules that establish price and time priority for submitted orders. RMs and MTFs are required to publish pre-trade quotes and report details of executed trades to the market (CFA Institute, 2011). Both RMs and MTFs are allowed to organize primary listing, however, they differ in that. RMs are legally authorised to list regulated financial instruments, while financial products listed by MTFs are considered to be unregulated instruments. In practice, only RMs offer primary listing service. MTFs prefer not to do so and they may be viewed as equivalent to electronic trading networks (ECNs) in the US (Gresse, 2017). A firm choose on which RM to list, and once listed, MTFs may decide to organize trading in that firm as well. SIs are investment firm that internalise order flow to deal on their own account on an 'organised, frequent and systematic basis'. SIs are required to report trades to the market and to publicise pre-trade transparency information under certain conditions. Trades executed through SIs are reported as the over-the-counter (OTC) trades.

The largest RMs include the LSE Group (operator of the London Stock Exchange and Borsa Italiana), NYSE Euronext (which operates exchanges in France, Belgium, the Netherlands, Portugal, and the United Kingdom), and Deutsche Borse Group (operator of the Frankfurt Exchange and the Xetra trading system). In the midst of trading venue proliferation, CHIX, BATS, Turquoise and NASDAQ-OMX Europe were among the few MTFs which started operations at the beginning of the post-MiFID period in Europe, eventually, the latter closed the operations in 2010. At present, CHIX, BATS and Turquoise are the three leading alternative trading venues to execute more than one third

[2] In order to capture 'dark pool' operators and other alike trading systems, a new category of trading venue called Organised Trading Facility (OTF) is introduced for non-equity instruments in MIFID II which came into effect on 3 January 2018

of the European lit equity trading (see Table A.1).

Market design. The LSE runs electronic order books on which buy and sell orders are continuously matched from the open to the close according to the price-time priority rules. Automated trading sessions start at 8:00 and close at 16:30 in local time. MTFs also run transparent order books in which anonymous orders are matched continuously for the same trading hours relative to primary exchanges. MTFs differ in terms of the speed of execution, the number of securities traded, and trading fee structure (Degryse et al., 2015). Their market models are adapted to the needs of high-frequency traders by offering low-latency trading with high throughput rates. Most MTFs follow a so-called maker/taker fees model—offer a transaction rebate to those who provide liquidity (the market maker), while charging customers who take that liquidity. The LSE also followed the maker/taker fees model before switching back to a traditional fee schedule on September 1st, 2009.

Chapter 2

The Impact of High Frequency Trading and Market Fragmentation on Liquidity

2.1 Introduction

MiFID repealed the concentration rule[1] in November 2007 and paved the way for the electronic trading venues to compete with the traditional established exchanges. Afterwards, exchanges have been investing heavily to minimize the latency[2]and several alternative exchanges have been launched. Consequently, order flow has spread across many trading floors, creating a fragmented market place. The beneficiaries from this massive investment in technology appear to be a new breed of high frequency traders who implement low-latency strategies (Hasbrouck and Saar, 2013). MiFID has made the necessary breakthrough in spreading HFT market access across European equity markets. Menkveld (2016) points to this reality, "...the two most salient trends in securities markets since the turn of the century—order flow fragmentation and HFT entry—are intimately related and both driven by technology and regulation."

[1]The concentration rule led to a situation where a single stock exchange dominated each member state in EU.

[2]According to Hasbrouck and Saar (2013), latency is viewed as the time it takes to learn about an event, generate a a response, and have the exchange act on the response.

The motivation of this chapter is centred on the view that HFT and fragmentation of orderflows are closely related and should be examined together. A research design based upon both HFT and market fragmentation avoids over simplification of market complexities, and is methodologically more sensible than the one where either HFT or fragmentation is addressed alone. My thesis contributes to this research niche in this chapter which is relatively new and so far unexplored in the context of the European equity market.

The goal here is to examine the impact of HFT and market fragmentation on market liquidity within a market. Thereby, I investigate several aspects: i) how HFT impacts market liquidity; ii) how order flow fragmentation impacts market liquidity; iii) how the interaction of HFT and order flow fragmentation impacts market liquidity; iv) whether the impact changes over time; and v) the determinants of HFT and market fragmentation.

To investigate these questions, I use millisecond time-stamped TRTH dataset for the period 2005–2016, from which I exploit both HFT footprint and order flow fragmentation across exchanges in European equity markets. The dataset includes 132 large capitalized stocks, primarily listed in the LSE and also traded across main alternative exchanges/MTFs—CHIX, BATS and Turquoise. The dataset provides the coverage for both pre and post-MiFID eras, and is the largest dataset employed in the HFT research to date. I take the data for LSE (a traditional exchange) and develop daily measures for liquidity, HFT and fragmentation from the millisecond records and use them in the analyses. I specify six models and estimate them using OLS with proper control. To check the robustness of OLS estimates, I take two alternative approaches which tackle the possible endogeneity that could arise from simultaneity, by applying i) IV-GMM estimation and ii) simultaneous equations model. I use the whole sample as well as its suitable subsamples in all analyses performed in this chapter.

The results suggest that HFT improves liquidity, whereas higher fragmentation is detrimental to liquidity. The interaction between HFT and fragmentation shows that some of the possible benefits of HFT on market liquidity is offset by the extra cost of market making in the fragmented market. Conversely, some extra cost of market fragmentation is also neutralized by the benefits derived from HFT. It appears that in the absence of HFT a fragmented market would be more detrimental to liquidity.

Analyses expanded on large and small stocks show that both HFT and market fragmentation have impacted the liquidity of large stocks more positively. In other words, higher HFT is associated with more liquidity to large stocks whereas higher market fragmentation in small stocks appears more detrimental to liquidity. The interaction between HFT and market fragmentation across large and small stocks provides significant insight. An increase in HFT in large stocks is associated with higher liquidity when fragmentation is greater and an increase in fragmentation is associated with less harm to liquidity when HFT is greater. On the contrary, for small cap stocks, an increase in HFT activities is associated with lower liquidity if fragmentation is greater, and an increase in fragmentation is more detrimental to liquidity when HFT is higher. The findings show that fragmentation has affected the liquidity of small stocks more negatively than that of large stocks.

In the quest for detecting the possible sources of changes in liquidity, I extend the analysis by decomposition the effective spread into realized spreads and price impacts. The results show that HFT has contributed in both better execution and less adverse selection whereas fragmentation appears to be detrimental to both.

Time-varying analyses show that HFT and market fragmentation have time-varying impacts; surprisingly (against a general belief), HFT appears to provide liquidity even during a crisis period. Market fragmentation has increased over time, and increasingly harmed liquidity over the period. Time-varying impacts in HFT and market fragmentation provide evidence that increased fragmentation has offset, at least partially, the additional liquidity which has been generated by higher HFT during the latter period of the sample.

The expanded analysis on alternative HFT proxies (based upon the quotes update at different depths of the limit order book) show that HFT actively contributes beyond the best imit price (BBO) in the order book but not much outside the best five limit prices.

The alternative estimation methods, IV-GMM and simultaneous equations model, confirm that estimates obtained in OLS are robust. The use of simultaneous equations model extends understanding on the reverse causality between HFT, market fragmentation and liquidity. It appears that simultaneity exists, the level of fragmentation and HFT intensity both affect each other, and which in turn affect liquidity. However, the effect of HFT seems stronger.

The rest of the chapter is structured as follows. Section 2.2 associates this study with the similar literature. Section 2.3 describes the data and relevant measures of liquidity, HFT and market fragmentation, and reports descriptive results. Section 2.4 explains the research strategies and main results for three alternative estimation methods: OLS, IV-GMM and simultaneous equations model. Section 2.5 concludes.

2.2 Relevant literature

This section relates the chapter with the present body of literature reviewed in the section 1.2 (chapter 1). The study I present in this chapter is relevant to papers examining the causality between HFT and market quality, or market fragmentation and market quality, and concentrate on a single market HFT analysis. The literature like Hendershott et al. (2011), Hasbrouck and Saar (2013), and Boehmer et al. (2015) provide evidence on the causality between HFT and market quality, and O'Hara and Ye (2011), Gresse (2017), Degryse et al. (2015) assess the impact of market fragmentation on market quality. The originality of my study is that I address both HFT and market fragmentation and investigate their relation with market liquidity. Unlike these papers, my sample has a wider coverage and higher granularity of data. The study bears a close resemblance to papers examining interaction between HFT and market fragmentation like Brogaard, Hendershott and Riordan (2014b) and Aitken, Harris and Harris (2015) and contributes to this research area which is relatively unexplored in the literature and particularly in European markets.

2.3 Data and measures

2.3.1 Data

It is widely believed and also demonstrated by previous studies that large-cap stocks are highly liquid, attract more HFT, and also extremely fragmented. To study the high frequency trading and market fragmentation phenomenon simultaneously, I primarily choose large market capitalized common stocks in European equity markets. Generally, papers studying HFT and fragmentation have a small data coverage, ranging from couple of months to few years, due to huge management and computational burden. Managing a small data set covering short time periods is relatively easy but has caveats—results are more likely to be specific to the period specific factors. I attempt to get a wider coverage of the data with two aims: i) it should provide enough variation in both HFT and fragmentation measures, and ii) it should also cover pre and post-MiFID regulation periods, i.e. both the period before and the period after November 2007.

I select the period January 2005–December 2016 which gives the most recent and longest coverage of data in HFT research. The dataset has the footprint of all economic shocks and recoveries of the last decade (see Figure A.1), and provides enough data points extraordinarily to implement sound econometric tools. The post-MiFID period is prominently characterized by the proliferation of low-latency based modern trading venues, through which markets have been experiencing a large influx of HFT investment.

In constructing the sample, I take the STOXX 800 as the benchmark index which constitutes the largest 800 market capitalized stocks in Europe. Table A.1 (Panel A) shows the STOXX 800 composition at the end of year 2016. It reports that top 50% stocks of the list are coming from only three primary trading venues, the London Stock Exchange, Deutche Boerse (Xetra) and Euronext Paris, of which LSE listed stocks are more than 50%. Table A.1 (Panel B) shows the market share of both primary and alternative lit trading venues in European equity markets. Among the trading venues, CHIX, BATS and Turquoise facilitate most of the lit trading besides the primary platforms. Remarkably, the present market share of CHIX exceeds that of any other trading venues. These three alternative trading venues are selected to supplement the sample.

The primary source of my data is Thomson Reuters Tick History (TRTH) ,[3] a product of the Securities Industry Research Centre of Asia-Pacific (SIRCA), which is compiled from the Global Thomson Reuters exchange feeds. Two resilient London based recording devices provide the millisecond time stamp to each recorded message. The primary analysis of TRTH data structure reveals that time synchronization of trades and respective quote messages is not uniform across trading venues. TRTH provides better quotes and trades time synchronization for the trading venues which are physically closer to the IDN Collection LAN in London (e.g. LSE, CHIX, Bats, Turquoise) than for those which are not (e.g. Deutche Boerse (Xetra) , Euronext Paris). This issue raises some real challenges in determining trades and quotes based measure of transaction cost, which is particularly true for the effective spread.

Considering the TRTH time synchronization issue, I narrow down the sample choice only to the UK-based LSE listed stocks included in STOXX 800. To address the fragmented environment of these stocks appropriately, I select CHIX, BATS, and Turquoise as their alternative venue counterparts. These four trading venues facilitated around 99% of lit trading during the period 2014–2016 for the stocks that are primarily listed in LSE, and this pattern is quite regular over the sample period (see Table A.1, Panel C). The trades and quotes data is available from TRTH since 1996 for most of the primary trading venues, which for alternative trading venues in MiFID zone started to be available from mid of 2008. Among the 220 primarily selected stocks from the LSE, TRTH provides data support only for 204 stocks. Table A.2 shows the TRTH data availability for the sample stocks across trading venues.

TRTH supplies quotes and trades records through two main files, the Time and Sales (TS) and the Market Depth (MD). The time and sales file provides transaction records and the best quote updates, and the market depth file comes with the queue of bid and ask limit prices and respective quantities (displayed in the limit order book). The records

[3] I acknowledge gratefully the support of Prof. Riccardo Palumbo, Professor of Accounting at University of Chieti-Pescara and CEO at European Capital Markets CRC for allowing me to use the TRTH database during the period of my Ph.D. study.

in market depth can be extracted to 25 best limit prices (based upon their availability) of which I extract the best 10 levels. I download and process these two files for all stocks primarily selected in the sample (Table A.2, Panel A). This requires me to dedicate substantial computing resources and data processing time. I processed about 885 files of market depth data of 70 million records each and 300 files of Time and Sales data of 110 million records each, on average, in several phases before harvesting the usable output. The unzipped physical size of the data is around 20 TB.

At this point, a primary analysis shows that among those 204 securities, some are not compatible for further analysis due to reasons like delisting, takeovers or mergers with other firms or liquidation at some point or do not have enough data coverage for all four trading venues for unknown reasons, and I set a final filter to ensure uniform data coverage of the selected stocks and exclude them if they do not satisfy the following conditions: (i) data availability in LSE at least from 2006; (ii) data availability in alternative trading venues from 2008 or at least from the 1st quarter of 2009. Table A.2 (Panel B) shows the reduced list of quarterly data used to construct the panel for the period 2005–2016. I also rely on the Thomson Reuters's Datastream for the relevant data, which are not supported from TRTH but used in this study (e.g. daily market capitalization).

2.3.2 Measures

This section describes the liquidity measures, HFT proxies, and market fragmentation proxies that are used in the analyses.

2.3.2.1 Liquidity

Foucault et al. (2013) define liquidity as the degree to which an order can be executed within a short time frame at a price closer to the security's consensus value. Conversely, if a price deviates substantially from the consensus value, there is illiquidity. Liquidity providers and seekers interact through the trading processes which are available in market places and differ a lot across their structures and also contribute differently to liquidity. The literature classifies liquidity into several dimensions, like transaction costs, quantity, time, and so on. Measurement of liquidity is of paramount importance to practitioners, regulators and academics. Investors and intermediaries frame strategies with aim to minimize the effect of liquidity shortage on their investment performance. Regulators and researchers try to understand the relationship between market structure and liquidity to work out and recommend the right policies.

In this study, I compute several standard measures of liquidity from the literature. I use the relative quoted spread ($spread\_bps_{it}$) and the effective-half spread ($espread_{it}$) which measure the transaction costs dimension of liquidity. The effective half-spread is decomposed into price impacts ($price\_impact_{it}$) and realized spreads ($rspread_{it}$). Price impact or adverse selection cost measures gross losses to liquidity demanders and realized spreads measures revenue to liquidity providers. To reflect the price pressure reasonably, an appropriate time gap (Δ) between transaction and post transaction quote adjustment is sought. In transparent and active markets adjustment is generally fast, so a modest value of Δ is appropriate (Foucault et al., 2013). In a low-latency environment, HFT traders should have the capacity to close their position in a very short period of time. A range of values are chosen—10 seconds, 30 seconds, 1 minute and 5 minutes—for Δ which likely fit both HFT and non-HFT trading environments.

To address the quantity dimension of liquidity, the quoted depth is measured at and beyond best limit prices. The recent evidence (Financial Markets Regulator (France), 2017) suggests that HFTs' liquidity supply surrounds several depth levels of the limit order book. The quoted depth refers to the available buy and sell quantities at prices close to the best bid and offer (BBO) price, practically at mid price. The measures, $depth1_{it}$ and $depth3_{it}$, refer the average quoted depth and cumulative depth at BBO and upto the best three limit price respectively.

For the t^{th} quote in stock j, quoted spread in basis points (bps) ($spread\_bps_{it}$) is defined as

$$spread\_bps_{it} = ((ask_{it} - bid_{it})/mp_{it}) * 10000,$$

where ask_{it} is the best quoted ask price, bid_{it} is the best quoted bid price, and mp_{it} is quote midpoint at BBO calculated as $(ask_{it} + bid_{it})/2$. For the t^{th} trade in stock i, the effective half-spread in bps, $espread_{it}$, is defined as

$$espread_{it} = (d_{it} * (p_{it} - mp_{it})/mp_{it}) * 10000,$$

where d_{it} is an indicator variable that equals $+1$ if the k^{th} trade is a liquidity demander's buy and -1 if the k^{th} trade is a liquidity demander's sell, p_{it} is the trade price, and mp_{it} is the quote midpoint prevailing at the time of the k^{th} trade.

Academics often find studying effective spreads challenging due to the trade signing requirements. This becomes even more complicated in the low-latency environment due to the involvement of enormous masses of quote updates, even within a millisecond. Empirical studies usually use the readily avilable trade signing approaches from literature, for example, Lee and Ready (1991). Signing a trade with these methods come with a high price of inaccuracy due to the fact that exchange platforms and data providers do not follow a uniform data synchronization system. In contrast to common practices, I develop algorithms which are capable of signing trade precisely in TRTH European data structure and give this study a unique advantage over others that rely on effective spreads in their analysis. The algorithms match every trade price with the immediate prevailing quotes, both bid and ask, and define k^{th} trade as liquidity demander's buy if it matches quoted ask price and as liquidity demander's sell if it matches quoted bid price.

The trade signing methodology adopted in this study goes as follows. In a first phase, algorithms filter all trades not sourcing from the automatic session and then accumulate trades executed on the same milliseconds with the same price. The problem arising from accumulating all trades indiscriminately executed in the same millisecond is carefully avoided. Generally, trade records delivered with the same time-stamp include both buy and sell trades. So, it is important to distinguish them as buyer or seller initiated trades before accumulating them. The second phase is bit more complex and time consuming where algorithms match trade price with the relevant quotes, both bid and ask , considering several "if and then" conditions. The algorithms attempt to match a trade price with the immediately available prior quotes (either bid or sell), if they find a match with bid then provide a seller initiated trade flag or a buyer initiated flag when find a match with ask. If the algorithms do not find a match with the immediate quotes, then they look for a match to the one before the immediate one and so on. In contrast, a traditional trade signing approach compares changes in trade price with the changes in mid price to ascertain whether an executed trade is buyer or seller initiated, and does not seem to fit a dynamic low-latency environment where quote update speed is very high and the time synchronization between trades and quotes updates is not quite orderly. The algorithms used in this study can assign a trade sign as accurate as more than 99%.

For the k^{th} trade in stock i, the percentage realized half-spread in bps, $rspread_{it}$ is defined as

$$rspread_{it} = (d_{it} * (p_{it} - mp_{i,t+\Delta})/mp_{it}) * 10000,$$

where d_{it} is an indicator variable that equals $+1$ if the k^{th} trade is a liquidity demander's buy and -1 if the k^{th} trade is a liquidity demander's sell, p_{it} is the trade price, mp_{it} is the quote midpoint prevailing at the time of the k^{th} trade, and $mp_{i,t+\Delta}$ is the quote midpoint after Δ. The gross losses to liquidity demanders, $price\_impact_{it}$, due to adverse selection (using the same variable) is defined as

$$price\_impact_{it} = ((mp_{i,t+\Delta} - mp_{it})/mp_{it}) * 10000.$$

A two-way decomposition of the effective spread, as an identity, can be defined as

$$espread_{it} = rspread_{it} + price\_impact_{it}.$$

The average quoted depth can be decomposed into offer depth (the specified quantity that a liquidity supplier is willing to sell at ask price), and bid depth (the specified quantity that a liquidity supplier is willing to buy at the bid price). The first of the two quantity based liquidity measures, $depth_l$, refers to the average offer and bid quantity available at the best bid and offer prices (BBO), or at best depth level. At time t, for stock i, average market depth $adepth_{it}$ is defined as

$$depth1_{it} = (Offer \quad depth * offer\ price + bid\ depth * bid\ price) * 0.5.$$

To measure the depth available beyond BBO, the average cumulative depth ($depth3_{it}$) is defined, which measures cumulative depth up to three best limit prices using the similar procedure defined for $depth1_{it}$.

2.3.2.2 High frequency trading

In HFT research, identification of HFT is critical. The literature has limited choices and depends on either one or both of the two approaches i) to use an exchange provided HFT flag dataset and ii) to define a proxy which tracks the footprint of HFT. Some of the possible disadvantages (Conrad, Wahal and Xiang, 2015) of using exchange-identified HFT data are that exchange houses select samples according to their own criteria and this may not be free from the possible conflict of interest, which prevails among the users of HFT flagged data, HFT firms and trading platforms. Other than that, sample firms appear to be large and specialized in HFT often operate in several exchanges across countries. There are various reasons why there could be a non-random distribution of trades across trading venues due to their heterogeneity in liquidity, fee structure etc. Therefore, sometimes, drawing inferences from these datasets may not reflect the true HFT behavior. In contrast, a proxy tracks the predominant market making nature of HFT through posting and renewing quotes.

The literature has introduced several HFT proxies defined on: (i) daily net position—intermediaries with high volume trades and low intraday and overnight position considered as HFT (Kirilenko, Kyle, Samadi and Tuzun, 2017); ii) 'Strategic Runs' of linked messages (Hasbrouck and Saar, 2013) where the proxy exploits the particular order sending and cancelling pattern of low-latency traders ; iii) electronic message traffic rate/normalized electronic message traffic rate (Hendershott et al., 2011), quote updates (Conrad et al., 2015), message-to-trade ratios (Friederich and Payne, 2015; Frino et al., 2017). TRTH datasets do not support the measure in (i) and (ii), as no ‘order id’ field is provided with the supplied data of European markets.

To exploit the benefit of the rich sample dataset, a set of HFT proxies are defined. This study mainly relies on HFT proxy defined on the message traffic rate (per unit of time), and it will be explained later why this measure is preferred over others. The principal HFT proxy is defined on three different depth levels—from the narrowest to the widest depth—of a limit order book. The first, $hft1$, the second, $hft2$ and the third, $hft3$, HFT proxies measure the per-minute message traffic in the best limit prices (BBO), the five best limit prices and the ten best limit prices respectively. These alternative proxies defined on different depth levels are motivated from the recent evidence reported in a HFT case study (Financial Markets Regulator (France), 2017) which shows that HFTs actively participate beyond the best limit prices. It also reports that the average market share of HFT in the BBO, the two best limit prices and the three best limit prices are 70.8%, 77.3% and 79.3% respectively.

In defining HFT proxies, I track every millisecond record to detect any change in all ten best price limits due to: i) order execution, ii) arrival of new limit order, ii) quote cancellation, and iv) quote modification; then aggregate them to daily sum divided by the number of minutes allocated for each daily automated trading session (8.00 to 16.30).

I also develop proxies similar to Hendershott et al. (2011), which are $hft1h$, $hft2h$, and $hft3h$ based on BBO, five best price limits and ten best price limits respectively, and order to trade ratio (OTR). Boehmer et al. (2015) also use the measure similar to Hendershott et al. (2011) calculated for the only best quotes as they did not have access beyond the best limit prices in the dataset. Though OTR is used in many countries (e.g. Italy, Germany , EU) as a benchmark to impose financial tax or other regulatory measure on HFT activities, it has flaws. European Securities and Markets Authority (2014) describes OTR as a useful metric to assess potential risks linked to trading system overload rather a method to identify firms carrying out HFT activities.

The electronic message rate ($hft1$) is defined as

$$hft1_{it} = qupdate1_{it}/T,$$

where $qupdate1_{it}$ is the aggregate quote update at 10 best limit prices for stock i on day t and T is the length of trading sessions (in minutes). $hft2$ and $hft3$ are also defined in the same way on the relevant depth of the order book. $hft1h$ is defined as

$$hft1h_{it} = value_{it}/(qupdate1_{it} * 100) * (-1),$$

where $value_{it}$ is the value of the trading volume of stock i on day t. Finally, $ordtotrd1$ is defined as

$$ordtotrd1_{it} = qupdate1_{it}/ntrades_{it},$$

where $ntrades_{it}$ is the number of executed trade of stock i on day t.

One of the widely used measures of HFT proxy in the empirical literature is $hft1h$, based on Hendershott et al. (2011) which measures the number of electronic messages per $\$100$ of trading volume. The measure is normalized in the original literature to adjust for an upward trend of trading volume which was particular to NYSE's and the employed sample period (2001–2005). To put the measure in the same spirit of the message traffic per unit of time ($hft1$), they use the dollar volume per electronic message time (-1) which practically measures the negative of traded value per message. An increase in the ratio or a smaller absolute value of the ratio stands for the increase in HFT intensity over time. A problem with this measure is that this interpretation does not hold when this ratio is used to compare two groups of stocks, or in other words, comparing this measure across stocks does not provide the same interpretation.

Let us imagine, for stock X, the value of $|hft1h|$ is v_1 and that for stock Y is v_2 where $v_1 < v_2$. It can't be said that HFT traders are more intense in stocks X (lower $v1$) than Y, practically, it is the reverse. Let us imagine again, for any stock, the value of $|hft1h|$ at time t_1 and t_2 are v_{1t1} and v_{1t2} respectively . In this setup, $v_{2t2} < v_{1t2}$ implies that HFT intensity is more at t_2 than t_1 which is consistent with the definition. Figures A.3 and

A.2 show that both the small-cap and large-cap quintiles have the same HFT time trends which can be interpreted as the increase in HFT intensity over time. If the same measure is compared across quintiles, one has to be careful in interpreting the values. The higher value of HFT proxy in small cap quintile (Figure A.3b) does not necessarily mean that HFTs/ATs are more intense in small-cap stocks; rather the opposite is the case. This is also true in interpreting $ordtotrd1$.

Moreover, the period 2005–2016 which is used to construct the sample is not associated with any increasing volume trend. Conversely, the LSE has lost market share to the competing alternative trading venues during the same period (see Figure A.4). Besides, both the normalized measures, $hft1h$ and $ordtotrd1$, do not incorporate the technological aspects of HFT which is revealed through the speed of message trafficking. If Figures A.4a and A.2 are compared, it can be seen that over the sample period $hft1$ increased monotonically though the trading value declined. Proxies defined on the speed aspect of HFT ($hft1$) do not encounter the explained issues as observed in $hft1h$.

To avoid these pitfalls, I only employ the HFT proxies developed on electronic message rate ($hft1$, $hft2$ and $hft3$) in the regression estimates, but use them all in descriptive analyses. One should also be aware of the fact that use of message traffic intensity per unit of time as a HFT proxy has its own limitation in tracking HFTs' footprint. It does not reflect the activity of a particular HFT, rather it should reveal a mixture of strategies adopted by HFTs. The analyses should contemplate the impact of dominant strategy because the mixture of strategies in actual markets may overwhelm the effect of one strategy or the other (Hasbrouck and Saar, 2013).

2.3.2.3 Market fragmentation

I use the Herfindahl-Hirschman Index (HHI)—the most commonly used definition of market concentration in literature—as the proxy of market fragmentation. Market fragmentation refers to the extent to which order flow of a security is split across exchanges. Fragmentation measures used in this study refer to the fragmentation of order flows across lit trading venues (where HFTs only participate). Though market fragmentation commonly refers to the volume fragmentation, recently, some literature came up with the idea of quote fragmentation. For instance, Madhavan (2012) argues volume fragmentation does not incorporate the HFT aspect which is demonstrated through huge quote update across exchanges. One of the limitations of traditional volume based measure is that it shows the end results of the traders routing decision across trading venues but fails to incorporate the huge quoting activities devoted by HFT behind every trade. Looking at quote based measure, one can learn more cross markets footprint of HFT. I define and use both volume and quote based fragmentation measures. Stock level volume fragmentation, $HHItrd_{it}$, is defined as

$$HHItrd_{it} = 1/\sum_{j}^{n} v_{ij},$$

where v_{ij} is the square of the trading volume share on venue j among n for security i at day t. This is a normalized measure which ranges from 1 to n, where 1 stands for no fragmentation or full concentration in the primary venue and n for evenly distributed order flow across n exchanges. Since this study considers one primary venue (LSE) and three alternative exchanges ($n = 4$) , the range is defined as, $1 <= HHItrd_{it} <= 4$. Some literature (Degryse et al., 2015) also use the non-inverted form of the same measure that is referred to as $HHItrd2$ here. A quote fragmentation proxy ($HHIqu_{lit}$) can be defined in a similar fashion by replacing trade volume with quote update. $HHIqu_{lit}$ is defined as

$$HHIqu_{lit} = 1/\sum_{j}^{n} q_{lij},$$

where q_{ij} is the square of the quote update share for the depth level l in the limit order book at venue j (among n), for security i at day t.

2.3.3 Descriptive statistics

This section presents the descriptive evidence associating HFT, market fragmentation and market liquidity using the full sample of 149 stocks for the period December 2005–December 2016 after winsorizing extreme 1% values on both tails. To facilitate cross sectional comparison, the full sample is divided into 5 equal quintiles based on market capitalization. Descriptive analyses show that both aggregate and quintile based measures (mean, median and standard deviations), monthly trends over the period 2005–2016, and pre and post-MiFID periods averages. All measures are calculated on LSE data[4].

High frequency trading

Tables A.3, A.4, and A.5 present the descriptive summary of the three HFT proxies measured at three depth levels. The first proxy (hft), for its three variants $hft1$, $hft2$ and $hft3$, measures the per-minute electronic message update in the ten best price limits, five best price limits and BBO respectively. The second proxy, $hfth$, for its two variants $hft1h$ and $hft2h$, measures the HFT intensity in the best 10 price limits and 5 price limits respectively. The third, $ordtotrd$, is shown for only the ten best price limits. Figures A.2 and A.3 present the monthly time series evolution of the three HFT proxies during the period 2005–2016.

These measures show that HFT intensity increased remarkably over the years. The trend reflects the phenomenon which started to explode in European markets at the beginning of 2008 and has grown monotonically over the periods except a small bump in 2012, though $ordtotrd$ shows almost no growth starting from 2011. As can be seen, HFT intensity increased significantly across cross sections but in different magnitude. Table A.5 shows that average message traffic rate ($hft1$) in 2005 starts at 9 messages per-minute and rises to 176 messages in 2016. It can also be seen that HFT activities are more intensive in the large stocks. For example, the quintile of the highest capitalized stocks (Large) starts from 22 messages in 2005 (whereas the smallest starts from below 5 messages) and rises to around 650 messages per-minute at the beginning of 2016 (whereas the smallest-cap rises to 60 messages) before falling to around 340 messages at the end of 2016 (the smallest falls to 37 messages).

Figures A.1, A.2, and A.3 also show that the general trend of rising HFT during this

[4] Table A.2 (Panel B) shows the periodical coverage of stocks in detail.

long period has not been interrupted by any mid or long-term economic shocks through which European markets have undergone during the last decade, rather, these shocks have magnified the HFT intensity manifold, instead.

Market fragmentation

Figure A.5a shows the fragmentation index for the period 2005–2016. All measures reflect the fragmentation of LSE listed stocks across LSE, CHIX, BATS and Turquoise. Table A.1 (Panel C), shows that these four trading venues share around 100% market share of lit turnover of the stocks which have primary listing in LSE and also fragmented in other alternative trading venues. So, not including the other alternative trading venue's except these four should not make any difference in the fragmentation scenario as we see in Figure A.5a.

Since the starting of the post-MiFID era in 2008, we see stocks, both large and small, started to fragment sharply across alternative trading venues. Consistently, Figure A.4a shows that LSE's trading volume declined in the same period, which indicates that LSE lost its substantial market share to the competing trading venues (see also Figure A.4b). As can be seen (also from Figure A.5a), large stocks are relatively more fragmented. Over the sample period, the fragmentation gap between large-cap and small-cap stocks decreased but did not close completely. Until 2011, the split in venue market share grew steeply and started to slow down afterwards. With few small bumps, stocks across the quintiles reached the maximum fragmentation point at the end of 2015, before starting to fall at the beginning of 2016. The fragmentation index never reached the theoretical maximum (which is four(4)), that suggests, the gain from alternative floors has not come to substitute for the heterogenous users' needs that can be fulfilled by a primary trading floor as LSE.

Table A.6 and A.7 present the yearly and quintile based market fragmentation statistics respectively. The evolution of market fragmentation is shown in Figures A.5a and A.5b, where the former shows the traditional volume based measure and the latter a measure of quote fragmentation. The two figures reveal a notable difference between quote and volume fragmentation that quotes' fragmentation ($HHIqu5$) was more symmetrical across markets and followed a steeper path to reach its peak. Table A.6 provides evidence that markets fragmented more and faster in quoting activities than trade execution. Quotes' fragmentation is almost symmetrical across markets, though 75% of trades, on average, are executed only in CHIX and LSE.

Market liquidity

Tables A.8 and A.9 report the yearly pre and post-MiFID liquidity descriptive summary for the period 2005–2016. The liquidity measures presented in these tables are the absolute quoted spread ($spread\_abs$), the relative quoted spread ($spread\_bps$), the relative effective half-spread ($espread$), the 5-minute realized half-spread ($rspread$), the 5-minute price impact ($price\_impact$), the average best depth level at the best limit prices ($depth1$), and the average cumulative depth for the three best prices ($depth3$). Figures A.6, A.7 and A.8a present liquidity trends observed during the period 2005-2016.

All measures, except $depth1$ and $depth3$, reveal substantial improvement of liquidity. For instance, overall spreads decreased by more than one-half to 30 bps at the end of 2016 compared to 13.5 bps in 2005. Spreads decreased asymmetrically across the quintiles. Quoted spreads in the largest-cap group decreased around 70% (from 17 bps in 2005 to 5.5 bps in 2016) and 90% (from 48 bps in 2005 to 10 bps in 2016) for the smallest stock.

To make the percentage quoted spread comparable to other three liquidity measures—effective half-spreads, realized-half spreads and price-impacts—the former is to be scaled down by a factor of 2. This makes it equivalent to a half-round trip cost of a hypothetical transaction. As can be seen, effective-half spreads also decreased across quintiles. The average effective spread, for the whole sample, is 60% lower in 2016 than in 2015. Tables A.8 and A.9 show that effective spreads are smaller than quoted spreads (20 bps vs. 28 bps in the pre-MIFID period and 13 bps vs. 18 bps in the post-MIFID), which is evidently reflecting the hidden liquidity and also indicative of within-quote trading.

Tables A.10 and A.11 report the effective-half spread decomposition into realized-half spreads and price impacts. I develop four measures, $rspread1/prce\_impact1$, $rspread2/price\_impact2$, $rspread3/price\_impact3$ and $rspread4/price\_impact4$ based on 10 seconds, 30 seconds, 1-minute and 5-minute post-trade quote adjustment periods (Δ), respectively. The literature has mostly used the 5-minute quote update gap in doing the same decomposition. The idea of using the lower granularity of time in measuring realized spreads and price impacts is consistent with the recent proliferation of low-latency trading environments across European markets.

Between 2005–2016, both realized spreads and price impacts declined, which explains the same trend observed in effective spreads. One of the possible reasons for declining realized spreads is that, in the presence of HFT the market has achieved better execution quality over the period. This evidence of decreased realized spread differs from Hendershott et al. (2011), who report the opposite impact that was observed after the introduction of auto quote in NYSE (which they also refer to as an unexpected finding). Table A.10 shows that all the variants of realized half-spreads across quintiles turn negative in the post-MIFID period, whereas only 5-minute realized half-spread for mid-cap and large-cap stocks are seen to be negative in the pre-MIFID era. The higher the time gap between trade execution and post-trade quote update, the less the realized spreads are, and conversely for price impact measures. Generally, larger cap stocks have smaller realized spreads and price impacts—a well evident stylized fact commonly observed in liquidity analysis.

The study of the time-series evolution of liquidity, along with general equity market trend (Figure A.1) reveals that liquidity as measured by quoted and effective spreads is sensitive to economic and financial shocks. Events such as the financial crisis in 2008-2009, the euro-area sovereign debt crisis in 2011, the Chinese market crash in 2015, Brexit, US presidential election in 2015-2016 explain the extreme illiquidity conditions illustrated in the graph during the same periods.

Figure A.8a, shows that BBO level's depth started to decrease steeply for large-cap stocks since 2005, long before the implementation of MIFID. This might be attributable to the common driving forces of the similar trend that is also observed in average trade sizes (Figure A.8b), and a monotonically rising message traffic rate (FigureA.2) during the same period. The trend showing decreasing depth might be attributable to the narrowed spreads as market makers have less incentive to offer a large depth as usually done for wider spreads. The literature (Hendershott et al., 2011; Gresse, 2017; Aitken et al., 2014) argues for the possible link between the rising HFT intensity and decreasing trade sizes around the global financial marketplaces, especially in the last decade. HFTs typically trade in small lots and prominently use slice-and-dice strategy in executing the large order. Aitken et al. (2014) provide evidence that substantial changes in trade size is linked to the rise of HFT.

Trends revealed in trade sizes and quoted depths are consistent with the argument that recent changes in market microstructure, particularly HFT, have impacted trade sizes, which in turn have affected the average depth level. For example, Table A.9 shows that in the pre-MiFID period , 2005-2007, the average depth level in best price limit, $depth1$, and three best price limits, $depth3$, for all stocks are GBP854 and GBP3047 respectively and which decrease to GBP260 and GBP1331 respectively in the post MiFD period. Both periods are also associated with higher depth in large-cap stocks. Figure A.8a also reports that depth level offered for the best limit prices in large stocks appears to have increased steadily since 2009.

Correlations analysis

Table A.13 shows the correlation coefficients between liquidity, HFT, and market fragmentation measures respectively. Correlation between liquidty, HFT proxy ($hft2$) and market fragmentation proxy ($HHItrd$) show that HFT intensity is negatively correlated with quoted spreads (-0.78), effective spreads (-0.77), realized spreads (-0.15), price impacts (-0.58), and positively correlated with fragmentation proxy (0.55), BBO level depth (0.38), and cumulative depths (0.47). On the other hand, fragmentation proxy (HHItrd) is negatively correlated with quoted spreads (-0.40), effective spreads (-0.45), BBO level depth (-0.09), realized spreads (-0.04), price impacts (-0.58) and positively correlated with cumulative depths (0.08). These estimates show a positive association between liquidity and HFT and also between liquidity and market fragmentation; however, the correlation between HFT and liquidity is much stronger. In the existence of highly positive correlation between HFT and fragmentation, the apparently visible correlation among variables might be dubious and necessarily not indicate the causality. I examine the associations among these variable in the next section more systematically.

2.4 Research strategies, results and discussions

The setup and identification of the regressions models, empirical findings and their analysis are presented in this section. I present the basic setup and results using OLS in the section 2.4.1. In sections 2.4.2 and 2.4.3, I control for the possible endogeneity issues that could affect the OLS inferences through two alternative approaches, IV-GMM (H2SLS) and simultaneous equations estimation (H3SLS).

2.4.1 Basic setup and identification

The relationships among HFT, market fragmentation and liquidity are examined in six alternative linear specifications. The potential unobserved heterogeneity across firms is addressed by introducing stock fixed effects and time effects, in all six specifications. The basic specification starts with the HFT proxy, the main regressor of interest. The first specification is:

$$MQ_{it} = \alpha_i + \gamma_t + \beta_1 HFT_{it} + \sigma' X_{it} + \epsilon_{it}, \tag{2.1}$$

where MQ_{it} represents one of the daily (t) market quality measures ($spread\_bps$, $espread$, $depth1$, $depth3$, $rspread$ or $price\_impact$) for stock i, HFT_{it} represents one of the HFT proxies ($hft1$, $hft2$ or $hft3$), the vector X_{it} includes three control variables, log normalized market capitalization ($Log(mktcap)$), log normalized intraday mid price volatility ($Log(voltintra)$) and inverse of daily average prices ($invprice$), which are commonly evident as liquidity determinant in empirical market microstructure literature, α_1 is the firm fixed effects, and γ_t is the time fixed effects.

The expanded second specification on (2.1) includes the daily market fragmentation proxy, $Mfrag_{it}$, with HFT_{it}, so that the impact of both HFT and market fragmentation can be assessed in the same model. The second specification is :

$$MQ_{it} = \alpha_i + \gamma_t + \beta_1 HFT_{it} + \beta_2 MFrag_{it} + \sigma' X_{it} + \epsilon_{it}, \tag{2.2}$$

where $MFrag_{it}$ is the market fragmentation proxy, measured by $HHItrd$.

The idea of the interaction effect between HFT and market fragmentation arises from the fact that the level of fragmentation and HFT are likely to influence each other. For example, if HFT firms wish to engage in market making across markets they must rely on stocks which are fragmented. Consequently, the level of HFT participation across markets is likely to impact the fragmentation level itself. It can be argued that if there exists an impact of HFT and fragmentation on market quality, then the level of each on which the other ride might also play a role in determining the extent of impact. It might be the case that the effect of HFT on the market quality of a low fragmented stock is different from that of a high fragmented stock, or that the level of HFT participation also determines the extent to which market fragmentation impacts stock's liquidity or both. At this point, I expand the model (2.2) to include an additional interaction term between HFT and market fragmentation proxies. The third specification is:

$$\begin{aligned} MQ_{it} = {} & \alpha_i + \gamma_t + \beta_1 HFT_{it} + \beta_2 MFrag_{it} + \beta_3 HFT_{it} * MFrag_{it} \\ & + \sigma' X_{it} + \epsilon_{it}. \end{aligned} \tag{2.3}$$

Models defined so far are time-invariant. If the impact of HFT and market fragmentation is not static then the above three models only provide an over the period average estimates. To see the estimates' dynamics, model (2.1)–(2.3) are extended to include period dummies. The fourth, fifth and sixth specifications address the time-varying impact of HFT, market fragmentation and their interaction. The expanded fourth specification on (2.2) includes three additional time interaction dummies, $DYr(8, 9, 10)$, $DYr(11, 12, 13)$ and $DYr(14, 15, 16)$ which represent the respective period dummy, with HFT_{it}. The fourth specification is :

$$\begin{aligned}MQ_{it} &= \alpha_i + \gamma_t + \beta_1 HFT_{it} + \beta_2 MFrag_{it} + \beta_3 HFT_{it} * MFrag_{it} \\ &\quad + \beta_4 HFT_{it} * DYr(8,9,10) + \beta_5 HFT_{it} * DYr(11,12,13) \\ &\quad + \beta_6 HFT_{it} * DYr(14,15,16) + \sigma' X_{it} + \epsilon_{it}.\end{aligned} \tag{2.4}$$

The model (2.4) compares the HFT impact on liquidity in three equally divided post-MiFID eras, (2008–2010), (2011–2013), (2014–2016) with that averagely observed in the pre-MiFID period (2005–2007).

The expanded fifth specification on (2.2) includes two additional time interaction dummies, $DYr(11,12,13)$ and $DYr(14,15,16)$, with market fragmentation proxy ($MFrag_{it}$), which represent two evenly divided terms of the period 2011-2016. Essentially, the fifth specification is :

$$\begin{aligned}MQ_{it} &= \alpha_i + \gamma_t + \beta_1 HFT_{it} + \beta_2 MFrag_{it} + \beta_3 HFT_{it} * MFrag_{it} \\ &\quad + \beta_4 MFrag_{it} * DYr(11,12,13) + \beta_5 MFrag_{it} * DYr(14,15,16) + \\ &\quad + \sigma' X_{it} + \epsilon_{it}.\end{aligned} \tag{2.5}$$

The model (2.5) compares the market fragmentation impact in (2011-2013), (2014-2016) with that observed, on average, in the initial three years (2008-2010) of market fragmentation.

Finally, the expanded sixth specification on specification (2.3) includes two more HFT_{it} and $MFrag_{it}$ interaction terms interacted with the two time dummies, $DYr(14,15,16)$ and $DYr(11,12,13)$ respectively. The model is expected to assess the time-varying impacts of market fragmentation and HFT interaction on market quality. The sixth and final specification is :

$$\begin{aligned}MQ_{it} &= \alpha_i + \gamma_t + \beta_1 HFT_{it} + \beta_2 MFrag_{it} \\ &\quad + \beta_3 HFT_{it} * MFrag_{it} + \beta_4 HFT_{it} * MFrag_{it} * DYr(11,12,13) \\ &\quad + \beta_4 HFT_{it} * MFrag_{it} * DYr(14,15,16) + \sigma' X_{it} + \epsilon_{it}.\end{aligned} \tag{2.6}$$

Results and discussions

To avoid the econometric pitfalls due to unbalanced panel estimation, a balanced panel is constructed from the primarily selected sample with 132 stocks and 2624 trading days (for the period December 2005–December 2016). For cross sectional comparison, the sample is also classified into two equal quintiles of 66 stocks each, based on market capitalization. To facilitate time-varying analyses, the sample is further divided into four periods: (i) 2005–2007, represents the pre-MiFID era; and (ii) three equally divided periods, 2008–2010, 2011-2013 and 2014-2016, represent the post-MiFID distinct phases. Table A.12, calculated from the balanced panel for the full sample and subsamples, shows the relevant descriptive statistics of regression variables which should be used as average reference value for all regression estimates. All measures, other than realized-half spreads, are natural log transformed. Realized-half spreads, on average, are negative and not transformable. The message traffic rate developed for the five best limit prices ($hft2$) is used as the HFT proxy in all analyses otherwise it is not stated since a comparison across three HFT measures ($hft1$, $hft2$ and $hft3$) shows that the impact of $hft2$ is the strongest, which is also presented in a later section (2.4.1.6).

Models (2.1)–(2.3) are estimated employing both full sample and subsamples of large and small stocks for different liquidity measures. Models (2.4)–(2.6) are estimated employing the four periodical subsamples. The coefficient estimates in all models are OLS, and the standard errors are computed using the Newey-West HAC estimator, a heteroscedasticity and autocorrelation consistent covariance matrix estimator (lags for autocorrelation are optimally determined). All estimates include both time (daily) and stock fixed effects.

The remainder of this section is organized as follows. The estimates for Models (2.1)–(2.3) for different liquidity measures employing the full sample are presented in the subsections 2.4.1.1 and 2.4.1.2. Subsection 2.4.1.3 presents the estimates for the same models employing large and small stocks's subsamples. Subsection 2.4.1.4 presents an analysis on the estimates for realized half-spreads and price-impacts (a decompositions of effective half-spreads). Subsection 2.4.1.5 presents the estimates for Models (2.4)–(2.6) employing four periodical subsamples for three groups of stocks—All, large and small stocks. Subsection 2.4.1.6 presents the analysis comparing alternative HFT proxies employing the full sample.

2.4.1.1 HFT and market fragmentation

The regression results for the liquidity measures employing the whole sample are reported in Table A.14. Upper panel (Panel A) reports the result for the first three liquidity measures—$spread\_bps_{it}$, $espread_{it}$ and $depth_{it}$— and the lower Panel (Panel B) reports the rest—$depth3_it$, $rspread\_5min_{it}$, and $price\_impact\_5min_{it}$. Sub columns (1), (2), (3) report the results for the models (1), (2) and (3) respectively. The coefficients of $Log(hft2)$ and $HHItrd$ measure the association of HFT intensity and market fragmentation with the market quality measures respectively. We observe from the coefficient of $hft2_{it}$ that higher HFT intensity is associated with lower quoted and effective spreads, lower depth (in the both $depth1$ and $depth3$), lower realized spreads and price impacts. On the contrary, the coefficients of $MFrag_{it}$ show that higher fragmentation is associated with higher quoted and effective spreads, lower BBO depth, higher depth in the deep of the order book (not significant), lower realized spreads and higher price impacts.

We must be careful in the interpretation of the estimated coefficient since the variables used in the estimation have different measurement scales. For example, to interpret the coefficient of $hft2$, we would recall that $hft2$ measures the per-minute electronic message trafficking rate and the liquidity measures—other than the $depth1$ and $depth3$—are measured in basis point and $depth1$ and $depth3$ are measured in GBP100. Since both $hft2$ and market quality measures are log normalized, the interpretation becomes easier and does not depend any more on the unit of measurement. Thus, the estimate of -0.288 of $Log(hft2)$ (Panel A, column I) means that, ceteris paribus, 1% increase in the HFT is associated with 0.288% decrease in quoted spread. For instance, a one standard deviation increase in HFT from its sample mean of 95 messages/per-minute to 220 ($Log(220/95) \approx 88\%$) would narrow quoted spreads by 25% (88*0.288) i.e.

the sample mean of quoted spreads would go down from 18.37 bps to 13.77 bps (for descriptive statistics, see Table A.12). The coefficient of $HHItrd$ is not log transformed, so the estimate of 0.052 (Panel A, column II) against the log transformed quoted spreads means that, ceteris paribus, a unit increase in $HHItrd$, for example, from the sample mean of 2.17 to 3.17 is associated with 5.2% increase in quoted spreads.

The coefficients of $hft2$ and $HHItrd$ on $rspread\_5min$ have different interpretations due to the use of a different measurement scale. The estimate of -1.871 of $hft2$ (Panel B, column V) means that a one percent increase (decrease) in the average $hft2$ would decrease (increase) the average realized half-spreads by 0.0187 bps. Since both the $HHItrd$ and $rspread\_5min$ are employed in the regression in their original unit of measurement, the estimate of -0.245 (Panel B, column V) means that a unit increase (decrease) in $HHItrd$ is associated with a decrease (increase) of 0.245 bps in the average realized half-spreads.

The impact of $hft2$ and $HHItrd$ on the average quoted depth is negative, which implies that both the HFT and fragmentation are associated with less average quoted depth. The positive sign of the coefficient $HHItrd$ in column II (Panel B) implies that more fragmentation is associated with more quoted depth in the deeper level of an limit order book, however the coefficient is not significant. One might argue that the depleted market liquidity through the quoted depth is likely to outweigh the benefit of the liquidity added through the narrower quoted and effective spreads. Hendershott et al. (2011) perform a calibration exercise to overcome the doubt and concluded that the depth reduction is small relative to the narrowing of the spread.

The coefficients of the control variables have the expected signs and are significant at the 1% level. The large market capitalized stocks are associated with lower quoted and effective spreads, lower price impacts, greater depth and higher realized spreads. The inverse price coefficient implies that a stock with higher price is associated with lower quoted and effective spreads, higher depth and lower realized spreads and price impacts. The positive estimate of the volatility coefficient implies that a higher intraday volatility increases quoted and effective spreads, provides greater depth in the best price, and is associated with higher price impact and lower realized spread.

2.4.1.2 Interaction between HFT and market fragmentation

I now turn to the most interesting specification (2.3) and the corresponding estimates for β_3 (column III, VI, IX of Panel A and B) presented in the Table A.14. The coefficient β_3 measures the interaction effect between HFT and market fragmentation. The coefficients associated to both main effects, $hft2$ and $HHItrd$, are all significant at the 1% level other than in column III (Panel A), which is only significant at the 10% level, and in column III (Panel B), which is not significant at all. All the interactions signs are negative except that for the realized-half spread. A negative sign (column III and VI) for the interaction estimate implies that an increase in HFT is associated with higher liquidity when fragmentation is greater, and an increase in fragmentation is associated with less harm to liquidity when HFT is higher.

Since in the presence of interaction term ($Log(hft2) * HHItrd$) the interpretation of the coefficient estimates of the variables $hft2$ and $HHItrd$ become tricky, it is useful here to explain the interpretation of the parameters in model (2.3). In the given setting, the partial effect of $hft2$ on liquidity depends on the average level of $HHItrd$, and vice versa. Let us define a general expression for the partial effect. Partial effects of $hft2$ and $HHItrd$ (on liquidity) can be defined as $\Delta MQ_{it}/\Delta(HFT_{it}) = \beta_1 + \beta_3 * MFrag_{it}$ and $\Delta MQ_{it}/\Delta(MFrag_{it}) = \beta_2 + \beta_3 * HFT_{it}$ respectively. To interpret partial effects, these expressions should be evaluated at some interesting values. I evaluate them at sample mean, which is a well-practised norm among academics.

For instance, the estimated coefficient on the interaction term between HFT and market fragmentation against the response variable effective-half spread is -0.02 (Table A.14, Panel A, Column VI), and the respective estimates for the partial effects of HFT ($\Delta espread/\Delta(Log(hft2))$ and market fragmentation ($\Delta espread/\Delta(HHItrd)$ are ($-0.282-0.02*2.17 \approx$) -0.32 and ($0.144-0.02*Log(96) \approx$) 0.053 respectively[5]. These estimates are close to the estimated coefficients on the same variables where the interaction

[5]The general expressions for the partial effect are evaluated at the sample means of HFT (96) and market fragmentation (2.17) as reported in Table A.12.

effects are not introduced (column V of Panel A). I do not report joint significance tests for partial effects, which are trivial as both interaction and main effects are significant.

The model (3) greatly expands the understanding of the relationship of HFT_{it} and $MFrag_{it}$. The results indicate that models without interaction effects could lead to a misspecification as the effect is statistically significant, and one could argue that the model (2.1) and model (2.2) are poorly specified as no interaction terms are included. The marginal effect of $hft2$ and $HHItrd$ in model (2.3) are generally smaller than observed in (2.2) and (2.1), which may imply that some of the possible benefits of HFT intensity on market liquidity is offset by the extra cost of market making in a fragmented market. Conversely, some extra cost of market fragmentation is also offset by the benefits derived from HFT . It seems that a fragmented market would be more detrimental to liquidity if there was no HFT.

2.4.1.3 Large and small stocks

Table A.15 reports the estimates of the model (2.1)–(2.3) for large and small stocks. For brevity, I do not report the estimates for control variables which are significant, at the 1% level, and have the same signs as they appear in the estimated coefficients for the full sample in Table A.14. Panel A1 and B1 report the estimates for large stocks and Panel A2 and B2 do the same for small stocks.

The estimated coefficients for $Log(hft2)$ and $HHItrd$ are significant (at the 1% level) and larger than the respective estimates for small stocks, which implies that both HFT and market fragmentation impacted the liquidity of large stocks more (column I, II, IV, V of both Panel A1 and A2). This difference in estimates between large and small stocks also may imply that HFT is associated with more liquidity for large stocks whereas market fragmentation is more liquidity detrimental for small ones.

Table A.15 shows that the estimated interaction coefficients on HFT and fragmentation for large stocks are significant, at the 1% level (column III, VI and IX in both Panel A1 and B1), and that for small stocks show a few exceptions (column III, VI and IX in both Panel A2 and B2). The results of the Wald test, a test for joint hypothesis, are reported where marginal effect of market fragmentation ($HHItrd$), or interaction between market fragmentation and HFT ($HFT * Mfrag$), or both are not significant.

As can be seen from Table A.15, the estimated interaction coefficients on HFT and fragmentation against quoted and effective spreads for large stocks are negative, but those for small stocks are positive. This contradiction in the direction of interaction effect between large and small stocks hint to two different implications. An increase in HFT for a higher fragmentation or an increase in fragmentation for a higher HFT activity is associated with a higher liquidity for large stocks whereas the same appears detrimental to liquidity for small stocks, which suggests that fragmentation is more liquidity harming for small stocks.

Table A.15 also reports an interesting finding regarding the effect of HFT and market fragmentation on the quoted depth. The estimated coefficients on fragmentation ($HHItrd$) and HFT ($hft2$) variables against the average quoted depth at best price ($depth1$) appear to have the same sign in all three group of samples—Full, large-cap and small-cap. The estimated coefficient on fragmentation ($HHItrd$) is positive and significant at the 1% level (Panel B1 , column III), which implies that market fragmentation contributed more liquidity to the deeper of the order book in contrast to the HFT which decreased quoted depth at both best price($depth1$) and beyond that ($depth3$).

2.4.1.4 Sources of liquidity supply

The results reported in Table A.14 and A.15 for the effective-half spread decomposition into realized-half spreads and price impacts are discussed in this section. As reported earlier in the previous section, column (IV)–(VI) in Panel A of Table A.14 and the same columns in Panel A1 and A2 of Table A.14 show that an increase in HFT is associated with lower effective spreads whereas an increase in market fragmentation results in greater effective spreads. The column (IV)–(IX) in Panel B of Table A.14 and the same column in Panel B1 and B2 of Table A.15 reports the coefficients associated with realized-half spreads and price impacts.

A narrower (greater) effective spread implies either less (more) revenue per trade for liquidity providers, or smaller (larger) gross loss due to informed liquidity demanders, or both (Hendershott et al., 2011). Respective columns in both tables show that both HFT and market fragmentation positively impacted the realized-half spread and the price impacts measured at 5-minute intervals and the results are robust across all stocks group. Apparently, HFT has improved the execution quality, and it can be seen in the descriptive statistics (Table 2.23) that the realized spread has turned to almost negative in the post-MiFID era. The negative effects on realized-half spreads indicate that liquidity providers are earning less revenue per trade than before. On the other hand, liquidity providers are also losing less to liquidity demanders (less adverse selection). But the significant interaction effect between $HHItrd$ and $hft2$ implies that higher HFT in more fragmented market, or more fragmented market with higher HFT is detrimental to both realized-half spreads and price impacts. The reported effects of HFT on the realized spread and price impacts are quite expected in HFT environments. It seems that, the low-latency environment increases HFTs' market making capability in every side of the order book which in turn reduces both realized spreads and price impacts. If we consider HFT as a market making agent, one would rationally expect that HFTs try to design algorithm such as to minimize the adverse selection cost and HFT as a liquidity demander would try to do the same to get better execution. I excerpt few relevant lines from a recent trading magazine's article by Rick Baert (http://www.pionline.com).

"Money managers and internally managed pension funds are expected to follow the lead of T. Rowe Price Group Inc. in sending direct equity order flow to high-frequency trading firms, which could chip away at the established use of brokers to direct institutional trades, sources said...Added Valerie Bogard, equity analyst at TABB Group LLC, New York: 'The buy side has gotten more comfortable with high-frequency trading firms, and they weren't before. A lot of their strategies used to make the buy side uncomfortable. But now the buy side understands much better how those firms work and they're ramping up their transaction cost analysis, and HFT firms are providing the liquidity they need.'...T. Rowe Price executives said in the past two years the real success of the program has been in finding liquidity—a growing problem for institutions as fewer stocks trade on public markets and more institutional investors move toward passive investing...' **The execution quality has been solid,'** said Mehmet Kinak, vice president and head of global equity market structure and electronic trading at T. Rowe Price, Baltimore..."

2.4.1.5 Time-varying impact

This section presents the regression results for the specification (2.4)–(2.6). Model (2.4) examines the incremental impact of HFT for the last three three-year consecutive periods. Model (2.4) essentially examines the incremental HFT impact on the last 9 years sample period, in three equally divided blocks, with the beginning three years. Model (2.5) examines the incremental impact of market fragmentation for the last two three-year consecutive periods. Model (2.6) examines the incremental interaction effect of HFT and market fragmentation in the last 6 years of the sample, divided into two equally divided three-year blocks, with that of the first three years since 2008. Table A.16 reports the results for the full sample and Table A.17 and A.18 report the same for both large and small stocks groups. For brevity, Table A.17 and A.18 do not report coefficient estimates of control variable which are significant and have the same expected sign as in Table A.16.

All coefficients in Table A.16 are significant at the 1% level, except one (panel A, column I) and show that incremental effects of HFT ($hft2$), market fragmentation ($HHItrd$) and interaction of market fragmentation and HFT ($hft2 * HHItrd$) are highly significant. This implies that HFT and market fragmentation impacted market liquidity differently in different periods. The average impact of HFT and market fragmentation (as reported previously in Table A.14) remained the same. For example, the coefficient of $Log(hft2) * DYr8, 9, 10$ (Panel A, column I) is -0.03 which implies that compared to the period 2005–2007, a 1% increase in HFT during 2008-2010 is associated with 0.03% extra narrower spreads. We also observe a similar effect during this period in effective spread, realized-half spreads and price impacts. This result is very significant in the sense that the period 2008-2010 is a period of high illiquidity and high volatility due to global financial crisis (as can be seen in Figures A.6a or A.6b). The evidence provided in Hasbrouck and Saar (2013) also confirm the similar effect.

There might be different reasons why HFT could create positive externality in the market at the time that the market needs it the most. Higher volatility might create more profitable opportunity for HFT traders. In a high illiquid period, HFT even might find small stocks—which are generally less liquid—profitable (Hasbrouck and Saar, 2013). Table A.14 also shows that in the subsequent periods, since 2008 compared to 2005–2007, higher HFT in small cap stocks is associated with less narrower quoted and effective spreads. But among them, the incremental impact was the lowest during 2008–2010. As long as quoted depth is concerned, HFT has impacted the depth mostly during 2008–2010. One of the possible explanations might be that after the enactment of MiFID, the influx of HFT impacted the trade sizes and quoted depths significantly. The coefficients of $HHItrd * DYr11, 12, 13$ and $HHItrd * DYr14, 15, 16$ (panel A, column II,IV) are positive which confirm the implication of the result provided in Table A.14. This means that increased market fragmentation has depleted more liquidity during this period as measured by quoted and effective spreads.

The previous results on HFT and market fragmentation show that HFT has improved more liquidity in the latter periods (2011-2016) whereas in the same period market fragmentation has been seen to be more detrimental to liquidity. The columns (III) and (VI) of Panel A (Table A.16) show how the coefficient of HFT and market fragmentation interaction changed over the periods. As can be seen, the sign of the interaction effect has turned to be less intense during 2011–2013 and 2014–2016 compared to the period 2008–2010 , though the interaction effect is still negative and significant. It seems that the increased fragmentation has offset, at least partially, the additional liquidity which has been generated by the higher HFT during the latter periods. Column (II) of Panel B again confirms that higher fragmentation is associated with higher quoted depth in the deeper of the order book whereas at best price both HFT and fragmentation decreased the quoted depth. Column (VI-IX) of Panel B shows that both HFT and market fragmentation improved price impacts positively over the period.

It is noteworthy to recall that all period-dummy interactions show the incremental effect which should be compared with the beginning base period. For example, in

Table A.17, the coefficient of $Log(hft2) * DYr11, 12, 13$ (Column I, Panel A) is -0.112 (incremental effect) that has to be compared with -0.279 (the base coefficient of $Log(hft2)$). The base coefficient of $Log(hft2)$ during 2011–2013 can be found by adding the incremental coefficient with the base coefficient which is 0.391 ($-0.279 - 0.112$). Thus, the incremental coefficient -0.112 implies that an increase in 1% HFT during 2011-2013 is associated with 0.112% more narrower spread compared to the base period 2005-2007.

I now examine the period-dummy effects in different stock groups. Table A.17 shows that since 2008, higher HFT in both small and large stocks are associated with narrower quoted and effective spreads, on the contrary, higher fragmentation is more liquidity detrimental to small stocks. The incremental interaction effect of HFT and market fragmentation in large stocks (column III, VI in Panel A) has become stronger (more negative). For large stocks, it seems that in the latter periods, a greater fragmentation when HFT is higher, or a higher HFT when fragmentation is greater is associated with more liquidity, whereas that for small stocks is more detrimental to liquidity (column III, VI in Panel B). Table A.17 and A.18 also show that since 2008, an increase in HFT and fragmentation, both are associated with far less quoted depth in large stocks whereas more fragmentation has improved the quoted depth of smaller stocks but not the HFT. Turning to the impact on realized spreads and price impacts, it appears that the impact of HFT on realized spread has decreased, in both small and large stocks, over the period, while the effect on price impacts has increased. Among all the models, realized half-spread shows the worst fit. The effect of fragmentation on price impacts, in both small and large stocks, has increased over the period, but that on realized spreads is the reverse. The interaction effect on realized spreads, in both large and small stocks, is positive, but stronger in small stocks. The same effect on price impacts is negative for both groups of stocks.

2.4.1.6 A comparison of alternative HFT proxies

In this section, I estimate Models (2.1)–(2.3) by employing three alternative HFT proxies tracking the HFT footprint for three different depth levels of the limit order book. So far, $hft2$, a measure based on quotes updates for five best limit prices, is used for all estimates. As mentioned in section 2.3.2, the remaining two proxies, $hft1$ and $hft3$ measure the message traffic rate for the best 10 depth levels and the first best level (BBO) respectively. Usually, the literature uses the maximum depth levels provided in the dataset, however, it is not a common practice [6]. The exercise I perform here, is likely to expand the understanding on the HFTs liquidity supply in the deeper level of a limit order book.

Table A.19 reports the regression estimates of $hft1$, $hft2$, and $hft3$ for different liquidity measures. We see that all the estimates across the three HFT proxies are significant and $hft2$ is showing more reliable estimates. For example, column (III) of Panel A reports the estimates -0.275, -0.289 and -0.265 for $hft1$, $hft2$ and $hft3$ respectively. In the same column, it can also be seen that HFT and fragmentation interaction coefficients are not significant except for $hft2$ (subpanel A2). These results are robust across all liquidity measures.

These results provide some insight on the soundness of HFT proxies used in this study. Apparently, the average level of $hft1$ is greater than $hft2$, and $hft1$ is likely to provide the most reliable estimates. But, the analysis of HFT on alternative depth levels does not confirm this outward observation. It appears that neither $hft1$ nor $hft3$ provides as much variations in regressions as supplied by $hft2$, which is indicative of HFTs participation beyond the first best limit price, but not too far from the five best limit prices, and consitent with the evidence in Financial Markets Regulator (France) (2017). These results also support the literature evidence that HFTs provide both tight (marketable) and wider (non-marketable) quotes.

2.4.2 A two-stage optimal IV-GMM regression (H2SLS) approach

Endogeneity issues are commonly admitted in the HFT and fragmentation literature (e.g. Hendershott et al. (2011); Hasbrouck and Saar (2013); Degryse et al. (2015); Gresse (2017)). In the presence of endogeneity, establishing a causal link between HFT and

[6] Boehmer et al. (2015) use only the BBO level depth (as datasets do not provided access beyond the first best limit price), a few studies employ 10 best limit prices, and Financial Markets Regulator (France) (2017) considers the three best limit prices.

market quality or fragmentation and market quality is challenging. It might be the case that HFT and market quality, or market fragmentation and market quality, or both are simultaneously determined in equilibrium, so that there may be bi-directional causality. For example, as Hasbrouck and Saar (2013) explain "an exogenous drop in spreads might establish a more attractive environment for, and lead to increase in, low-latency activity." This mechanism would induce correlation between HFT proxy and the error terms, which makes OLS estimates inconsistent. It might also be the case that a highly liquid stock is more fragmented than an illiquid one. There is also the possibility that fragmentation is driven by the HFT, or a highly fragmented market attracts more HFT, and eventually both impact market quality. If this is the case then it challenges all the previous specifications. A simultaneous structural equations estimates is more appropriate, and indeed this is how the endogeneity problem is addressed in the next section.

To tackle the possible biases arising from endogeneity, one solution could be adopting an IV approach where one needs to find at least one or more distinct instruments, at least one for each of the two endogenous variables HFT_{it} and $MFrag_{it}$, which should be correlated with HFT_{it}/ $MFrag_{it}$ but not with the error terms (ϵ_{it}). There should be instruments which meet the above criteria and consistently show the same association over the whole sample period (December 2005–December 2016).

Hendershott et al. (2011) use the event of NYSE's quote automation in 2003 that increases AT as an exogenous instrument where they cover a relatively short post event period, 2003-2005. Boehmer et al. (2015) use the starting date of colocation hosting by the exchanges across countries as an instrument in their sample. Neither of the approaches is well suited in my case. The reason is LSE underwent several market structural changes throughout the period 2005–2016 to enhance its low-latency environment. For example, LSE launched accelerated hosting service (in its limited form) in September 2008 which full service form was again launched in September 2009. In one of their latest attempts, LSE migrated its UK cash markets to a new ultra low-latency trading platform, Millennium Exchange, in February 2011. Further, LSE introduced sponsoring access to non-member clients from June 2011. LSE also signed contracts with third parties which also provide low-latency access. If we analyse these market structure changes, we would see that no particular event explains the HFT intensity observed in the LSE (see Fig.A.2).

One of the potential problems of identifying HFT/AT through this approach (colocation instrumenting) is that it does not go with the idea that presence of HFT in market leads to the introduction of colocation service not the opposite. Exchanges choose to offer colocation services in reponse to the low-latency demand of HFT (Aitken et al., 2014). Menkveld (2016) argues that there is a bi-directional loop between a modern trading platform and HFT requirements. Aitken et al. (2014) also shows that HFT activities have been documented quite earlier on average than the date of first colocation hosting event by an exchange. Historically, trading firms seeking speed located themselves next to or across the street from the exchange well before the colocation hosting within the exchange.

I choose a similar approach to Hasbrouck and Saar (2013), Degryse et al. (2015) and Gresse (2017) where they developed instruments from the existing available measures of HFT_{it} and $MFrag_{it}$. Let us consider, any of the specifications in (2.1)–(2.6) on which OLS coefficents are estimated in the last section. A sound instrument, for HFT intensity, should be (i)correlated with HFT_{it}; (ii) and not correlated with the ϵ_{it}. Hasbrouck and Saar (2013) argue that if low-latency activity has a significant market wide component then a market wide average of HFT_{it} is likely to satisfy the first requirement. They explain that funding constraints or inventory risk management might be the causes for which HFT participate across stocks, which means that HFT_{it} is related with HFT_{-it}, and determined by the market wide HFT factors.

This exercise is performed for the fist five specifications to show that my previous OLS estimates are robust even in IV setting due to employing large samples and implemented control. The basic three specifications (2.1)–(2.3) used in the previous section are redefined in the following instrumental variable modelling setup. The redefined specifications are:

$$HFT_{it} = a_{1,i} + \sum_{t=1}^{W} \omega_{1,t} + b_1' Z_{it} + c_1' W_{it} + \varepsilon_{1,it}, \tag{2.7}$$

$$MFrag_{it} = a_{2,i} + \sum_{t=1}^{W} \omega_{2,t} + b_2' Z_{it} + c_2' W_{it} + \varepsilon_{2,it}, \tag{2.8}$$

$$HFT * MFrag_{it} = a_{3,i} + \sum_{t=1}^{W} \omega_{3,t} + b_3' Z_{it} + c_3' W_{it} + \varepsilon_{3,it}, \tag{2.9}$$

$$MQ_{it} = \alpha_i + \sum_{t=1}^{W} \gamma_t + \theta_1 \widehat{HFT_{it}} + \mu' W_{it} + \eta_{1,it}, \tag{2.10}$$

$$MQ_{it} = \alpha_i + \sum_{t=1}^{W} \gamma_t + \theta_1 \widehat{HFT_{it}} + \theta_2 \widehat{MFrag_{it}} + \mu' W_{it} + \eta_{2,it}, \tag{2.11}$$

$$MQ_{it} = \alpha_i + \sum_{t=1}^{W} \gamma_t + \theta_1 \widehat{HFT_{it}} + \theta_2 \widehat{MFrag_{it}} + \theta_3 \widehat{HFT * MFrag_{it}} + \mu' W_{it} + \eta_{3,it}, \tag{2.12}$$

where the vector Z_{it} represents the instrumental variables which are excluded from the second-stage regressions (2.10), (2.11), and (2.12), the vector X_{it} includes control variables, $\widehat{HFT_{it}}$, $\widehat{MFrag_{it}}$ and $\widehat{HFT * MFrag_{it}}$ representing the predicted values of HFT_{it}, $MFrag_{it}$ and $HFT * MFrag_{it}$, respectively, generated from the three first-stage regression equations (2.7) (2.8) and (2.9) and used in the respective second-stage regression equations (2.10) (2.11) and (2.12), $\sum_{t=1}^{W} \omega_t$, and $\sum_{t=1}^{W} \gamma_t$ are the time fixed- effect, a_i and α_i are the firm-fixed effect in the first-stage and second-stage regression respectively. To ease the computational burden, I use weekly time dummies for each of the 591 weeks instead of using daily. The following section describes the vector for instrumental variables and control variables.

X_{it} includes log market capitalization ($Log(mktcap)$), log intraday mid price range volatility ($Log(voltintra)$), price inverse ($invprice$) and the average degree of liquidity of stocks in the same size group excluding stock i ($\overline{MQ_{-it}}$), calculated from the four equally divided firm size group based on market capitalization. I follow Degryse et al. (2015) to include, $\overline{MQ_{-it}}$, in the vector of control variables in addition to the control variables used for OLS estimation. Anyway, using this control variable also in OLS estimation does not change any of the existing estimates.

In the specification (2.7) (2.8) (2.9), HFT_{it} and $MFrag_{it}$ and $HFT * MFrag_{it}$ are the potential endogenous variables. The first set of instruments I consider for these three variables are the daily average of each variables over all stocks in the same size group excluding stock i. Hasbrouck and Saar (2013) adopted this approach which was latter also applied in Degryse et al. (2015) and Buti, Rindi and Werner (2011).

One of the short comings of using only this instrument is that by construction it decreases both between and within variations of HFT intensity which are observed in the original dataset. Consequently, dataset losses the original panel's inherent power of distinguishing a high HFT intensed stock from its low counterpart. This problem is more visible in a panel where apparent heterogeneity across stock is substantial. This problem also persists in my dataset.

At this point, I proceed to look for relevant instruments with a motivation from Gresse (2017). I add three more instruments, log trading volumes ($Log(value)$), average trade sizes ($Log(size)$), and relative tick sizes ($rtick$) in the first stage regression to increase the predictive power of $\widehat{HFT_{it}}$, $\widehat{MFrag_{it}}$ and $\widehat{HFT * MFrag_{it}}$ in my instrument set. I explain here the rationale of choosing these variables as instruments. Evidence shows that recent influx of HFT across financial market places is associated with the changes of some particular market microstructures. For example, O'Hara, Saar and Zhong (2014) and Hagströmer and Nordén (2013) report how variations in relative tick sizes affects the HFT activities. In a large relative tick size environment, HFTs leave orders in the book longer, trade more aggressively, and have higher profit margins than a small one.

Trading volume is also directly related to both HFT intensity and market fragmentation. Gresse (2017) explains that the portion of AT corresponding to HFT market making is expected to be more profitable and thus more developed for heavily traded stocks. The evidence in Hendershott et al. (2011) and Brogaard, Hendershott and Riordan (2014a) also show that large stocks attract more HFTs.

HFTs typically trade in small lots and prominently use slice-and-dice strategy in executing the large order. Hendershott and Riordan (2013) show that by splitting large orders into smaller slices, algorithmic traders reduce their own market impact but also the volatility of liquidity in general. The literature supports the link between the rising HFT intensity and decreasing trade size in the recent years and attribute the causality from HFT to trade sizes (Hendershott et al., 2011; Aitken et al., 2014).

Results and discussion

To estimate the Models (2.8)–(2.12), I use the same panel dataset employed in OLS estimation, with 132 stocks and 2624 days for the period December 2005–December 2016. I use two-stage optimal IV-GMM (H2SLS) estimator. The inference is based on standard errors that are robust to heteroskedasticity and autocorrelation (Newey-West HAC, based on 5 lags). Similarly, I also estimate the models (2.4) and (2.5) in the IV-GMM setting for which I do not mention the IV specifications explicitly. I also repeat the estimations for large and small stock groups. All specifications include stock and time (weekly) fixed effects.

Table A.20 reports the estimates for IV-GMM models (2.10)–(2.12) (equivalent to (2.1), (2.2) and (2.3) respectively in OLS) and similar IV estimation for OLS model (2.4) and (2.5) employing full sample, and Table A.21 shows the same for large and small stocks. For conserving space, estimates for control variables are not reported. To keep consistency with the previous table, sub columns(1)–(5) represent the equivalent IV-GMM estimation of OLS models (2.1)–(2.5). For a comparison between OLS and IV-GMM estimates, sub columns (1)–(3) of table A.20 should be matched with the respective sub columns of Table A.14, and sub columns (4)–(5) with the same sub columns of Table A.16. For small and large stocks, sub columns of Table A.21 should be compared with the same sub columns in Table A.15 and Table A.17. To avoid repetition, I do not discuss the results which are mostly similar to OLS.

The results in Table A.20 suggest that higher HFT intensity is associated with lower quoted and effective spreads, while higher fragmentation is associated with higher quoted and effective spreads. In fact, the estimates appear to be stronger for IV-GMM. Hasbrouck and Saar (2013) also find similar magnification when they switched from OLS to IV which they attributed to market wide role of HFT. I find only non-matching exception with OLS in sub column (3), but the partial effect generates the same sign as is observed in OLS. For example, coefficient of $HHItrd$ (column IX) shows a non-significant negative sign (-0.019), and if the partial effect is evaluated at average log normalize value of HFT ($Log(95.91) = 4.56$), then the estimate becomes 0.045, and is consistent with the previous OLS results. The control variables appear with same signs as in OLS and the sign in the additional control variable, average market liquidity, is positive as expected and significant.

It seems that instruments do not provide the estimation with enough variations to assess the interaction effect or complex specifications like (2.4)–(2.6). The incremental effects of both HFT and Market fragmentation are significant in the latter year of the sample periods but coefficients of $HHItrd * Dyr11, 12, 13$ and $HHItrd * Dyr14, 15, 16$ are not consistent with OLS estimates.

Table A.21 reports the same estimates for large and small stocks samples. It can be seen that models (1)-(5) for both small and large stocks confirm the same signs of OLS estimates more consistently than full sample. The only contradiction I find is that HFT and market fragmentation appear to have much stronger effect on small-cap stocks where OLS estimates reported the same for large-cap stocks. This apparent contradiction might be attributed to the weakness of the instrument, a classical problem in implementing IV.

In conclusion, the IV-GMM estimates confirm most of the results documented in the previous section, though there are some limitations to the identification strategy. To overcome these limitations, the next section turns to a simultaneous equations model approach to address the endogeneity problem.

2.4.3 A simultaneous equations model approach

I turn to a simultaneous equations model estimation approach to tackle the possible endogeneity among the market quality, HFT and market fragmentation. As explained in the previous section (2.4.2), there exist at least two possible mechanisms through which market liquidity and HFT affect each other: (i) the long trend of declining spread based transaction measures in the financial market places might be attributable to the low market making cost of HFTs; and (ii) the rising competition among HFT firms through huge investment in high speed trading technology. There is also evidence that both the level of liquidity or volatility affect the level of HFTs participation in market. On the other hand, if we consider HFT from the demand perspective, then the proliferation of modern low-latency-based trading venues should be attributed to response by the supply side (exchanges).

Market liquidity seems to play a role, being one of the determinants of order flow fragmentation where we see that more liquid stocks are more fragmented. It is more likely that liquidity affects the fragmentation decision than the other way round, though it is commonly accepted in the literature that market fragmentation also impacts liquidity. Over the last few years, the most active channel which affected the quoting and trading activities across markets is HFT. In this connection, the recent responses of the supply side on the rising HFT demand have rapidly increased the number of electronic exchanges with low-latency technology across the European equity market. It is apparent that there are simultaneity among HFT, market fragmentation and liquidity.

To tackle the simultaneity among HFT, market fragmentation and market quality, I consider simultaneous equations model which is relatively a new approach in market microstructure research. Buti et al. (2011) and Aitken et al. (2014) use similar arguments that market quality, fragmentation and HFT are jointly determined in equilibrium and they used simultaneous equations model in their setting. Hasbrouck and Saar (2013) also use the same approach in a more simple setting where an attempt has been made to determine the impact of low-latency on market quality.

I consider market quality (MQ_{it}), HFT (HFT_{it}) and market fragmentation ($MFrag_{it}$) are determined in equilibrium, and three equations are defined accordingly, one for each of the variables. I include the variables on the right hand side of MQ_{it}, HFT_{it} and $MFrag_{it}$ which are found to be determinants of each of the variables by the literature, and also explained in the IV-GMM section (2.4.2). This setup should overcome the limitations of IV-GMM specifications. The three-equation simultaneous model is:

$$
\begin{aligned}
MQ_{it} = {} & \alpha_{i_{(mq)}} + \sum_{m=1}^{M} \gamma_{(mq)m} + \beta_{1_{(mq)}} HFT_{it} + \beta_{2_{(mq)}} MFrag_{it} + \beta_{3_{(mq)}} \overline{MQ}_{-it} \\
& + \beta_{4_{(mq)}} Log(mktcap)_{it} + \beta_{5_{(mq)}} Log(voltintra)_{it} + \beta_{6_{(mq)}} inv(price)_{it} + \epsilon_{it_{(mq)}},
\end{aligned}
\tag{2.13}
$$

$$
\begin{aligned}
HFT_{it} = {} & \alpha_{i_{(hft)}} + \sum_{m=1}^{M} \gamma_{(hft)m} + \beta_{1_{(hft)}} MQ_{it} + \beta_{2_{(hft)}} MFrag_{it} + \beta_{3_{(hft)}} \overline{HFT_{-it}} \\
& + \beta_{4_{(hft)}} Log(size)_{it} + \beta_{5_{(hft)}} Log(value)_{it} + \beta_{6_{(hft)}} rtk_{it} + \beta_{7_{(hft)}} Log(mktcap)_{it} \\
& + \beta_{8_{(hft)}} Log(voltintra)_{it} + \epsilon_{it_{(hft)}},
\end{aligned}
\tag{2.14}
$$

$$
\begin{aligned}
MFrag_{it} = {} & \alpha_{i_{(frg)}} + \sum_{m=1}^{M} \gamma_{(frg)m} + \beta_{1_{(frg)}} HFT_{it} + \beta_{2_{(frg)}} MQ_{it} + \beta_{3_{(hft)}} \overline{MFrag_{-it}} \\
& + \beta_{4_{(frg)}} Log(value)_{it} + \beta_{5_{(frg)}} Log(mktcap)_{it} + \beta_{6_{(frg)}} Log(voltintra)_{it} + \epsilon_{it_{(frg)}},
\end{aligned}
\tag{2.15}
$$

where indices i and t represent stocks and day respectively, MQ_{it} represents one of the two log normalized market liquidity measures ($spread\_bps$, $espread$), HFT_{it} represents the HFT proxy ($hft2$) , $MFrag_{it}$ represents the market fragmentation proxy ($HHItrd$), $\overline{MQ}_{-it}$ represents average market liquidity level over all stocks in the same size group excluding stock i, $\overline{MFrag}_{-it}$ represents the average market fragmentation level over all stocks in the same size group excluding stock i, $\overline{HFT}_{-it}$ represents the average HFT intensity over all stocks in the same size group excluding stock i, $Log(mktcap)$ is the log normalized market capitalization, $Log(voltintra)$ is the log normalized intraday mid price range volatility, $invprice$ is the inverse of daily average price, $Log(size)$ is the log normalized trade sizes, $Log(value)$ is the log normalized trading volumes, rtk_{it} is the relative tick size, α_i is the firm fixed effects, $\sum_{m=1}^{M} \gamma_m$ is the time (month) fixed effects, index $_{(mq)}$, $_{(hft)}$, $_{(frg)}$ refer the respective coefficient of the equations MQ_{it}, HFT_{it} and $MFrag_{it}$ respectively.

I employ a balanced panel of 132 stocks (that are also used in OLS and IV-GMM estimations) for the period 2008–2016 (2240 days). The selected period essentially represents the fragmented era of European equity market. I also estimate them both for large and small stocks separately and also for three equally divided periods (2008-2010), (2011-2013) and (2014-2016).

I use GMM approach (H3SLS) to estimate the simultaneous equations model, an approach that is robust to unknown heteroskedastic error structure. This a three-stage estimation and asymptotically the same as 3SLS if the disturbance are homoscedastic (Greene, 2003). This derives estimation efficiency over two-stage (Zellner and Theil, 1962) and which is relevant to my study in at least two ways. The first, the European equity market structure which rationalises the use of simultaneous equations model also gives rise to the probability that existing simultaneity among MQ_{it}, HF_{it} and $MFrag_{it}$ might produce non-zero contemporaneous covariance in the structural disturbances. And the second, the use of disproportionate instruments for HFT_{it} , $MFrag_{it}$ and MQ_{it} produces both identified and over-identified equations in the system. In both cases, 3SLS has full information characteristics. The coefficient estimated through three-stage least squares are also reported to compare the robustness of the estimates. All estimations include monthly time-fixed effect for each of the 108 months for the period January 2008– December 2016, and stock level fixed effects.

Results and discussion

Table A.22, A.23 and A.24 report the estimates for the whole sample, large and small stocks, three sub-sample periods respectively where only quoted spreads and effective half-spreads are used as dependent variables. Table A.22 presents two set of estimates using GMM and 3SLS for each liquidity measure and the others only report GMM estimates. The discussions on estimates for the model (2.13), (2.14) and (2.15) are presented one after another.

Market liquidity (MQ_{it})

Table A.22 shows that all the coefficients estimated through GMM and 3SLS for each model are highly significant, and as expected, 3SLS's estimates are much stronger than those obtained by using GMM. However, only one exception is seen where the coefficient of MQ_{it} in column (III) is not significant but the corresponding 3SLS estimates are highly significant. The model in which we are most interested is (2.13), Columns I, IV and VII, X report the estimated coefficient for quoted spreads and effective half-spreads respectively. We see that higher HFT is associated with narrower quoted and effective spreads whereas higher fragmentation is associated with wider quoted and effective spreads. Other estimates show that average liquidity level in the same group of other stocks, firm sizes measured by market capitalization, volatility and price level are also the determinants of liquidity, where higher market liquidity in the same size group and higher volatilities are associated with wider spreads, conversely, larger firm sizes (market capitalization), and higher price levels are associated with narrower spreads. The results confirm the evidence documented in OLS with stronger estimates (see section 2.4.1).

The estimates in A.23 (column I, IV, VII, X) confirm the same sign of the estimates, as observed in full sample, across large and small stock groups with different magnitude which implies that both HFT and fragmentation in smaller cap stocks seems to have more striking effect. The coefficients in Table A.24 (column I, IV, VII) show that both the impact of HFT and market fragmentation on liquidity has narrowed and even turn out to non significant during 2014-2016 for market fragmentation. This might be due the fact that fragmentation has reached its saturation stage for the employed stocks where variations in fragmentation does not create enough space to explain the changes in liquidity econometrically.

High frequency trading (HFT_{it})

The coefficient estimates of the Model (2.14) are reported in Table A.22 (column II, V, VIII, XI). The coefficient estimates of the Model HFT_{it} necessarily explain the factors influencing the HFT intensity, and extend the understanding about the bi-direction causality between HFT, market fragmentation and liquidity. Now, it can be seen that there are some indirect impacts which channelise to liquidity through HFT. The coefficient of MQ_{it} is positive which implies that there is one or more mechanisms which associate wider quoted and effective spreads with higher HFT. This is likely to indicate the phenomenon where HFTs post non-marketable limit order as a part of their regular market making activities. Aitken et al. (2015) has found similar results and argued accordingly. Evidence suggests HFT participation is not only limited to BBO (Financial Markets Regulator (France), 2017). HFTs also post quotes around the BBO and even in the deeper levels of the order book depending on the market conditions, consistent with the evidence provided in section 2.4.1.6.

The estimate of $MFrag_{it}$ implies that a higher fragmentation level is also associated with a higher HFT intensity as expected. As can be seen, market wide factors ($\overline{HFT}_{-it}$) play a good role in determining HFT which supports the argument and methodology of HFT instruments development on market level HFT activities (Hasbrouck and Saar, 2013). Among others, larger firm size and larger trading volume are associated with higher HFT. On the contrary, higher volatility, higher relative tick sizes and higher trade sizes indicative of lower HFT . All the findings support the hypothesis and evidences on which HFT instruments are developed in the previous IV-GMM section (2.4.2).

Table A.23 provides more insight on the determinants of HFT activities across stocks. A higher estimate of the coefficient MQ_{it} for small stocks may indicate that spreads in non-marketable limit orders become wider when HFTs post them for relatively illiquid stocks. Other results also show that larger stocks are associated with more intense market wide factors, relative tick sizes and trade sizes. Volatility in small and large stocks appears to have different impact on HFT, though estimates seem not significant in GMM. Higher volatility in large stocks tends to reduce the HFT intensity, which is consistent with the explanation of Hasbrouck and Saar (2013). It is argued that during the period of high illiquidity HFT creates externalities by participating more in illiquid stocks. A similar evidence is also observed in OLS results (section 2.4.1) that HFT provided more liquidity during 2008-2009 while it was scarce. But the same estimates for effective spreads are negative to imply that mere HFT participation by providing non-marketable quotes may not benefit to reduce the actual trading cost.

Table A.24 (column III, VI, IX) shows the similar effect observed in Table A.22 over the periods other than few exceptions. HFT appears to provide more intense non-marketable quotes during 2014-2016 which is consistent with less intense HFT impact on liquidity (column VII) in the same period.

Market fragmentation ($MFrag_{it}$)

The estimates for Model (2.14) are likely to explain the factors those determine the market fragmentation which in turn affect HFT and liquidity indirectly. The coefficient estimates of Model (2.14) are reported in Table A.22 (column III, VI, IX, XII).

The result shows that HFT, market wide factor of fragmentation, firm sizes, trade volumes and volatility have statistically significant impact on liquidity. Stocks with a higher HFT participation, larger market capitalization and wider spreads seem to be related to a higher fragmentation. A positive association between liquidity and fragmentation may indicate that a higher fragmented stock is exposed to a higher market making cost across markets.

2.5 Conclusion

The debate regarding the social benefit of HFT is far from closed, and the evaluation of HFT effects is becoming more complex due to its multifaceted exposure in the financial marketplace. I investigate the impact of HFT on European equity market environments by adopting a relatively new approach and unique dataset.

The results suggest that the decreasing trend of spread based liquidity measures in the European equity market is attributable to the rising HFT intensity over the period, but the tendency of higher order flow fragmentation appears to harm liquidity. My analyses show that a higher fragmentation in order flows imposes extra cost on the ability of HFT market making, and offset some potential liquidity benefits otherwise that could have been derived from this new market maker. It seems that some extra cost imposed through higher market fragmentation is also neutralized by the benefits derived from intense HFT .

The results support the general view documented in the HFT literature (Hendershott et al., 2011; Hasbrouck and Saar, 2013; Boehmer et al., 2015; Aitken et al., 2015) that HFT improves liquidity, however, unlike the existing literature, evidence provided in this chapter helps to explain the mechanism through which the surplus is generated or offset.

The evidence provided in this paper has strong policy implications due to the rising concern of curving HFT. MiFID II has recently come out with strict HFT monitoring rules without implementing any direct measures which seems to provide European market with a good device to trade-off between the benefit and concern of HFT.

In the next chapter, I expand the analyses to a multi-venue setting which should address the limitations of studying HFT within a market.

Chapter 3

High Frequency Trading, Market Fragmentation and Liquidity: A Cross-Market Analysis

3.1 Introduction

The current marketplace is highly fragmented, and market participants can employ smart order routing (SOR) techniques to find liquidity across multiple trading venues. The potential counterparties for HFT market-makers have a large selection of trading venues on which they can trade. To interact with this order flow, HFTs must be present on all these trading venues (The Netherlands Authority for the Financial Markets, 2016). O'Hara (2015) argues HFT is strategic because it maximizes against market design, other HFTs, and other traders, and HFTs need to optimize in a market that contains other HFT players. The cross-market HFT presence makes limit order books linked across markets, and so, too, order flows and price behaviour. I address this added cross-market complexity to HFT research in this chapter in analysing the impact of HFT and market fragmentation on market liquidity. In doing so, I extend the simultaneous equations model in chapter 2 to incorporate alternative exchanges (MTFs) and introduce a novel approach to creating a macro view of cross-market HFT analysis. To the best of my knowledge, no literature to date attempts a similar analysis capable of taking account of the HFT activities simultaneously across markets.

I primarily examine how HFT and fragmentation affect market liquidity in a cross-market setting. The research setup I use to examine the primary research question also allows me to investigate some other related issues like the drivers of HFT within and across markets, nature of exchange competitions etc. To answer these questions, I use millisecond time-stamped TRTH data on the LSE and three alternative electronic exchanges. The dataset covers the whole post-MiFID period until 2016. I develop daily measures for liquidity, HFT and fragmentation across four markets included in the sample. I also develop some consolidated measures to reflect the level and evolution of exchange competition over the time. I estimate the simultaneous equations model using the three-stage least squares method for the full sample as well as its suitable subsamples classified on both cross sections and time-series dimensions.

The results suggest that HFT improves liquidity across markets and exchange level latency has significant impact on liquidity. Among the exchanges, CHIX is highly competitive and attracts more HFT. Furthermore, market fragmentation harms liquidity in the primary exchange while it improves that in alternative exchanges.

The analyses on HFT drivers provide evidence that HFTs' market making activities are linked across markets and HFTs provide liquidity when spreads are wider. A wider/narrower spread in CHIX and the LSE appear to affect the HFT activities across markets. Among others, fragmentation, order sizes, relative tick sizes and volatilities have significant impact on HFT activities. Besides, HFTs concentrate in the primary exchange during the period of a higher volatility.

The analyses extended on large and small stocks provide evidence that HFT remains active in highly liquid stocks even when spreads are narrow. The time-varying analysis shows that the direction of the association between HFT and liquidity and market fragmentation and liquidity appear almost stable across markets over the sample period with time-varying impact.

The rest of the chapter is structured as follows. Section 3.2 links this chapter with the existing body of literature. Section 3.3 describes data and measures, and presents descriptive evidence. Section 3.4 explains the research strategies, and discusses the main results. Section 3.5 concludes.

3.2 Relevant literature

This chapter covers three related aspects : i) HFT, its drivers and speed competition across exchanges; ii) market fragmentation and iii) their impact on market liquidity. A detailed literature review was provided in the section 1.2 (chapter 1). I briefly mention here some of them which are more relevant to this chapter.

The evidence provided in several studies (Hendershott et al., 2011; Hasbrouck and Saar, 2013; Boehmer et al., 2015) on the relation between AT/HFT and market quality show that AT/HFT improves liquidity. The papers studying the impact of market fragmenation on market quality (O'Hara and Ye, 2011; Gresse, 2017; Degryse et al., 2015) mostly support the liquidity improving view of market fragmentation. The novelty of this chapter is that I study both HFT and market fragmentation across markets using a panel dataset for a relatively long period compared to those mostly used in the literature.

The spirit of this chapter is close to the papers which study the HFT and market fragmentation across markets like Upson and Van Ness (2017), Brogaard, Hendershott and Riordan (2014b) but the approach and measures I use are different from those that they used in their research. This chapter also joins the strands of HFT literature: i) examining the HFT liquidity supply and demand within and across markets (Hendershott and Riordan, 2013; Carrion, 2013; Menkveld, 2013); ii) studying HFT on LSE listed stocks (Brogaard, Hendershott, Hunt and Ysusi, 2014; Jarnecic and Snape, 2014), and iii) studying exchange competition (He et al., 2015; Riordan et al., 2011). The motivation of papers (Riordan and Storkenmaier, 2012; Frino et al., 2014; Murray et al., 2016; Frino et al., 2017; Brogaard et al., 2015) examining the impact of speed on market environments supports the analysis conducted in this chapter.

3.3 Data, measures and descriptive statistics

3.3.1 Data

The dataset includes 149 large capitalized stocks, primarily listed on the London Stock Exchange (LSE) and also traded across three alternative exchanges (MTFs)—CHIX, BATS and Turquoise. Almost 100% of LSE listed stocks' lit trading volumes are concentrated in these four trading venues (see Table A.1). One of the challenges of HFT and fragmentation research across markets is to identify the same security across trading venues. TRTH provides unique identification symbology known as Reuters Instrument Code (RIC). RIC structure is pretty complex where several parameters—defined on a stock's primary listing venue, trading venues, currency denominations etc.—are arranged in a particular order to form a RIC. International Securities Identification Number (ISIN) provides the unique identification of a stock across exchanges. I use ISIN and RIC to identify sample stocks across exchanges[1]. Since the analysis is extended across exchanges in this chapter, the sample period covers only the post-MiFID period (October 2008–December 2016) for three and starts from the earliest month from which the widest coverage of data support is available from TRTH across three alternative trading venues (see Table A.2). The section 2.3.1 (chapter 2) presented the data preparation in details.

3.3.2 Measures and descriptive statistics

I use the same measures defined on market liquidity ($spread_{it}$, $espread_{it}$, $rspread_{it}$, $espread_{it}$ and $depth1_{it}$), HFT ($hft1$, $hft2$, $ordtotrd$ and $hft1h$) and market fragmentation ($HHItrd$) in chapter 2 (see section 2.3.2) and expand these measures for all three alternative trading venues included in the sample. All measures are developed using intraday millisecond trades and quotes records for the automated trading sessions (8.00–16.30/London time) of the respective exchanges. Besides, I develop some consolidated measures across trading venues, which are explained below.

EBBO. European Best Bid and Offer (EBBO) is a hypothetical aggregate measure of

[1]I refer Table B.1 for a better explanation, which illustrates how a stock with unique ISIN but different RICs is identified across exchanges.

the best bid and offer prices for LSE listed stocks across trading venues, which can be seen to be equivalent to the NBBO (National Best Bid and Offer), the US counterpart. I take snapshots of the transparent limit order books of all four trading venues at each 500 millisecond interval for the trading hours between 8.10–16.25. The first 10 minutes and the last 5 minutes of the automated trading sessions are excluded to avoid the undue price pressure from opening and closing sessions. At each snapshot, the best bid (the highest among the four local bid prices) and the best offer (the lowest among the four local offer prices) are defined, and both do not necessarily have to come from the same trading venue.

%EBBO. The %EBBO measures the frequency by which a trading venue uniquely or jointly contributes in the EBBO. A trading venue's contribution for both the lowest ask price and the highest bid price is included in the %EBBO. The joint/simultaneous trading venue participation rate (single/double/triple/quadruple) refers the number of trading venues contributing in the EBBO each time. For a unique contribution, the unique venue participation rate measures which exchange contributes in the EBBO. In the presence of HFTs, these measures are expected to reveal the order flows competition across trading venues.

Unlike RegNMS, MiFID does not impose a consolidated tape and trade through rules for European markets, rather it allows some aspects to be decided in the market. For example, MiFID directives details the 'obligation to execute orders on terms most favourable to the client'. This provision requires that firms take relevant steps to ensure the best possible execution for clients and consider 'price, costs, speed, likelihood of execution and settlement, size, nature or any other consideration relevant to the execution of the order'. The % EBBO shows the extent to which limit order books are linked across European markets in providing competitive quotes.

Quotes update speed. The average quotes update speed shows the average time between two quotes updates, and is measured by dividing the number of quotes updates by the length of the automated trading sessions (measured in seconds). The measure is expected to reflect the speed aspect of exchange competition.

Descriptive statistics

This section presents the descriptive evidence regarding HFT and market liquidity across four markets for the full sample of 149 stocks[2] for the period December 2005– December 2016 after winsorizing extreme 1% values on both tails. To facilitate cross sectional comparison, the full sample is divided into 5 equal quintiles based on market capitalization. Descriptive analyses show both aggregate and quintile based measures (mean, median an standard) and monthly trends over the sample period. The descriptive evidence provided on the EBBO, %EBBO and quotes update speed are based on a subsample of 45 stocks which became fragmented across main four trading venues in initial post-MiFID period, and on which the maximum data support from TRTH is available for the period thereafter.

Tables B.2 and B.3 report the quarterly summary of the %EBBO for the unique and joint venue participation rate in the EBBO and Figures B.1a and B.1b show the quarterly trends of the same measures respectively. Initiating from a rate of 100% in the 1st quarter of 2008, the unique trading venue participation rate started to decline afterwards and the level of joint participation in the %EBBO increased over the period. During the period 2008–2010, the average single, double, triple and quadruple venue participation rate were 54%, 23%, 13% and 10% respectively and remained perfectly symmetrical in both sides of the order book throughout the years. It is quite apparent that the joint participation rate in the %EBBO increased over the period but never exceeded 50%. The trends of venue participation rate indicate that order flows competition in European markets got intense over the years.

The rivalry between the LSE and CHIX can be imagined clearly from Table B.3 which reports how the LSE lost its market share to the alternative exchanges over the sample period. Since the competition for order flows in European equity markets started at the end of 2007, CHIX dominated the position of providing the best bid and ask prices. To remain competitive for HFTs, the LSE made huge investment in low-latency technology

[2]The details of stocks coverage across four trading venues over the sample period can be seen in Panel B of Table A.2

and upgraded the trading system in several phases during the period 2006–2011. As can be seen in Table B.3, the LSE started to regain some of its lost market share starting from the year 2013. Among the competing venues, CHIX dominated the position in contributing the EBBO. Turquoise was the next after the LSE and CHIX to contribute to the EBBO. The EBBO participation rate of the exchanges was apparently symmetrical in both sides of the order book.

Figure B.1c depicts the trends of quotes update speed across four markets. The trends of both quotes update speed and %EBBO moved together consistently throughout the sample period—the higher the quotes update speed the more the exchanges share were in the %EBBO—which outwardly indicates that exchanges providing a better low-latency technology attract more traders/market-makers relying on the speed. Since the enactment of MiFID, the quotes update speed in CHIX was the highest until the year 2012, and the LSE started to take the lead back thereafter. Turquoise also seems to become more competitive over the years in contrast to BATS which lost its competitiveness in the same period.

The summary statistics of HFT proxies across four markets are presented in Tables B.4 and B.5, and Figures B.3b and B.3a show the trends of the respective measures over the period. Among the exchanges, the average HFT intensity measured by all proxies are the highest in CHIX. The average per-minute message rate ($hft2$) for the LSE, CHIX, BATS and Turquoise are 84, 116, 69 and 67 respectively. The evidence also shows, HFTs predominantly relied on large stocks, a common feature to observe across exchanges. The rising trends of HFT was consistent across exchanges throughout the sample period. CHIX was found to be more competitive than the LSE in the initial post-MiFID period (2008–2011), and in the latter period, the LSE appeared to regain the position. This phenomenon is also consistent to the %EBBO pattern which is mentioned already.

The descriptive statistics of liquidity measures are presented in Tables B.6 and B.7, and Figures B.4 and B.5 show the trends of the same measures across four markets. Among the exchanges, the LSE provided the most tightest quotes, particularly for small stocks over the period. Moreover, both quoted and effective spreads in large stocks found a low in the

LSE in the initial period of the market fragmentation which eventually disappeared in the latter period due to fierce exchange competition. Large stocks are the highest fragmented stocks and it seems that the competition in order flows impacted them the most. As can be seen, effective half-spreads for the large stocks found a low in almost all trading venues other than BATS, especially in the latter period of the sample, 2013-2016. Figure B.9 also shows that trends in quotes update speed and quoted spreads across trading venues moved together consistently, particularly in large stocks—the higher the quotes update speed the lower the spreads were.

Figures B.7 and B.8 show the trends in average quoted depth and trade size across four markets. For the LSE, both the quoted depth and trade sizes started to decrease sharply in the pre-MiFID period (2005–2008) particularly in large stocks and the trend continued throughout the post-MiIFD period. For alternative trading venues, both quoted depth and trade sizes were consistently smaller than those of the LSE, and declined throughout the period. Figure B.8 depicts the evolution of trade size in large stocks and show that over the years trade size has been declining monotonically across trading venues which essentially indicates the increasing HFT intensity over the years and also consistent to the observations mentioned in the previous chapter (section 2.4.2).

Tables B.8, B.9, B.10 , and B.11 show the decomposition of effective spreads into realized spreads and price impacts across four trading venues and Figure B.6 depicts the trend of these measures. The decomposition is based on four hypothetical post-trade quotes adjustment intervals (10 seconds, 30 seconds, 1 minute and 5 minute). As can be seen, the price impacts and realized spreads decreased across trading venues over the period and realized spreads were negative in all markets for all measures other than in BATS. The evidence suggests that trade execution quality improved across exchanges over the years. I mentioned the same only for the LSE in chapter 2 (see section 2.4.1.4).

The overall descriptive evidence shows that over the years, quoted and effective spreads narrowed and HFT intensity increased across exchanges. It appears that HFT played a substantial role to integrate the fragmented European market using the available low-latency structure. As a result, both quoted and effective spreads converges to low across trading venues over the post-MiFID period. The subsequent sections address the issue more systematically.

3.4 Research strategies, results and discussions

3.4.1 Methodology

I use a cross-market simultaneous equations model approach to examine the relation between HFT, market fragmentation and liquidity. A multi-market setup should overcome the endogeneity arising from simultaneity within and across markets. The endogeneity within a market is well-acknowledged in the HFT literature (Hendershott et al., 2011; Hasbrouck and Saar, 2013; Boehmer et al., 2015). The idea of endogeneity within a market is that an exogenous shock in liquidity might establish a more (less) attractive environment for, and lead to an increase (decrease) in HFT activities. However, the same argument can be made for the endogeneity across markets, and seems more intuitive, as far as the existing equity market structure is concerned, specifically in Europe. My research design agrees with the recommendations and evidence of recent HFT literature (O'Hara, 2015; The Netherlands Authority for the Financial Markets, 2016)[3].

I expand and redefine the simultaneous equations model estimated in chapter 2 (section 2.4.3) to incorporate all trading venues (LSE, CHIX, BATS and TURQ) included in the extended sample. The original model was comprised of three equations (2.13–2.15) each representing one of the three endogenous variables—liquidity (MQ_{it}), HFT (HFT_{it}) and market fragmentation ($MFrag_{it}$). The cross-market approach which I adopt here integrates the fragmented markets for LSE listed stocks altogether by defining models on each venue. So, I drop the fragmentation equation (2.15) and redefine the system across trading venues. The redefined models for four trading venues include eight equations, two for each market to represent the equations for endogenous variables—market liquidity and HFT. The same rationale I followed in chapter 2 (sections 2.4.2 and 2.4.3) justifying the determinants of HFT and market liquidity also motivate

[3]Figure B.2 presents some of the evidence provided in The Netherlands Authority for the Financial Markets (2016) regarding the cross-market HFT activity in European markets

the specifications in this chapter. The cross-market simultaneous equations model is :

$$MQ_{1it} = \alpha_{i_{(mq)1}} + \sum_{m=1}^{M} \gamma_{(mq)_{1m}} + \beta_{1_{(mq)1}} HFT_{1it} + \beta_{2_{(mq)1}} HHItrd_{it} + \beta_{3_{(mq)1}} \overline{MQ}_{-1it}$$
$$+ \beta_{4_{(mq)1}} ln(mktcap)_{it} + \beta_{5_{(mq)1}} ln(voltintra)_{1it} + \beta_{6_{(mq)1}} inv(price)_{it} + \epsilon_{it_{(mq)1}}, \tag{3.1}$$

$$MQ_{2it} = \alpha_{i_{(mq)2}} + \sum_{m=1}^{M} \gamma_{(mq)_{2m}} + \beta_{1_{(mq)2}} HFT_{2it} + \beta_{2_{(mq)2}} HHItrd_{it} + \beta_{3_{(mq)2}} \overline{MQ}_{-2it}$$
$$+ \beta_{4_{(mq)2}} ln(mktcap)_{it} + \beta_{5_{(mq)2}} ln(voltintra)_{2it} + \beta_{6_{(mq)2}} inv(price)_{it} + \epsilon_{it_{(mq)2}}, \tag{3.2}$$

$$MQ_{3it} = \alpha_{i_{(mq)3}} + \sum_{m=1}^{M} \gamma_{(mq)_{3m}} + \beta_{1_{(mq)3}} HFT_{3it} + \beta_{2_{(mq)3}} HHItrd_{it} + \beta_{3_{(mq)3}} \overline{MQ}_{-3it}$$
$$+ \beta_{4_{(mq)3}} ln(mktcap)_{it} + \beta_{5_{(mq)3}} ln(voltintra)_{3it} + \beta_{6_{(mq)3}} inv(price)_{it} + \epsilon_{it_{(mq)3}}, \tag{3.3}$$

$$MQ_{4it} = \alpha_{i_{(mq)4}} + \sum_{m=1}^{M} \gamma_{(mq)_{4m}} + \beta_{1_{(mq)4}} HFT_{4it} + \beta_{2_{(mq)4}} HHItrd_{it} + \beta_{3_{(mq)4}} \overline{MQ}_{-4it}$$
$$+ \beta_{4_{(mq)4}} ln(mktcap)_{it} + \beta_{5_{(mq)4}} ln(voltintra)_{3it} + \beta_{6_{(mq)4}} inv(price)_{it} + \epsilon_{it_{(mq)4}}, \tag{3.4}$$

$$HFT_{1it} = \alpha_{i_{(hft)1}} + \sum_{m=1}^{M} \gamma_{(hft)_{1m}} + \sum_{v=1}^{4} \beta_{v_{(hft)v}} MQ_{vit} + \beta_{5_{(hft)1}} \overline{HFT_{-1it}} + \beta_{6_{(hft)1}} HHItrd_{1it}$$
$$+ \beta_{7_{(hft)1}} ln(size)_{1it} + \beta_{8_{(hft)}} ln(volume)_{1it} + \beta_{9_{(hft)1}} rtk_{1it} + \beta_{10_{(hft)1}} ln(mktcap)_{it}$$
$$+ \beta_{11_{(hft)1}} ln(voltintra)_{1it} + \epsilon_{it_{(hft)1}}, \tag{3.5}$$

$$HFT_{2it} = \alpha_{i_{(hft)2}} + \sum_{m=1}^{M} \gamma_{(hft)_{2m}} + \sum_{v=1}^{4} \beta_{v_{(hft)v}} MQ_{vit} + \beta_{5_{(hft)2}} \overline{HFT_{-2it}} + \beta_{6_{(hft)2}} HHItrd_{2it}$$
$$+ \beta_{7_{(hft)2}} ln(size)_{2it} + \beta_{8_{(hft)2}} ln(volume)_{2it} + \beta_{9_{(hft)2}} rtk_{2it} + \beta_{10_{(hft)2}} ln(mktcap)_{it}$$
$$+ \beta_{11_{(hft)2}} ln(voltintra)_{2it} + \epsilon_{it_{(hft)2}}, \tag{3.6}$$

$$HFT_{3it} = \alpha_{i_{(hft)3}} + \sum_{m=1}^{M} \gamma_{(hft)_{3m}} + \sum_{v=1}^{4} \beta_{v_{(hft)v}} MQ_{vit} + \beta_{5_{(hft)3}} \overline{HFT_{-3it}} + \beta_{6_{(hft)3}} HHItrd_{3it}$$
$$+ \beta_{7_{(hft)3}} ln(size)_{3it} + \beta_{8_{(hft)3}} ln(volume)_{3it} + \beta_{9_{(hft)3}} rtk_{3it} + \beta_{10_{(hft)3}} ln(mktcap)_{it}$$
$$+ \beta_{11_{(hft)3}} ln(voltintra)_{3it} + \epsilon_{it_{(hft)3}}, \tag{3.7}$$

$$HFT_{4it} = \alpha_{i_{(hft)4}} + \sum_{m=1}^{M} \gamma_{(hft)_{4m}} + \sum_{v=1}^{4} \beta_{v_{(hft)v}} MQ_{vit} + \beta_{5_{(hft)4}} \overline{HFT_{-4it}} + \beta_{6_{(hft)4}} HHItrd_{4it}$$
$$+ \beta_{7_{(hft)4}} ln(size)_{4it} + \beta_{8_{(hft)4}} ln(volume)_{4it} + \beta_{9_{(hft)4}} rtk_{4it} + \beta_{10_{(hft)4}} ln(mktcap)_{it}$$
$$+ \beta_{11_{(hft)4}} ln(voltintra)_{4it} + \epsilon_{it_{(hft)4}}, \tag{3.8}$$

where indices i, t, v represent stocks, time (days) and trading venues respectively, v takes the value $1, 2, 3, 4$ for the LSE, CHIX, BATS and Turquoise respectively, MQ_{vit} represents one of the two log normalized liquidity measures (quoted spreads/$spread\_bps$, effective half-spreads/$espread$), HFT_{vit} represents the HFT proxy ($hft2$), $HHItrd_{it}$ represents the market fragmentation proxy, $\overline{MQ}_{-vit}$ represents the average market liquidity level over all stocks in the same size group excluding stock i at venue v, $\overline{HFT}_{-vit}$ represents the average HFT intensity over all stocks in the same size group excluding stock i at venue v, $ln(mktcap)$ is the log normalized market capitalization, $ln(voltintra)_{vit}$ is the log normalized intraday mid price range volatility, $invprice$ is the inverse of daily average prices, $ln(size)_{vit}$ is the log normalized trade size, $ln(value)_{vit}$ is the log normalized trading volume, $rtick_{vit}$ is the relative tick size, α_i is the firm fixed effect, $\sum_{m=1}^{M} \gamma_m$ is the time (month) fixed effect, $_{(mq)}$ and $_{(hft)_v}$ index the coefficient of the equations MQ_{vit}, HFT_{vit} respectively. Market wide measures on liquidity ($\overline{MQ}_{-vit}$) and HFT ($\overline{HFT}_{-vit}$) for each venue are based on four equal size stocks groups, classified on market capitalization.

Model identification and the order condition. To meet the order condition, the number of exogenous variables that appear elsewhere in the equation system must be at least as large as the number of endogenous variables in the equation. The number of endogenous variables in equations (3.1)–(3.4) and (3.5)–(3.8) are two and five respectively. The control variables which are specified in the model (3.1)–(3.8) should be considered exogenous. Model (3.1)–(3.4) and (3.5)–(3.8) use the same control variables as specified in the section 2.4.3 for Models (2.13) and (2.14) respectively. All models share three common exogenous variables, $ln(mktcap)$ and $inv(prce)$, $HHItrd_{it}$ and the rest—$\overline{MQ_{-vit}}$, $\overline{HFT_{-vit}}$, $ln(size)$, $ln(voltintra)$, $ln(volume)$, $rtick$—are based on the respective market and different from each other. The system has, in aggregate, more excluded exogenous variables than required by the order conditions and meet the order condition. The rank condition ensures that there is a unique solution to this set of equations. In practical terms, the rank condition is difficult to establish in large equation systems. Practitioners typically take it as given (Greene, 2003).

To estimate the models in equations (3.1)–(3.8) as a system, I use a panel dataset with 4 markets/trading venues (LSE, CHIX, BATS and Turquoise), 149 stocks and 2060 days, from October 2008–December 2016. I also use its suitable subsamples classified on both cross-section and time-series dimensions. I use the three-stage least squares method, an approach which derives estimation efficiency over two-stage (Zellner and Theil, 1962) and also has full information characteristics when the use of disproportionate instruments produces both identified and over-identified equations in the system.

3.4.2 Results and discussions

The impact of HFT and market fragmentation on market quality

The results of the simultaneous equations model estimation for the full sample are presented in Table B.12. Panel A reports the results for market quality equations

(3.1–3.4). Columns I–IV and V–VIII of Panel A report the results estimated for quoted spreads and effective half-spreads respectively.

The main variables of interest are HFT (HFT_{it}) and market fragmentation ($HHItrd_{it}$). All of the estimates for HFT are highly statistically significant across equations and have the same sign—which is negative—for both liquidity measures. Among the trading venues, CHIX appears to have the strongest estimates for HFT. The estimates for the market fragmentation variable are also highly statistically significant across equations and produce different signs across trading venues. All of the signs in alternative venue equations—$MQ_{(chix)it}$, $MQ_{(bats)it}$, $MQ_{(turq)it}$—are negative while the same for the primary venue equation ($MQ_{(lse)it}$) are positive. The estimates reported in the previous chapter (see Table A.14) regarding HFT and market fragmentation for both quoted and effective spreads have similar signs with slightly weaker estimates for HFT and almost similar estimates for market fragmentation.

These results suggest that HFT improves liquidity across trading venues and low latency sophistication at exchange level appears to have significant impact on liquidity—the lower the latency the narrower the quoted and effective spread. Among the exchanges, CHIX was highly competitive in attracting HFT due to its better low latency technology sophistication since it started operations. The evidence I provide here regarding HFT is consistent with the findings of Hendershott et al. (2011), Hasbrouck and Saar (2013), Boehmer et al. (2015) who examine the causality between HFT and market quality. The findings also support the evidence provided in many papers which assess how external HFT shocks affect market liquidity[4]. The results also suggest that market fragmentation harms primary exchange's liquidity while it improves that of alternative trading exchanges. For the primary venue LSE, it appears that scale economics and network externality arguments given in favor of a concentrated market dominates the alternative view of exchange competition favouring the fragmented markets and the reverse holds for alternative exchanges. The evidence regarding the fragmentation effect on alternative trading venues implies that trader preferences over technology

[4]Frino et al. (2014); Murray et al. (2016); Frino et al. (2017); Riordan and Storkenmaier (2012); Brogaard et al. (2015).

differentiation are sufficiently important in modern marketplaces. From a multi-venue perspective, the evidence regarding market fragmentation I provide here is striking. To the best of my knowledge, no research to date provides similar evidence that reconciles the trade-off between the two opposite effects of exchange co-existence. The evidence does not support the liquidity improving view of market fragmentation, particularly in primary exchanges, provided in O'Hara and Ye (2011) and Gresse (2017). However, the set up that I use in my approach is different from those that they adopt in their research.

All of the estimates regarding control variables are highly statistically significant and have the expected signs. The average exchange level liquidity appears to have varying positive impacts on stock level liquidity across exchanges—the less the exchange level liquidity the stronger the effect is. Among the others, a lower inverse price level and intraday volatility and a higher market capitalization seem to improve liquidity. The estimates are consistent in signs with the estimates reported in chapter 2 for the same control variables.

Drivers of HFT

Columns I-IV and V-VIII of Panel B (Table B.12) report the results for HFT equations (3.5)—(3.8) estimated for quoted spreads and effective half-spreads respectively. All estimates for market quality (MQ_{vit}), average market wide HFT ($\overline{HFT}_{-vit}$), market fragmentation ($HHItrd$), trade sizes ($ln(size)_{vit}$) , relative tick sizes ($rtick_{vit}$), trade volume ($ln(volume)_{vit}$), market capitalization ($ln(mktcap)_{vit}$) and intraday mid price range volatility ($ln(voltintra)_{vit}$) are highly statistically significant across equations other than two reported in column V. The main variables of interest are market liquidity (MQ_{vit}) and market fragmentation ($HHItrd$). These variables explain the links among HFT, market quality and fragmentation that arguably exist across markets. The reported estimates for market liquidity as measured by quoted and effective spreads for the LSE ($MQ_{(lse)_{it}}$) and CHIX ($MQ_{(chix)_{it}}$) have positive signs in equations specified for the respective markets and negative signs in equations for the others. On the contrary, all reported estimates for market liquidity in the other equations, BATS ($MQ_{(bats)_{it}}$) and Turquoise ($MQ_{(turq)_{it}}$), are positive across exchanges. Among the rest, the reported estimates for average market wide HFT, market fragmentation and trade volume are positive and that for market capitalization, trade sizes and relative tick sizes are negative across equations.

These results regarding cross-market liquidity and HFT activities suggest that both the liquidity level in the LSE and CHIX determine the HFT activity across trading venues. For the LSE and CHIX, HFTs supply liquidity in the LSE when quoted and effective spreads are wider in the LSE and narrower in CHIX, and conversely for CHIX. The wider spreads in BATS and Turquoise also appear to increase HFT activities in both the LSE and CHIX as well. For BATS and Turquoise, HFTs supply liquidity in both markets when quoted and effective spreads are narrower in the LSE and CHIX and wider in the respective markets. The results have at least two implications: firstly, HFTs' market making activities are linked across markets; secondly, HFTs provide liquidity when spreads are wider. The evidence I provide here is consistence to the finding of Hendershott and Riordan (2013) and Carrion (2013) who also provide evidence that HFT supply liquidity when its scarce and demand liquidity when its cheap. They explain that when spreads are narrow HFTS/ATs are less likely to submit new orders, less likely to cancel their orders, and more likely to initiate trades. HFTs/ATs react more quickly to events and even more so when spreads are wide. The findings also confirm the cross market HFT strategies portrayed in The Netherlands Authority for the Financial Markets (2016) and Menkveld (2013).

The results in B.12 also suggest that a statistically significant market wide HFT component exists and affects HFT activities positively across markets. Among others determinants, fragmentation affects HFT positively—the higher the fragmentation the more the HFT activity. The order size is negatively associated with HFT—the smaller the order size the higher the HFT activity is. The evidence agrees with the arguments provided in Hendershott et al. (2011) and Aitken et al. (2014) regarding the association between HFT and smaller order size. The relative tick size seems to affect HFT activities significantly—the lower the relative tick size the higher the HFT activity is. The findings in O'Hara (2015) support the indicated association I find between HFT and relative tick sizes. They show that HFTs leave orders in the book longer when relative tick size is larger and trade more aggressively. Higher volatilities in alternative exchanges seem to lower the HFT activity there but that of primary exchange increases the HFT activity. It seems that during high volatile periods HFTs find a primary exchange more favourable to execute their strategies. This evidence support the findings in He et al. (2015) who show that trading tends to concentrate on primary exchanges during market stress. These results also support the evidence provided in the section 2.4.3 (chapter 2) about the association between HFT, relative tick sizes, order sizes, order volume and market capitalization.

Large and small stocks

Since descriptive analyses demonstrate significant differences in liquidity and HFT across quintiles, I divide here the full sample (149 stocks) into two equal subsamples—small-cap group (which comprises 75 stocks below the median market capitalization) and large-cap group (which comprises 74 stocks above the median market capitalization)—to examine how different firm sizes affect the results obtained for the full sample. I estimate the system of equations (3.1)—(3.8) for both groups of stocks as performed on the full sample using quoted and effective half-spreads as dependent variables in the previous section. The results are reported in Table B.13 and Table B.14 for large and small stocks respectively. Panel A and Panel B, in both tables, report the results for market quality equations (3.1)–(3.4) and HFT equations (3.5)–(3.8) respectively where columns I–IV show the estimates for quoted spreads and V–VIII show that for effective half-spreads. To avoid repetition, I only discuss the results which do not agree with those presented in the case of full sample.

All of the estimates are highly statistically significant for both groups of stocks with almost similar estimated coefficients, particularly in signs, as obtained for the full sample. The estimates for market quality equations, both for large and small stocks, do not show any notable differences regarding the main variables of interest— HFT and market fragmentation—from those of the full sample. The estimates for large stocks against quoted spreads in Panel B (Table B.13) for the LSE and CHIX are different from those of the full sample (column I and II), which have the same sign i.e. negative across equations. The estimates against the other liquidity measure, effective half-spreads, are still consistent with those of the full sample.

The results could suggest that HFT remains active in highly liquid stocks even when spreads are narrow. The descriptive evidence shows that the average quoted spreads in large stocks is around 60% smaller than that in small stocks. To remain competitive, HFTs must update quotes in liquid stocks, which could require them to supply liquidity on large stocks even when it is less profitable. The results also suggest stronger HFT and market fragmentation effects on liquidity for small stocks, and consistent to the findings in the previous chapter.

It is useful to recall that stocks included in my samples are the highest market capitalized stocks in the LSE which are mostly included in FTSE 100 and FTSE 250 indices. Small stocks, as are classified in subsamples, do not necessarily hold the characteristics of a typical small stock that we see in the literature. This might be the reason to produce similar estimates across stock groups.

Time-varying effects

I extend the analyses to see whether the effects of high frequency trading and market fragmentation on liquidity and factors determining the liquidity supply of HFT vary over time. I divide the original sample into three subsamples (2008–2010, 2011–2013, and 2014–2016), each with three years and 149 stocks. I estimate the system of equations (3.1)–(3.8) for each subsample using the same liquidity measures used for the estimates in previous sections. These subsamples necessarily do not reflect any motivation other than a uniform classification of the sample period. The results are reported in Table B.15. Panel A and Panel B report the estimates for market quality equations (3.1)–(3.4) and HFT equations (3.5)–(3.8) respectively, in which columns I–IV and V–VIII present the estimates on quoted spreads and effective half-spreads respectively. To conserve space, I report the estimates only for the variables of main interest and the rests that I do not report are also significant at the 1% level and produce the expected signs.

The coefficient estimates on equations (3.1)–(3.4) in Panel A for three subperiods are all statistically significant at the 1% level and have the same signs as reported in Table B.12 for the full sample other than in column II, VI. Estimates showing the exceptions are estimated for the period 2008–2010 (the initial period of alternative trading venues' operations) where one is only significant at the 5% level (for CHIX). This exception should not affect the general findings revealed in all others estimates. For HFT, the estimates appear stronger in the latter period of sample (2014–2016) whereas that for fragmentation appear weaker in the same period across markets.

These results suggest that the general direction of the association between HFT and liquidity and market fragmentation and liquidity that are reported for the full sample appears almost stable across markets over the sample period with time-varying impact. Panel B reports the coefficient estimates for HFT equations across markets and a few estimates appear time-varying: for the LSE, the estimates for $MQ_{(lse)it}$ for the period 2014–2016 (column I) ; for CHIX, the estimates for $MQ_{(lse)it}$ (column II) and $MQ_{(chix)it}$ (column II, VI) for the period 2008–2010 and the estimates for $MQ_{(chix)it}$ (column II) for the period 2011–2013; for BATS, the estimates for $MQ_{(lse)it}$ (column III, VII) for the period 2008-2010 and the same estimates (column VII) for the period 2011-2013 and 2014–2016; and for Turquoise, the estimates for $MQ_{(chix)it}$ (column VIII) for the period 2011—2013. These estimates are mostly associated with the LSE and CHIX which suggest that varying spreads level in main two competitive exchanges have time-varying impact on HFTs' liquidity supply across markets.

The estimates for the period 2008–2010 seem to reflect some historical facts about the fierce market competition in European equity markets immediately after the adoption of MiFID. Between 2008–2010, CHIX, BATS and Turquoise emerged as alternative trading venues and started to compete with the incumbent exchange LSE. Among the MTFs, CHIX was the most advanced in providing low latency trading platforms. Descriptive evidence shows that spreads in CHIX was the narrowest during that period for large stocks and remained at that level until the end of 2013 after which the LSE took the lead. Between 2007–2011, the LSE undertook several initiatives (that I also mentioned in explaining the descriptive evidence) to upgrade its trading system with the aim to regain the market share which it started to lose at the beginning of 2008. Accordingly, the results suggest that between 2008–2010, HFT liquidity supply was positively associated with narrower spreads in CHIX across markets. On the contrary, during the same period the liquidity in the LSE , as measured by quoted spreads, does not seem to have a similar effect on HFT liquidity supply across markets. On the other hand, between 2011–2013, the narrower spreads in the LSE affected the HFT liquidity supply positively across alternative trading venues, and the effect of CHIX's quoted spreads started to become less strong across other trading venues. These results further suggest, between 2014–2016, the narrower spreads in both LSE and CHIX were associated with a higher HFT liquidity supply across trading venues.

3.5 Conclusion

HFT strategies are generally implemented across markets. To obtain an encompassing view of HFT, it is necessary to incorporates all markets in the analysis whose order books are believed to be linked due to HFT activity. This chapter uses a unique approach of cross-market simultaneous equations model to tackle the possible endogeneity between HFT, market fragmentation and market quality across markets and estimates the specified model employing a rich panel dataset. I provide evidence that HFT improves liquidity across markets and fragmentation harms the primary exchange's liquidity while improves that of alternative exchanges. The direction of association between HFT , market fragmentation and liquidity are almost stable across markets with time-varying impact. I also show that HFT activities are linked across markets and HFT activities is higher when spread is wider, however, HFT remain active in highly liquid stocks even when spread is narrow.

HFTs are diverse in their use of trading strategies (Biais and Foucault, 2014; Hagströmer and Nordén, 2013). The HFT proxy used in this thesis does not necessarily reflect the activity of a particular HFT, It's rather reveal a mixture of strategies. The evidence I provided in this chapter could indicate the relative dominance of a subset of HFT who pursue certain strategies that strengthen market environments. The findings are more indicative of market making HFTs than other HFT types whose activity improves liquidity across markets.

The results of this chapter have important implications for both regulators and trading platforms providers, particularly in Europe. Any regulation aimed at hindering HFT activities indiscriminately may have serious negative consequences for market environments and thus market participants. Encouraging exchanges to provide better HFT-friendly platforms and their high-speed connecting channels may help to increase market competitions and decrease trading costs. Providing a better low-latency technology may help exchanges to gain market share.

Appendix A

Appendix - Chapter 2

Table A.1 The universe of sample stocks

This table shows the decomposition of large-cap stocks across European countries (Panel A) as presented in the STOXX 800 index at the end of year 2016, and relative position of the European lit trading venues based upon total European equity trading volumes (Panel B) and trading volumes of LSE listed stocks (Panel C).

Panel A: STOXX 800's composition

Country (Primary listing venue)	No. of Instruments	(%)
The United Kingdom (LSE)	220	27.5
France (Euronext (Paris))	95	11.88
Germany (Xetra)	84	10.5
Switzerland (Six Swiss)	61	7.63
Sweden	60	7.5
Italy	47	5.88
Spain	37	4.63
The Netherlands	28	3.5
Denmark	25	3.13

Panel B: Market share of European lit trading venues (European equities)

1/1/2014 to 31/12/2016 EURbn	Lit	
	Turnover	%
Bats CXE	5,500.47	18.48
LSE	3,514.16	11.81
Paris	3,239.61	10.89
Deutsche Börse	3,060.42	10.28
Turquoise	3,051.43	10.25
Milan	2,111.71	7.10
Bats BXE	1,732.77	5.82
SIX Swiss	1,531.50	5.15
Amsterdam	1,474.64	4.96
Madrid	1,020.24	3.43
Stockholm	954.50	3.21
Swiss Exchange	424.97	1.43
Copenhagen	352.76	1.19
Brussels	322.49	1.08

Panel C: Market share of European lit trading venues (LSE listed stocks)

1/1/2014 to 31/12/2016 GBPbn	Lit	
	Turnover	%
LSE	2,736.05	56.19
Bats CXE	1,003.66	20.61
Turquoise	729.17	14.97
Bats BXE	355.05	7.29
Aquis	30.96	0.64
Equiduct	8.04	0.17
ICAP Securities	6.40	0.13
Nyse Arca	0.00	
Total	**4,869.33**	**100.00**

source: Fidessa

Table A.2: TRTH's data support over the sample period

This table shows the data availability (from TRTH) for the LSE listed stocks primarily selected for the sample (across major trading venues). Panel A reports the data availability for the LSE and other three alternative trading venues, BATS, CHIX, Turquoise (TURQ), since 2008. Panel B reports the data availability for the stocks finally selected for the sample.

PANEL A

Post MiFID TRTH DATA availability for LSE stocks included in STOXX 800				
Year	BATS	CHIX	LSE	TURQ
Jan05–Dec07	0	0	180*	0
Jan-08	n.a.**	n.a.	182	n.a.
Jan-09	159	159	184	156
Jan-10	160	160	185	162
Jan-11	165	166	190	167
Jan-12	183	183	192	171
Jan-13	189	189	194	174
Jan-14	197	197	198	197
Jan-15	204	203	204	203
Jan-16	205	204	205	204

* availability varies over the period
** not available

PANEL B

Unbalanced panel constructed by taking eligible stocks from PANEL A					
year	Qtr	LSE	CHIX	BATS	TURQ
2005	1	118	n.a	n.a	n.a
2005	2	122	n.a	n.a	n.a
2005	3	131	n.a	n.a	n.a
2005	4	136	n.a	n.a	n.a
2006	1	138	n.a	n.a	n.a
2006	2	139	n.a	n.a	n.a
2006	3	140	n.a	n.a	n.a
2006	4	142	n.a	n.a	n.a
2007	1	143	n.a	n.a	n.a
2007	2	146	n.a	n.a	n.a
2007	3	149	n.a	n.a	n.a
2007	4	149	n.a	n.a	n.a
2008	1	149	n.a	n.a	n.a
2008	2	149	138	n.a	n.a
2008	3	149	143	n.a	70
2008	4	149	149	136	75
2009	1	149	149	148	85
2009	2	149	149	149	147
2009	3	149	149	149	148
2009	4	149	149	149	148
2010	all	149	149	149	149
2011	all	149	149	149	149
2012	all	149	149	149	149
2013	all	149	149	149	149
2014	all	149	149	149	149
2015	all	149	149	149	149
2016	1	149	149	149	149
2016	2	149	149	149	149
2016	3	147	147	147	147
2016	4	146	146	146	146

Table A.3: Descriptive statistics for HFT proxies: full sample (2005–2016) and post-MiFID period (2008–2016)

This table presents the descriptive statistics for the daily HFT proxies for the full sample (2005-2016) and post-MiFID period (2008-2016) calculated on LSE data. The sample comprises 149 stocks divided into 5 equal quintile based on market capitalization. Table A.2/Panel B) shows the periodical data coverage in details.

			2005-2016						2008-2016/ post MiFID					
			All	Small	Q2	Q3	Q4	Large	All	Small	Q2	Q3	Q4	Large
$hft1$	messages per minute (best 10 depth levels)	Mean	100.36	23.77	38.97	72.39	103.80	266.31	129.13	30.75	49.15	91.62	131.36	342.73
		Median	47.62	17.95	30.57	56.44	90.99	171.75	72.11	25.01	39.19	73.91	113.69	250.90
		StdDev	157.77	23.22	38.25	67.59	84.81	271.25	174.25	23.62	39.34	67.91	81.09	274.27
$hft2$	messages per minute (best 5 depth levels)	Mean	83.72	19.35	32.61	60.80	89.36	219.40	107.16	25.02	40.87	76.42	112.49	281.00
		Median	41.10	14.42	26.09	49.34	80.49	149.97	62.47	20.31	33.25	63.67	99.17	212.50
		StdDev	126.10	19.37	30.18	52.20	68.56	213.53	138.85	19.86	30.70	51.40	64.17	214.14
$hft3$	messages per minute (BBO)	Mean	39.32	9.57	15.67	28.33	43.50	100.91	49.61	12.27	19.36	34.99	53.91	127.52
		Median	20.38	7.48	12.87	23.91	40.04	76.56	30.31	10.28	16.10	30.21	48.41	102.10
		StdDev	54.32	8.98	13.09	21.76	30.81	88.30	59.51	9.10	13.08	21.04	28.54	86.95
$ordtotrd$	number of messages per executed order (order to trade ratio)	Mean	22.02	28.05	19.51	20.45	21.18	20.73	27.03	34.35	23.64	25.12	26.28	25.79
		Median	19.06	19.97	17.28	18.46	20.40	19.83	23.07	24.32	20.58	22.43	24.59	24.12
		StdDev	21.25	36.67	15.59	14.76	13.57	13.84	22.40	40.98	16.09	14.30	12.03	12.39
$hft1h$	GBP volume (100) per message (best 10 depth levels) time (-1)	Mean	-6.93	-2.24	-3.73	-5.14	-7.51	-16.21	-2.11	-0.89	-1.40	-1.76	-2.34	-4.15
		Median	-1.96	-0.88	-1.37	-1.77	-2.38	-3.60	-1.41	-0.63	-1.04	-1.36	-1.83	-2.68
		StdDev	15.51	4.19	5.87	8.42	10.93	29.12	2.80	0.94	1.25	1.43	1.78	5.03
$hft2h$	GBP volume (100) per message (best 5 depth levels) time (-1)	Mean	-7.54	-2.80	-4.17	-5.53	-8.04	-17.37	-2.43	-1.12	-1.65	-2.02	-2.63	-4.72
		Median	-2.33	-1.10	-1.64	-2.05	-2.72	-4.19	-1.68	-0.79	-1.24	-1.59	-2.11	-3.18
		StdDev	16.22	5.30	6.32	8.78	11.36	30.28	3.01	1.15	1.41	1.55	1.89	5.32
	observations (stock*day)		439583	90046	88374	87239	86786	87138	324230	64760	65142	64919	64459	64950

Table A.4: Descriptive statistics for HFT proxies: pre-MiFID (2005–2007) and post-MiFID (2008–2016)

This table presents the descriptive statistics for the daily HFT proxies for the pre-MiFID (2005-2007) and post-MiFID period (2008-2016) calculated on LSE data. The sample comprises 149 stocks divided into 5 equal quintile based on market capitalization. Table A.2/Panel B) shows the periodical data coverage in details.

			2005–2007/pre MiFID						2008-2016/post MiFID					
			All	Small	Q2	Q3	Q4	Large	All	Small	Q2	Q3	Q4	Large
$hft1$	messages per minute (best 10 depth levels)	Mean	19.47	5.91	10.44	16.44	24.25	42.61	129.13	30.75	49.15	91.62	131.36	342.73
		Median	12.09	3.42	7.15	11.70	17.58	28.89	72.11	25.01	39.19	73.91	113.69	250.90
		StdDev	25.37	6.84	11.01	15.37	21.35	40.08	174.25	23.62	39.34	67.91	81.09	274.27
$hft2$	messages per minute (best 5 depth levels)	Mean	17.84	4.83	9.47	15.36	22.60	39.10	107.16	25.02	40.87	76.42	112.49	281.00
		Median	11.26	2.62	6.62	11.18	16.61	27.33	62.47	20.31	33.25	63.67	99.17	212.50
		StdDev	22.72	5.75	9.82	13.86	19.53	34.94	138.85	19.86	30.70	51.40	64.17	214.14
$hft3$	messages per minute (BBO)	Mean	10.42	2.65	5.31	8.97	13.44	23.01	49.61	12.27	19.36	34.99	53.91	127.52
		Median	6.82	1.61	3.93	6.82	10.12	17.02	30.31	10.28	16.10	30.21	48.41	102.10
		StdDev	12.65	2.94	5.14	7.63	11.09	18.62	59.51	9.10	13.08	21.04	28.54	86.95
$ordtotrd$	number of messages per executed order (order to trade ratio)	Mean	7.92	11.90	7.92	6.85	6.45	5.95	27.03	34.35	23.64	25.12	26.28	25.79
		Median	6.38	9.25	7.03	6.22	5.87	5.17	23.07	24.32	20.58	22.43	24.59	24.12
		StdDev	6.33	11.17	4.03	2.82	2.39	3.17	22.40	40.98	16.09	14.30	12.03	12.39
$hft1h$	GBP volume (100) per message (best 10 depth levels) time (-1)	Mean	-20.47	-5.70	-10.27	-14.97	-22.45	-51.53	-2.11	-0.89	-1.40	-1.76	-2.34	-4.15
		Median	-12.52	-4.13	-8.31	-12.64	-20.28	-40.56	-1.41	-0.63	-1.04	-1.36	-1.83	-2.68
		StdDev	25.42	6.61	8.30	11.89	12.43	39.78	2.80	0.94	1.25	1.43	1.78	5.03
$hft2h$	GBP volume (100) per message (best 5 depth levels) time (-1)	Mean	-21.91	-7.10	-11.22	-15.74	-23.66	-54.40	-2.43	-1.12	-1.65	-2.02	-2.63	-4.72
		Median	-13.58	-5.15	-9.11	-13.26	-21.39	-43.24	-1.68	-0.79	-1.24	-1.59	-2.11	-3.18
		StdDev	26.41	8.41	8.87	12.41	12.78	40.95	3.01	1.15	1.41	1.55	1.89	5.32
	observations (stock*day)		115353	25286	23232	22320	22327	22188	324230	64760	65142	64919	64459	64950

Table A.5: Yearly descriptive statistics for HFT proxies

This table presents the yearly summary statistics of the constructed Panel (unbalanced, Table A.2/Panel B) for daily HFT proxies based on daily LSE data. The panel comprises 149 stocks, divided into 5 stock groups based on market capitalization.

			2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
$hft1$	messages per minute (top 10 depth levels)	Mean	9.33	12.70	25.34	57.50	70.65	122.37	136.35	123.59	140.97	134.81	168.49	176.69
		Median	7.67	9.49	18.79	40.01	44.11	70.31	81.43	66.01	80.39	77.25	93.77	106.13
$hft2$	messages per minute (top 5 depth levels)	Mean	8.79	11.73	23.23	51.98	61.80	105.51	116.69	100.10	112.90	107.23	136.06	147.76
		Median	7.42	8.93	17.58	35.30	38.23	61.67	72.23	58.84	69.65	64.32	78.09	90.69
$hft3$	messages per minute (BBO)	Mean	5.58	7.29	13.15	30.90	32.49	43.93	51.82	47.11	49.94	48.84	61.47	69.99
		Median	4.77	5.64	10.05	20.19	19.37	26.31	32.13	30.43	34.97	31.18	37.21	43.91
$ordtotrd$	number of messages per executed order	Mean	7.04	7.13	7.68	12.25	18.09	33.65	31.95	28.80	33.69	28.00	26.20	25.42
		Median	5.73	5.79	6.59	10.47	14.62	26.35	26.34	23.27	28.73	24.68	24.16	23.63
$hft1h$	GBP volume (100) per message (top 10 depth levels) time (-1)	Mean	-26.68	-23.89	-17.24	-6.94	-3.19	-1.80	-1.52	-1.54	-1.35	-1.60	-1.50	-1.34
		Median	-17.25	-15.50	-11.38	-4.76	-2.46	-1.39	-1.17	-1.29	-1.17	-1.29	-1.23	-1.10
$hft2h$	GBP volume (100) per message (top 5 depth levels) time (-1)	Mean	-28.41	-25.55	-18.43	-7.66	-3.66	-2.09	-1.74	-1.80	-1.61	-1.95	-1.81	-1.57
		Median	-18.47	-16.73	-12.20	-5.32	-2.88	-1.62	-1.36	-1.49	-1.38	-1.60	-1.51	-1.31

Table A.6: Yearly descriptive statistics for market fragmentation proxies

This table presents the yearly summary statistics of the constructed Panel (unbalanced, Table A.2/Panel B) for daily market fragmentation proxies based on daily volume data of LSE, CHIX, BATS, Turquoise. The panel comprises 149 stocks divided into 5 stock groups based on market capitalization.

			2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
$HHItrd$	Herfindahl index (proxy for trade volume fragmentation)	Mean	1	1	1	1.254	1.85	2.35	2.50	2.43	2.62	2.68	2.69	2.89
		Median	1	1	1	1.18	1.83	2.38	2.52	2.47	2.66	2.70	2.71	2.91
		StdDev	0	0	0	0.27	0.47	0.44	0.37	0.34	0.39	0.35	0.39	0.35
$HHIqu5$	Herfindahl index (proxy for best 5-level quote fragmentation)	Mean	1	1	1	1.66	3.08	3.32	3.21	3.55	3.53	3.50	3.50	3.54
		Median	1	1	1	1.70	3.22	3.35	3.30	3.58	3.57	3.54	3.54	3.58
		StdDev	0	0	0	0.64	0.60	0.37	0.52	0.28	0.28	0.31	0.27	0.25
$HHIqu10$	Herfindahl index (proxy for best 10- level quote fragmentation)	Mean	1	1	1	1.65	3.02	3.26	3.18	3.50	3.47	3.42	3.43	3.46
		Median	1	1	1	1.66	3.16	3.31	3.25	3.52	3.50	3.46	3.47	3.51
		StdDev	0	0	0	0.64	0.61	0.38	0.51	0.29	0.28	0.33	0.30	0.29
$HHItrd2$	alternative measure of Herfindahl index (1- inverse of HHItrd)	Mean	0	0	0	0.16	0.42	0.56	0.59	0.58	0.61	0.62	0.62	0.65
		Median	0	0	0	0.15	0.45	0.58	0.60	0.60	0.62	0.63	0.63	0.66
		StdDev	0	0	0	0.15	0.16	0.10	0.07	0.07	0.07	0.06	0.06	0.05
	observations (stock*day)		29778	34774	36971	37818	37686	37543	37251	37546	37697	37692	37696	37131

Table A.7: Descriptive statistics for market fragmentation proxies

This table presents the quintile based summary statistics of the constructed Panel (unbalanced, Table A.2/Panel B) for daily market fragmentation proxies based on on daily volume data of LSE, CHIX, BATS, Turquoise. The panel comprises 149 stocks, divided into 5 stock groups based on market capitalization.

			All	Small	Q2	Q3	Q4	Large
$HHItrd$	Herfindahl index (proxy for trade volume fragmentation)	Mean	2.05	1.84	1.99	2.09	2.15	2.16
		Median	2.26	1.85	2.14	2.31	2.44	2.48
		StdDev	0.79	0.74	0.77	0.79	0.79	0.79
$HHIqu5$	Herfindahl index (proxy for best 5-level quote fragmentation)	Mean	2.7	2.59	2.68	2.69	2.73	2.8
		Median	3.26	3.13	3.26	3.21	3.29	3.37
		StdDev	1.12	1.14	1.13	1.08	1.1	1.13
$HHIqu10$	Herfindahl index (proxy for best 10- level quote fragmentation)	Mean	2.66	2.52	2.63	2.65	2.71	2.77
		Median	3.19	2.98	3.16	3.16	3.25	3.32
		StdDev	1.09	1.1	1.1	1.06	1.08	1.11
$HHItrd2$	alternative measure of Herfindahl index (1- inverse of HHItrd)	Mean	0.41	0.36	0.4	0.42	0.44	0.44
		Median	0.56	0.46	0.53	0.57	0.59	0.6
		StdDev	0.27	0.26	0.27	0.27	0.27	0.27
	observations (stock*day)		439583	90046	88374	87239	86786	87138

Table A.8: Yearly descriptive statistics for liquidity measures

This table presents the yearly summary statistics of the constructed Panel (unbalanced, Table A.2/Panel B) for the daily liquidity measures based on LSE data. The panel comprises 149 stocks, divided into 5 stock groups based on market capitalization.

			2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
$spread\_abs$	absloute quoted spread (GBX)	Mean	1.86	1.92	1.62	1.64	1.47	1.14	1.43	1.18	1.32	1.40	1.37	1.42
		Median	0.95	1.09	1.00	0.99	0.77	0.74	0.77	0.78	0.92	0.91	0.93	0.90
$spread\_bps$	percentage quoted spread (bps)	Mean	29.78	29.00	21.67	29.67	29.92	19.85	17.80	16.68	13.99	13.13	12.21	13.45
		Median	17.95	17.88	15.19	21.37	18.84	13.54	13.48	12.67	11.98	11.04	10.71	11.46
$espread$	effective half-spread (bps)	Mean	10.65	10.47	7.79	10.59	10.46	6.94	6.08	5.83	5.00	4.68	4.35	4.02
		Median	6.79	6.64	5.79	7.60	6.66	5.13	4.89	4.49	4.35	4.15	4.15	3.72
$rspread$	share volume wighted 5-minute realized half-spread (bps)	Mean	1.90	1.66	0.07	-1.19	-1.42	-0.64	-0.17	-0.29	-0.45	-0.47	-0.46	-0.39
		Median	0.42	0.17	-0.14	-0.87	-1.47	-0.72	-0.11	-0.31	-0.27	-0.23	-0.24	-0.22
$price\_impact$	share volume wighted 5-minute price_impact (bps)	Mean	8.70	8.80	7.72	11.77	11.88	7.59	6.32	6.13	5.45	5.16	4.81	4.41
		Median	6.36	6.50	6.02	8.50	8.62	5.92	4.92	4.77	4.50	4.21	4.15	3.78
$depth1$	avergae depth (BBO level/GBP 100)	Mean	1100.07	953.33	769.59	300.56	207.66	271.41	248.06	272.87	278.58	279.73	266.33	230.32
		Median	448.10	437.74	388.39	187.12	130.78	157.20	135.85	140.06	168.58	154.91	146.80	128.20
$depth3$	average cumulative depth (best three levels/GBP 100)	Mean	3820.29	3447.01	2811.00	1015.18	784.24	1294.38	1255.92	1402.60	1594.38	1644.42	1564.53	1307.23
		Median	1405.16	1393.99	1254.77	576.39	439.98	598.33	596.39	638.68	904.13	909.68	880.79	743.25
	observations (stock*day)		29778	34774	36971	37818	37686	37543	37251	37546	37697	37692	37696	37131

Table A.9: Descriptive statistics for liquidity measures: pre-MiFID (2005–2008) and post-MiFID (2008-2016) periods

This table presents the quintile based post-MiFID (2008–2016) and pre-MiFID (2005–2007) summary statistics of the constructed Panel (unbalanced, Table A.2/Panel B) for the daily liquidity measures based on LSE data. The panel comprises 149 stocks, divided into 5 stock groups based on market capitalization.

			pre-MiFID (2005–2007)						post-MiFID (2008-2016)					
Variable			All	Small	Q2	Q3	Q4	Large	All	Small	Q2	Q3	Q4	Large
$spread\_abs$	absloute quoted spread (GBX)	Mean	1.84	4.01	1.62	1.24	1.08	0.97	1.29	1.97	1.32	1.09	1.22	0.87
		Median	1.03	1.78	1.08	0.93	0.81	0.72	0.84	0.99	0.90	0.81	0.80	0.79
		StdDev	3.49	6.66	1.73	1.09	0.87	0.65	1.71	2.93	1.25	1.04	1.46	0.76
$spread\_bps$	percentage quoted spread (bps)	Mean	27.96	64.63	27.10	18.73	14.10	10.28	17.70	36.12	20.43	13.79	10.79	7.34
		Median	17.60	46.90	23.13	16.98	12.92	9.63	12.83	26.20	16.97	12.47	10.39	6.43
		StdDev	32.82	52.35	15.40	8.35	4.70	3.37	18.74	31.55	12.28	7.14	4.13	5.03
$espread$	effective half-spread (bps)	Mean	10.01	21.96	9.34	6.97	5.64	4.55	6.15	11.98	6.85	4.89	4.07	2.96
		Median	6.54	14.96	8.16	6.30	5.15	4.28	4.71	8.49	5.72	4.49	4.01	2.49
		StdDev	11.99	20.69	4.75	2.92	1.96	1.53	6.55	11.35	4.54	2.66	1.51	2.34
$depth1$	average depth (BBO level/GBP 100)	Mean	853.79	88.52	228.01	447.18	849.93	2794.03	260.22	63.51	101.63	185.02	368.39	583.24
		Median	369.38	75.07	203.47	387.36	714.46	1532.31	148.72	55.80	88.08	153.04	305.11	457.90
		StdDev	2413.17	77.28	196.06	299.00	582.49	4978.89	313.74	36.07	63.79	116.82	239.16	479.78
$depth3$	average cumulative depth (best three levels/GBP 100)	Mean	3047.92	287.73	734.87	1440.15	2876.36	10405.37	1331.01	256.23	452.21	916.58	1966.30	3067.78
		Median	1182.28	245.62	658.37	1230.41	2382.55	5282.24	685.35	221.92	355.49	685.15	1594.45	2401.87
		StdDev	9406.68	251.98	634.30	977.94	2063.30	19576.90	1643.40	144.98	305.47	687.65	1388.86	2337.90
observations (stock*day)			115353	25286	23232	22320	22327	22188	324230	64760	65142	64919	64459	64950

Table A.10: Descriptive statistics for realized half-spread measures

This table presents the quintile based post-MiFID (2008–2016) and pre-MiFID (2005–2007) summary statistics of the constructed Panel (unbalanced, Table A.2/Panel B) for the various daily volume weighted realized half-spreads measures based on intraday LSE data. The panel comprises 149 stocks, divided into 5 stock groups based on market capitalization.

			2005–2007/pre MiFID						2008–2016/post MiFID					
			All	Small	Q2	Q3	Q4	Large	All	Small	Q2	Q3	Q4	Large
	share volume	Mean	3.07	9.03	2.19	1.32	1.16	0.90	0.24	1.40	0.12	-0.09	-0.12	-0.10
$rspread1$	wighted 10-sec	Median	1.32	4.29	1.80	1.18	1.02	0.82	-0.13	0.39	-0.05	-0.19	-0.19	-0.18
	realized spread (bps)	StdDev	8.99	17.48	3.42	2.07	1.43	0.95	3.52	7.04	2.51	1.50	0.99	1.02
	share volume	Mean	2.27	7.76	1.33	0.59	0.54	0.42	-0.23	0.53	-0.49	-0.51	-0.39	-0.27
$rspread2$	wighted 30-sec	Median	0.74	3.31	1.07	0.57	0.49	0.38	-0.34	-0.18	-0.46	-0.45	-0.37	-0.26
	realized half-spread (bps)	StdDev	8.96	17.55	3.62	2.25	1.57	1.02	3.48	6.95	2.59	1.64	1.12	1.00
	share volume	Mean	1.82	7.09	0.81	0.17	0.18	0.18	-0.47	0.07	-0.82	-0.74	-0.53	-0.34
$rspread3$	wighted 1-minute	Median	0.42	2.82	0.65	0.21	0.17	0.17	-0.44	-0.47	-0.66	-0.58	-0.45	-0.28
	realized half-spread (bps)	StdDev	9.08	17.82	3.85	2.43	1.68	1.07	3.61	7.13	2.79	1.82	1.28	1.04
	share volume	Mean	1.12	5.79	-0.10	-0.42	-0.24	0.00	-0.68	-0.71	-1.13	-0.82	-0.49	-0.25
$rspread4$	wighted 5-minute	Median	0.07	2.06	-0.04	-0.26	-0.16	0.06	-0.41	-0.77	-0.74	-0.54	-0.35	-0.17
	realized half-spread (bps)	StdDev	9.55	18.85	4.67	2.97	1.99	1.29	4.12	7.87	3.51	2.35	1.69	1.29
	observations (stock*day)		115353	25286	23232	22320	22327	22188	324230	64760	65142	64919	64459	64950

Table A.11: Descriptive statistics for price impact/adverse selection cost measures

This table presents the quintile based post-MiFID (2008–2016) and pre-MiFID (2005–2007) summary statistics of the constructed Panel (unbalanced, Table A.2/Panel B) for the various daily volume weighted price impacts measures based on intraday LSE data. The panel comprises 149 stocks, divided into 5 stock groups based on market capitalization.

			2005–2007/pre-MiFID						2008-2016/post MiFID					
			All	Small	Q2	Q3	Q4	Large	All	Small	Q2	Q3	Q4	Large
	share volume	Mean	6.91	12.87	7.12	5.63	4.48	3.64	5.91	10.57	6.73	4.99	4.20	3.06
$price\_impact1$	wighted 10-sec	Median	5.06	9.74	6.27	5.10	4.15	3.41	4.65	8.11	5.73	4.53	4.02	2.68
	price impact (bps)	StdDev	6.71	11.43	3.85	2.59	1.63	1.24	5.18	8.41	4.06	2.56	1.66	1.90
	share volume	Mean	7.72	14.14	8.00	6.36	5.09	4.12	6.38	11.44	7.35	5.41	4.47	3.24
$price\_impact2$	wighted 30-sec	Median	5.68	10.81	7.05	5.73	4.73	3.87	4.94	8.71	6.15	4.81	4.21	2.80
	price impact (bps)	StdDev	7.38	12.54	4.40	3.04	1.89	1.37	5.73	9.19	4.63	2.96	1.92	2.17
	share volume	Mean	8.18	14.84	8.51	6.79	5.45	4.36	6.62	11.89	7.68	5.64	4.61	3.30
$price\_impact3$	wighted 60-sec	Median	6.04	11.36	7.55	6.15	5.06	4.10	5.08	9.07	6.37	4.95	4.30	2.83
	price impact (bps)	StdDev	7.80	13.28	4.73	3.27	2.06	1.46	6.04	9.64	4.99	3.22	2.10	2.30
	share volume	Mean	8.87	16.12	9.43	7.38	5.87	4.55	6.83	12.68	7.98	5.72	4.57	3.21
$price\_impact4$	wighted 5-min	Median	6.51	12.45	8.34	6.67	5.45	4.25	5.04	9.53	6.50	4.92	4.18	2.70
	price impact (bps)	StdDev	8.68	14.73	5.53	3.72	2.39	1.72	6.87	11.06	5.73	3.71	2.39	2.50
observations (stock*day)			115353	25286	23232	22320	22327	22188	324230	64760	65142	64919	64459	64950

Table A.12: Descriptive statistics for regression variables

This table presents the descriptive statistics calculated using the balanced sample of 132 stocks for all variables employed in the regression analyses divided into four periods: 2005–2016 represents the full sample; 2005–2007 represents the pre-MiFID era and 2008–2010, 2011-2013, 2014–2016 represent the three equally divided post-MiFID periods. $hft2$ represents the per minute quote update for the best 5 depth levels in the limit order book, $HHItrd$ is the Herfindhal-Hirchman index (HHI) showing the degree of market fragmentation, $spread\_bps$ is the time weighted daily quoted spread in basis point, $espread$ is the volume weighted effective half-spread in basis point, $rspread$ is the volume weighted 5-minute realized half-spreads in basis point, $p.impact$ is the volume weighted 5-minute price impacts in basis point, $depth1$ is the average BBO level quoted depth measured in GBP100, $depth3$ is the average cumulative depth upto three best limit price measured in GBP100, $voltintra$ is the intraday mid price range volatility measured in basis point, $mktcap$ is the average market capitalization measured in million GBP, $price$ is the daily average price level measured in GBX, $value$ is the average daily trading volume in GBP1000, $rtick$ is the relative tick size, and $size$ is the average trade size. All are daily measures constructed from the intraday millisecond records. This table presents the reference value (mean) for all regression estimates.

Year	Observations		$hft2$	$HHItrd$	$spread\_bps$	$espread$	$rspread$	$p.impacts$	$depth1$	$depth3$	$voltintra$	$mktcap$	$price$	$value$	$rtick$	$size$
							All stocks									
2005-2016	346368	Mean	95.91	2.17	18.37	6.41	-0.47	6.89	358.15	1636.19	286.23	9666	934.35	19.25	0.0006	5159.34
	346368	StdDev	135.38	0.74	19.83	6.92	4.93	6.61	869.90	3613.90	542.32	17853	0.00	973.37	39.3631	0.00
2005-2007	50688	Mean	20.08	1.00	23.20	8.38	0.46	7.93	867.15	3175.77	252.56	9187	847.28	31.46	0.0008	11298.30
	50688	StdDev	20.06	0.00	25.53	9.33	7.62	7.04	2072.99	8391.49	190.24	18521	0.00	726.75	60.9117	0.00
2008-2010	97284	Mean	76.92	1.83	24.64	8.65	-1.11	9.76	272.68	1087.78	412.36	8113	697.27	22.63	0.0008	5097.49
	97284	StdDev	108.80	0.61	27.07	9.47	6.38	8.92	334.17	1480.89	961.92	16814	0.00	734.27	47.8480	0.00
2011-2013	99528	Mean	113.88	2.52	15.46	5.36	-0.36	5.74	278.93	1488.97	242.37	9926	928.25	13.54	0.0005	3761.03
	99528	StdDev	144.43	0.37	12.26	3.99	3.12	4.84	313.53	1745.04	214.38	18218	0.00	923.22	24.0534	0.00
2014-2016	98868	Mean	135.39	2.76	12.63	4.27	-0.43	4.70	261.04	1534.68	223.52	11179	1218.41	15.42	0.0005	3480.48
	98868	StdDev	161.76	0.38	8.94	2.53	2.13	3.35	288.70	1699.08	181.64	17990	0.00	1232.89	23.5760	0.00
							Large stocks									
2005-2016	173184	Mean	155.91	2.27	10.57	4.07	-0.44	4.52	589.80	2742.64	278.33	17694	1153.90	35.05	0.0006	7390.96
	173184	StdDev	168.21	0.73	6.63	2.51	1.78	3.21	1179.95	4838.50	707.51	22524	0.00	1169.33	50.8014	0.00
2005-2007	25344	Mean	30.40	1.00	13.16	5.27	-0.27	5.54	1491.49	5565.18	232.64	17077	960.57	57.75	0.0007	16670.92
	25344	StdDev	22.80	0.00	8.03	2.51	2.02	3.04	2788.67	11357.79	152.99	23684	0.00	722.96	77.4727	0.00
2008-2010	48642	Mean	125.88	1.99	12.97	5.10	-1.04	6.15	432.65	1794.20	428.60	15155	881.05	41.61	0.0007	7142.37
	48642	StdDev	135.03	0.61	8.71	3.46	2.18	4.31	407.33	1821.45	1299.02	21581	0.00	871.62	61.9440	0.00
2011-2013	49764	Mean	184.08	2.65	9.42	3.59	-0.15	3.75	460.11	2529.89	218.59	18184	1176.01	24.53	0.0005	5403.99
	49764	StdDev	175.06	0.28	3.81	1.41	1.59	2.14	354.29	1933.14	174.88	22943	0.00	1121.60	30.1011	0.00
2014-2016	49434	Mean	221.47	2.80	8.03	2.92	-0.24	3.16	412.69	2442.97	214.02	20015	1499.25	27.56	0.0005	4878.09
	49434	StdDev	189.75	0.32	3.84	1.27	1.12	1.72	334.27	1936.97	139.65	22123	0.00	1515.32	28.3290	0.00
							Small stocks									
2005-2016	173184	Mean	35.91	2.08	26.16	8.76	-0.50	9.27	126.50	529.73	294.13	1639	714.80	3.45	0.0007	2927.72
	173184	StdDev	34.06	0.74	24.93	8.86	6.74	8.11	117.72	510.99	295.84	1125	0.00	656.64	4.3365	0.00
2005-2007	25344	Mean	9.75	1.00	33.24	11.49	1.20	10.31	242.80	786.37	272.48	1297	733.99	5.17	0.0008	5925.68
	25344	StdDev	8.44	0.00	32.21	12.19	10.54	8.85	196.13	646.99	219.51	801	0.00	712.75	5.9952	0.00
2008-2010	48642	Mean	27.96	1.66	36.30	12.20	-1.18	13.38	112.70	381.36	396.12	1072	513.49	3.65	0.0008	3052.61
	48642	StdDev	25.45	0.56	33.42	11.92	8.76	10.70	79.02	265.26	403.26	729	0.00	501.06	4.5993	0.00
2011-2013	49764	Mean	43.68	2.40	21.50	7.13	-0.58	7.73	97.75	448.05	266.16	1667	680.49	2.55	0.0005	2118.07
	49764	StdDev	34.89	0.41	14.61	4.86	4.11	5.85	73.73	431.59	245.37	1000	0.00	569.12	3.0905	0.00
2014-2016	49434	Mean	49.32	2.72	17.24	5.62	-0.62	6.25	109.40	626.38	233.01	2343	937.58	3.29	0.0006	2082.86
	49434	StdDev	38.87	0.42	10.13	2.74	2.78	3.83	94.69	609.85	215.18	1305	0.00	765.59	3.8267	0.00

Table A.13: Pearson correlation coefficient matrix for regression variables

This table presents the Pearson correlation coefficient matrix for the variables employed in the regression analysis, calculated on the same balanced sample employed in the regression analyses. The sample comprises 132 stocks for the period December 2005–December 2016 (2624 days). $hft2$ represents the per minute quote update for the best 5 depth levels in the limit order book, $HHItrd$ is the Herfindhal-Hirchman index (HHI) showing the degree of market fragmentation, $spread\_bps$ is the time weighted daily quoted spread in basis point, $espread$ is the volume weighted effective half-spread in basis point, $realizedspread\_5min$ is the volume weighted 5-minute realized half-spreads in basis point, $priceimpact\_5min$ is the volume weighted 5-minute price impacts in basis point, $depth1$ is the average BBO level quoted depth measured in GBP100, $depth3$ is the average cumulative depth upto three best limit price measured in GBP100, $voltintra$ is the intraday mid price range volatility measured in basis point, $mktcap$ is the average market capitalization measured in million GBP, $invprice$ is the inverse of daily average price level measured in GBX, value is the average daily trading volume in GBP1000, $rtick$ is the relative tick size, and $size$ is the value of the average executed trade, $spread\_bps\_mktvag$ represents the average $spread\_bps$ level over all stocks in the same size group excluding stock i, $espread\_mktvag$ represents the average $espread$ level over all stocks in the same size group excluding stock i, $depth1\_mktvag$ represents the average $depth1$ level over all stocks in the same size group excluding stock i, $hft2\_mktavg$ represents the average HFT intensity over all stocks in the same size group excluding stock i, $HHItrd\_mktavg$ represents the average market fragmentation level over all stocks in the same size group excluding stock i.

ID	Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	$log(hft2)$	1.00																	
2	$HHItrd$	0.55																	
3	$Log(spread\_bps)$	-0.78	-0.40																
4	$Log(espread)$	-0.77	-0.45	0.95															
5	$Log(depth1)$	0.38	-0.09	-0.52	-0.41														
6	$Log(depth3)$	0.47	0.08	-0.59	-0.49	0.97													
7	$Log(voltintra)$	-0.03	-0.24	0.36	0.37	-0.21	-0.28												
8	$Log(mktcap)$	0.75	0.27	-0.83	-0.77	0.72	0.77	-0.23											
9	$Log(value)$	0.72	-0.01	-0.74	-0.66	0.78	0.76	0.01	0.85										
10	$rtick$	-0.39	-0.29	0.46	0.54	0.10	0.05	0.19	-0.28	-0.16									
11	$Log(size)$	0.28	-0.35	-0.45	-0.34	0.87	0.79	-0.06	0.67	0.82	0.03								
12	$realizedspread\_5min$	-0.15	-0.04	0.10	0.13	0.00	0.01	-0.18	-0.01	-0.10	0.09	0.00							
13	$Log(priceimpacts\_5min)$	-0.58	-0.36	0.77	0.79	-0.36	-0.43	0.44	-0.65	-0.52	0.40	-0.29	-0.32						
14	$invprice$	-0.11	-0.09	0.24	0.26	-0.16	-0.19	0.21	-0.16	-0.12	0.41	-0.15	0.10						
15	$hft2\_mktavg$	0.87	0.60	-0.68	-0.66	0.40	0.51	-0.11	0.73	0.59	-0.24	0.22	-0.05	-0.53	-0.04				
16	$HHItrd\_mktavg$	0.52	0.91	-0.34	-0.40	-0.20	-0.03	-0.21	0.19	-0.08	-0.30	-0.45	-0.02	-0.33	-0.05	0.58			
17	$Log(spread\_bps\_mktavg)$	-0.73	-0.47	0.79	0.75	-0.58	-0.67	0.32	-0.82	-0.66	0.25	-0.42	-0.01	0.63	0.11	-0.84	-0.42		
18	$Log(espread\_mktavg)$	-0.74	-0.55	0.78	0.75	-0.52	-0.63	0.33	-0.78	-0.60	0.27	-0.33	-0.01	0.64	0.11	-0.85	-0.50	0.98	
19	$Log(depth1\_mktavg)$	0.41	-0.14	-0.58	-0.49	0.81	0.79	-0.18	0.76	0.78	-0.05	0.81	0.03	-0.43	-0.06	0.47	-0.24	-0.71	-0.63

all coefficients are significant at the 1% level

Table A.14: The effects of HFT and market fragmentation on liquidity

This table presents the panel regression results of Models (2.1)–(2.3) where various liquidity measures are regressed on HFT ($hft2$), market fragmentation ($HHItrd$) proxy. $hft2$ represents the per minute quote update for the best 5 depth levels in the limit order book. $HHItrd$ is the Herfindhal-Hirchman index (HHI), shows the degree of market fragmentation. The liquidity measures are time weighted quoted spread ($spread\_bps$), volume weighted effective-half spread ($espread$), average quoted depth at best limit price ($depth1$), accumulated average quoted depth up to best three limit price ($depth3$), volume weighted 5-minute realized-half spread ($rspread\_5min$), and volume weighted 5-minute price impacts ($price\_impact\_5min$). All dependent variables are log transformed except $rspread\_5min$, all spreads based measures are in basis point and depth in 100GBP. Panel A shows the result for $spread\_bps$, $espread$ and $depth1$ and Panel B shows the rest. Control variables are log market capitalization ($Log(mktcap)$), log normalized intraday mid-price volatility ($Log(voltintra)$) and inverse of average daily price level ($invprice$). The regression is based on a balanced panel of 132 stocks and 2624 days (December2005–December2016) and have both time (daily) and stock fixed effects. Coefficient estimates are OLS , t-statistics shown in the parentheses below the coefficient, calculated using Newey-West (HAC) standard errors (lags are optimally determined). ***, **, * denote significance at 1%, 5%, and 10% levels, respectively.

PANEL A

	I	II	III	IV	V	VI	VII	VIII	IX
		Log(spread_bps)			Log(espread)			Log(depth1)	
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Log(hft2)	-0.288***	-0.289***	-0.283***	-0.322***	-0.324***	-0.282***	-0.25***	-0.247***	-0.155***
	(-92.96)	(-93.21)	(-54.74)	(-99.98)	(-100.38)	(-51.56)	(-40.36)	(-40.14)	(-14.76)
HHItrd		0.052***	0.065***		0.065***	0.144***		-0.105***	0.068***
		-17.72	-9.09		-20.72	-18.81		(-19.51)	-4.5
Log(hft2)*HHItrd			-0.003*			-0.02***			-0.044***
			(-1.96)			(-11.24)			(-11.27)
Log(mktcap)	-0.198***	-0.201***	-0.202***	-0.127***	-0.131***	-0.139***	0.811***	0.818***	0.801***
	(-49.14)	(-49.93)	(-49.97)	(-29.43)	(-30.41)	(-32.17)	-85.82	-86.18	-86.22
Log(voltintra)	0.167***	0.169***	0.168***	0.21***	0.213***	0.209***	0.029***	0.025***	0.018***
	-58.38	-59.99	-59.69	-63.96	-65.83	-64.35	-7.69	-6.42	-4.59
Inv(price)	17.159***	17.205***	17.335***	19.847***	19.904***	20.694***	1.584	1.492	3.227*
	-22.65	-22.68	-22.7	-21.99	-22.04	-22.83	-0.82	-0.77	-1.69
stock fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
time fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	346368	346368	346368	346368	346368	346368	346368	346368	346368
R-Square	0.87	0.87	0.87	0.84	0.84	0.84	0.82	0.82	0.82

PANEL B

	I	II	III	IV	V	VI	VII	VIII	IX
	Log(depth3)			rspread_5min			Log(price_impact_5min)		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Log(hft2)	-0.268***	-0.268***	-0.264***	-1.871***	-1.866***	-4.251***	-0.239***	-0.242***	-0.08***
	(-39.84)	(-39.78)	(-22.66)	(-16.17)	(-16.18)	(-19.34)	(-65.29)	(-66)	(-13.69)
HHItrd		0.003	0.01		-0.245***	-4.711***		0.133***	0.433***
		-0.52	-0.62		(-3.86)	(-20.12)		(33.12)	(50.34)
Log(hft2)*HHItrd			-0.002			1.125***			-0.076***
			(-0.43)			(20)			(-38.8)
Log(mktcap)	0.863***	0.863***	0.862***	0.472***	0.488***	0.929***	-0.112***	-0.121***	-0.15***
	-82.97	-82.75	-83.75	(8.3)	(8.3)	(15.55)	(-24.13)	(-26.01)	(-33.11)
Log(voltintra)	-0.007*	-0.007*	-0.007*	-1.44***	-1.451***	-1.272***	0.399***	0.405***	0.393***
	(-1.71)	(-1.68)	(-1.75)	(-33.23)	(-34.24)	(-29.22)	(69.93)	(72.22)	(69.25)
Inv(price)	0.173	0.176	0.249	212.544***	212.328***	167.51***	10.844***	10.963***	13.937***
	-0.09	-0.09	-0.13	(13.19)	(13.2)	(11.77)	(13.7)	(13.84)	(18.23)
stock fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
time fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	346368	346368	346368	346368	346368	346368	346368	346368	346368
R-Square	0.81	0.81	0.81	0.16	0.16	0.19	0.64	0.64	0.64

Table A.15: The effects of HFT and market fragmentation on liquidity: large and small stocks

This table presents the panel regression results of Models (2.1)–(2.3) for large and small stocks in the panels A1, B1 and A2, B2 respectively, where various liquidity measures are regressed on HFT ($hft2$), market fragmentation ($HHItrd$) proxy. $hft2$ represents the per minute quote update for the best 5 depth levels in the limit order book. $HHItrd$ is the Herfindhal-Hirchman index (HHI), shows the degree of market fragmentation. The liquidity measures are time weighted quoted spreads ($spread\_bps$), volume weighted effective-half spreads ($espread$), average quoted depths at best limit price ($depth1$), accumulated average quoted depths up to best three limit price ($depth3$), volume weighted 5-minute realized-half spreads ($rspread\_5min$), and volume weighted 5-minute price impacts ($price\_impact\_5min$). All dependent variables are log transformed except $rspread\_5min$, all spreads based measures are in basis point and depth in GBP100. Panel (A1, A2)show the result for $spread\_bps$, $espread$ and $depth1$ and Panel (B1, B2) show the rests. Control variables are log market capitalization ($Log(mktcap)$), log normalized intraday mid-price volatility ($Log(voltintra)$) and inverse of average daily price level ($invprice$) are significant with expected sign but not shown. The regression is based on a balanced panel of 66 stocks and 2624 days (December2005–December2016) for each group and have both time (daily) and stock fixed effects. Coefficient estimates are OLS, t-statistics shown in the parentheses below the coefficient, calculated using Newey-West (HAC) standard errors (lags are optimally determined). ***, **, * denote significance at 1%, 5%, and 10% levels, respectively.

PANEL A1 Large-cap stocks

	I	II	III	IV	V	VI	VII	VIII	IX
	Log(spread_bps)			Log(espread)			Log(depth1)		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Log(hft2)	-0.392***	-0.391***	-0.283***	-0.455***	-0.454***	-0.248***	-0.649***	-0.65***	-0.099***
	(-84.62)	(-84.42)	(-30.17)	(-87.43)	(-87.11)	(-25.48)	(-49.23)	(-49.49)	(-3.28)
HHItrd		0.035***	0.241***		0.033***	0.427***		-0.044***	1.01***
		(8.29)	(15.52)		(6.89)	(25.6)		(-4.78)	(21.74)
Log(hft2)*HHItrd			-0.043***			-0.083***			-0.222***
			(-13.64)			(-24.78)			(-22.81)
Observations	173184	173184	173184	173184	173184	173184	173184	173184	173184
R-Square	0.83	0.84	0.84	0.81	0.81	0.81	0.75	0.75	0.76

PANEL A2 Small-cap stocks

	I	II	III	IV	V	VI	VII	VIII	IX
	Log(spread_bps)			Log(espread)			Log(depth1)		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Log(hft2)	-0.237***	-0.239***	-0.282***	-0.251***	-0.254***	-0.294***	-0.087***	-0.084***	-0.03**
	(-57.62)	(-58.09)	(-36.79)	(-59.59)	(-60.2)	(-36.41)	(-15.96)	(-15.46)	(-2.49)
HHItrd		0.076***	-0.004		0.086***	0.011		-0.093***	0.009
		(20.02)	(-0.34)		(21.6)	(1)		(-16.83)	(0.53)
Log(hft2)*HHItrd			0.023***			0.022***			-0.03***
			(7.87)			(7.01)			(-5.93)
Wald test ($\beta_2 + \beta_3 = 0$)			5.85**			15.09***			2.87*
Observations	173184	173184	173184	173184	173184	173184	173184	173184	173184
R-Square	0.80	0.81	0.81	0.78	0.78	0.78	0.72	0.73	0.73

PANEL B1 Large-cap stocks

	I	II	III	IV	V	VI	VII	VIII	IX
	Log(depth3)			rspread_5min			Log(price_impact_5min)		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Log(hft2)	-0.664***	-0.662***	-0.09***	-0.074***	-0.081***	-0.178**	-0.431***	-0.428***	-0.172***
	(-47.96)	(-47.88)	(-2.88)	(-2.7)	(-2.93)	(-2.45)	(-74.74)	(-73.99)	(-15.37)
HHItrd		0.061***	1.156***		-0.217***	-0.404***		0.101***	0.59***
		(5.93)	(23.69)		(-9.66)	(-3.79)		(15.96)	(31.41)
Log(hft2)*HHItrd			-0.231***			0.039*			-0.103***
			(-22.68)			(1.87)			(-27.15)
Observations	173184	173184	173184	173184	173184	173184	173184	173184	173184
R-Square	0.72	0.72	0.74	0.20	0.20	0.20	0.61	0.61	0.62

PANEL B2 Small-cap stocks

	I	II	III	IV	V	VI	VII	VIII	IX
	Log(depth3)			rspread_5min			Log(price_impact_5min)		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Log(hft2)	-0.117***	-0.117***	-0.135***	-2.488***	-2.475***	-6.718***	-0.146***	-0.152***	-0.039***
	(-18.58)	(-18.59)	(-10.1)	(-16.3)	(-16.31)	(-20.97)	(-33.17)	(-34.26)	(-4.43)
HHItrd		0.01	-0.023		-0.436***	-8.366***		0.178***	0.387***
		(1.56)	(-1.17)		(-5.23)	(-21.8)		(35.03)	(28.56)
Log(hft2)*HHItrd			0.01*			2.318***			-0.061***
			(1.69)			(21.57)			(-17.14)
Wald test ($\beta_2 + \beta_3 = 0$)			0.85						
Observations	173184	173184	173184	173184	173184	173184	173184	173184	173184
R-Square	0.69	0.69	0.69	0.20	0.21	0.24	0.53	0.54	0.54

Table A.16: The time-varying effects of HFT and market fragmentation on liquidity

This table presents the panel regression results of Models (2.4)–(2.6) where various liquidity measures are regressed on HFT ($hft2$), market fragmentation ($HHItrd$) proxy. $hft2$ represents the per minute quote update for the best 5 depth levels in the limit order book. $HHItrd$ is the Herfindhal-Hirchman index (HHI) , shows the degree of market fragmentation. The liquidity measures are time weighted quoted spreads ($spread\_bps$), volume weighted effective-half spreads ($espread$), average quoted depths at best limit price ($depth1$), accumulated average quoted depths up to best three limit price ($depth3$), volume weighted 5-minute realized-half spreads ($rspread\_5min$), and volume weighted 5-minute price impacts ($price\_impact\_5min$). All dependent variables are log transformed except $rspread\_5min$, all spreads based measures are in basis point and depth in GBP100. Panel A shows the result for $spread\_bps$, $espread$ and $depth1$ and Panel B shows the rest. $DYr8, 9, 10$, $DYr11, 12, 13$, $DYr14, 15, 16$ represent the period dummies for the three three-year blocks 2008–2010 , 201–2013 and 2014–2016, respectively. Control variables are log market capitalization ($Log(mktcap)$), log normalized intraday mid-price volatility ($Log(voltintra)$) and inverse of average daily price level ($invprice$). Models (4), (5), (6) examine the time-varying effect of HFT, market fragmentation and their interaction in the latter three three-year block periods of the sample, respectively. The regression is based on a balanced panel of 132 stocks and 2624 days (December 2005–December 2016) and have both time (daily) and stock fixed effects. Coefficient estimates are OLS, t-statistics shown in the parentheses below the coefficient, calculated using Newey-West (HAC) standard errors (lags are optimally determined). ***, **, * denote significance at 1%, 5%, and 10% levels, respectively.

PANEL A

	I	II	III	IV	V	VI	VII	VIII	IX
	Log(spread_bps)			Log(espread)			Log(depth1)		
	(4)	(5)	(6)	(4)	(5)	(6)	(4)	(5)	(6)
Log(hft2)	-0.278***	-0.288***	-0.252***	-0.294***	-0.323***	-0.268***	-0.004	-0.245***	-0.081***
	(-56.67)	(-94.17)	(-46.5)	(-54.92)	(-101.01)	(-46.63)	(-0.44)	(-39.83)	(-7.25)
HHItrd	0.056***	-0.024***	0.122***	0.067***	0.021***	0.166***	-0.034***	-0.195***	0.208***
	-19.63	(-3.8)	-15.65	-21.74	-3.06	-20.17	(-6.83)	(-18.8)	-12.66
Log(hft2)*DYr8,9,10	-0.03***			-0.034***			-0.353***		
	(-6.92)			(-7.26)			(-39.26)		
Log(hft2)*DYr11,12,13	0.005			-0.012***			-0.232***		
	-1.17			(-2.74)			(-25.83)		
Log(hft2)*DYr14,15,16	-0.013***			-0.052***			-0.246***		
	(-3.09)			(-11.1)			(-25.81)		
HHItrd*DYr11,12,13		0.106***			0.072***			0.136***	
		-13.46			-8.85			-11.5	
HHItrd*DYr14,15,16		0.112***			0.054***			0.119***	
		-14.57			-6.68			-9.08	
Log(hft2)*HHItrd			-0.029***			-0.032***			-0.105***
			(-12.52)			(-13.21)			(-21.07)
Log(hft2)*HHItrd*DYr11,12,13			0.018***			0.013***			0.039***
			-17.17			-12.35			-23.37
Log(hft2)*HHItrd*DYr14,15,16			0.015***			0.003***			0.038***
			-13.15			-2.98			-18.84
Log(mktcap)	-0.201***	-0.2***	-0.198***	-0.143***	-0.13***	-0.14***	0.791***	0.819***	0.813***
	(-48.23)	(-49.51)	(-48.19)	(-32.24)	(-30.2)	(-31.83)	-84.74	-86.04	-85.69
Log(voltintra)	0.17***	0.169***	0.168***	0.212***	0.213***	0.21***	0.02***	0.024***	0.017***
	-61.17	-59.04	-59.15	-65.73	-65.23	-64.05	-5.76	-6.34	-4.46
Inv(price)	16.421***	16.784***	17.052***	19.32***	19.611***	20.132***	-0.699	0.941	2.878
	-21.55	-21.81	-22.09	-21.24	-21.53	-21.9	(-0.39)	-0.48	-1.51
stock fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
time fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	346368	346368	346368	346368	346368	346368	346368	346368	346368
R-Square	0.87	0.87	0.87	0.84	0.84	0.84	0.83	0.82	0.82

PANEL B

	I	II	III	IV	V	VI	VII	VIII	IX
	Log(depth3)			rspread_5min			price_impact_5min		
	(4)	(5)	(6)	(4)	(5)	(6)	(4)	(5)	(6)
Log(hft2)	-0.058***	-0.264***	-0.169***	-3.953***	-1.844***	-4.345***	-0.107***	-0.242***	-0.072***
	(-5.65)	(-39.33)	(-13.49)	(-18.84)	(-16.24)	(-19.46)	(-18.43)	(-66.16)	(-11.58)
HHItrd	0.074***	-0.18***	0.191***	-0.599***	-1.217***	-4.888***	0.153***	0.163***	0.447***
	-12.59	(-14.6)	-10.47	(-12.6)	(-6.75)	(-20.21)	-37.84	-21.8	-46.72
Log(hft2)*DYr8,9,10	-0.361***			1.931***			-0.126***		
	(-37.5)			-10.88			(-25.42)		
Log(hft2)*DYr11,12,13	-0.165***			2.415***			-0.144***		
	(-17.2)			-15.69			(-28.66)		
Log(hft2)*DYr14,15,16	-0.195***			2.691***			-0.179***		
	(-19)			-17.49			(-36.22)		
HHItrd*DYr11,12,13		0.258***			1.417***			-0.038***	
		-18.33			-7.5			(-3.95)	
HHItrd*DYr14,15,16		0.26***			1.341***			-0.044***	
		-16.92			-7.48			(-4.61)	
Log(hft2)*HHItrd			-0.082***			1.204***			-0.083***
			(-14.76)			-17.91			(-30.86)
Log(hft2)*HHItrd*DYr11,12,13			0.055***			-0.053***			0.007***
			-26.87			(-2.72)			-6.05
Log(hft2)*HHItrd*DYr14,15,16			0.047***			-0.047**			0.002*
			-19.7			(-2.25)			-1.84
Log(mktcap)	0.856***	0.866***	0.876***	1.122***	0.505***	0.915***	-0.163***	-0.121***	-0.151***
	-83.24	-82.91	-83.74	-17.76	-8.46	-14.56	(-34.74)	(-26.14)	(-32.64)
Log(voltintra)	-0.005	-0.007*	-0.007*	-1.313***	-1.453***	-1.272***	0.397***	0.405***	0.393***
	(-1.34)	(-1.74)	(-1.81)	(-31.5)	(-34.82)	(-29.16)	-71.75	-72.56	-69.08
Inv(price)	-3.629**	-0.856	-0.493	201.111***	206.635***	168.227***	11.336***	11.111***	13.636***
	(-2.01)	(-0.44)	(-0.26)	-12.89	-13.14	-11.81	-14.52	-14.05	-17.66
stock fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
time fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	346368	346368	346368	346368	346368	346368			
R-Square	0.82	0.81	0.81	0.19	0.16	0.19	0.64	0.64	0.64

Table A.17: The time-varying effects of HFT and market fragmentation on liquidity: Large and small stocks (part-1/2)

This table presents the panel regression results of Models (2.4)–(2.6) for large and small stocks where various liquidity measures are regressed on HFT ($hft2$), market fragmentation ($HHItrd$) proxy. $hft2$ represents the per minute quote update for the best 5 depth levels in the limit order book. $HHItrd$ is the Herfindhal-Hirchman index (HHI), shows the degree of market fragmentation. The liquidity measures are time weighted quoted spread s($spread\_bps$), volume weighted effective-half spreads ($espread$), average quoted depths at best limit price ($depth1$). All dependent variables are log transformed, all spreads based measures are in basis point and depth in GBP100. Panel A and Panel B show the result for large and small stocks respectively. $DYr8,9,10$, $DYr11,12,13$, $DYr14,15,16$ represent the period dummies for the three three-year blocks 2008–2010 , 201–2013 and 2014–2016, respectively. Control variables are log market capitalization ($Log(mktcap)$), log normalized intraday mid-price volatility ($Log(voltintra)$) and inverse of average daily price level ($invprice$). Models (4), (5), (6) examine the time-varying effect of HFT, market fragmentation and their interaction in the latter three three-year block periods of the sample, respectively. The regression is based on a balanced panel of 66 stocks and 2624 days (December 2005–December 2016) for each group and have both time (daily) and stock fixed effects. Coefficient estimates are OLS, t-statistics shown in the parentheses below the coefficient , calculated using Newey-West (HAC) standard errors (lags are optimally determined). ***, **, * denote significance at 1%, 5%, and 10% levels, respectively.

PANEL A: Large stocks

	I	II	III	IV	V	VI	VII	VIII	IX
	Log(spread_bps)			Log(espread)			Log(depth1)		
	(4)	(5)	(6)	(4)	(5)	(6)	(4)	(5)	(6)
Log(hft2)	-0.279***	-0.392***	-0.305***	-0.29***	-0.455***	-0.275***	-0.171***	-0.652***	-0.108***
	(-31.93)	(-84.55)	(-30.83)	(-33.36)	(-87.47)	(-26.96)	(-5.71)	(-49.82)	(-3.33)
HHItrd	0.033***	0.045***	0.194***	0.029***	0.08***	0.368***	-0.052***	0.135***	0.984***
	-7.8	-5.3	-11.45	-5.94	-8.28	-20.53	(-5.77)	-8.23	-19.44
Log(hft2)*DYr8,9,10	-0.083***			-0.121***			-0.455***		
	(-9.66)			(-13.62)			(-16.64)		
Log(hft2)*DYr11,12,13	-0.112***			-0.156***			-0.457***		
	(-13.8)			(-18.96)			(-17.24)		
Log(hft2)*DYr14,15,16	-0.145***			-0.219***			-0.583***		
	(-17.08)			(-24.93)			(-21.55)		
HHItrd*DYr11,12,13		0.01			-0.035***			-0.216***	
		-0.98			(-2.9)			(-10.12)	
HHItrd*DYr14,15,16		-0.035***			-0.096***			-0.296***	
		(-3.19)			(-7.56)			(-13.11)	
Log(hft2)*HHItrd			-0.026***			-0.062***			-0.215***
			(-6.24)			(-14.01)			(-17.85)
Log(hft2)*HHItrd*DYr11,12,13			-0.005***			-0.003*			0.013***
			(-3.07)			(-1.76)			-3.74
Log(hft2)*HHItrd*DYr14,15,16			-0.016***			-0.022***			-0.02***
			(-8.32)			(-10.58)			(-4.88)
Observations	173184	173184	173184	173184	173184	173184	173184	173184	173184
R-Square	0.84	0.84	0.84	0.81	0.81	0.81	0.76	0.75	0.76

PANEL B : Small stocks

	I	II	III	IV	V	VI	VII	VIII	IX
	Log(spread_bps)			Log(espread)			Log(depth1)		
	(4)	(5)	(6)	(4)	(5)	(6)	(4)	(5)	(6)
Log(hft2)	-0.292***	-0.238***	-0.251***	-0.303***	-0.253***	-0.271***	0.107***	-0.083***	0.057***
	(-45.63)	(-58.57)	(-30.11)	(-46.16)	(-60.7)	(-30.81)	-10.07	(-15.2)	-4.37
HHItrd	0.067***	0.019*	0.042***	0.077***	0.024**	0.045***	-0.062***	-0.146***	0.133***
	-18	-1.93	-3.52	-19.84	-2.37	-3.71	(-11.81)	(-12.73)	-7.3
Log(hft2)*DYr8,9,10	0.05***			0.049***			-0.291***		
	-6.93			-6.6			(-26.07)		
Log(hft2)*DYr11,12,13	0.074***			0.07***			-0.27***		
	-11.13			-10.22			(-24.08)		
Log(hft2)*DYr14,15,16	0.082***			0.073***			-0.116***		
	-11.61			-9.94			(-9.19)		
HHItrd*DYr11,12,13		0.058***			0.071***			0.018	
		-5.05			-5.98			-1.34	
HHItrd*DYr14,15,16		0.091***			0.089***			0.118***	
		-8.19			-7.84			-8.24	
Log(hft2)*HHItrd			-0.003			0.001			-0.098***
			(-0.74)			-0.32			(-14.21)
Log(hft2)*HHItrd*DYr11,12,13			0.015***			0.012***			0.017***
			-6.62			-5.36			-5.71
Log(hft2)*HHItrd*DYr14,15,16			0.018***			0.013***			0.061***
			-7.61			-5.44			-18.25
Observations	173184	173184	173184	173184	173184	173184	173184	173184	173184
R-Square	0.81	0.81	0.81	0.78	0.78	0.78	0.74	0.73	0.73

Table A.18: The time-varying effects of HFT and market fragmentation on liquidity: Large and small stocks (part-2/2)

This table presents the panel regression results of Models (2.4)–(2.6) for large and small stocks where various liquidity measures are regressed on HFT ($hft2$), market fragmentation ($HHItrd$) proxy. $hft2$ represents the per minute quote update for the best 5 depth levels in the limit order book. $HHItrd$ is the Herfindhal-Hirchman index (HHI), shows the degree of market fragmentation. The liquidity measures are accumulated average quoted depths up to best three limit price ($depth3$), volume weighted 5-minute realized-half spreads ($rspread\_5min$), and volume weighted 5-minute price impacts ($price\_impact\_5min$). All dependent variables are log transformed except $rspread\_5min$, all spreads based measures are in basis point and depth in GBP100. Panel A and Panel B show the result for large and small stocks respectively. $DYr8, 9, 10$, $DYr11, 12, 13$, $DYr14, 15, 16$ represent the period dummies for the three three-year blocks 2008–2010 , 201–2013 and 2014–2016, respectively. Control variables are log market capitalization ($Log(mktcap)$), log normalized intraday mid-price volatility ($Log(voltintra)$) and inverse of average daily price level ($invprice$). Models (4), (5), (6) examine the incremental effect of HFT, market fragmentation and their interaction in the latter three three-year block periods of the sample, respectively. The regression is based on a balanced panel of 66 stocks and 2624 days (December 2005–December 2016) for each group and have both time (daily) and stock fixed effects. Coefficient estimates are OLS, t-statistics shown in the parentheses below the coefficient, calculated using Newey-West (HAC) standard errors (lags are optimally determined). ***, **, * denote significance at 1%, 5%, and 10% levels, respectively.

PANEL A: Large stocks

	I	II	III	IV	V	VI	VII	VIII	IX
		Log(depth3)			rspread_5min		Log(price_impact_5min)		
	(4)	(5)	(6)	(4)	(5)	(6)	(4)	(5)	
Log(hft2)	-0.154***	-0.664***	-0.08**	-0.466***	-0.079***	-0.407***	-0.209***	-0.429***	-0.185***
	(-5.12)	(-48.23)	(-2.41)	(-8.27)	(-2.88)	(-5.65)	(-22.01)	(-74.51)	(-15.53)
HHItrd	0.05***	0.274***	1.167***	-0.211***	-0.294***	-0.874***	0.1***	0.179***	0.56***
	-4.87	-14.36	-22.04	(-9.49)	(-5.26)	(-8.18)	-15.64	-14.98	-27.07
Log(hft2)*DYr8,9,10	-0.515***			0.74***			-0.197***		
	(-18.34)			-13.38			(-22.2)		
Log(hft2)*DYr11,12,13	-0.467***			0.301***			-0.223***		
	(-17.15)			-6.39			(-24.6)		
Log(hft2)*DYr14,15,16	-0.614***			0.316***			-0.261***		
	(-22.11)			-6.75			(-28.51)		
HHItrd*DYr11,12,13		-0.261***			0.055			-0.067***	
		(-10.51)			-0.85			(-4.08)	
HHItrd*DYr14,15,16		-0.348***			0.158**			-0.15***	
		(-13.59)			-2.57			(-9.44)	
Log(hft2)*HHItrd			-0.239***			0.223***			-0.092***
			(-18.73)			-9.15			(-17.94)
Log(hft2)*HHItrd*DYr11,12,13			0.026***			-0.118***			0
			-6.03			(-12.61)			-0.14
Log(hft2)*HHItrd*DYr14,15,16			-0.013***			-0.115***			-0.013***
			(-2.74)			(-12.29)			(-5.23)
Observations	173184	173184	173184	173184	173184	173184	173184	173184	173184
R-Square	0.74	0.73	0.74	0.21	0.2	0.21	0.62	0.61	0.62

PANEL B: Small stocks

	I	II	III	IV	V	VI	VII	VIII	IX
		Log(depth3)			rspread_5min		Log(price_impact_5min)		
	(4)	(5)	(6)	(4)	(5)	(6)	(4)	(5)	
Log(hft2)	0.054***	-0.112***	0.027*	-5.339***	-2.433***	-6.39***	-0.086***	-0.152***	-0.047***
	-4.72	(-17.98)	-1.82	(-18.21)	(-16.42)	(-19.42)	(-10.71)	(-34.34)	(-4.78)
HHItrd	0.037***	-0.18***	0.21***	-0.945***	-2.722***	-7.876***	0.189***	0.216***	0.376***
	-5.85	(-13.41)	-10.11	(-13.22)	(-9.26)	(-19.71)	-37.37	-20.25	-25.51
Log(hft2)*DYr8,9,10	-0.312***			2.267***			-0.048***		
	(-26.23)			-6.79			(-5.77)		
Log(hft2)*DYr11,12,13	-0.249***			4.215***			-0.084***		
	(-20.1)			-15.77			(-9.87)		
Log(hft2)*DYr14,15,16	-0.005			4.892***			-0.131***		
	(-0.34)			-18.25			(-14.97)		
HHItrd*DYr11,12,13		0.136***			3.097***			-0.05***	
		-8.4			-10.03			(-3.82)	
HHItrd*DYr14,15,16		0.359***			2.847***			-0.049***	
		-21.31			-9.77			(-3.83)	
Log(hft2)*HHItrd			-0.12***			2.018***			-0.055***
			(-14.89)			-14.34			(-10.95)
Log(hft2)*HHItrd*DYr11,12,13			0.041***			0.218***			0
			-10.81			-3.98			(-0.1)
Log(hft2)*HHItrd*DYr14,15,16			0.11***			0.169***			-0.006**
			-27.47			-3			(-2.31)
Observations	173184	173184	173184	173184	173184	173184	173184	173184	173184
R-Square	0.71	0.7	0.7	0.24	0.21	0.24	0.54	0.54	0.54

Table A.19: The effects of HFT and market fragmentation on liquidity: a comparison of alternative HFT proxies

This table presents the panel regression results of Models (2.1)–(2.3) for the three alternative HFT proxies ($hft1$, $hft2$, $hft3$) where various liquidity measures are regressed on each HFT proxy and market fragmentation proxy ($HHItrd$). $hft1$, $hft2$, $hft3$ represent the per minute quote update for the best 10 depth levels in the limit order book, 5 depth levels in the limit order book, and at BBO respectively. $HHItrd$ is the Herfindhal-Hirchman index (HHI), shows the degree of market fragmentation. The liquidity measures are time weighted quoted spreads ($spread\_bps$), volume weighted effective-half spreads ($espread$), average quoted depths at best limit price ($depth1$), accumulated average quoted depths up to best three limit price ($depth3$), volume weighted 5-minute realized-half spreads ($rspread\_5min$), and volume weighted 5-minute price impacts ($price\_impact\_5min$). All dependent variables are log transformed except $rspread\_5min$, all spreads based measures are in basis point and depth in GBP100. Panel A1, A2, A3 present the proxies $hft1$, $hft2$, $hft3$ respectively and show the result for $spread\_bps$, $espread$ and $depth1$ and Panel B1, B2, B3 present the proxies $hft1$, $hft2$, $hft3$ respectively and show the result for the rest of the liquidity measures. Control variables are log market capitalization ($Log(mktcap)$), log normalized intraday mid-price volatility ($Log(voltintra)$) and inverse of average daily price level ($invprice$) are not shown, all are significant with expected sign. The regression is based on a balanced panel of 132 stocks and 2624 days (December 2005–December 2016) and have both time (daily) and stock fixed effects. Coefficient estimates are OLS , t-statistics shown in the parentheses below the coefficient, calculated using Newey-West (HAC) standard errors (lags are optimally determined). ***, **, * denote significance at 1%, 5%, and 10% levels, respectively.

PANEL A

	I	II	III	IV	V	VI	VII	VIII	IX
	Log(spread_bps)			Log(espread)			Log(depth1)		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
A1: hft1 (best 10 depth levels)									
Log(hft1)	-0.278***	-0.279***	-0.275***	-0.32***	-0.321***	-0.279***	-0.284***	-0.282***	-0.183***
	(-89.68)	(-90.04)	(-50.21)	(-98.95)	(-99.46)	(-48.42)	(-45.35)	(-45.25)	(-16.13)
HHItrd		0.049***	0.057***		0.061***	0.139***		-0.107***	0.078***
		(15.9)	(7.21)		(18.99)	(16.6)		(-19.94)	(4.76)
Log(hft1)*HHItrd			-0.002			-0.019***			-0.045***
			(-1.15)			(-10.19)			(-10.99)
R-Square	0.87	0.87	0.87	0.84	0.84	0.84	0.82	0.82	0.83
A2: hft2 (best 5 depth levels)									
Log(hft2)	-0.288***	-0.289***	-0.283***	-0.322***	-0.324***	-0.282***	-0.25***	-0.247***	-0.155***
	(-92.96)	(-93.21)	(-54.74)	(-99.98)	(-100.38)	(-51.56)	(-40.36)	(-40.14)	(-14.76)
HHItrd		0.052***	0.065***		0.065***	0.144***		-0.105***	0.068***
		-17.72	-9.09		-20.72	-18.81		(-19.51)	-4.5
Log(hft2)*HHItrd			-0.003*			-0.02***			-0.044***
			(-1.96)			(-11.24)			(-11.27)
R-Square	0.87	0.87	0.87	0.84	0.84	0.84	0.82	0.82	0.82
A3: hft3 (best price limit/BBO)									
Log(hft3)	-0.269***	-0.269***	-0.265***	-0.291***	-0.292***	-0.254***	-0.148***	-0.146***	-0.07***
	(-81.77)	(-81.87)	(-48.72)	(-84.2)	(-84.41)	(-42.99)	(-24.55)	(-24.45)	(-6.76)
HHItrd		0.047***	0.054***		0.059***	0.118***		-0.112***	0.007
		-15.35	-8.35		-17.78	-17.05		(-20.37)	-0.52
Log(hft3)*HHItrd			-0.002			-0.018***			-0.037***
			(-1.08)			(-9.68)			(-9.4)
R-Square	0.87	0.87	0.87	0.83	0.83	0.83	0.82	0.82	0.82

PANEL B

	I	II	III	IV	V	VI	VII	VIII	IX
	Log(depth3)			rspread_5min			price_impact_5min		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
B1: hft1 (best 10 depth levels)									
Log(hft1)	-0.295***	-0.295***	-0.285***	-1.662***	-1.657***	-4.25***	-1.601***	-1.61***	-3.217***
	(-43.27)	(-43.22)	(-22.7)	(-16.45)	(-16.43)	(-19.5)	(-28.06)	(-28.28)	(-27.31)
HHItrd		0.001	0.019		-0.273***	-5.116***		0.507***	-2.496***
		(0.21)	(1.05)		(-4.14)	(-20.25)		(9.08)	(-15.18)
Log(hft1)*HHItrd			-0.004			1.17***			0.725***
			(-0.98)			(19.86)			(20.9)
R-Square	0.81	0.81	0.81	0.16	0.16	0.18	0.54	0.54	0.54
B2: hft2 (best 5 depth levels)									
Log(hft2)	-0.268***	-0.268***	-0.264***	-1.871***	-1.866***	-4.251***	-1.988***	-2***	-3.615***
	(-39.84)	(-39.78)	(-22.66)	(-16.17)	(-16.18)	(-19.34)	(-30.64)	(-30.9)	(-29.57)
HHItrd		0.003	0.01		-0.245***	-4.711***		0.543***	-2.48***
		-0.52	-0.62		(-3.86)	(-20.12)		(10.11)	(-15.79)
Log(hft2)*HHItrd			-0.002			1.125***			0.762***
			(-0.43)			(20)			(22.07)
R-Square	0.81	0.81	0.81	0.16	0.16	0.19	0.54	0.54	0.55
B3: hft3 (best price limit/BBO)									
Log(hft3)	-0.184***	-0.184***	-0.188***	-2.184***	-2.18***	-4.901***	-1.779***	-1.788***	-3.053***
	(-27.66)	(-27.63)	(-16.19)	(-17.55)	(-17.57)	(-22.06)	(-26.39)	(-26.53)	(-24.68)
HHItrd		-0.004	-0.01		-0.253***	-4.505***		0.508***	-1.469***
		(-0.57)	(-0.74)		(-4.15)	(-23.38)		-9.94	(-11.33)
Log(hft3)*HHItrd			0.002			1.312***			0.61***
			-0.5			-23.87			-18.2
R-Square	0.81	0.81	0.81	0.21	0.21	0.25	0.56	0.56	0.57

Table A.20: The effects of HFT and market fragmentation on liquidity: a two-stage IV-GMM estimation

This table presents the second stage result of the two-stage optimal IV-GMM estimation of the Models (2.10)–(2.13) for time weighted quoted spreads ($spread\_bps$) and volume weighted effective half-spreads ($espread$). Sub columns (1)–(5) represent the equivalent IV-GMM estimation of Models (2.1)–(2.5) in OLS. The three suspected endogenous variables HFT_{it}, $MFrag_{it}$, and $HFT * MFrag_{it}$ are predicted by the three first-stage Models (2.7), (2.8), and (2.9) respectively. $hft2$ represents the per minute quote update for the best 5 depth levels in the limit order book. $HHItrd$ is the Herfindhal-Hirchman index (HHI), shows the degree of market fragmentation. The liquidity measures are time weighted quoted spreads ($spread\_bps$), and volume weighted effective-half spreads ($espread$). All dependent variables are log transformed and liquidity measures are in basis point. Control variables are log market capitalization ($Log(mktcap)$), log intraday mid price range volatility ($Log(voltintra)$), price inverse ($invprice$) and the average degree of liquidity of stocks in the same size group excluding stock i ($\overline{MQ}_{-it}$), calculated from four equally divided firm size group based on market capitalization. The regression is based on a balanced panel of 132 stocks and 2624 days (December 2005–December 2016), have both time (weekly time dummy for each of 591 weeks) and stock fixed effects. Coefficient estimates are IV-GMM, t-statistics shown in the parentheses below the coefficient, calculated using Newey-West (HAC) standard errors (based on 5 day lags). ***, **, * denote significance at 1%, 5%, and 10% levels, respectively.

	I	II	III	IV	VI	VII	VIII	IX	X	XI
			Log(spread_bps)					Log(espread)		
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Log(hft2)	-0.442***	-0.413***	-0.456***	-0.402***	-0.407***	-0.425***	-0.418***	-0.457***	-0.409***	-0.415***
	(-127.12)	(-116.04)	(-78.56)	(-64.16)	(-112.43)	(-115.36)	(-111.42)	(-79.64)	(-66.25)	(-105.87)
HHItrd		0.254***	0.147***	0.27***	0.161***		0.085***	-0.019	0.077***	0.056***
		(25.78)	(10.86)		-8.82		(8.24)	(-1.42)	-6.67	-2.86
Log(hft2)*HHItrd			0.016***					0.015***		
			(7.72)					(7.7)		
Log(hft2)*DYr8,9,10				-23.86					-0.017***	
				-0.037***					(-3)	
Log(hft2)*DYr11,12,13				(-6.42)					0.018***	
				0.003					-3.46	
Log(hft2)*DYr14,15,16				-0.47					-0.031***	
				-0.012**					(-5.99)	
HHItrd*DYr11,12,13				(-2.31)	0.07**					0.01
					-2.57					-0.34
HHItrd*DYr14,15,16					0.269***					0.101***
					-8.14					-2.78
$Log(\overline{MQ}_{-it})$	0.367***	0.391***	0.38***	0.365***	0.338***	0.249***	0.252***	0.25***	0.262***	0.235***
	(42.53)	(45.56)	(49.48)	-39.46	-30.4	(24.85)	(25.16)	(28.81)	-25.1	-18.38
Log(mktcap)	-0.117***	-0.152***	-0.137***	-0.148***	-0.152***	-0.09***	-0.101***	-0.087***	-0.107***	-0.101***
	(-29.98)	(-37.23)	(-37.71)	(-33.94)	(-36.72)	(-21.66)	(-23.06)	(-22.09)	(-23.05)	(-22.88)
Log(voltintra)	0.225***	0.232***	0.232***	0.233***	0.237***	0.266***	0.27***	0.268***	0.269***	0.271***
	(78.56)	(82.4)	(91.63)	-84.82	-82.24	(82.51)	(82.01)	(88.38)	-83.03	-82.62
Inv(price)	9.968***	10.507***	9.655***	9.973***	10.212***	9.588***	9.705***	8.772***	8.587***	9.633***
	(14.26)	(15.01)	(16.71)	-14.05	-14.36	(11.72)	(11.85)	(12.41)	-10.48	-11.69
stock fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
time fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	346368	346368	346368	346368	346368	346368	346368	346368	346368	346368
adj_Rsqr	0.85	0.85	0.85	0.85	0.85	0.81	0.82	0.81	0.82	0.82

Table A.21: The effects of HFT and market fragmentation on liquidity: a two-stage IV-GMM estimation for large and small stocks

This table presents the second stage result of the two-stage optimal IV-GMM estimation of the Models (2.10)–(2.13) employing large (Panel A) and small (Panel B) stocks for time weighted quoted spreads ($spread\_bps$) and volume weighted effective half-spreads ($espread$). Sub columns (1)–(5) represent the equivalent IV-GMM estimation of Models (2.1)–(2.5) in OLS. The three suspected endogenous variables HFT_{it}, $MFrag_{it}$, and $HFT * MFrag_{it}$ are predicted by the three first-stage equations (2.7), (2.8), and (2.9) respectively. $hft2$ represents the per minute quote update for the best 5 depth levels in the limit order book. $HHItrd$ is the Herfindhal-Hirchman index (HHI), shows the degree of market fragmentation. The liquidity measures are time weighted quoted spread ($spread\_bps$), and volume weighted effective-half spread ($espread$). All dependent variables are log transformed and liquidity measures are in basis point. Control variables are log market capitalization ($Log(mktcap)$), log intraday mid price range volatility ($Log(voltintra)$), price inverse ($invprice$) and the average degree of liquidity of stocks in the same size group excluding stock i ($\overline{MQ}_{-it}$), calculated from four equally divided firm size group based on market capitalization. Estimates on the coefficient of control variables are not shown, all are significant, and have the expected sign. The regression is based on a balanced panel of 66 stocks and 2624 days (December 2005–December 2016) for each group of stock, have both time (weekly time dummy for each of 591 weeks) and stock fixed effects. Coefficient estimates are IV-GMM, t-statistics shown in the parentheses below the coefficient , calculated using Newey-West (HAC) standard errors (based on 5 day lags). ***, **, * denote significance at 1%, 5%, and 10% levels, respectively.

Panel A : Large stocks

	I	II	III	IV	VI	VII	VIII	IX	X	XI
	Log(spread_bps)					Log(espread)				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Log(hft2)	-0.408***	-0.398***	-0.293***	-0.312***	-0.409***	-0.376***	-0.406***	-0.201***	-0.282***	-0.437***
	(-85.82)	(-80.29)	(-25.19)	(-29.39)	(-69.08)	(-71.27)	(-70.1)	(-16.18)	(-23.9)	(-55.45)
HHItrd		0.086***	0.31***	0.108***	0.231***		-0.155***	0.289***	-0.121***	0.242***
		(6.14)	(12.66)	-7.62	-8.37		(-9.07)	(10.43)	(-7.31)	-7.21
Log(hft2)*HHItrd			-0.039***					-0.08***		
			(-8.93)					(-16.68)		
Log(hft2)*DYr8,9,10				-0.082***					-0.096***	
				(-7.28)					(-7.59)	
Log(hft2)*DYr11,12,13				-0.061***					-0.082***	
				(-5.87)					(-7.13)	
Log(hft2)*DYr14,15,16				-0.126***					-0.203***	
				(-11.68)					(-16.95)	
HHItrd*DYr11,12,13					0.021					-0.129**
					-0.41					(-2.02)
HHItrd*DYr14,15,16					-0.612***					-1.502***
					(-9.66)					(-17.29)
Observations	173184	173184	173184	173184	173184	173184	173184	173184	173184	173184
adj_Rsqr	0.81	0.81	0.81	0.81	0.79	0.77	0.77	0.78	0.78	0.65

Panel B : Small stocks

	I	II	III	IV	VI	VII	VIII	IX	X	XI
	Log(spread_bps)					Log(espread)				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Log(hft2)	-0.507***	-0.448***	-0.508***	-0.464***	-0.459***	-0.49***	-0.475***	-0.591***	-0.469***	-0.464***
	(-105.95)	(-86.75)	(-42.39)	(-35.72)	(-88.45)	(-98.78)	(-90.09)	(-52.8)	(-39.92)	(-77.82)
HHItrd		0.378***	0.226***	0.321***	-0.423***		0.086***	-0.172***	0.047***	-1.109***
		(25.33)	(8.86)	-21.13	(-6.98)		(5.76)	(-7.04)	-3.08	(-14.58)
Log(hft2)*HHItrd			0.025***					0.059***		
			(4.68)					(11.52)		
Log(hft2)*DYr8,9,10				-0.042***					-0.086***	
				(-2.77)					(-6.21)	
Log(hft2)*DYr11,12,13				0.041***					0.034***	
				-2.88					-2.67	
Log(hft2)*DYr14,15,16				0.014					0.029**	
				-1	1.002***				-2.19	1.626***
HHItrd*DYr11,12,13					-11.84					-15.49
					1.082***					1.498***
HHItrd*DYr14,15,16					-14.29					-16.65
Observations	173184	173184	173184	173184	173184	173184	173184	173184	173184	173184
adj_Rsqr	0.74	0.73	0.74	0.74	0.67	0.71	0.72	0.71	0.71	0.57

Table A.22: The effects of HFT and market fragmentation on liquidity: a simultaneous equations model estimation

This table presents the simultaneous equations model estimation of the equations (2.13)–(2.14) using both GMM (H3SLS) and 3SLS estimations for time weighted quoted spreads ($spread\_bps$) and volume weighted effective half-spreads ($espread$). Indices i and t represent stocks and day respectively, MQ_{it} represents one of the two log normalized market quality (liquidity) measures ($spread\_bps$, $espread$), HFT_{it} represents the HFT proxy ($hft2$), $MFrag_{it}$ represents the market fragmentation proxy ($\underline{HHItrd}$), $\overline{MQ}_{-it}$ represents the average market liquidity level over all stocks in the same size group excluding stock i, $\overline{MFrag}_{-it}$ represents the average market fragmentation level over all stocks in the same size group excluding stock i, $\overline{HFT}_{-it}$ represents the average HFT intensity over all stocks in the same size group excluding stock i, $Log(mktcap)$ is the log normalized value of market capitalization, $Log(voltintra)$ is the log normalized value of intraday mid price range volatility, $invprice$ is the inverse of daily average price, $Log(size)$ is the log normalized average value of trade size , $Log(value)$ is the log normalized value of trading volume, indices $(_{(mq)})$, $(_{(hft)})$, $(_{(frg)})$ refer the respective coefficient of the equations MQ_{it}/(2.13), HFT_{it}/(2.14) and $MFrag_{it}$/(2.15) respectively. The regression is based on a balanced panel of 132 stocks and 2240 days (January 2008–December 2016), have both time (monthly time dummy for each of the 108 months) and stock fixed effects. Coefficient estimates are GMM (H3SLS) and 3SLS, t-statistics shown in the parentheses below the coefficient . ***, **, * denote significance at 1%, 5%, and 10% levels, respectively.

	$spread\_bps$						$espread$					
	H3SLS (GMM)			3SLS			H3SLS (GMM)			3SLS		
	$Log(MQ)_{it}$	$Log(HFT)_{it}$	$Mfrag_{it}$	$Log(MQ)_{it}$	$Log(HFT)_{it}$	$Mfrag_{it}$	$Log(MQ)_{it}$	$Log(HFT)_{it}$	$Mfrag_{it}$	$Log(MQ)_{it}$	$Log(HFT)_{it}$	$Mfrag_{it}$
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
$Log(HFT)_{it}$	-0.317***			-0.376***			-0.311***			-0.384***		
	(-87.92)			(-249.1)			(-80.6)			(-242.44)		
$MFrag_{it}$	0.189***			0.231***			0.073***			0.063***		
	-22.9			-56.24			-8.35			-14.54		
$log(\overline{MQ})_{-it}$	0.45***			0.372***			0.311***			0.239***		
	-54.94			-100.48			-31.2			-51.98		
$Log(mktcap)_{it}$	-0.187***			-0.182***			-0.143***			-0.118***		
	(-44.18)			(-104.03)			(-30.2)			(-63.84)		
$Log(volintra)_{it}$	0.183***			0.185***			0.219***			0.208***		
	-64.82			-170.98			-65.67			-182.68		
$inv(price)_{it}$	14.086***			14.979***			16.373***			18.568***		
	-17.36			-67.92			-17.13			-80.84		
$Log(MQ)_{it}$		0.726***			0.685***			1.253***			1.378***	
		-21.87			-47.3			-15.69			-39.02	
$MFrag_{it}$		0.149***			0.106***			0.138***			0.113***	
		-8.14			-10.61			-6.13			-9.45	
$Log(\overline{HFT)_{-it}}$		0.429***			0.463***			0.371***			0.51***	
		-48.19			-113.05			-26.12			-82.05	
$Log(mktcap)_{it}$		0.296***			0.317***			0.405***			0.458***	
		-26.43			-69.65			-20.03			-53.06	
rtk_{it}		-462.761***			-386.573***			-712.829***			-582.874***	
		(-29.26)			(-115.61)			(-23.14)			(-55.95)	
$Log(size)_{it}$		-0.762***			-0.774***			-0.949***			-0.954***	
		(-44.64)			(-120.39)			(-27.92)			(-67.74)	
$Log(volume)_{it}$		0.702***			0.691***			0.769***			0.797***	
		-58.18			-136.6			-37.48			-90.91	
$Log(volintra)_{it}$		-0.049***			-0.043***			-0.142***			-0.172***	
		(-6.85)			(-14.19)			(-9.01)			(-26.01)	
$Log(MQ)_{it}$			0.01			0.248***			0.11***			0.241***
			-1.04			-51.29			-12.03			-53.1
$Log(HFT)_{it}$			0.209***			0.26***			0.246***			0.284***
			-41.53			-85.6			-46.27			-85.4
$\overline{MFrag_{-it}}$			0.629***			0.613***			0.627***			0.624***
			-71.1			-124.25			-70.69			-123.35
$Log(mktcap)_{it}$			0.216***			0.262***			0.238***			0.261***
			-50.01			-126.24			-55.27			-126.12
$Log(volume)_{it}$			-0.242***			-0.218***			-0.249***			-0.242***
			(-81.98)			(-146.77)			(-86.01)			(-156.93)
$Log(volintra)_{it}$			-0.005*			-0.049***			-0.024***			-0.051***
			(-1.83)			(-31.54)			(-8.65)			(-32.45)
observations	295680	295680	295680	295680	295680	295680	295680	295680	295680	295680	295680	295680
adjrsq	0.87			0.86			0.83			0.83		
adjrsq		0.91			0.91			0.85			0.83	
adjrsq			0.78			0.77			0.78			0.78

Table A.23: The effects of HFT and market fragmentation on liquidity (large and small stocks): a simultaneous equations model estimation

This table presents the simultaneous equations model estimation of the equations (2.13)–(2.14) for large and small stocks using GMM(H3SLS)estimation for log normalized time weighted quoted spreads ($spread\_bps$) and volume weighted effective half-spreads ($espread$). Indices i and t represent stocks and day respectively, MQ_{it} represents one of the two market quality (liquidity) measures ($spread\_bps, espread$), HFT_{it} represents HFT proxy ($hft2$), $MFrag_{it}$ represents market fragmentation proxy ($HHItrd$), $\overline{MQ}_{-it}$ represents average market liquidity level over all stocks in the same size group excluding stock i, $\overline{MFrag}_{-it}$ represents average market fragmentation level over all stocks in the same size group excluding stock i, $\overline{HFT}_{-it}$ represents average HFT intensity over all stocks in the same size group excluding stock i, $Log(mktcap)$ is the log normalized value of market capitalization, $Log(voltintra)$ is the log normalized value of intraday mid price range volatility, $invprice$ is the inverse of daily average price, $Log(size)$ is the log normalized average value of trade size , $Log(value)$ is the log normalized value of trading volume , indices $_{(mq)}$, $(_{(hft)})$, $_{(frg)}$ refer the respective estimates of the equations MQ_{it}/(2.13), HFT_{it}/(2.14) and $MFrag_{it}$/(2.15) respectively. The regression is based on a balanced panel of 6 stocks and 2240 days (January 2008–December 2016) for each group of stock, have both time (monthly time dummy for each of 108 months) and stock fixed effects. Coefficient estimates are GMM (H3SLS), t-statistics shown in the parentheses below the coefficient . ***, **, * denote significance at 1%, 5%, and 10% levels, respectively.

	$spread\_bps$						$espread$					
	LARGE			SMALL			LARGE			SMALL		
	$Log(MQ)_{it}$	$Log(HFT)_{it}$	$Mfrag_{it}$	$Log(MQ)_{it}$	$Log(HFT)_{it}$	$Mfrag_{it}$	$Log(MQ)_{it}$	$Log(HFT)_{it}$	$Mfrag_{it}$	$Log(MQ)_{it}$	$Log(HFT)_{it}$	$Mfrag_{it}$
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
$Log(HFT)_{it}$	-0.277***			-0.379***			-0.266***			-0.382***		
	(-60.69)			(-71.68)			(-51.04)			(-68.43)		
$MFrag_{it}$	0.081***			0.268***			-0.074***			0.101***		
	-7.31			-22.35			(-5.92)			-8.19		
$Log(\overline{MQ})_{-it}$	0.585***			0.365***			0.333***			0.316***		
	-41.71			-31.72			-16.35			-25.97		
$Log(mktcap)_{it}$	-0.189***			-0.176***			-0.141***			-0.119***		
	(-35.43)			(-27.3)			(-23.25)			(-16.9)		
$Log(volintra)_{it}$	0.13***			0.239***			0.161***			0.284***		
	-35.56			-73.69			-35.27			-79.64		
$inv(price)_{it}$	13.963***			10.738***			12.239***			20.848***		
	-13.42			-10.32			-10.98			-15.82		
$Log(MQ)_{it}$		0.355***			0.535***			0.563***			0.689***	
		-12.33			-10.64			-8.27			-9.17	
$MFrag_{it}$		0.094***			0.325***			0.097***			0.36***	
		-5.29			-9.08			-4.78			-10.41	
$Log(\overline{HFT})_{-it}$		0.532***			0.325***			0.523***			0.281***	
		-64.65			-22.82			-48.58			-14.94	
$Log(mktcap)_{it}$		0.335***			0.202***			0.406***			0.219***	
		-28.89			-12.52			-17.8			-12.4	
rtk_{it}		-292.35***			-484.626***			-409.578***			-597.873***	
		(-13.84)			(-32.05)			(-10.88)			(-25.72)	
$Log(size)_{it}$		-0.89***			-0.525***			-1.009***			-0.541***	
		(-44.52)			(-26.5)			(-25.1)			(-23.59)	
$Log(volume)_{it}$		0.577***			0.661***			0.585***			0.666***	
		-56.94			-38.08			-40.42			-33.15	
$Log(volintra)_{it}$		-0.01			0.018			-0.029***			-0.022	
		(-1.54)			-1.64			(-2.81)			(-1.28)	
$Log(MQ)_{it}$			0.113***			0.063***			0.141***			0.158***
			-12.3			-3.78			-15.77			-10.75
$Log(HFT)_{it}$			0.186***			0.272***			0.214***			0.306***
			-33.78			-32.31			-36.38			-35.43
$\overline{MFrag_{-it}}$			0.661***			0.373***			0.653***			0.372***
			-57.33			-29.69			-56.71			-29.33
$Log(mktcap)_{it}$			0.219***			0.194***			0.23***			0.216***
			-43.02			-30.62			-44.74			-33.55
$Log(volume)_{it}$			-0.228***			-0.255***			-0.249***			-0.259***
			(-62.03)			(-60.57)			(-65.82)			(-66.41)
$Log(volintra)_{it}$			-0.016***			-0.025***			-0.015***			-0.051***
			(-4.49)			(-5.06)			(-4.37)			(-10.74)
observations	147840	147840	147840	147840	147840	147840	147840	147840	147840	147840	147840	147840
adjrsq	0.82			0.77			0.78			0.74		
adjrsq		0.9			0.82			0.88			0.8	
adjrsq			0.81			0.76			0.81			0.76

Table A.24: The time-varying effects of HFT and market fragmentation on liquidity: a simultaneous equations model estimation

This table presents the simultaneous equations model estimation of the equations (2.13)–(2.14) using GMM (H3SLS) estimation for the sub periods (2008-2010), (2011-2013) and (2014-2016) for log normalized time weighted quoted spreads ($spread\_bps$) and volume weighted effective half-spreads ($espread$). Indices i and t represent stocks and days respectively, MQ_{it} represents one of the two market quality (liquidity) measures ($spread\_bps, espread$), HFT_{it} represents HFT proxy ($hft2$) , $MFrag_{it}$ represents market fragmentation proxy ($HHItrd$), $\overline{MQ}_{-it}$ represents average market liquidity level over all stocks in the same size group excluding stock i, $\overline{MFrag}_{-it}$ represents average market fragmentation level over all stocks in the same size group excluding stock i, $\overline{HFT}_{-it}$ represents average HFT intensity over all stocks in the same size group excluding stock i, $Log(mktcap)$ is the log normalized value of market capitalization, $Log(voltintra)$ is the log normalized value of intraday mid price range volatility, $invprice$ is the inverse of daily average price, $Log(size)$ is the log normalized average value of trade size , $Log(value)$ is the log normalized value of trading volume , indices ($_{mq}$), ($_{hft}$), ($_{frg}$) refer the respective estimates of the equations MQ_{it}/(2.13), HFT_{it}/(2.14) and $MFrag_{it}$/(2.15) respectively. The regression is based on a balanced panel of 132 stocks and 2240 days (January 2008–December 2016), have both time (monthly time dummy for each of the 36 months in each sub period) and stock fixed effects. Coefficient estimates are GMM (H3SLS), t-statistics shown in the parentheses below the coefficient . ***, **, * denote significance at 1%, 5%, and 10% levels, respectively.

$effective \quad half-spread$

	2008-2010			2011-2013			2014-2016		
	$Log(MQ)_{it}$	$Log(HFT)_{it}$	$Mfrag_{it}$	$Log(MQ)_{it}$	$Log(HFT)_{it}$	$Mfrag_{it}$	$Log(MQ)_{it}$	$Log(HFT)_{it}$	$Mfrag_{it}$
	I	II	III	IV	V	VI	VII	VIII	IX
$Log(HFT)_{it}$	-0.251***			-0.253***			-0.221***		
	(-47.54)			(-45.09)			(-31.81)		
$MFrag_{it}$	0.239***			0.095***			0.01		
	-18.62			-9.94			-0.95		
$Log(\overline{MQ})_{-it}$	0.391***			0.477***			0.486***		
	-30.13			-31.7			-28.47		
$Log(mktcap)_{it}$	-0.222***			-0.25***			-0.231***		
	(-25.81)			(-28.15)			(-22.63)		
$Log(volintra)_{it}$	0.204***			0.175***			0.151***		
	-41.98			-56.27			-44.7		
$inv(price)_{it}$	19.067***			6.573***			46.527***		
	-16.63			-9			-16.81		
$Log(MQ)_{it}$		0.297***			0.297***			0.39***	
		-6.36			-6.46			-10.47	
$MFrag_{it}$		0.206***			0.304***			0.249***	
		-9.85			-10.57			-11.34	
$Log(\overline{HFT})_{-it}$		0.514***			0.587***			0.457***	
		-46.12			-55.7			-52.42	
$Log(mktcap)_{it}$		0.237***			0.175***			0.18***	
		-14.22			-10.42			-10.48	
rtk_{it}		-221.216***			-1006.162***			-875.426***	
		(-17.64)			(-30.09)			(-24.12)	
$Log(size)_{it}$		-0.618***			-0.314***			-0.517***	
		(-33.04)			(-24.33)			(-37.47)	
$Log(volume)_{it}$		0.527***			0.452***			0.552***	
		-42.38			-45.08			-66.02	
$Log(volintra)_{it}$		0.053***			0.041***			0.001	
		-5.87			-5.09			-0.15	
$Log(MQ)_{it}$			0.199***			0.105***			0.17***
			-15.88			-6.89			-13.73
$Log(HFT)_{it}$			0.212***			0.273***			0.382***
			-29.38			-35.21			-40.26
$\overline{MFrag}_{-it}$			0.685***			0.544***			0.557***
			-59.75			-36.63			-36.51
$Log(mktcap)_{it}$			0.2***			0.219***			0.357***
			-23.71			-19.94			-33.92
$Log(volume)_{it}$			-0.185***			-0.3***			-0.376***
			(-47.39)			(-75.46)			(-74.72)
$Log(volintra)_{it}$			-0.054***			0.01**			-0.011***
			(-12.48)			-2.53			(-3.28)
observations	97284	97284	97284	99528	99528	99528	98868	98868	98868
adjrsq	0.85			0.84			0.83		
adjrsq		0.92			0.93			0.95	
adjrsq			0.83			0.44			0.46

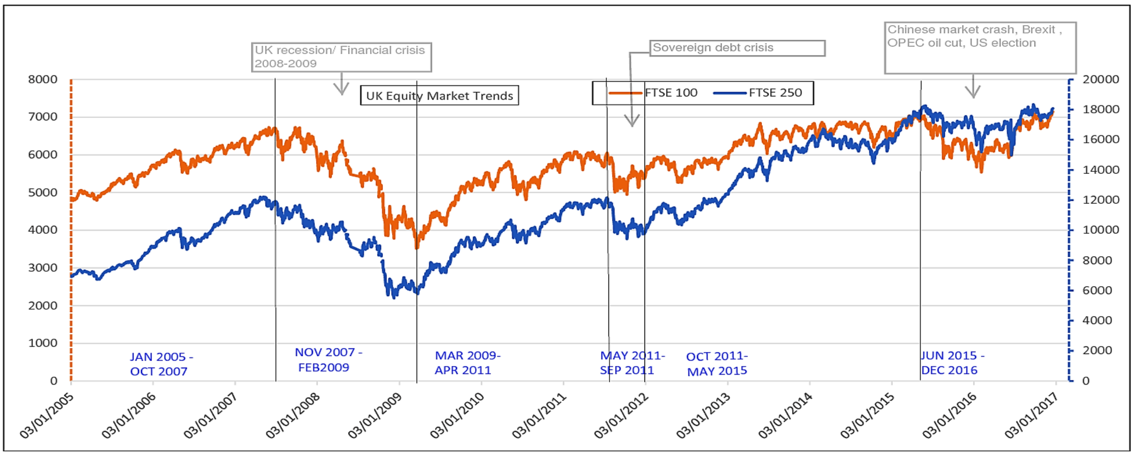
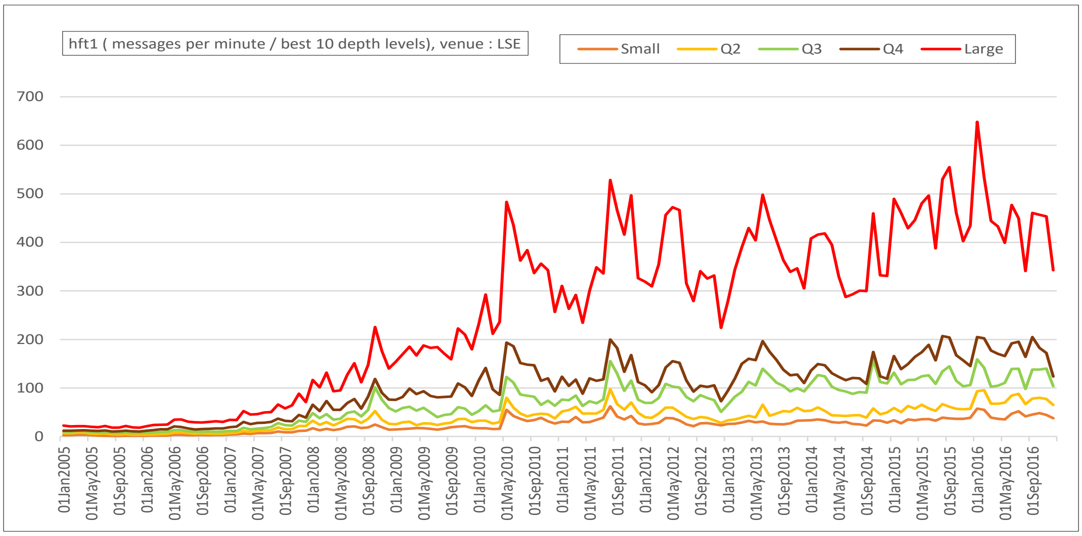
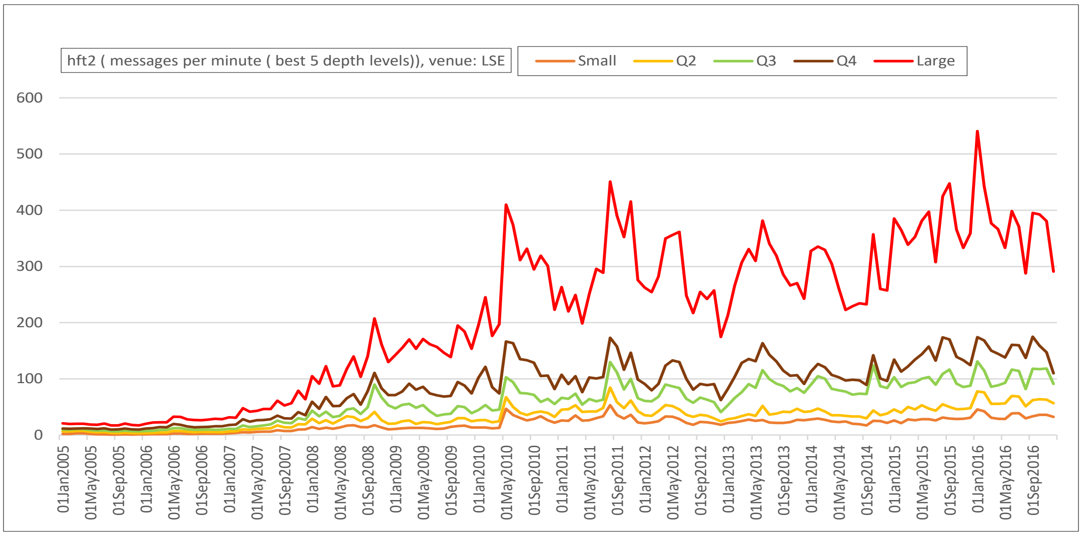


Fig. A.1: Market trends: 2005–2016

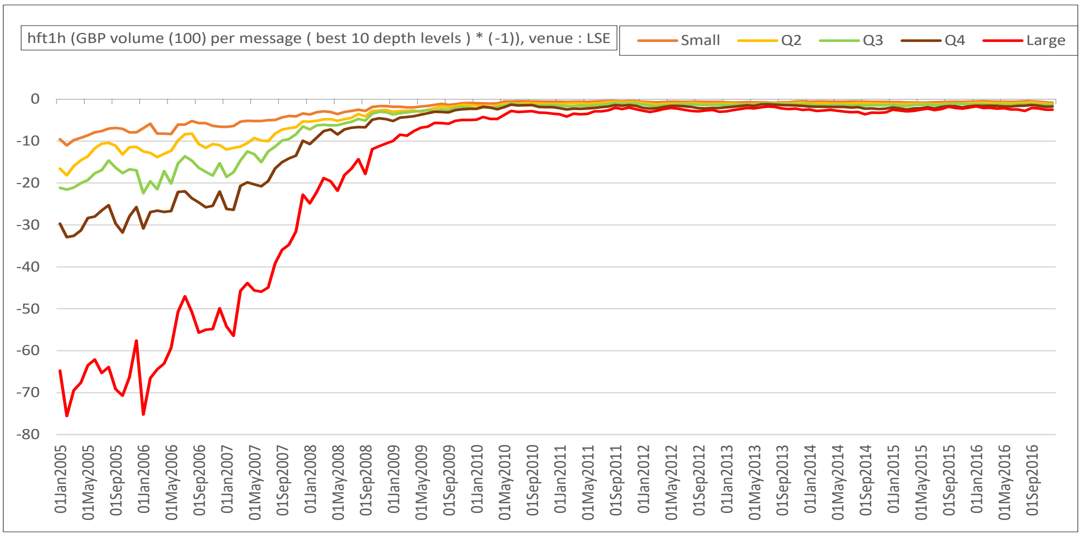


(a) HFT proxy: $hft1$ (10 depth levels)

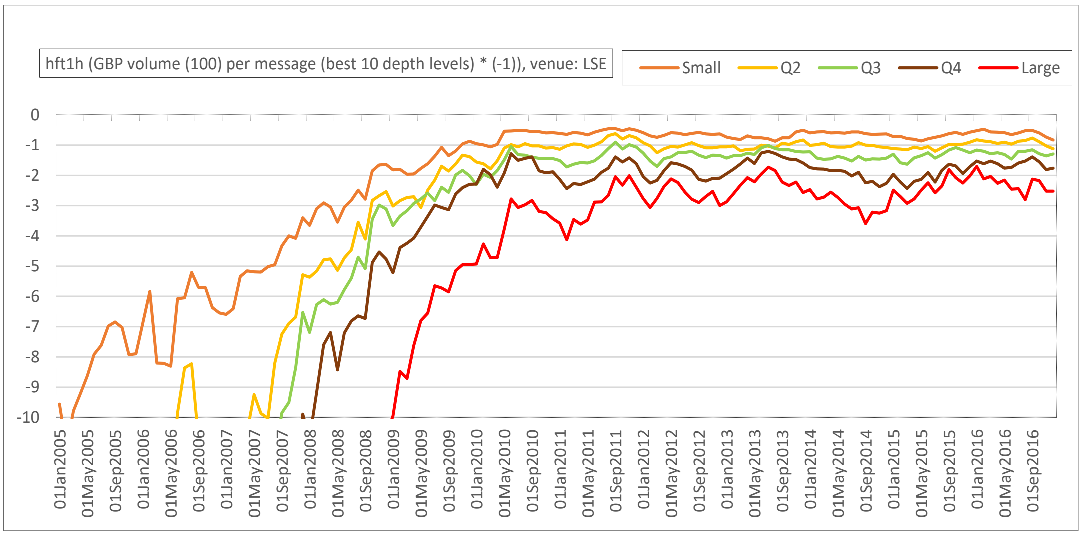


(b) HFT proxy: $hft2$ (5 depth levels)

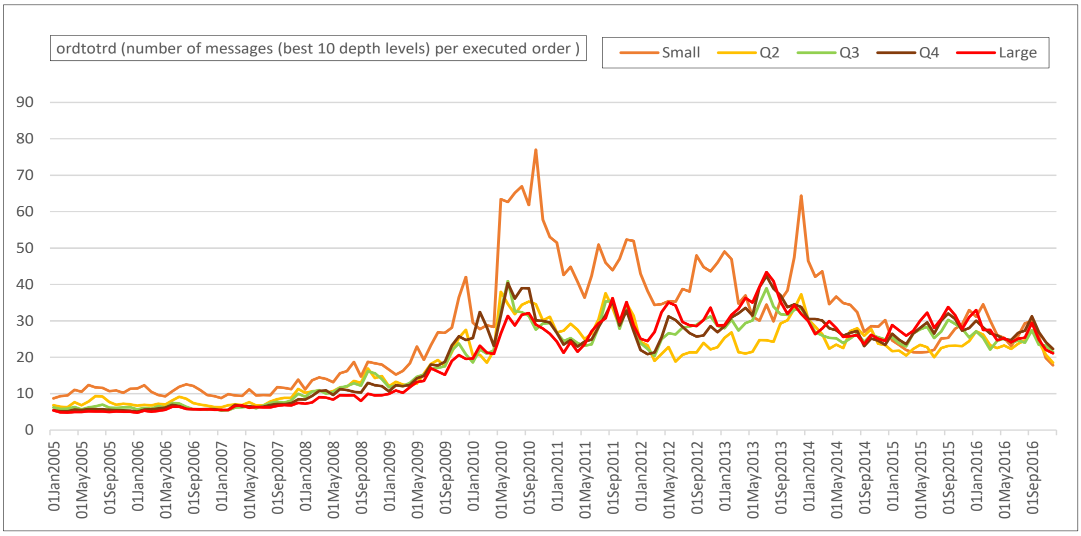
Fig. A.2: HFT proxies: electronic message rate



(a) HFT proxies: $hft1h$ (Hendershott et al., 2011)

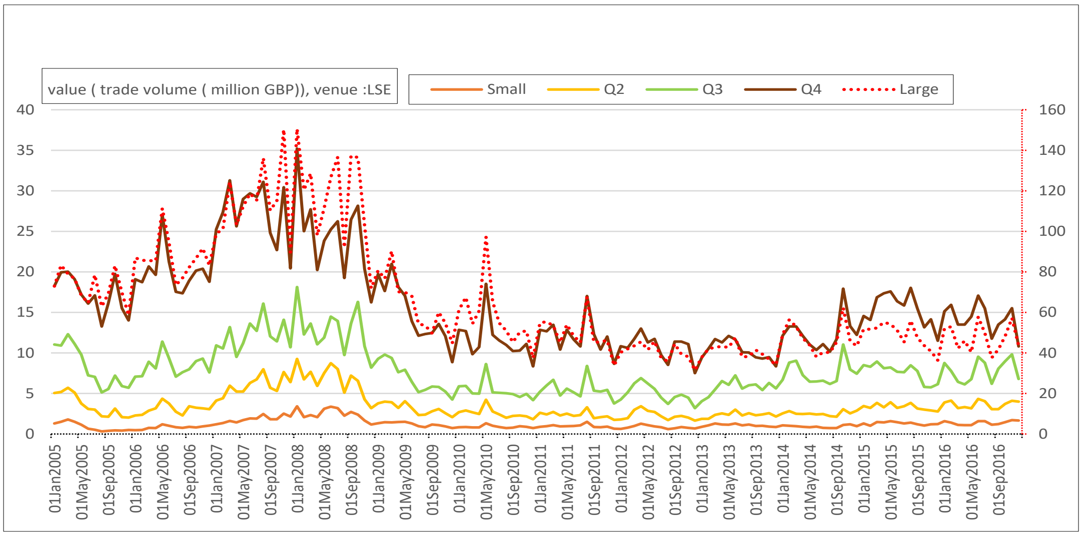


(b) HFT proxies: $hft1h$ (zoomed in)

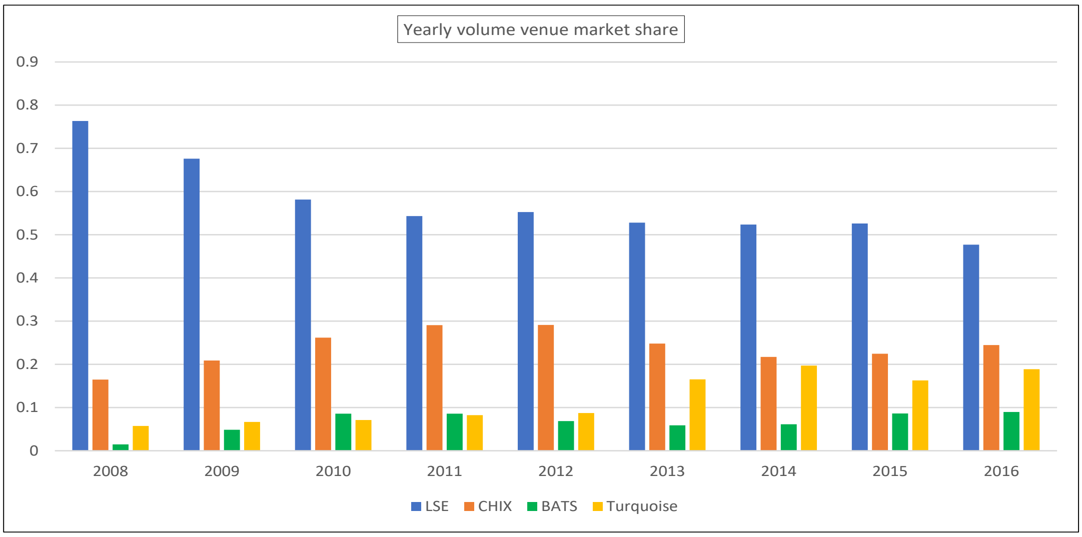


(c) HFT proxies: $ordtotrd$

Fig. A.3: HFT proxies: Hendershott et al. (2011)'s measure and order to trade ratio

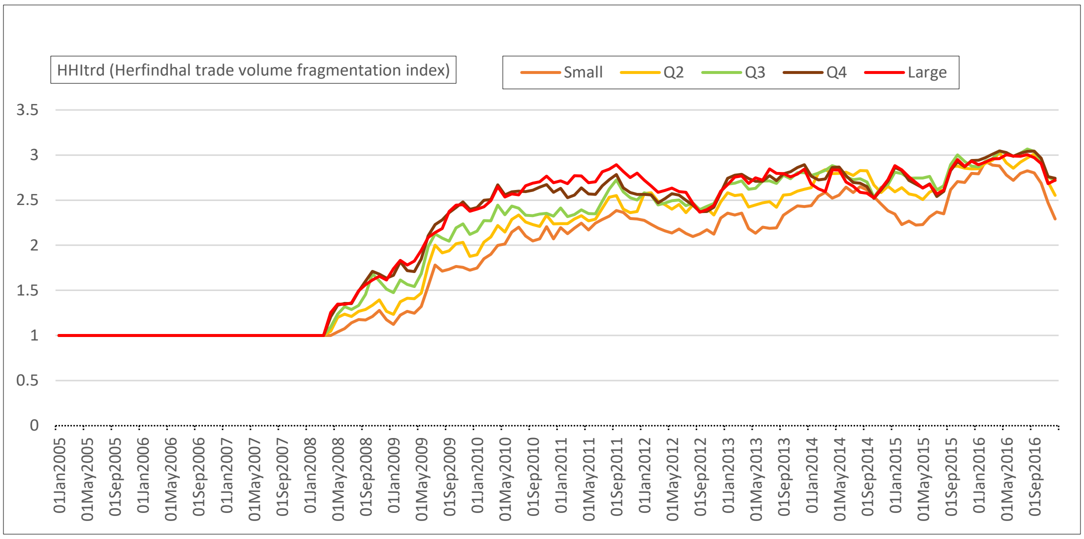


(a) Trends in trading volumes across quintiles

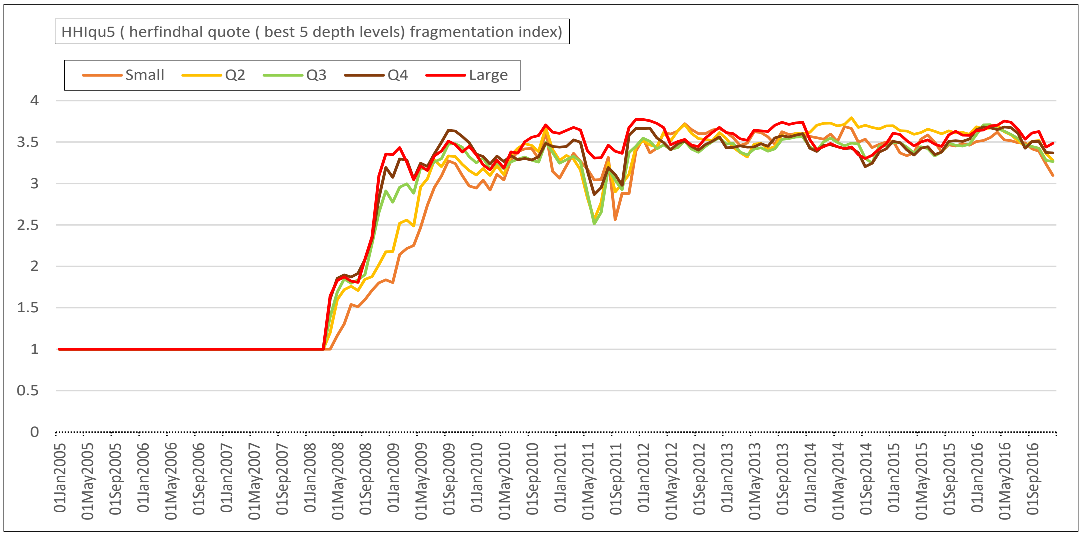


(b) Trading volumes' market share across exchanges

Fig. A.4: LSE listed stocks: Trends in trading volumes and venue market share

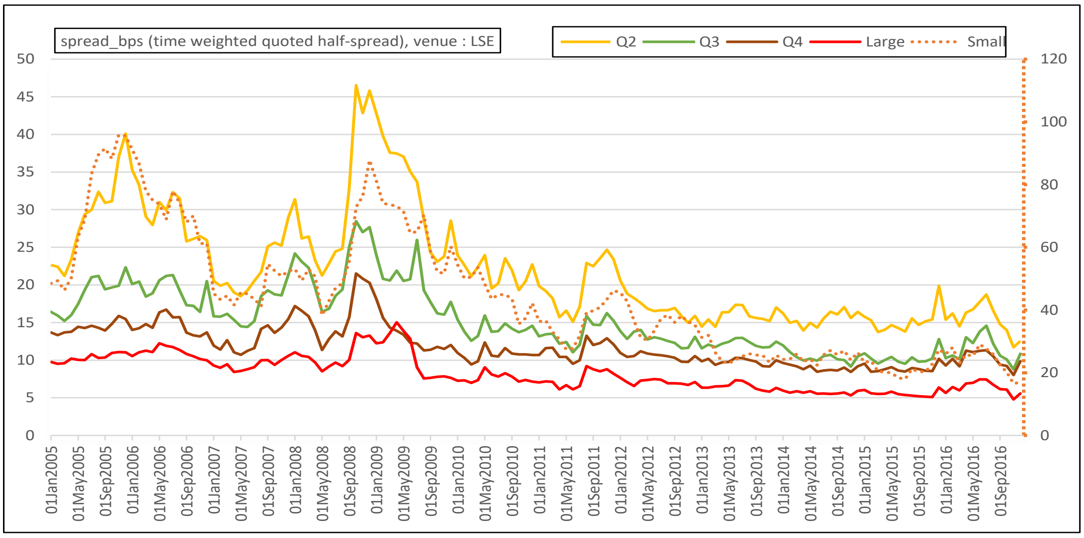


(a) The volume fragmentation proxy ($HHItrd$)

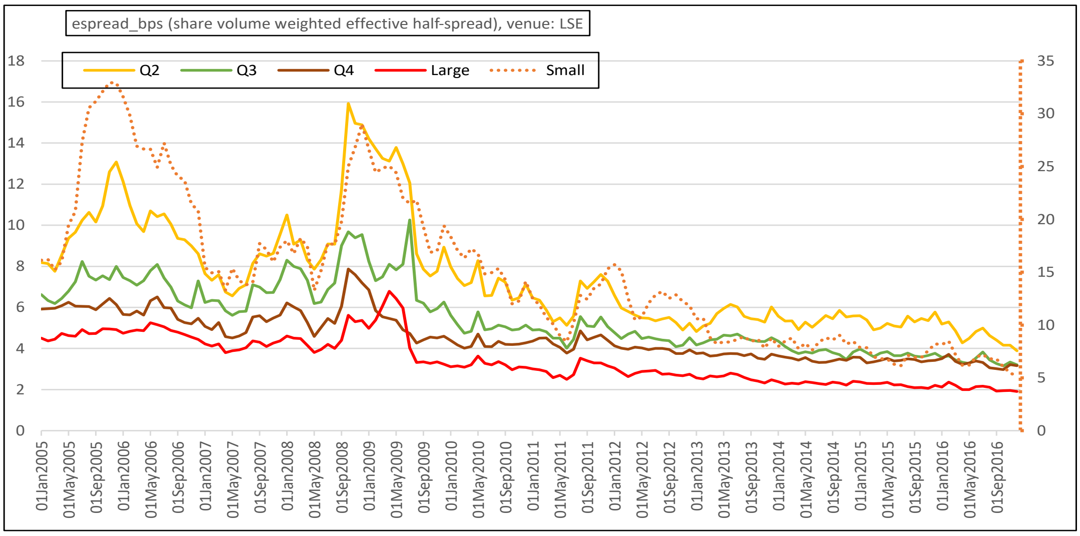


(b) The quote fragmentation proxy ($HHIqu5$)

Fig. A.5: Trends in market fragmentation proxies

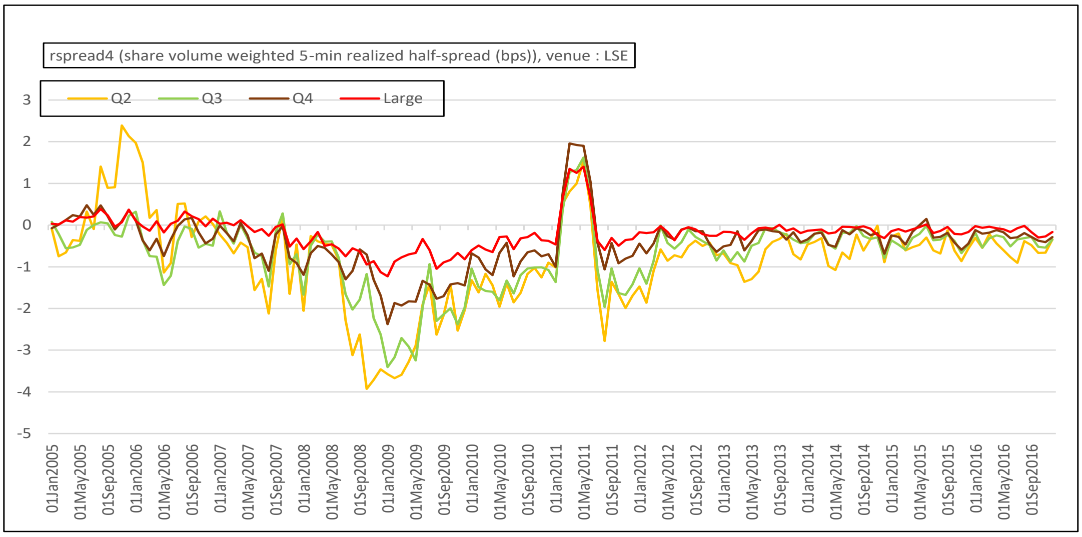


(a) Time weighted quoted spreads ($spread\_bps$)

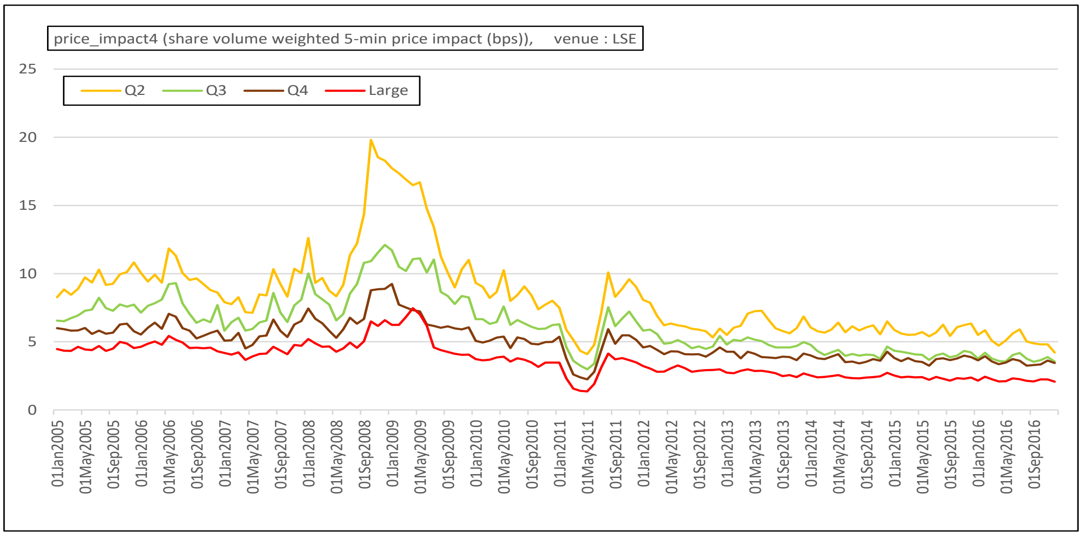


(b) Volume weighted effective spreads ($espread$)

Fig. A.6: Trends in average quoted spreads and effective half-spreads

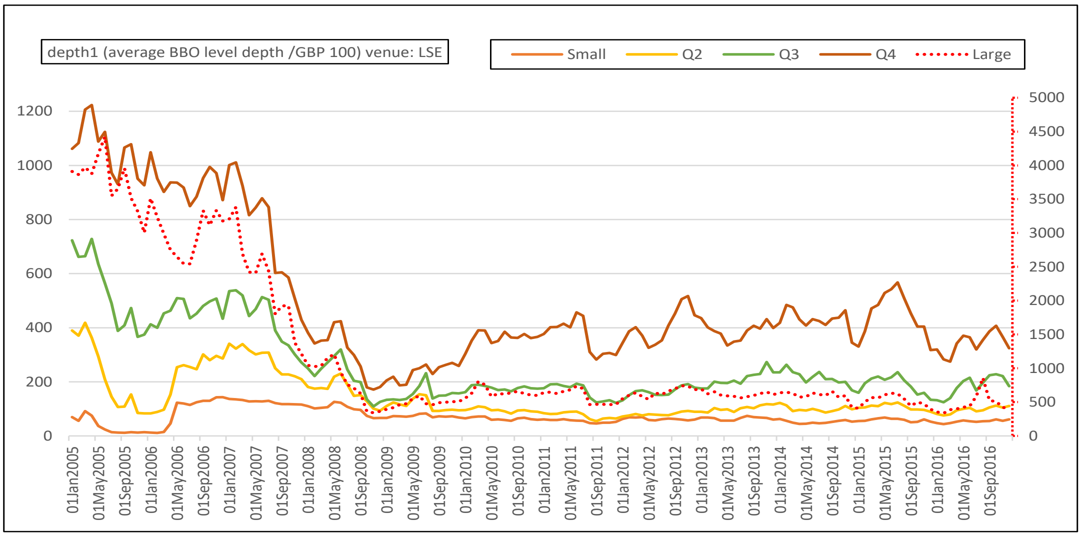


(a) 5-minute realized half-spreads ($rspread$)

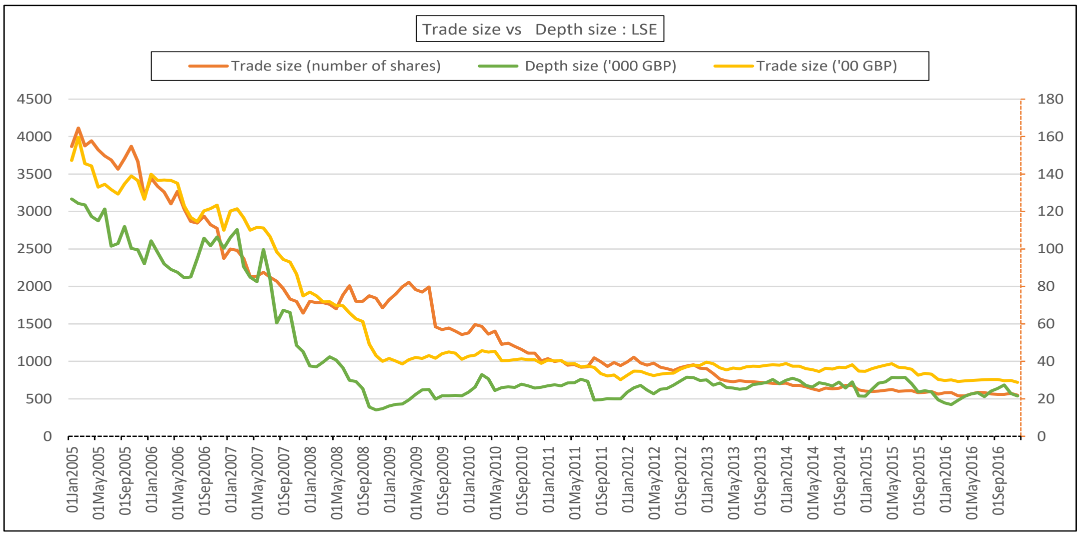


(b) 5-minute price impacts ($price\_impact$)

Fig. A.7: Trends in realized half-spreads and price impacts



(a) Average quoted depths at best price ($depth1$)



(b) Average trade sizes (number of shares)

Fig. A.8: Trends in average quoted depths and trade sizes

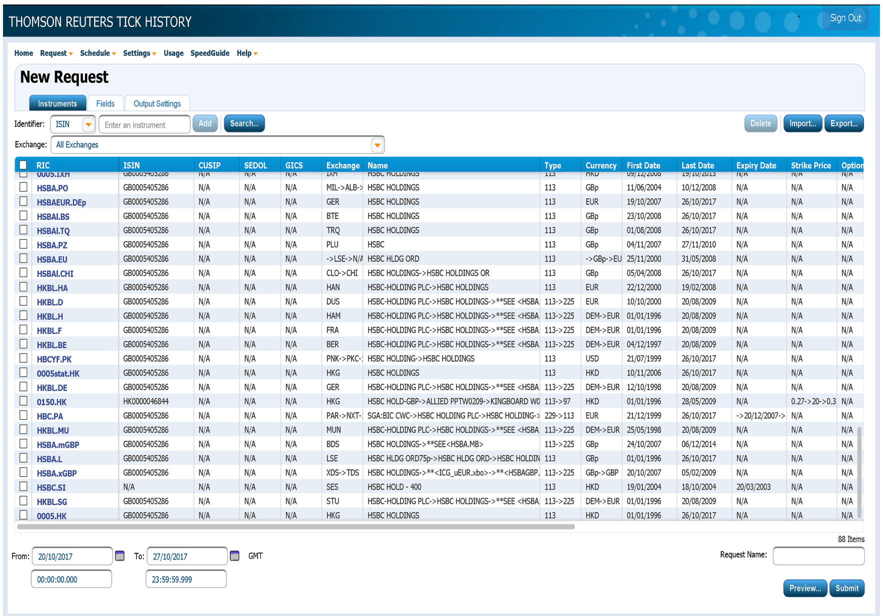
Appendix B

Appendix - Chapter 3

Table B.1: Reuters Instrument Code (RIC) structure

This table explains how a stock is identified across exchanges in the TRTH data request environment. This illustration is based on an excerpted TRTH data request snapshot of HSBC HOLDINGS, an LSE listed UK based company. Under uniform symbology, RIC structure of a stock comprises two parts: the unique root part which is 'HSBA' in this example and the listing/trading venue extension (upper case 'L','BS', 'TQ' and 'CHI' for LSE, BATS, Turquoise and CHIX respectively) that comes after a period '.'. If a stock is traded in the primary exchange, the first part of the RIC only includes the ticker root while an additional lower case letter referring an unique primary venue (lower case 'l' for LSE) is added with the root if it is traded on any other exchanges. Accordingly, 'HSBA.L', refers the RIC of the primary exchange LSE, and 'HSBAl.CHI', 'HSBAl.BS' and 'HSBAl.TQ' refer that for alternative exchange CHIX, BATS and Turquoise respectively. The ISIN is unique for a stock and can be used to link all RICs defined against a stock. The lower section of this table shows a real TRTH data request environment.

RIC	ISIN	Exchange	Name	First Date	Last Date	Underlying RIC
HSBAl.BS	GB0005405286	BTE	HSBC HOLDINGS	23/10/2008	26/10/2017	HSBA.L
HSBAl.TQ	GB0005405286	TRQ	HSBC HOLDINGS	1/8/2008	26/10/2017	HSBA.L
HSBAl.CHI	GB0005405286	CHI	HSBC HOLDINGS	5/4/2008	26/10/2017	HSBA.L
HSBA.L	GB0005405286	LSE	HSBC HOLDINGS	1/1/1996	26/10/2017	HSBA.L



RIC	ISIN	CUSIP	SEDOL	GICS	Exchange	Name	Type	Currency	First Date	Last Date	Expiry Date	Strike Price	Option
HSBA.PO	GB0005405286	N/A	N/A	N/A	MIL->ALB->	HSBC HOLDINGS	113	GBp	11/06/2004	10/12/2008	N/A	N/A	N/A
HSBAEUR.DEp	GB0005405286	N/A	N/A	N/A	GER	HSBC HOLDINGS	113	EUR	19/10/2007	26/10/2017	N/A	N/A	N/A
HSBAl.BS	GB0005405286	N/A	N/A	N/A	BTE	HSBC HOLDINGS	113	GBp	23/10/2008	26/10/2017	N/A	N/A	N/A
HSBAl.TQ	GB0005405286	N/A	N/A	N/A	TRQ	HSBC HOLDINGS	113	GBp	01/08/2008	26/10/2017	N/A	N/A	N/A
HSBA.PZ	GB0005405286	N/A	N/A	N/A	PLU	HSBC	113	GBp	04/11/2007	27/11/2010	N/A	N/A	N/A
HSBA.EU	GB0005405286	N/A	N/A	N/A	->LSE->N/A	HSBC HLDG ORD	113	->GBp->EU	25/11/2000	31/05/2008	N/A	N/A	N/A
HSBAl.CHI	GB0005405286	N/A	N/A	N/A	CLO->CHI	HSBC HOLDINGS->HSBC HOLDINGS OR	113	GBp	05/04/2008	26/10/2017	N/A	N/A	N/A
HKBL.HA	GB0005405286	N/A	N/A	N/A	HAN	HSBC-HOLDING PLC->HSBC HOLDINGS	113	EUR	22/12/2000	19/02/2008	N/A	N/A	N/A
HKBL.D	GB0005405286	N/A	N/A	N/A	DUS	HSBC-HOLDING PLC->HSBC HOLDINGS->**SEE <HSBA	113->225	EUR	10/10/2000	20/08/2009	N/A	N/A	N/A
HKBL.H	GB0005405286	N/A	N/A	N/A	HAM	HSBC-HOLDING PLC->HSBC HOLDINGS->**SEE <HSBA	113->225	DEM->EUR	01/01/1996	20/08/2009	N/A	N/A	N/A
HKBL.F	GB0005405286	N/A	N/A	N/A	FRA	HSBC-HOLDING PLC->HSBC HOLDINGS->**SEE <HSBA	113->225	DEM->EUR	01/01/1996	20/08/2009	N/A	N/A	N/A
HKBL.BE	GB0005405286	N/A	N/A	N/A	BER	HSBC-HOLDING PLC->HSBC HOLDINGS->**SEE <HSBA	113->225	DEM->EUR	04/12/1997	20/08/2009	N/A	N/A	N/A
HBCYF.PK	GB0005405286	N/A	N/A	N/A	PNK->PKC-:	HSBC HOLDING->HSBC HOLDINGS	113	USD	21/07/1999	26/10/2017	N/A	N/A	N/A
0005stat.HK	GB0005405286	N/A	N/A	N/A	HKG	HSBC HOLDINGS	113	HKD	10/11/2006	26/10/2017	N/A	N/A	N/A
HKBL.DE	GB0005405286	N/A	N/A	N/A	GER	HSBC-HOLDING PLC->HSBC HOLDINGS->**SEE <HSBA	113->225	DEM->EUR	12/10/1998	20/08/2009	N/A	N/A	N/A
0150.HK	HK0000046844	N/A	N/A	N/A	HKG	HSBC HOLD-GBP->ALLIED PPTW0209->KINGBOARD W0	113->97	HKD	01/01/1996	28/05/2009	N/A	0.27->20->0.3	N/A
HBC.PA	GB0005405286	N/A	N/A	N/A	PAR->NXT-:	SGA:BIC CWC->HSBC HOLDING PLC->HSBC HOLDING->	229->113	EUR	21/12/1999	26/10/2017	->20/12/2007->	N/A	N/A
HKBL.MU	GB0005405286	N/A	N/A	N/A	MUN	HSBC-HOLDING PLC->HSBC HOLDINGS->**SEE <HSBA	113->225	DEM->EUR	25/05/1998	20/08/2009	N/A	N/A	N/A
HSBA.mGBP	GB0005405286	N/A	N/A	N/A	BDS	HSBC HOLDINGS->**SEE<HSBA.MB>	113->225	GBp	24/10/2007	06/12/2014	N/A	N/A	N/A
HSBA.L	GB0005405286	N/A	N/A	N/A	LSE	HSBC HLDG ORD75p->HSBC HLDG ORD->HSBC HOLDIN	113	GBp	01/01/1996	26/10/2017	N/A	N/A	N/A
HSBA.xGBP	GB0005405286	N/A	N/A	N/A	XDS->TDS	HSBC HOLDINGS->**<ICG_uEUR.xbo>->**<HSBAGBP.	113->225	GBp->GBP	20/10/2007	05/02/2009	N/A	N/A	N/A
HSBC.SI	N/A	N/A	N/A	N/A	SES	HSBC HOLD - 400	113	HKD	19/01/2004	18/10/2004	20/03/2003	N/A	N/A
HKBL.SG	GB0005405286	N/A	N/A	N/A	STU	HSBC-HOLDING PLC->HSBC HOLDINGS->**SEE <HSBA	113->225	DEM->EUR	01/01/1996	20/08/2009	N/A	N/A	N/A
0005.HK	GB0005405286	N/A	N/A	N/A	HKG	HSBC HOLDINGS	113	HKD	01/01/1996	26/10/2017	N/A	N/A	N/A

Table B.2: The simultaneous trading venue participation rate (quarterly) in the EBBO

This table shows the joint venue participation rate (%) in the European Best Bid and Offer (EBBO), a hypothetical aggregate measure of limit order books across LSE, CHIX, BATS and Turquoise, builded on 500 milliseconds snapshots. The single, double, triple and quadruple refer the number of venue(s) which each time contributes in the EBBO. The EBBO is measured on a subsample of 45 stocks which were fragmented across four main exchanges immediately after the event of MiFID, and on which TRTH provides the maximum data support for the period 2008–2016.

		% EBBO (The highest bid price)					% EBBO (The lowest ask price)				
year	qtr	single	Double	triple	quadruple	total	single	Double	triple	quadruple	total
2008	1	100.00	-	-	-	100	100.00	-	-	-	100
2008	2	83.83	16.17	-	-	100	83.48	16.52	-	-	100
2008	3	73.56	25.78	0.66	-	100	73.34	25.88	0.78	-	100
2008	4	69.58	24.27	5.49	0.67	100	68.13	24.75	6.29	0.83	100
2009	1	59.43	25.11	12.53	2.93	100	57.37	25.47	13.63	3.53	100
2009	2	50.22	27.49	16.50	5.79	100	49.30	26.97	17.34	6.40	100
2009	3	56.77	23.18	13.36	6.70	100	56.39	23.18	13.58	6.85	100
2009	4	55.96	23.53	13.44	7.07	100	55.79	23.53	13.48	7.19	100
2010	1	50.19	26.46	16.12	7.24	100	50.02	26.41	16.14	7.44	100
2010	2	47.54	26.07	16.92	9.47	100	47.24	26.10	16.96	9.70	100
2010	3	48.13	24.26	15.95	11.66	100	48.13	24.25	15.93	11.69	100
2010	4	50.80	23.26	15.32	10.62	100	50.65	23.22	15.35	10.78	100
2011	1	51.63	22.83	14.69	10.86	100	51.52	22.84	14.68	10.96	100
2011	2	56.56	22.10	11.79	9.56	100	56.66	22.08	11.71	9.56	100
2011	3	58.02	22.22	10.61	9.16	100	58.01	22.21	10.59	9.19	100
2011	4	51.71	24.71	13.59	9.99	100	51.83	24.66	13.53	9.99	100
2012	1	42.60	26.72	18.90	11.79	100	42.45	26.71	18.94	11.90	100
2012	2	48.61	23.65	14.09	13.65	100	48.71	23.65	14.07	13.56	100
2012	3	49.92	22.90	15.14	12.04	100	49.89	22.84	15.12	12.15	100
2012	4	49.15	23.48	14.47	12.90	100	49.33	23.45	14.41	12.82	100
2013	1	53.95	22.74	12.44	10.87	100	54.00	22.78	12.37	10.85	100
2013	2	52.58	23.08	12.92	11.42	100	52.81	23.13	12.80	11.26	100
2013	3	54.21	22.26	12.76	10.76	100	54.38	22.27	12.73	10.61	100
2013	4	52.25	22.93	13.86	10.96	100	52.29	22.97	13.88	10.85	100
2014	1	54.65	22.15	12.89	10.30	100	54.85	22.20	12.82	10.13	100
2014	2	53.54	22.29	13.85	10.31	100	53.47	22.31	13.89	10.33	100
2014	3	54.45	23.47	13.82	8.26	100	54.43	23.44	13.80	8.33	100
2014	4	51.73	24.21	15.06	9.00	100	51.70	24.21	15.04	9.06	100
2015	1	48.22	22.80	15.24	13.74	100	48.31	22.79	15.21	13.69	100
2015	2	47.91	23.19	15.06	13.84	100	47.97	23.16	15.03	13.84	100
2015	3	46.82	23.66	15.47	14.05	100	46.68	23.71	15.52	14.09	100
2015	4	45.05	23.65	15.38	15.92	100	44.91	23.66	15.45	15.98	100
2016	1	46.00	24.26	15.70	14.04	100	45.85	24.18	15.75	14.22	100
2016	2	46.25	23.80	15.69	14.26	100	46.34	23.80	15.66	14.20	100
2016	3	48.78	22.33	15.00	13.88	100	48.64	22.31	15.06	14.00	100
2016	4	48.23	22.24	14.55	14.99	100	48.11	22.23	14.53	15.14	100
mean		54	23	13	10		54	23	13	10	

Table B.3: The unique trading venue participation rate (quarterly) in the EBBO

This table shows the unique venue participation rate (%) in the European Best Bid and Offer (EBBO), a hypothetical aggregate measure of limit order books across LSE, CHIX, BATS and Turquoise, builded on 500 milliseconds snapshots. LSE, CHIX, BATS and TURQ refer the percentage of time each venue uniquely contributing in the consolidated best bid/offer (EBBO). The EBBO is measured on a subsample of 45 stocks which were fragmented across four main exchanges immediately after the event of MiFID, and on which TRTH provides the maximum data support for the period 2008–2016.

		% EBBO (The highest bid price)					% EBBO (The lowest ask price)				
year	qtr	LSE	CHIX	BATS	TURQ	total	LSE	CHIX	BATS	TURQ	total
2008	1	100.00	-	-	-	100	100	-	-	-	100
2008	2	68.88	31.12	-	-	100	67.94	32.06	-	-	100
2008	3	42.41	55.76	-	1.83	100	41.68	56.33	-	1.99	100
2008	4	34.66	53.39	1.85	10.10	100	34.08	53.58	1.92	10.42	100
2009	1	38.69	36.85	3.85	20.60	100	38.21	36.43	4.05	21.31	100
2009	2	48.15	36.79	11.05	4.01	100	47.05	37.18	11.54	4.23	100
2009	3	41.83	30.49	10.77	16.91	100	41.79	30.37	10.93	16.91	100
2009	4	32.44	37.33	9.86	20.37	100	32.39	37.32	9.87	20.43	100
2010	1	38.45	37.08	11.78	12.69	100	38.38	37.01	11.80	12.81	100
2010	2	36.98	41.19	10.65	11.18	100	36.59	41.33	10.72	11.35	100
2010	3	37.93	40.36	10.42	11.29	100	37.82	40.51	10.38	11.30	100
2010	4	34.27	39.79	16.72	9.22	100	34.27	39.95	16.60	9.17	100
2011	1	35.53	40.15	14.43	9.89	100	35.48	40.40	14.30	9.83	100
2011	2	15.00	71.67	7.14	6.18	100	14.98	71.73	7.12	6.17	100
2011	3	8.98	81.00	4.84	5.17	100	8.84	81.23	4.74	5.19	100
2011	4	13.71	61.93	13.54	10.82	100	13.54	62.12	13.54	10.79	100
2012	1	18.73	47.61	15.76	17.91	100	18.69	47.71	15.75	17.84	100
2012	2	32.88	40.89	11.57	14.67	100	32.87	40.91	11.54	14.69	100
2012	3	32.19	41.11	11.69	15.01	100	32.12	41.14	11.70	15.04	100
2012	4	37.11	32.85	13.19	16.85	100	37.25	32.79	13.15	16.81	100
2013	1	41.16	29.84	9.29	19.71	100	41.22	29.80	9.35	19.63	100
2013	2	47.44	27.07	8.15	17.35	100	47.47	27.13	8.18	17.21	100
2013	3	45.35	28.90	7.88	17.87	100	45.22	29.05	7.89	17.84	100
2013	4	43.85	27.65	9.08	19.41	100	43.83	27.72	9.05	19.39	100
2014	1	54.98	19.63	8.38	17.01	100	55.35	19.47	8.34	16.84	100
2014	2	57.89	18.15	9.12	14.84	100	58.13	17.98	9.12	14.77	100
2014	3	47.13	29.11	9.15	14.60	100	47.23	29.07	9.15	14.55	100
2014	4	48.22	25.80	8.60	17.39	100	48.25	25.72	8.60	17.43	100
2015	1	47.77	24.89	10.60	16.73	100	47.90	24.69	10.59	16.82	100
2015	2	50.70	22.98	10.78	15.54	100	51.06	22.74	10.77	15.44	100
2015	3	48.63	25.64	9.82	15.91	100	48.96	25.40	9.79	15.84	100
2015	4	47.14	28.25	7.93	16.69	100	46.98	28.36	7.99	16.67	100
2016	1	38.94	27.20	12.77	21.09	100	38.62	27.13	12.90	21.35	100
2016	2	41.42	23.34	13.91	21.33	100	41.45	23.30	13.93	21.33	100
2016	3	53.50	20.69	8.49	17.33	100	53.35	20.62	8.50	17.53	100
2016	4	57.91	17.93	7.49	16.66	100	57.92	17.89	7.51	16.69	100
mean		42	35	10	13		42	35	10	13	

Table B.4: Descriptive statistics for HFT proxies across venues: LSE and CHIX

This table presents the descriptive statistics for HFT proxies calculated on the millisecond-time stamped data for LSE and CHIX. The sample consists of 149 stocks divided into 5 equal quintiles based on market capitalization for the period 2005–2016. Table A.2 (Panel B) is refereed to see the periodical data coverage in the sample in detail.

			LSE						CHIX					
			All	Small	Q2	Q3	Q4	Large	All	Small	Q2	Q3	Q4	Large
$hft1$	messages per minute (best 10 depth levels)	Mean	100.36	23.77	38.97	72.39	103.80	266.31	136.59	27.78	47.24	87.99	137.37	382.44
		Median	47.62	17.95	30.57	56.44	90.99	171.75	69.65	20.18	34.02	69.81	113.63	276.53
		StdDev	157.77	23.22	38.25	67.59	84.81	271.25	208.13	27.99	44.18	66.96	93.11	343.19
$hft2$	messages per minute (best 5 depth levels)	Mean	83.72	19.35	32.61	60.80	89.36	219.40	115.73	24.98	41.32	75.05	118.97	318.24
		Median	41.10	14.42	26.09	49.34	80.49	149.97	60.85	18.06	30.11	59.78	97.67	236.66
		StdDev	126.10	19.37	30.18	52.20	68.56	213.53	168.72	24.96	37.16	55.60	80.33	272.31
$hft3$	messages per minute (BBO)	Mean	39.32	9.57	15.67	28.33	43.50	100.91	48.67	11.65	18.19	32.03	50.28	131.16
		Median	20.38	7.48	12.87	23.91	40.04	76.56	27.76	9.14	14.57	27.27	43.95	101.99
		StdDev	54.32	8.98	13.09	21.76	30.81	88.30	64.84	9.76	13.83	20.59	29.12	100.26
$ordtotrd$	number of messages per executed order (order to trade ratio)	Mean	22.02	28.05	19.51	20.45	21.18	20.73	53.33	92.32	49.77	41.92	41.65	41.07
		Median	19.06	19.97	17.28	18.46	20.40	19.83	35.99	52.67	34.10	31.26	33.96	35.34
		StdDev	21.25	36.67	15.59	14.76	13.57	13.84	91.07	182.09	63.23	34.24	27.06	22.51
$hft1h$	GBP volume (100) per message (best 10 depth levels) time (-1)	Mean	-6.93	-2.24	-3.73	-5.14	-7.51	-16.21	-0.83	-0.29	-0.55	-0.81	-1.04	-1.47
		Median	-1.96	-0.88	-1.37	-1.77	-2.38	-3.60	-0.70	-0.22	-0.48	-0.75	-0.96	-1.33
		StdDev	15.51	4.19	5.87	8.42	10.93	29.12	0.65	0.27	0.38	0.47	0.53	0.73
$hft2h$	GBP volume (100) per message (best 5 depth levels) time (-1)	Mean	-7.54	-2.80	-4.17	-5.53	-8.04	-17.37	-0.95	-0.32	-0.62	-0.93	-1.20	-1.71
		Median	-2.33	-1.10	-1.64	-2.05	-2.72	-4.19	-0.81	-0.24	-0.54	-0.87	-1.11	-1.56
		StdDev	16.22	5.30	6.32	8.78	11.36	30.28	0.74	0.29	0.41	0.52	0.60	0.83
	observations (stock*day)		439583	90046	88374	87239	86786	87138	324334	64743	65175	64952	64484	64980

Table B.5: Descriptive statistics for HFT proxies across venues: BATS and Turquoise

This table presents the descriptive statistics for HFT proxies calculated on the millisecond-time stamped data for BATS and Turquoise. The sample consists of 149 stocks divided into 5 equal quintiles based on market capitalization for the period 2005–2016. Table A.2 (Panel B) is refereed to see the periodical data coverage in the sample in detail.

			BATS						Turquoise					
			All	Small	Q2	Q3	Q4	Large	All	Small	Q2	Q3	Q4	Large
$hft1$	messages per minute (best 10 depth levels)	Mean	77.42	15.21	25.63	45.68	75.30	222.58	76.08	15.51	25.47	47.88	74.67	207.74
		Median	35.16	11.35	18.12	33.94	58.10	157.87	34.88	12.19	19.28	33.90	55.87	152.82
		StdDev	127.95	16.24	29.43	39.79	60.55	214.90	116.43	13.77	26.59	49.28	63.64	185.25]
$hft2$	messages per minute (best 5 depth levels)	Mean	68.98	14.66	24.02	41.17	67.87	194.84	67.60	14.89	23.50	42.74	66.39	182.53
		Median	32.46	11.07	17.36	31.18	52.43	141.80	32.04	11.77	18.18	30.85	50.16	134.98
		StdDev	108.71	15.17	25.77	34.37	54.25	178.95	101.24	12.92	22.32	43.51	55.22	160.56
$hft3$	messages per minute (BBO)	Mean	33.27	9.37	13.92	21.53	32.45	88.03	33.74	9.21	13.68	22.10	32.43	87.59
		Median	18.11	7.38	10.81	17.18	26.39	66.87	18.46	7.34	11.10	18.13	26.77	65.44
		StdDev	45.29	8.42	11.84	16.04	22.96	71.41	46.51	7.63	10.61	15.99	21.93	74.79
$ordtotrd$	number of messages per executed order (order to trade ratio)	Mean	98.21	152.48	94.92	87.01	80.26	78.06	59.31	83.30	45.63	54.46	58.14	56.23
		Median	54.01	72.88	56.53	48.44	48.12	52.01	34.94	39.06	29.16	31.99	36.30	39.50
		StdDev	200.02	332.55	161.99	166.68	140.96	118.31	99.78	171.83	64.94	85.77	78.36	56.24
$hft1h$	GBP volume (100) per message (best 10 depth levels) time (-1)	Mean	-0.47	-0.16	-0.26	-0.43	-0.63	-0.86	-0.79	-0.36	-0.58	-0.76	-0.95	-1.25
		Median	-0.33	-0.13	-0.21	-0.35	-0.51	-0.69	-0.61	-0.28	-0.50	-0.63	-0.78	-1.03
		StdDev	0.49	0.15	0.20	0.35	0.46	0.70	0.69	0.30	0.40	0.55	0.70	0.92
$hft2h$	GBP volume (100) per message (best 5 depth levels) time (-1)	Mean	-0.52	-0.17	-0.27	-0.47	-0.69	-0.96	-0.87	-0.37	-0.62	-0.83	-1.06	-1.40
		Median	-0.36	-0.13	-0.22	-0.38	-0.57	-0.78	-0.67	-0.29	-0.53	-0.69	-0.86	-1.16
		StdDev	0.54	0.15	0.22	0.38	0.51	0.76	0.77	0.32	0.43	0.60	0.78	1.01
	observations (stock*day)		301585	58860	60333	60950	60478	60964	298676	56255	57828	61216	61491	61886

Table B.6: Descriptive statistics for liquidity measures across venues: LSE and CHIX

This table presents the descriptive statistics for absolute quoted spread ($spread\_abs$) in GBX, relative quoted spreads ($spread\_bps$) in basis points (bps), volume weighted effective half-spreads ($espread$) in bps, average quoted depth at best price ($depth1$) in GBP100, average cumulative depth upto three best price ($depth3$) in GBP100. All measures are developed on millisecond time-stamped trades and quotes data for LSE and CHIX and appropriate averages are taken to convert them daily. The sample consists of 149 stocks divided into 5 equal quintiles based on market capitalization for the period 2005–2016. Table A.2 (Panel B) is refereed to see the periodical data coverage in the sample in detail.

			LSE						CHIX					
			All	Small	Q2	Q3	Q4	Large	All	Small	Q2	Q3	Q4	Large
$spread\_abs$	absloute quoted spread (GBX)	Mean	1.44	2.54	1.40	1.13	1.19	0.90	1.77	3.77	1.71	1.21	1.27	0.90
		Median	0.89	1.20	0.95	0.84	0.80	0.77	0.95	1.52	1.05	0.87	0.84	0.80
		StdDev	2.32	4.41	1.40	1.05	1.33	0.74	3.59	6.95	2.12	1.81	1.53	0.90
$spread\_bps$	percentage quoted spread (bps)	Mean	20.39	44.13	22.18	15.05	11.64	8.09	24.88	63.11	26.64	15.56	11.48	7.65
		Median	13.79	30.70	18.41	13.48	11.01	7.61	13.61	40.83	19.52	13.00	10.63	6.45
		StdDev	23.71	40.62	13.50	7.77	4.52	4.83	40.03	72.61	21.91	12.96	6.19	6.14
$espread$	effective half- spread (bps)	Mean	7.16	14.78	7.50	5.42	4.47	3.37	7.74	18.26	8.34	5.16	4.04	2.89
		Median	5.09	9.98	6.30	4.94	4.33	3.15	4.72	11.48	6.09	4.43	3.86	2.35
		StdDev	8.50	15.26	4.73	2.87	1.77	2.27	11.75	21.50	7.09	3.71	1.82	2.45
$depth1$	average depth (BBO level/GBP 100)	Mean	415.98	70.53	134.85	252.09	492.27	1146.17	161.02	33.88	58.69	110.17	217.45	385.17
		Median	182.62	58.05	98.32	189.16	383.08	620.19	89.78	30.87	51.56	93.69	186.67	303.01
		StdDev	1291.87	52.33	127.28	214.74	417.22	2722.34	201.63	17.65	34.72	65.53	129.04	313.04
$depth3$	average cumulative depth (best three levels/GBP 100)	Mean	1781.55	265.08	526.51	1050.53	2200.43	4936.16	679.75	98.80	195.02	429.48	934.68	1741.92
		Median	792.94	225.70	402.43	832.48	1843.02	3050.27	333.68	89.05	164.06	352.90	802.42	1422.42
		StdDev	5077.64	182.06	435.92	805.45	1638.94	10577.23	900.43	58.89	126.42	291.27	580.22	1334.61
	observations (stock*day)		439583	90046	88374	87239	86786	87138	324334	64743	65175	64952	64484	64980

Table B.7: Descriptive statistics for liquidity measures across venues: BATS and Turquoise

This table presents the descriptive statistics for absolute quoted spread ($spread\_abs$) in GBX, relative quoted spreads ($spread\_bps$) in basis points (bps), volume weighted effective half-spreads ($espread$) in bps, average quoted depth at best price ($depth1$) in GBP100, average cumulative depth upto three best price ($depth3$) in GBP100. These measures are developed on millisecond time-stamped trades and quotes data for BATS and Turquoise and appropriate averages are taken to convert them daily. The sample consists of 149 stocks divided into 5 equal quintiles based on market capitalization for the period 2005–2016. Table A.2 (Panel B) is refereed to see the periodical data coverage in the sample in detail.

			BATS						Turquoise					
			All	Small	Q2	Q3	Q4	Large	All	Small	Q2	Q3	Q4	Large
$spread\_abs$	absloute quoted spread (GBX)	Mean	2.34	4.81	2.54	1.66	1.68	1.12	1.93	3.76	2.00	1.47	1.49	1.06
		Median	1.24	2.03	1.50	1.13	1.01	0.98	1.12	1.69	1.21	1.05	0.97	0.91
		StdDev	4.37	8.50	2.86	1.71	2.08	1.08	3.11	5.89	2.18	1.50	1.81	0.97
$spread\_bps$	percentage quoted spread (bps)	Mean	31.78	77.77	37.90	21.06	14.21	9.44	26.28	60.27	31.14	20.23	13.65	9.38
		Median	17.95	53.65	28.70	17.47	12.93	8.04	15.51	41.82	21.83	15.21	11.95	7.50
		StdDev	45.62	79.20	30.53	14.70	6.54	7.66	33.90	54.89	27.88	19.01	8.19	10.24
$espread$	effective half- spread (bps)	Mean	9.75	23.15	11.46	6.53	4.73	3.34	8.01	17.10	9.32	6.45	4.70	3.36
		Median	5.72	15.65	8.52	5.54	4.47	2.84	5.05	11.44	6.41	4.92	4.22	2.68
		StdDev	13.59	23.62	9.69	4.39	1.95	2.58	10.07	16.69	8.75	5.71	2.82	3.56
$depth1$	average depth (BBO level/GBP 100)	Mean	97.15	26.85	38.18	61.87	126.54	229.51	100.19	31.55	45.74	73.48	132.56	207.71
		Median	51.55	24.70	34.01	50.27	101.42	164.98	64.27	28.81	40.76	60.96	112.43	167.32
		StdDev	175.23	12.58	20.51	40.55	88.18	337.85	130.39	15.28	21.52	42.86	80.88	229.62
$depth3$	average cumulative depth (best three levels/GBP 100)	Mean	376.07	75.79	123.44	228.73	505.75	934.64	353.85	75.34	129.33	240.67	485.90	797.58
		Median	177.62	63.46	104.49	176.92	399.14	682.28	196.13	70.59	108.61	192.07	406.58	634.22
		StdDev	589.82	65.73	87.53	170.21	383.39	1016.78	450.24	39.89	78.52	166.22	333.06	699.37
	observations (stock*day)		301585	58860	60333	60950	60478	60964	298676	56255	57828	61216	61491	61886

Table B.8: Descriptive statistics for realized half-spread measures across venues: LSE and CHIX

This table presents the descriptive statistics for realized half-spreads based on four hypothetical post-trade quote adjustment intervals. These measures are developed on millisecond time-stamped trades and quotes data for LSE and CHIX and weighted by the respective trade volume (number of shares) to convert them daily. The sample consists of 149 stocks divided into 5 equal quintiles based on market capitalization for the period 2005–2016. Table A.2 (Panel B) is refereed to see the periodical data coverage in the sample in detail.

			LSE						CHIX					
variable	Description (units)		All	Small	Q2	Q3	Q4	Large	All	Small	Q2	Q3	Q4	Large
	share volume	Mean	0.98	3.54	0.66	0.27	0.21	0.16	0.73	3.56	0.68	-0.08	-0.27	-0.23
$rspread1$	wighted 10-sec	Median	0.13	1.08	0.32	0.05	0.02	-0.04	-0.26	0.66	-0.08	-0.36	-0.36	-0.31
	realized spread (bps)	StdDev	5.65	11.54	2.93	1.78	1.25	1.09	7.22	15.07	4.02	2.16	1.02	1.27
	share volume	Mean	0.43	2.56	-0.02	-0.23	-0.15	-0.10	0.46	3.18	0.30	-0.37	-0.45	-0.35
$rspread2$	wighted 30-sec	Median	-0.15	0.41	-0.17	-0.28	-0.22	-0.16	-0.39	0.32	-0.37	-0.54	-0.47	-0.36
	realized half-spread (bps)	StdDev	5.59	11.48	3.01	1.88	1.32	1.05	7.06	14.74	3.99	2.10	1.03	1.21
	share volume	Mean	0.13	2.05	-0.39	-0.51	-0.35	-0.20	0.27	2.80	0.04	-0.54	-0.56	-0.39
$rspread3$	wighted 1-minute	Median	-0.28	0.08	-0.41	-0.44	-0.34	-0.20	-0.45	0.10	-0.51	-0.63	-0.53	-0.37
	realized half-spread (bps)	StdDev	5.68	11.65	3.18	2.03	1.43	1.07	7.02	14.67	3.99	2.12	1.11	1.21
	share volume	Mean	-0.21	1.12	-0.86	-0.72	-0.43	-0.18	0.08	1.99	-0.23	-0.58	-0.48	-0.29
$rspread4$	wighted 5-minute	Median	-0.32	-0.30	-0.60	-0.48	-0.31	-0.12	-0.36	-0.09	-0.52	-0.54	-0.39	-0.25
	realized half-spread (bps)	StdDev	6.09	12.36	3.87	2.53	1.77	1.30	7.20	14.98	4.50	2.50	1.50	1.34
	observations (stock*day)		439583	90046	88374	87239	86786	87138	324334	64743	65175	64952	64484	64980

Table B.9: Descriptive statistics for realized half-spread measures: BATS and Turquoise

This table presents the descriptive statistics for realized half-spreads based on four hypothetical post-trade quote adjustment intervals. These measures are developed on millisecond time-stamped trades and quotes data for BATS and Turquoise and weighted by the respective trade volume (number of shares) to convert them daily. The sample consists of 149 stocks divided into 5 equal quintiles based on market capitalization for the period 2005–2016. Table A.2 (Panel B) is refereed to see the periodical data coverage in the sample in detail.

			BATS						Turquoise					
			All	Small	Q2	Q3	Q4	Large	All	Small	Q2	Q3	Q4	Large
	share volume	Mean	1.43	5.40	1.65	0.30	0.03	-0.09	0.84	3.46	1.00	0.32	-0.18	-0.17
$rspread1$	wighted 10-sec	Median	0.11	1.86	0.63	0.07	-0.01	-0.13	-0.24	0.63	-0.05	-0.33	-0.36	-0.29
	realized spread (bps)	StdDev	8.91	18.08	6.72	2.84	1.30	1.31	6.92	13.34	6.36	3.95	1.69	2.12
	share volume	Mean	1.26	5.13	1.46	0.10	-0.11	-0.15	0.54	2.93	0.64	0.00	-0.39	-0.29
$rspread2$	wighted 30-sec	Median	0.04	1.61	0.47	-0.05	-0.09	-0.16	-0.36	0.26	-0.31	-0.50	-0.47	-0.33
	realized half-spread (bps)	StdDev	8.93	18.19	6.62	2.94	1.38	1.32	6.87	13.23	6.40	3.92	1.80	2.25
	share volume	Mean	1.11	4.82	1.27	-0.04	-0.20	-0.18	0.38	2.67	0.45	-0.19	-0.52	-0.31
$rspread3$	wighted 1-minute	Median	-0.01	1.42	0.35	-0.13	-0.14	-0.16	-0.41	0.08	-0.45	-0.58	-0.53	-0.33
	realized half-spread (bps)	StdDev	8.97	18.24	6.76	3.04	1.52	1.38	7.00	13.42	6.53	4.02	2.12	2.48
	share volume	Mean	0.90	3.82	1.04	-0.06	-0.14	-0.09	0.27	2.02	0.32	-0.23	-0.46	-0.14
$rspread4$	wighted 5-minute	Median	0.07	1.14	0.33	-0.07	-0.04	-0.05	-0.32	0.00	-0.40	-0.47	-0.40	-0.19
	realized half-spread (bps)	StdDev	9.13	18.30	7.48	3.78	2.14	1.76	8.04	14.34	7.78	5.21	4.19	4.54
observations (stock*day)			301585	58860	60333	60950	60478	60964	298676	56255	57828	61216	61491	61886

Table B.10: Descriptive statistics for price impact/adverse selection cost measures: LSE and CHIX

This table presents the descriptive statistics for price impacts/adverse selection costs based on four hypothetical post-trade quote adjustment intervals. These measures are developed on millisecond time-stamped trades and quotes data for LSE and CHIX and weighted by the respective trade volume (number of shares) to convert them daily. The sample consists of 149 stocks divided into 5 equal quintiles based on market capitalization for the period 2005–2016. Table A.2 (Panel B) is refereed to see the periodical data coverage in the sample in detail.

			LSE						CHIX					
			All	Small	Q2	Q3	Q4	Large	All	Small	Q2	Q3	Q4	Large
$price\_impact1$	share volume	Mean	6.17	11.21	6.84	5.16	4.27	3.21	7.00	14.66	7.65	5.25	4.30	3.12
	wighted 10-sec price	Median	4.74	8.52	5.86	4.67	4.06	2.96	4.87	10.09	6.21	4.66	4.04	2.69
	impact (bps)	StdDev	5.64	9.41	4.01	2.58	1.66	1.77	8.82	16.01	5.47	3.47	1.78	1.81
$price\_impact2$	share volume	Mean	6.73	12.20	7.52	5.66	4.63	3.46	7.27	15.06	8.03	5.54	4.49	3.24
	wighted 30-sec price	Median	5.12	9.23	6.38	5.04	4.35	3.21	5.08	10.46	6.51	4.89	4.20	2.78
	impact (bps)	StdDev	6.23	10.31	4.58	3.01	1.93	2.03	8.95	16.14	5.68	3.57	1.92	1.96
$price\_impact3$	share volume	Mean	7.03	12.72	7.90	5.93	4.83	3.57	7.46	15.45	8.29	5.71	4.60	3.28
	wighted 60-sec price	Median	5.32	9.63	6.67	5.26	4.50	3.31	5.18	10.77	6.69	5.00	4.27	2.81
	impact (bps)	StdDev	6.59	10.87	4.93	3.27	2.12	2.17	9.19	16.52	5.91	3.63	2.06	2.05
$price\_impact4$	share volume	Mean	7.37	13.65	8.36	6.14	4.91	3.55	7.66	16.32	8.56	5.75	4.52	3.18
	wighted 5-min price	Median	5.41	10.21	6.96	5.36	4.51	3.22	5.08	11.14	6.72	4.89	4.10	2.67
	impact (bps)	StdDev	7.44	12.30	5.71	3.78	2.46	2.40	10.17	18.33	6.70	4.18	2.35	2.23
observations (stock*day)			439583	90046	88374	87239	86786	87138	324334	64743	65175	64952	64484	64980

Table B.11: Descriptive statistics for price impact/adverse selection cost measures: BATS and Turquoise

This table presents the descriptive statistics for price impacts/adverse selection costs based on four hypothetical post-trade quote adjustment intervals. These measures are developed on millisecond time-stamped trades and quotes data for BATS and Turquoise and weighted by the respective trade volume (number of shares) to convert them daily. The sample consists of 149 stocks divided into 5 equal quintiles based on market capitalization for the period 2005–2016. Table A.2 (Panel B) is refereed to see the periodical data coverage in the sample in detail.

			BATS						Turquoise					
			All	Small	Q2	Q3	Q4	Large	All	Small	Q2	Q3	Q4	Large
	share volume	Mean	8.31	17.74	9.79	6.23	4.69	3.42	7.18	13.64	8.34	6.13	4.89	3.54
$price\_impact1$	wighted 10-sec price	Median	5.40	12.06	7.56	5.35	4.33	2.95	5.11	9.54	6.42	5.08	4.40	2.93
	impact(bps)	StdDev	10.89	19.46	8.29	3.92	2.02	2.21	7.87	13.21	7.26	4.28	2.61	2.69
	share volume	Mean	8.50	18.05	10.00	6.43	4.83	3.49	7.49	14.18	8.70	6.45	5.11	3.67
$price\_impact2$	wighted 30-sec price	Median	5.56	12.42	7.81	5.52	4.42	2.99	5.37	10.08	6.77	5.35	4.58	3.06
	impact(bps)	StdDev	11.11	19.81	8.53	4.10	2.17	2.32	8.15	13.68	7.31	4.56	2.87	2.95
	share volume	Mean	8.65	18.37	10.20	6.58	4.92	3.51	7.66	14.49	8.90	6.64	5.25	3.70
$price\_impact3$	wighted 60-sec price	Median	5.65	12.73	7.98	5.62	4.47	3.00	5.52	10.36	6.99	5.49	4.67	3.10
	impact(bps)	StdDev	11.31	20.11	8.72	4.30	2.32	2.39	8.42	14.10	7.48	4.73	3.22	3.24
	share volume	Mean	8.88	19.35	10.45	6.62	4.87	3.44	7.81	15.16	9.09	6.75	5.24	3.56
$price\_impact4$	wighted 5-min price	Median	5.55	13.30	8.07	5.52	4.34	2.87	5.53	10.79	7.09	5.43	4.52	3.00
	impact(bps)	StdDev	12.44	22.14	9.65	5.05	2.82	2.75	9.85	16.05	8.65	6.09	4.96	5.13
observations(stock*day)			301585	58860	60333	60950	60478	60964	298676	56255	57828	61216	61491	61886

Table B.12: The cross-market impact of high frequency trading and market fragmentation on liquidity: a simultaneous equations model estimation

This table presents the simultaneous equations model estimation for the system of equations (3.1)–(3.8) using the three-stage least squares methods for two liquidity measures: log normalized time weighted quoted spreads and log normalized volume weighted effective half-spreads. Indices i and t represent stocks and time (days) respectively, v represents one of the four venues: LSE, CHIX, BATS and Turquoise, HFT_{vit} represents the HFT proxy ($hft2$) developed on quotes update upto the fifth depth level, $HHItrd_{it}$ represents the market fragmentation proxy, $\overline{MQ}_{-vit}$ represents the average liquidity level over all stocks in the same size group excluding stock i at venue v, $\overline{HFT}_{-vit}$ represents the average HFT intensity at venue v over all stocks in the same size group excluding stock i, $ln(mktcap)$ is the log normalized market capitalization, $ln(voltintra)_{vit}$ is the log normalized intraday mid price range volatility, $invprice$ is the inverse of daily average price, $ln(size)_{vit}$ is the log normalized trade size, $ln(value)_{vit}$ is the log normalized trading volume, rtk_{vit} is the relative tick size. The estimation is based on a panel dataset of 149 stocks and 2060 days (October 2008–December 2016) and includes both time (the monthly time dummy for each of 99 months included in the panel) and stock fixed effects. Coefficient estimates are 3SLS , t-statistics shown in the parentheses below the coefficient. ***, **, * denote significance at 1%, 5%, and 10% levels, respectively. Panel A presents the estimates for market quality equations (3.1–3.4) and Panel B presents those for HFT equations (3.5–3.8).

	I	II	III	IV	V	VI	VII	VIII
	$MQ_{vit} = Log(quoted\ \ spreads)_{vit}$				$MQ_{(vit} = Log(effective\ \ half-spreads)_{vit}$			
	Panel A							
	$MQ_{(lse)it}$	$MQ_{(chix)it}$	$MQ_{(bats)it}$	$MQ_{(turq)it}$	$MQ_{(lse)it}$	$MQ_{(chix)it}$	$MQ_{(bats)it}$	$MQ_{(turq)it}$
$const$	4.134***	4.662***	3.876***	4.146***	3.241***	3.656***	3.198***	3.168***
	(198.68)	(211.7)	(150.98)	(165.36)	(159.75)	(168.25)	(128.16)	(128.45)
HFT_{vit}	-0.375***	-0.432***	-0.344***	-0.343***	-0.309***	-0.344***	-0.261***	-0.266***
	(-283.7)	(-320.37)	(-230.15)	(-231.81)	(-235.37)	(-255.86)	(-173.56)	(-177.83)
$HHItrd_{it}$	0.05***	-0.032***	-0.098***	-0.11***	0.056***	-0.036***	-0.1***	-0.103***
	(34.16)	(-18.69)	(-50.1)	(-57.85)	(36.05)	(-20.66)	(-50.53)	(-52.97)
$\overline{MQ}_{-vit}$	0.162***	0.186***	0.298***	0.274***	0.065***	0.125***	0.245***	0.233***
	(67.44)	(100.59)	(142.32)	(130.59)	(22.81)	(56.47)	(104.72)	(94.31)
$inv(price)_{it}$	13.148***	12.775***	15.018***	15.219***	15.8***	16.006***	17.75***	19.071***
	(61.9)	(52.1)	(52.66)	(54.88)	(70.04)	(61.86)	(60.31)	(66.32)
$ln(mktcap)_{it}$	-0.149***	-0.162***	-0.134***	-0.159***	-0.169***	-0.178***	-0.167***	-0.162***
	(-79.21)	(-76.77)	(-54.94)	(-67.11)	(-85.68)	(-81.85)	(-67.52)	(-66.14)
$ln(voltintra)_{vit}$	0.066***	0.028***	0.034***	0.026***	0.052***	0.02***	0.031***	0.02***
	(80.9)	(79.53)	(92.51)	(77.03)	(67.14)	(60.88)	(84.33)	(57.42)
	Panel B							
	$HFT_{(lse)it}$	$HFT_{(chix)it}$	$HFT_{(bats)it}$	$HFT_{(turq)it}$	$HFT_{(lse)it}$	$HFT_{(chix)it}$	$HFT_{(bats)it}$	$HFT_{(turq)it}$
$const$	6.604***	6.242***	7.687***	7.957***	6.079***	5.684***	5.97***	6.213***
	(191.96)	(152.71)	(164.69)	(178.3)	(167.18)	(142.15)	(124.56)	(129.78)
$MQ_{(lse)it}$	0.077***	-0.477***	-0.634***	-0.726***	0.359***	-1.281***	-1.246***	-1.843***
	(6.08)	(-40.99)	(-47.13)	(-56.45)	(15.53)	(-67.39)	(-62.92)	(-93.2)
$MQ_{(chix)it}$	-0.561***	0.071***	-0.662***	-0.36***	-0.782***	0.608***	-0.636***	0.031**
	(-78.77)	(6.13)	(-62.59)	(-36.39)	(-64.93)	(37.05)	(-41.72)	(2.07)
$MQ_{(bats)it}$	0.199***	0.156***	0.328***	0.3***	0.308***	0.268***	0.944***	0.494***
	(47.63)	(31.4)	(38.4)	(52.52)	(52.37)	(41.17)	(81.08)	(62.73)
$MQ_{(turq)it}$	0.179***	0.302***	0.332***	0.181***	0.259***	0.518***	0.458***	0.831***
	(38.55)	(53.87)	(49.1)	(21.02)	(39.46)	(67.19)	(48.95)	(67.26)
$\overline{HFT}_{-vit}$	0.36***	0.434***	0.428***	0.413***	0.372***	0.445***	0.506***	0.488***
	(183.93)	(199.33)	(183.41)	(186.26)	(161.51)	(196.87)	(203.51)	(199.9)
$HHItrd_{it}$	0.145***	0.114***	0.152***	0.086***	0.151***	0.244***	0.274***	0.29***
	(51.67)	(35.75)	(38.93)	(23.18)	(41.19)	(60.71)	(56.82)	(61.17)
$ln(mktcap)_{it}$	-0.074***	-0.106***	-0.275***	-0.321***	0	-0.083***	-0.228***	-0.268***
	(-19.17)	(-24.99)	(-60.39)	(-75.61)	(-0.09)	(-19.23)	(-46.76)	(-57.15)
$ln(volume)_{vit}$	0.506***	0.492***	0.349***	0.393***	0.552***	0.531***	0.412***	0.458***
	(172.27)	(198.29)	(160.08)	(172.89)	(162.7)	(226.49)	(186.08)	(192.92)
$ln(size)_{vit}$	-0.536***	-0.51***	-0.349***	-0.332***	-0.6***	-0.522***	-0.374***	-0.377***
	(-125.84)	(-120.14)	(-102.07)	(-89.74)	(-98.16)	(-127)	(-108.02)	(-93.64)
$rtick_{vit}$	-203.879***	-199.162***	-2.578***	-1.168***	-296.621***	-193.904***	-2.476***	-1.018***
	(-62.46)	(-60.71)	(-7.63)	(-12.38)	(-46.92)	(-49.98)	(-6.94)	(-9.88)
$ln(voltintra)_{vit}$	0.011***	-0.015***	-0.006***	-0.007***	-0.005	-0.023***	-0.022***	-0.025***
	(5.28)	(-19.85)	(-9.92)	(-11.83)	(-1.53)	(-29.69)	(-31.38)	(-37.24)
observations	277563	277563	277563	277563	277563	277563	277563	277563
second-stage adj_Rsqr (MQ_{vit})	0.84	0.86	0.84	0.83	0.79	0.81	0.80	0.79
second-stage adj_Rsqr (HFT_{vit})	0.90	0.88	0.82	0.83	0.89	0.87	0.80	0.80
system weighted Rsqr			0.80				0.76	

Table B.13: The cross-market impact of high frequency trading and market fragmentation on liquidity: a simultaneous equations model estimation for large stocks

This table presents the simultaneous equations model estimation for the system of equations (3.1)–(3.8) using the three-stage least squares methods for two liquidity measures: log normalized time weighted quoted spreads and log normalized volume weighted effective half-spreads. Indices i and t represent stocks and time (days) respectively, v represents one of the four venues: LSE, CHIX, BATS and Turquoise, HFT_{vit} represents the HFT proxy ($hft2$) developed on quotes update upto the fifth depth level, $HHItrd_{it}$ represents the market fragmentation proxy, $\overline{MQ}_{-vit}$ represents the average liquidity level over all stocks in the same size group excluding stock i at venue v, $\overline{HFT}_{-vit}$ represents the average HFT intensity at venue v over all stocks in the same size group excluding stock i, $ln(mktcap)$ is the log normalized market capitalization, $ln(voltintra)_{vit}$ is the log normalized intraday mid price range volatility, $invprice$ is the inverse of daily average price, $ln(size)_{vit}$ is the log normalized trade size, $ln(value)_{vit}$ is the log normalized trading volume, rtk_{vit} is the relative tick size. The estimation is based on a panel dataset of 74 large-cap stocks (above the median market capitalization stocks group) and 2058 days (October 2008–December 2016) and includes both time (the monthly time dummy for each of 99 months included in the panel dataset) and stock fixed effects. Coefficient estimates are 3SLS , t-statistics shown in the parentheses below the coefficient. ***, **, * denote significance at 1%, 5%, and 10% levels, respectively. Panel A presents the estimates for market quality equations (3.1–3.4) and Panel B presents those for HFT equations (3.5–3.8).

	I	II	III	IV	V	VI	VII	VIII
	$MQ_{vit} = Log(quoted\ \ spreads)_{vit}$				$MQ_{(vit} = Log(effective\ \ half-spreads)_{vit}$			
Panel A								
	$MQ_{(lse)it}$	$MQ_{(chix)it}$	$MQ_{(bats)it}$	$MQ_{(turq)it}$	$MQ_{(lse)it}$	$MQ_{(chix)it}$	$MQ_{(bats)it}$	$MQ_{(turq)it}$
$const$	4.221***	4.54***	4.203***	4.087***	3.617***	3.765***	3.527***	3.093***
	(163.13)	(181.53)	(150.67)	(150.65)	(144.1)	(148.97)	(131.66)	(111.24)
HFT_{vit}	-0.369***	-0.381***	-0.314***	-0.326***	-0.301***	-0.309***	-0.238***	-0.247***
	(-227.36)	(-263.81)	(-193.47)	(-215.16)	(-195.45)	(-216.11)	(-151.72)	(-160.68)
$HHItrd_{it}$	0.02***	-0.014***	-0.099***	-0.109***	0.018***	-0.014***	-0.091***	-0.097***
	(9.31)	(-6.33)	(-40.41)	(-44.86)	(7.83)	(-5.91)	(-36.15)	(-37.68)
$\overline{MQ}_{-vit}$	0.174***	0.143***	0.24***	0.232***	-0.059***	-0.01***	0.149***	0.16***
	(51.33)	(48.51)	(82.97)	(83.41)	(-13.56)	(-2.7)	(43.66)	(43.37)
$inv(price)_{it}$	14.125***	13.192***	15.231***	16.904***	16.078***	16.317***	16.835***	21.897***
	(53.81)	(48.39)	(49.3)	(57.79)	(57.03)	(55.27)	(53.46)	(66.47)
$ln(mktcap)_{it}$	-0.136***	-0.163***	-0.157***	-0.139***	-0.173***	-0.186***	-0.189***	-0.142***
	(-59.59)	(-70.89)	(-60.61)	(-55.03)	(-72.39)	(-75.48)	(-72)	(-52.14)
$ln(voltintra)_{vit}$	0.035***	0.024***	0.033***	0.019***	0.031***	0.012***	0.027***	0.013***
	(37)	(45.86)	(72.63)	(50.51)	(36.92)	(25.23)	(62.5)	(35.08)
Panel B								
	$HFT_{(lse)it}$	$HFT_{(chix)it}$	$HFT_{(bats)it}$	$HFT_{(turq)it}$	$HFT_{(lse)it}$	$HFT_{(chix)it}$	$HFT_{(bats)it}$	$HFT_{(turq)it}$
$const$	7.711***	7.683***	6.998***	8.94***	7.323***	7.266***	5.594***	6.753***
	(199.54)	(169.58)	(106.49)	(165.22)	(176.9)	(170.08)	(88.76)	(106.82)
$MQ_{(lse)it}$	-0.078***	-0.184***	-0.336***	-0.639***	0.188***	-1.366***	-1.329***	-2.598***
	(-5.05)	(-11.65)	(-14.16)	(-31.7)	(4.57)	(-39.7)	(-26.33)	(-52.69)
$MQ_{(chix)it}$	-0.673***	-0.47***	-1.252***	-0.476***	-0.909***	0.272***	-0.918***	0.598***
	(-54.48)	(-28.07)	(-57.02)	(-24.91)	(-32.07)	(8.78)	(-21.55)	(13.98)
$MQ_{(bats)it}$	0.204***	0.07***	0.758***	0.447***	0.214***	0.079***	1.243***	0.478***
	(30.81)	(9.26)	(48.52)	(45.78)	(24.38)	(8.55)	(63.94)	(33.06)
$MQ_{(turq)it}$	0.118***	0.291***	0.37***	-0.214***	0.136***	0.541***	0.648***	0.886***
	(17.57)	(37.82)	(32.12)	(-16.36)	(13.15)	(47.71)	(36.96)	(37.51)
$\overline{HFT}_{-vit}$	0.318***	0.321***	0.393***	0.305***	0.316***	0.321***	0.408***	0.377***
	(145.44)	(128.84)	(125.62)	(113.83)	(127.62)	(136.3)	(128.72)	(121.32)
$HHItrd_{it}$	0.108***	0.045***	0.132***	0.009*	0.102***	0.14***	0.233***	0.224***
	(31.76)	(12.08)	(23.53)	(1.74)	(23.59)	(31.3)	(33.61)	(32.06)
$ln(mktcap)_{it}$	-0.118***	-0.139***	-0.169***	-0.311***	-0.087***	-0.16***	-0.106***	-0.236***
	(-27.21)	(-28.31)	(-25.32)	(-60.82)	(-14.63)	(-34.47)	(-15.11)	(-39.66)
$ln(volume)_{vit}$	0.457***	0.481***	0.43***	0.403***	0.492***	0.502***	0.49***	0.516***
	(160.28)	(169.02)	(128.73)	(139.82)	(165.8)	(209.32)	(160.74)	(153.96)
$ln(size)_{vit}$	-0.511***	-0.488***	-0.393***	-0.322***	-0.563***	-0.479***	-0.414***	-0.398***
	(-103.73)	(-93.43)	(-70.99)	(-62.56)	(-78.29)	(-98.48)	(-76.34)	(-66.04)
$rtick_{vit}$	-148.319***	-152.959***	-124.478***	-1.398***	-163.34***	-49.812***	-59.925***	-1.037***
	(-24.89)	(-23.38)	(-15.06)	(-4.87)	(-14.1)	(-6.87)	(-5.67)	(-3.22)
$ln(voltintra)_{vit}$	0.022***	-0.003**	-0.02***	0.001	0.02***	0	-0.033***	-0.02***
	(10.91)	(-2.38)	(-16.9)	(1.57)	(7.54)	(-0.1)	(-26.66)	(-21.78)
observations	147501	147501	147501	147501	147501	147501	147501	147501
second-stage adj_Rsqr (MQ_{vit})	0.80	0.82	0.79	0.82	0.78	0.78	0.75	0.78
second-stage adj_Rsqr (HFT_{vit})	0.90	0.88	0.79	0.82	0.89	0.88	0.79	0.78
system weighted Rsqr			0.8				0.79	

Table B.14: The cross-market impact of high frequency trading and market fragmentation on liquidity: a simultaneous equations model estimation for small stocks

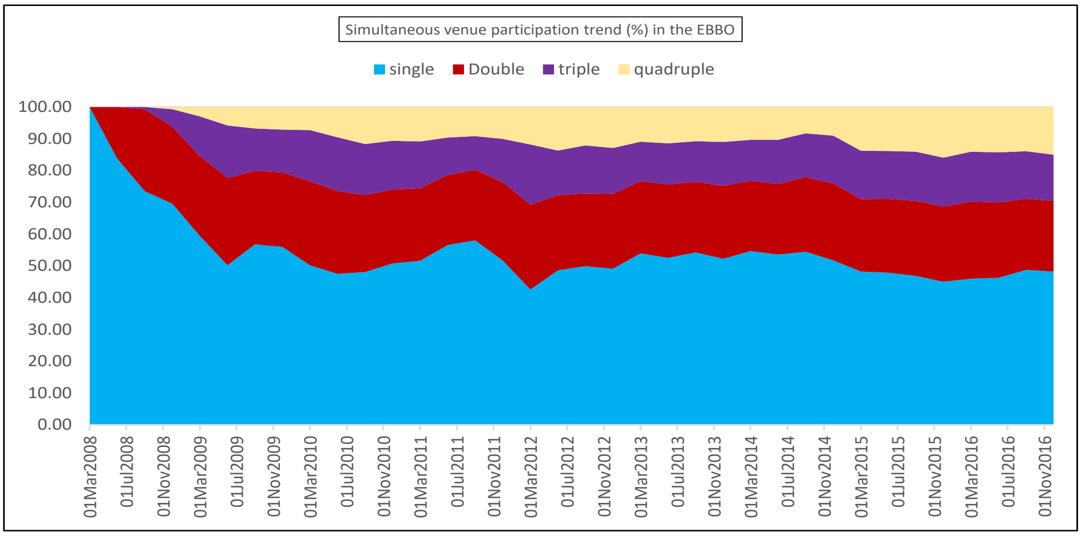
This table presents the simultaneous equations model estimation for the system of equations (3.1)–(3.8) using the three-stage least squares methods for two liquidity measures: log normalized time weighted quoted spreads and log normalized volume weighted effective half-spreads. Indices i and t represent stocks and time (days) respectively, v represents one of the four venues: LSE, CHIX, BATS and Turquoise, HFT_{vit} represents the HFT proxy ($hft2$) developed on quotes update upto the fifth depth level, $HHItrd_{it}$ represents the market fragmentation proxy, $\overline{MQ}_{-vit}$ represents the average liquidity level over all stocks in the same size group excluding stock i at venue v, $\overline{HFT}_{-vit}$ represents the average HFT intensity at venue v over all stocks in the same size group excluding stock i, $ln(mktcap)$ is the log normalized market capitalization, $ln(voltintra)_{vit}$ is the log normalized intraday mid price range volatility, $invprice$ is the inverse of daily average price, $ln(size)_{vit}$ is the log normalized trade size, $ln(value)_{vit}$ is the log normalized trading volume, rtk_{vit} is the relative tick size. The estimation is based on a panel dataset of 75 small-cap stocks (below the median market capitalization stocks group) and 2048 days (October 2008–December 2016) and includes both time (the monthly time dummy for each of 98 months included in the panel dataset) and stock fixed effects. Coefficient estimates are 3SLS , t-statistics shown in the parentheses below the coefficient. ***, **, * denote significance at 1%, 5%, and 10% levels, respectively. Panel A presents the estimates for market quality equations (3.1–3.4) and Panel B presents those for HFT equations (3.5–3.8).

	I	II	III	IV	V	VI	VII	VIII
	$MQ_{vit} = Log(quoted\ \ spreads)_{vit}$				$MQ_{(vit} = Log(effective\ \ half-spreads)_{vit}$			
	Panel A							
	$MQ_{(lse)it}$	$MQ_{(chix)it}$	$MQ_{(bats)it}$	$MQ_{(turq)it}$	$MQ_{(lse)it}$	$MQ_{(chix)it}$	$MQ_{(bats)it}$	$MQ_{(turq)it}$
$const$	4.333***	5.486***	4.398***	5.289***	3.311***	4.315***	3.771***	4.234***
	(124.24)	(134.9)	(91.27)	(115.87)	(97.39)	(111.59)	(80.82)	(97.37)
HFT_{vit}	-0.453***	-0.557***	-0.459***	-0.422***	-0.364***	-0.431***	-0.351***	-0.32***
	(-205.2)	(-227.73)	(-166.63)	(-152.8)	(-163.03)	(-178.48)	(-124.2)	(-114.83)
$HHItrd_{it}$	0.068***	-0.049***	-0.1***	-0.105***	0.071***	-0.063***	-0.12***	-0.111***
	(31.69)	(-18.14)	(-32.15)	(-35.5)	(31.8)	(-23.36)	(-37.87)	(-36.96)
$\overline{MQ}_{-vit}$	0.075***	0.119***	0.212***	0.166***	0.016***	0.069***	0.158***	0.121***
	(20.56)	(41.79)	(61.83)	(48.86)	(3.83)	(21.26)	(42.87)	(32.28)
$inv(price)_{it}$	10.013***	9.285***	13.389***	9.465***	13.21***	11.921***	15.391***	12.68***
	(29.04)	(20.96)	(25.5)	(18.8)	(35.38)	(26.41)	(28.64)	(25.19)
$ln(mktcap)_{it}$	-0.171***	-0.212***	-0.162***	-0.27***	-0.202***	-0.242***	-0.22***	-0.288***
	(-49.04)	(-49.97)	(-33.12)	(-57.68)	(-55.52)	(-57.28)	(-44.07)	(-61.44)
$ln(voltintra)_{vit}$	0.097***	0.028***	0.034***	0.028***	0.078***	0.021***	0.034***	0.02***
	(70.88)	(53.87)	(59.09)	(47.76)	(58.07)	(42.89)	(56.18)	(34.67)
	Panel B							
	$HFT_{(lse)it}$	$HFT_{(chix)it}$	$HFT_{(bats)it}$	$HFT_{(turq)it}$	$HFT_{(lse)it}$	$HFT_{(chix)it}$	$HFT_{(bats)it}$	$HFT_{(turq)it}$
$const$	6.245***	5.603***	6.025***	5.471***	5.633***	5.151***	4.566***	4.121***
	(114.4)	(85.62)	(76.31)	(71.03)	(102.54)	(83.47)	(57.29)	(52.82)
$MQ_{(lse)it}$	-0.114***	-0.377***	-0.68***	-0.735***	0.325***	-0.811***	-1.232***	-1.593***
	(-5.61)	(-22.62)	(-34.61)	(-38.59)	(9.79)	(-34.23)	(-47.37)	(-61.43)
$MQ_{(chix)it}$	-0.41***	0.112***	-0.563***	-0.343***	-0.669***	0.361***	-0.645***	-0.202***
	(-49.55)	(7.19)	(-42.12)	(-27.1)	(-49.46)	(18.91)	(-35.27)	(-11.77)
$MQ_{(bats)it}$	0.151***	0.079***	0.326***	0.223***	0.255***	0.138***	0.853***	0.427***
	(28.59)	(11.43)	(24.83)	(27.27)	(35.07)	(16.41)	(52.23)	(41.78)
$MQ_{(turq)it}$	0.197***	0.318***	0.54***	0.649***	0.269***	0.517***	0.778***	1.242***
	(30.24)	(37.19)	(51.32)	(43.38)	(29.12)	(47.31)	(55.84)	(65.32)
$\overline{HFT}_{-vit}$	0.253***	0.378***	0.414***	0.406***	0.286***	0.401***	0.507***	0.471***
	(75.77)	(100.3)	(93.86)	(94.98)	(67.89)	(104.61)	(107.11)	(102.93)
$HHItrd_{it}$	0.171***	0.105***	0.181***	0.152***	0.159***	0.2***	0.325***	0.348***
	(41.7)	(21.25)	(29.48)	(25.79)	(29.5)	(33.61)	(43.67)	(47.75)
$ln(mktcap)_{it}$	-0.058***	-0.072***	-0.194***	-0.143***	0.032***	-0.054***	-0.154***	-0.087***
	(-8.89)	(-10.44)	(-23.86)	(-17.53)	(3.57)	(-7.86)	(-17.71)	(-10.13)
$ln(volume)_{vit}$	0.486***	0.484***	0.34***	0.422***	0.56***	0.502***	0.386***	0.455***
	(96.66)	(119.98)	(92.23)	(101.81)	(97.4)	(146.9)	(107.23)	(114.38)
$ln(size)_{vit}$	-0.476***	-0.463***	-0.326***	-0.385***	-0.557***	-0.459***	-0.349***	-0.415***
	(-78.2)	(-73.8)	(-63.54)	(-65.05)	(-69.74)	(-80.65)	(-68.8)	(-70.17)
$rtick_{vit}$	-171.337***	-175.994***	-2.463***	-1.273***	-270.638***	-182.192***	-2.613***	-1.162***
	(-45.1)	(-45.87)	(-6.52)	(-11.77)	(-37.32)	(-40.91)	(-6.57)	(-10.13)
$ln(voltintra)_{vit}$	0.028***	-0.018***	-0.012***	-0.023***	-0.016***	-0.022***	-0.026***	-0.035***
	(7.36)	(-18.69)	(-12.91)	(-23.87)	(-2.74)	(-23.17)	(-25.35)	(-33.5)
observations	127864	127864	127864	127864	127864	127864	127864	127864
second-stage adj_Rsqr (MQ_{vit})	0.70	0.76	0.70	0.72	0.66	0.71	0.6	0.68
second-stage adj_Rsqr (HFT_{vit})	0.80	0.74	0.58	0.60	0.78	0.74	0.53	0.55
system weighted Rsqr		0.7				0.63		

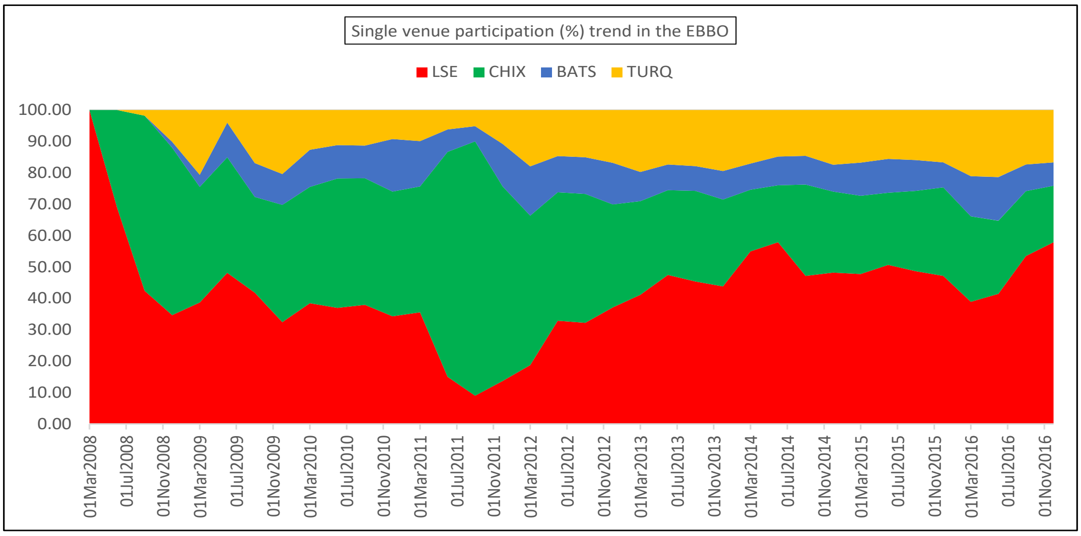
Table B.15: The cross-market time-varying impact of high frequency trading and market fragmentation on liquidity: a simultaneous equations model estimation

This table presents the simultaneous equations model estimation for the system of equations (3.1)–(3.8) using the three-stage least squares methods for two liquidity measures: log normalized time weighted quoted spreads and log normalized volume weighted effective half-spreads. Indices i and t represent stocks and time (days) respectively, v represents one of the four venues: LSE, CHIX, BATS and Turquoise, HFT_{vit} represents the HFT proxy ($hft2$) developed on quotes update upto the fifth depth level, $HHItrd_{it}$ represents the market fragmentation proxy. To conserve space, coefficients for $\overline{MQ_{-vit}}$, $\overline{HFT_{-vit}}$, $ln(mktcap)$, $ln(voltintra)_{vit}$, $ln(size)_{vit}$, $ln(value)_{vit}$, and rtk_{vit} are not presented. Estimations are based on three subsamples (2008–2010, 2011–2013 and 2014–2016) divided over the sample period with 149 stock each and include both time (the monthly time dummy for each months) and stock fixed effects. Coefficient estimates are 3SLS , t-statistics shown in the parentheses below the coefficient. ***, **, * denote significance at 1%, 5%, and 10% levels, respectively. Panel A presents the estimates for market quality equations (3.1–3.4) and Panel B presents those for HFT equations (3.5–3.8).

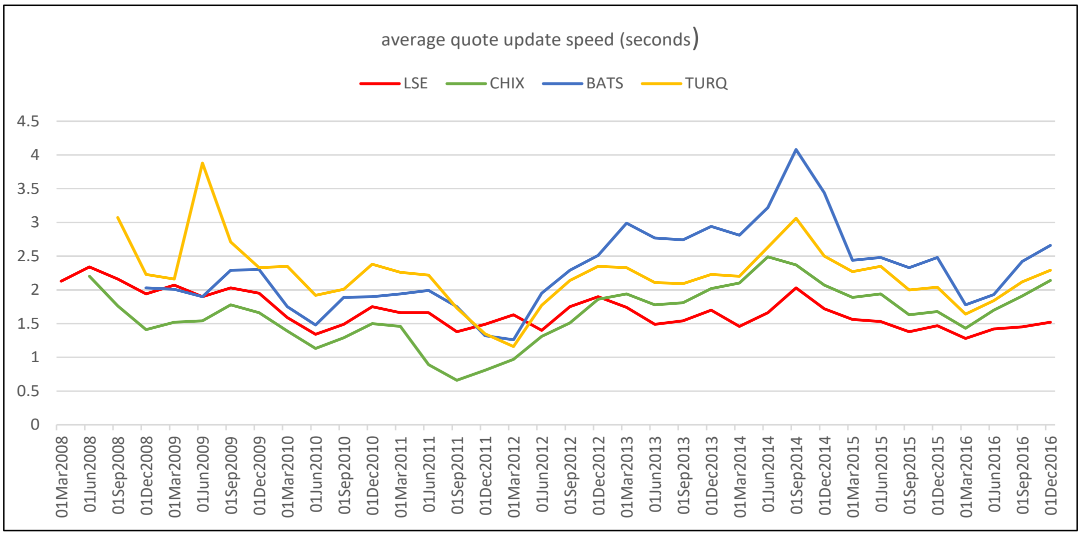
	I	II	III	IV	V	VI	VII	VIII
	$MQ_{vit} = Log(quoted\quad spreads)_{vit}$				$MQ_{(vit} = Log(effective\quad half-spreads)_{vit}$			
	Panel A							
	$MQ_{(lse)it}$	$MQ_{(chix)it}$	$MQ_{(bats)it}$	$MQ_{(turq)it}$	$MQ_{(lse)it}$	$MQ_{(chix)it}$	$MQ_{(bats)it}$	$MQ_{(turq)it}$
	2008-2010							
HFT_{vit}	-0.318***	-0.371***	-0.34***	-0.238***	-0.254***	-0.288***	-0.261***	-0.152***
	(-117.89)	(-146.07)	(-120.38)	(-80.5)	(-88.02)	(-110.07)	(-90.27)	(-50.96)
$HHItrd_{it}$	0.102***	0.034***	-0.047***	-0.096***	0.101***	0.008**	-0.056***	-0.098***
	(32.1)	(9.88)	(-11.07)	(-22.58)	(29.13)	(2.14)	(-12.74)	(-22.29)
	2011-2013							
HFT_{vit}	-0.288***	-0.366***	-0.322***	-0.295***	-0.234***	-0.265***	-0.235***	-0.227***
	(-147.18)	(-155.96)	(-115.06)	(-103.92)	(-108.78)	(-112.89)	(-80.68)	(-77.5)
$HHItrd_{it}$	0.064***	-0.062***	-0.097***	-0.114***	0.065***	-0.066***	-0.115***	-0.107***
	(30.97)	(-24.35)	(-30.05)	(-35.07)	(28.83)	(-25.68)	(-34.77)	(-32.35)
	2014-2016							
HFT_{vit}	-0.405***	-0.418***	-0.347***	-0.394***	-0.306***	-0.326***	-0.26***	-0.312***
	(-176.55)	(-207.51)	(-152.09)	(-202.31)	(-155.86)	(-164.86)	(-115.52)	(-161.38)
$HHItrd_{it}$	0.027***	-0.046***	-0.11***	-0.083***	0.016***	-0.058***	-0.117***	-0.09***
	(11.95)	(-18.79)	(-39.99)	(-34.44)	(7.27)	(-22.58)	(-41.65)	(-35.99)
	Panel B							
	$HFT_{(lse)it}$	$HFT_{(chix)it}$	$HFT_{(bats)it}$	$HFT_{(turq)it}$	$HFT_{(lse)it}$	$HFT_{(chix)it}$	$HFT_{(bats)it}$	$HFT_{(turq)it}$
	2008-2010							
$MQ_{(lse)it}$	0.152***	0.125***	0.387***	-0.732***	0.858***	-0.204***	0.202***	-1.498***
	(4.57)	(4.13)	(9.23)	(-17.99)	(11.67)	(-4.44)	(3.17)	(-22.55)
$MQ_{(chix)it}$	-1.075***	-0.848***	-1.617***	-0.86***	-1.953***	-0.886***	-2.415***	-1.08***
	(-39.52)	(-23.7)	(-37.47)	(-20.91)	(-37.96)	(-16.68)	(-36.2)	(-16.1)
$MQ_{(bats)it}$	0.39***	-0.002	-0.002	0.546***	0.715***	0.036*	0.654***	0.921***
	(29.3)	(-0.1)	(-0.06)	(27.1)	(32.53)	(1.67)	(16.14)	(28.56)
$MQ_{(turq)it}$	0.235***	0.46***	0.576***	-0.041**	0.209***	0.773***	0.977***	0.865***
	(23.12)	(40.03)	(36.66)	(-2.05)	(12.4)	(45.59)	(39.21)	(26.1)
	2011-2013							
$MQ_{(lse)it}$	0.611***	-0.225***	-0.927***	-0.495***	0.696***	-0.797***	-1.442***	-0.876***
	(23.67)	(-10.63)	(-35.84)	(-19.6)	(19.67)	(-23.91)	(-40.5)	(-23.6)
$MQ_{(chix)it}$	-0.474***	-0.059***	-0.065***	-0.015	-0.756***	0.366***	0.045	0.099***
	(-32.49)	(-2.77)	(-3.35)	(-0.79)	(-35.45)	(12.47)	(1.59)	(3.62)
$MQ_{(bats)it}$	0.067***	0.097***	0.688***	0.053***	0.138***	0.151***	1.035***	0.13***
	(8.93)	(12.5)	(47.26)	(5.52)	(15.48)	(15.6)	(63.52)	(11.13)
$MQ_{(turq)it}$	0.203***	0.23***	0.203***	0.416***	0.316***	0.379***	0.325***	0.689***
	(25.47)	(27.75)	(19.15)	(33.38)	(31.03)	(34.35)	(24.83)	(44.98)
	2014-2016							
$MQ_{(lse)it}$	-0.058***	-0.105***	-0.375***	-0.161***	0.967***	-1.167***	-1.705***	-1.133***
	(-4.73)	(-7.09)	(-22.07)	(-10.27)	(25.1)	(-36.46)	(-54.34)	(-34.95)
$MQ_{(chix)it}$	-0.284***	0.4***	-0.421***	-0.312***	-0.584***	1.139***	0.092***	-0.04**
	(-37.09)	(26.65)	(-31.46)	(-26.52)	(-35.8)	(48.66)	(4.35)	(-2.24)
$MQ_{(bats)it}$	0.154***	0.07***	0.021*	0.153***	0.328***	0.218***	0.728***	0.349***
	(29.89)	(9.95)	(1.84)	(20.63)	(35.64)	(22.49)	(42.94)	(34.51)
$MQ_{(turq)it}$	0.025***	-0.233***	0.011	0.363***	-0.048***	-0.045***	0.257***	0.892***
	(3.75)	(-25.05)	(0.99)	(27.61)	(-3.85)	(-3.25)	(15.21)	(46.14)



(a) The joint venue participation rate (% EBBO)



(b) The unique venue participation rate (% EBBO)



(c) The quotes update speed

Fig. B.1: Cross-market trends of quotes update speed and venue participation rate in the EBBO

Figure 5: *After one partial order of the investor hits Chi-X and leads to a transaction, the co-located HFT then reacts by cancelling duplicate orders on other trading venues. Because HFT have invested in ultrafast connections to trading venues, these cancellations arrive at these trading venues before the remaining partial orders of the investor do.*

(a) Cross-market quote updating

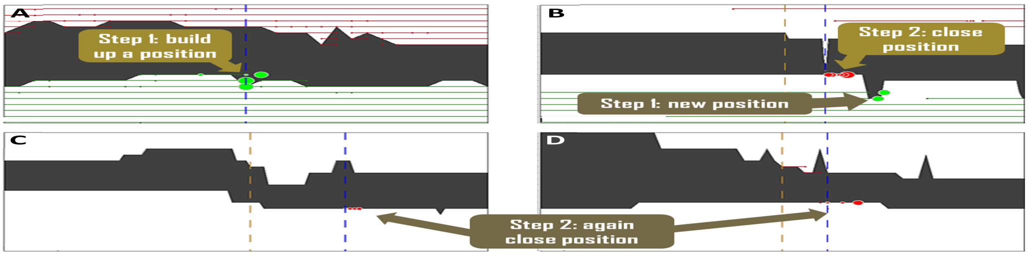


Figure 7a to 7d: *Cross-market trading visualization of one HFT firm. Each box shows the trading conduct of this HFT on one trading venue. The horizontal axis denotes time and the vertical axis the price (axis values not shown). The green/red bars represent buy/sell orders from start (left-hand side) to end (right-hand side), whereas the green/red dots represent buy/sell transactions. Larger dots represent larger sized transactions. We only show orders and transactions for the one HFT. The grey area represents the spread for the entire market. The blue, vertical lines in each box represent the time at which the HFT performs its first transaction on that specific trading venue. The orange vertical lines, on the other hand, represent the time of the first transaction over all trading venues (i.e., the first signal it can react to).*

(b) Cross-market trading visualization

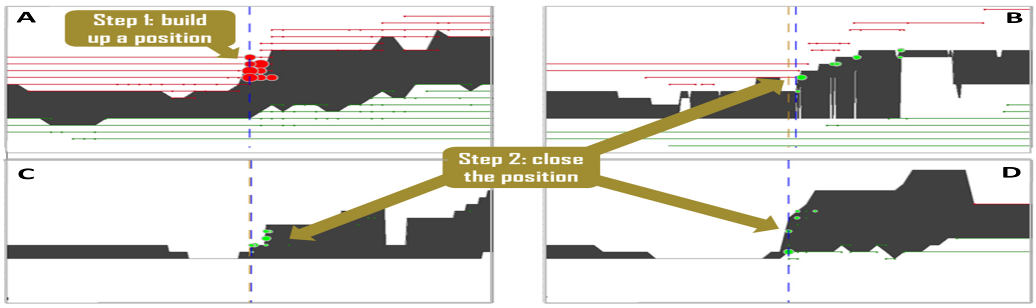


Figure 8a to 8d: *A more typical example (compared to the a-typical one in 7a to 7d) of HFT trading conduct. In Figure 8a the HFT passively builds up a position with the large investor being the counterparty. It then closes the position aggressively on other trading venues (Figures 8b to 8d), typically earning a few cents profit per share. There were no additional trades with the incoming (partial) orders of the investor.*

(c) Cross-market positioning

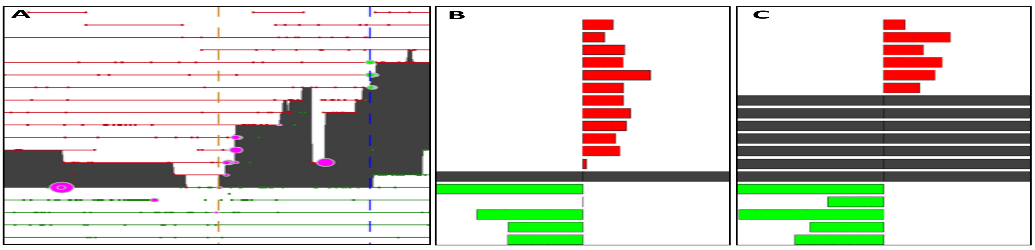
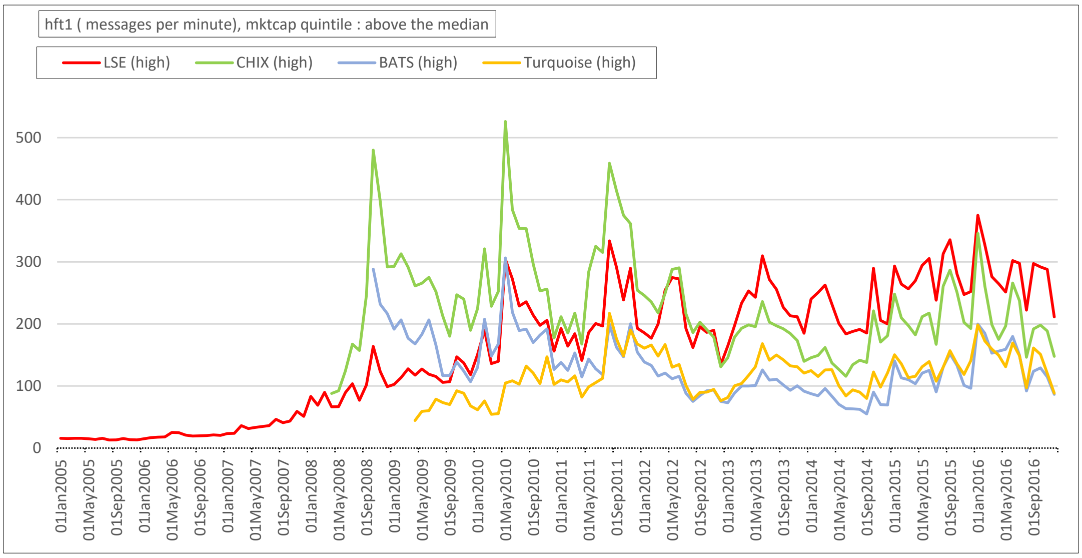


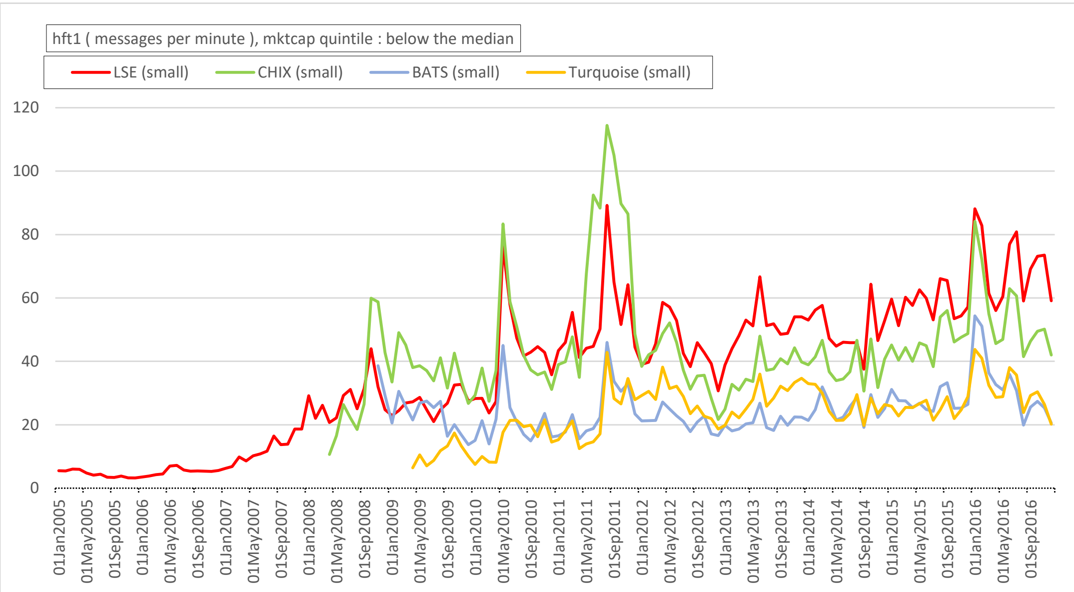
Figure 9a to 9c: *Figure 9a illustrates all orders and trades on one trading venue. The orange dotted line represents the time at which the first partial order of the investor is matched on another trading venue. The blue dotted line represents the exact time when the partial order of the investor hits this specific trading venue. The red/green bars represent sell/buy orders, from begin (left-hand side) to end (right-hand side). The purple dots represent trades by firms other than the investor, whereas the green dots represent the buy trades of the investor. Figure 9b and 9c respectively represent the order book during the orange dotted line and blue dotted line. Each bar represents the volume on a particular price level. The vertical axis denotes the price and the horizontal axis the volume. Red/green bars represent sell/buy liquidity, whereas the grey area represents the spread (which is similar to the grey area in Figure 9a).*

(d) Cross-market order matching time

Fig. B.2: A typical HFT firm's market making across markets (source: The Netherlands Authority for the Financial Markets (2016))

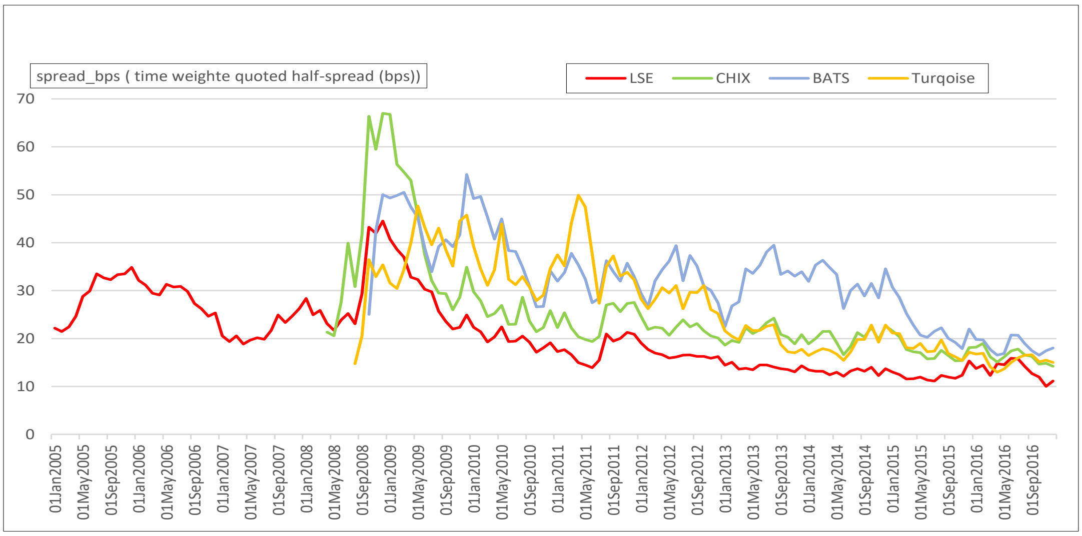


(a) Large stocks

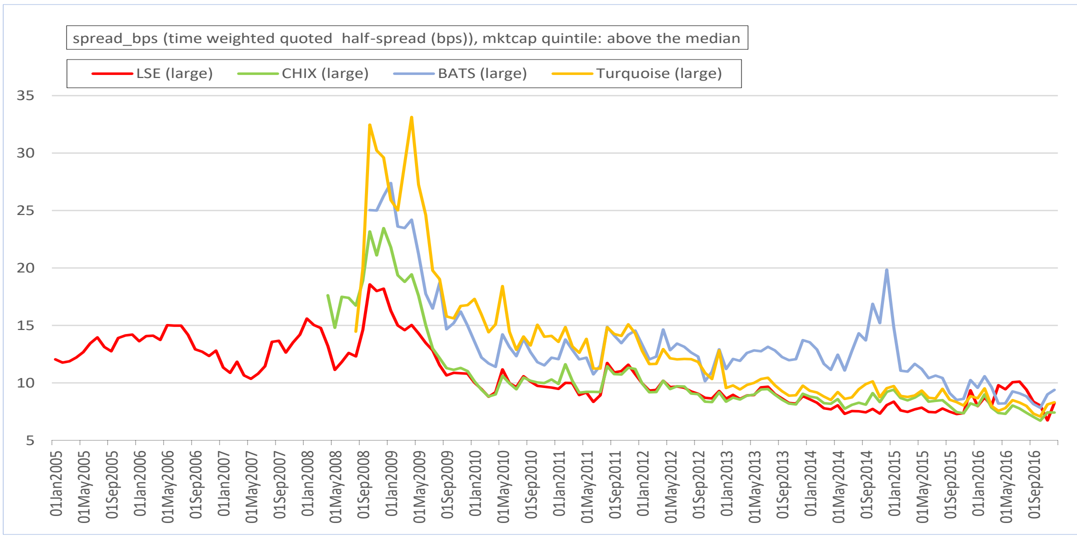


(b) Small stocks

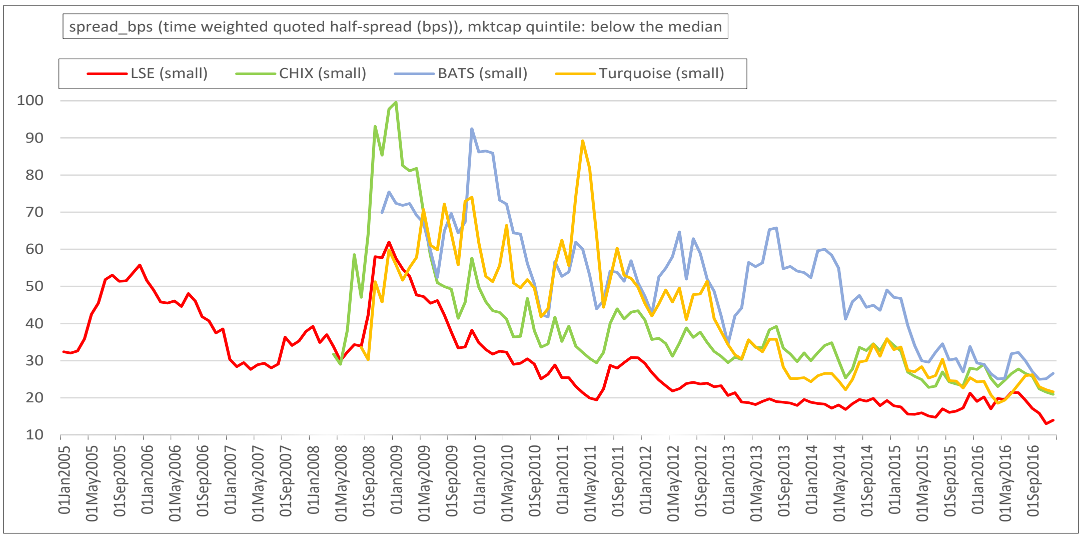
Fig. B.3: Cross market trends in average electronic message rate per-minute (for the best 10 depth levels)



(a) All stocks

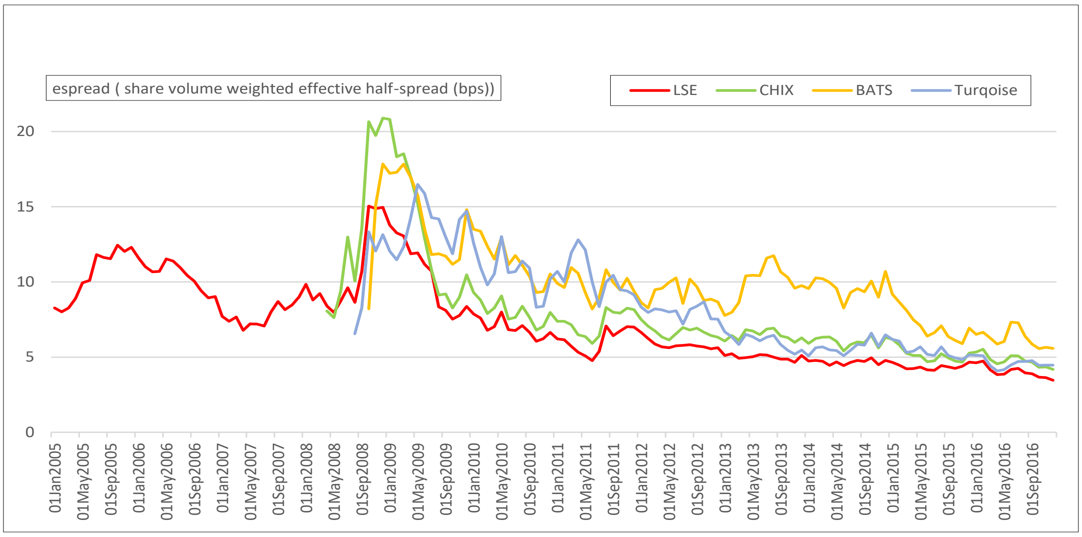


(b) Large stocks

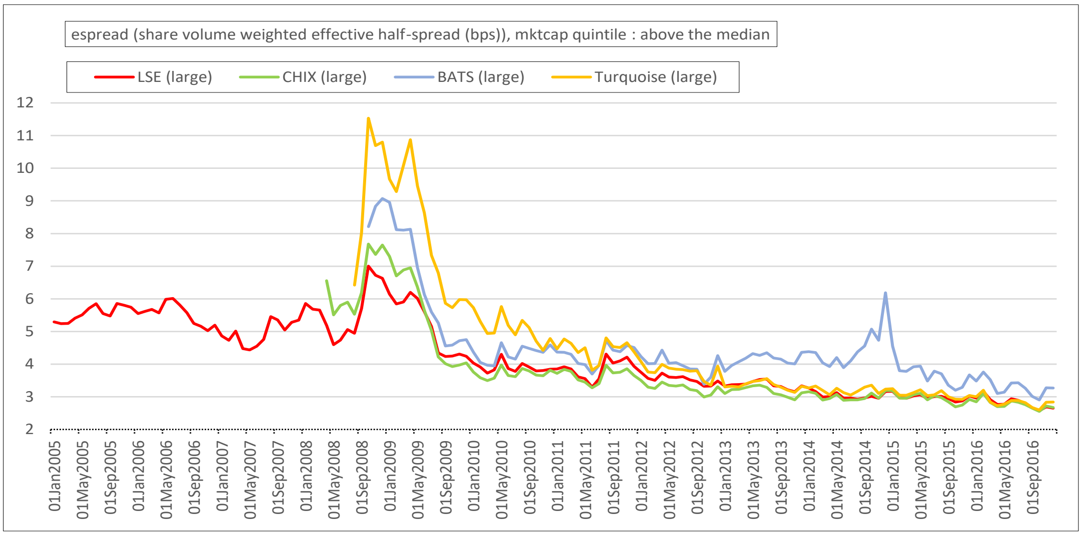


(c) Small stocks

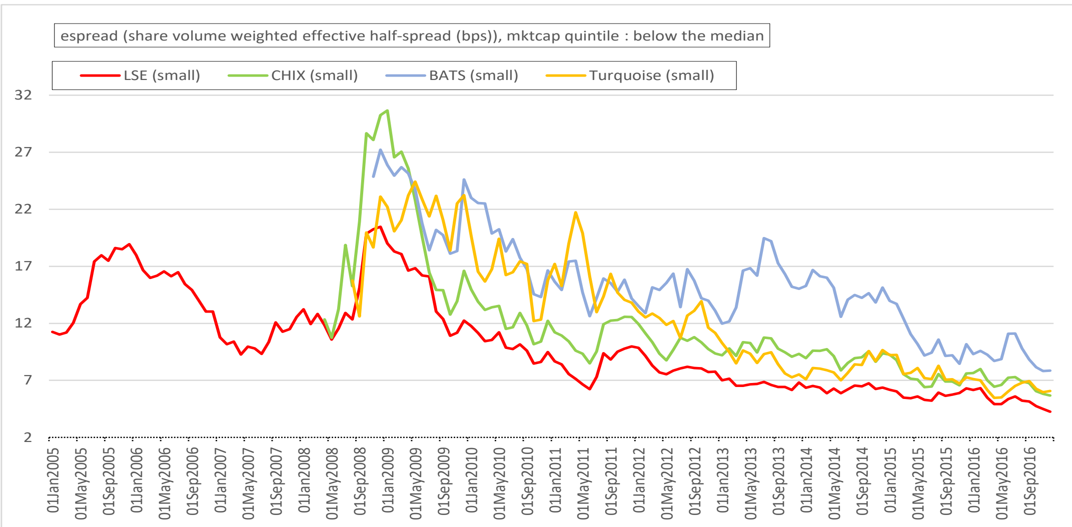
Fig. B.4: Trends in time weighted quoted spreads across markets



(a) All stocks

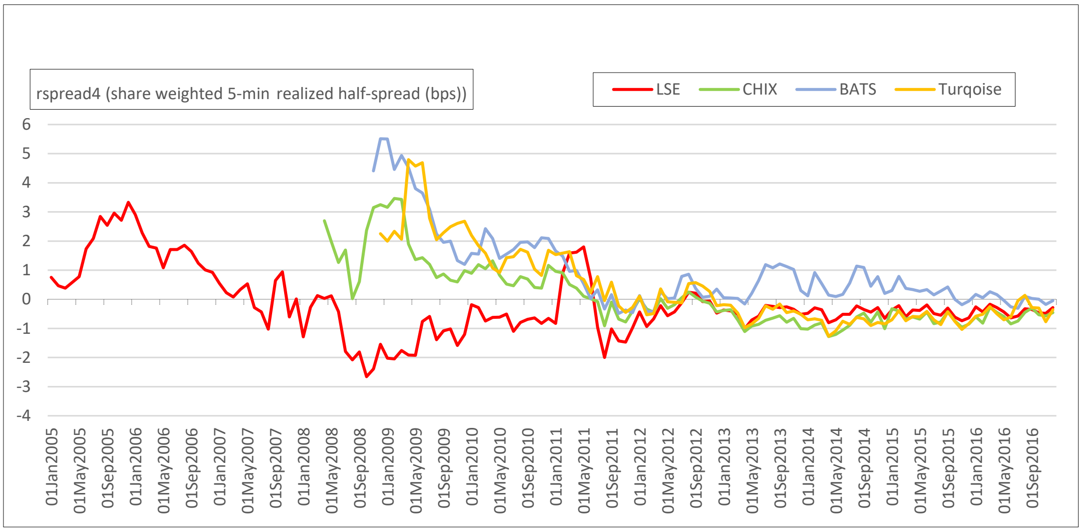


(b) Large stocks

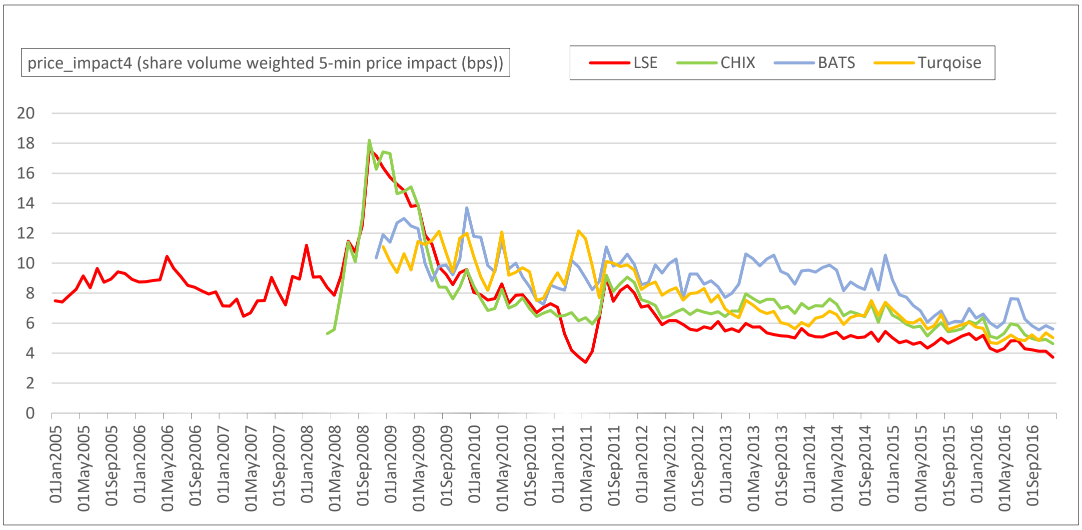


(c) Small stocks

Fig. B.5: Trends in volume weighted effective-half spreads across markets

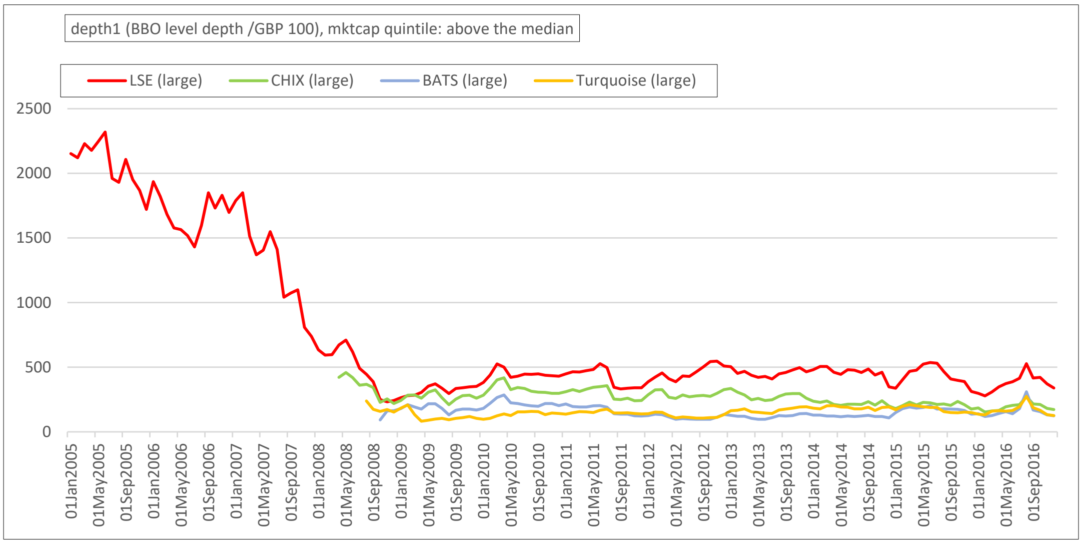


(a) realized half-spreads

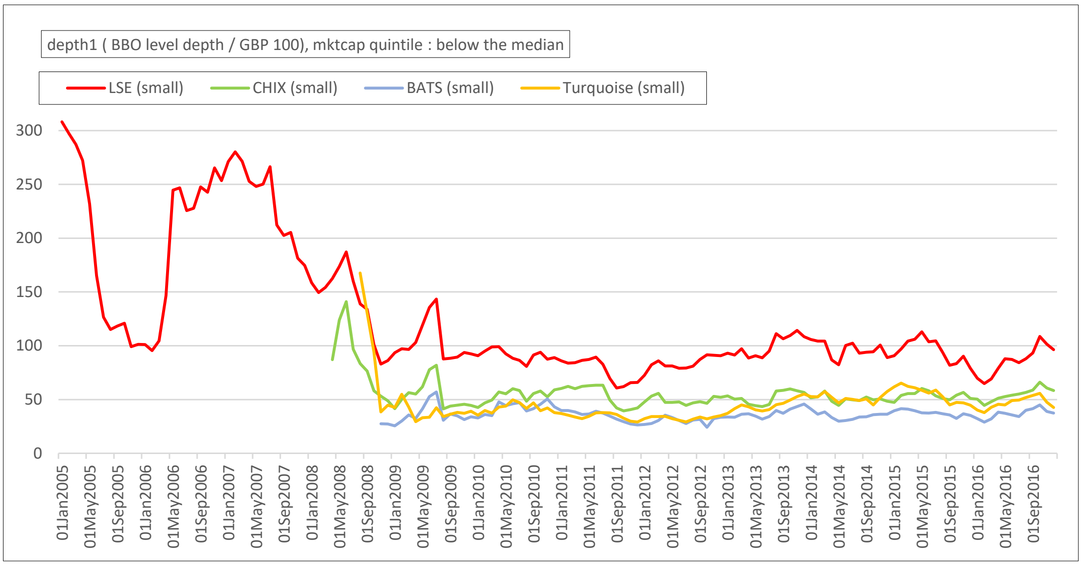


(b) price impacts

Fig. B.6: Trends in 5-minute realized half-spreads and price impacts across markets

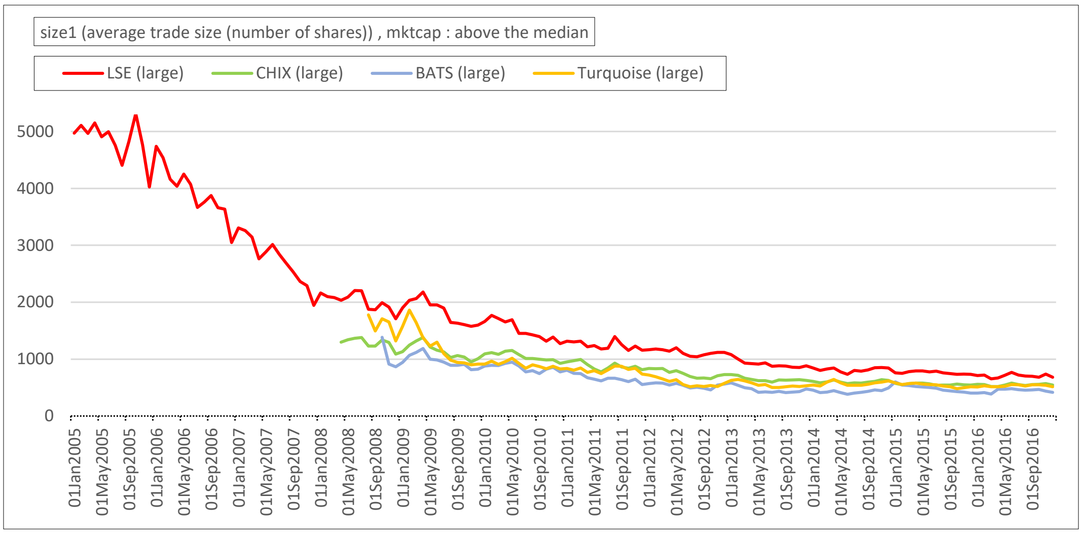


(a) Large stocks

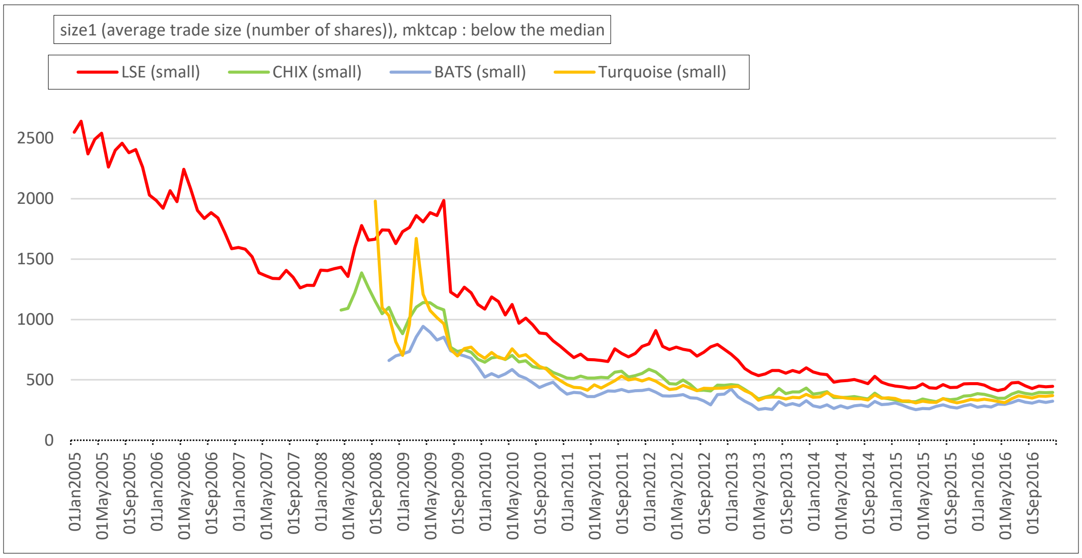


(b) Small stocks

Fig. B.7: Trends in average quoted depths (GBP100) at best limit price across markets

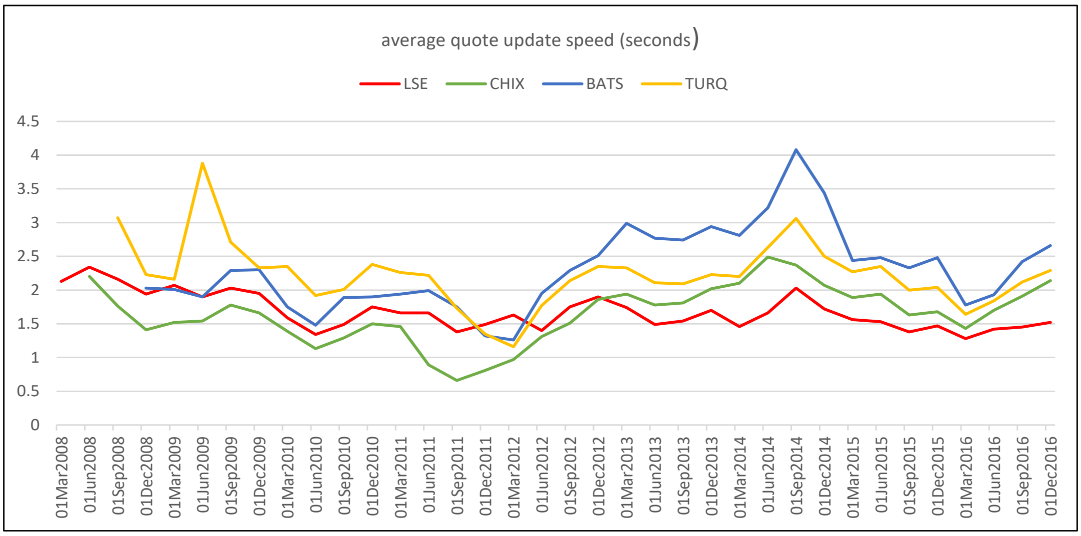


(a) Large stocks

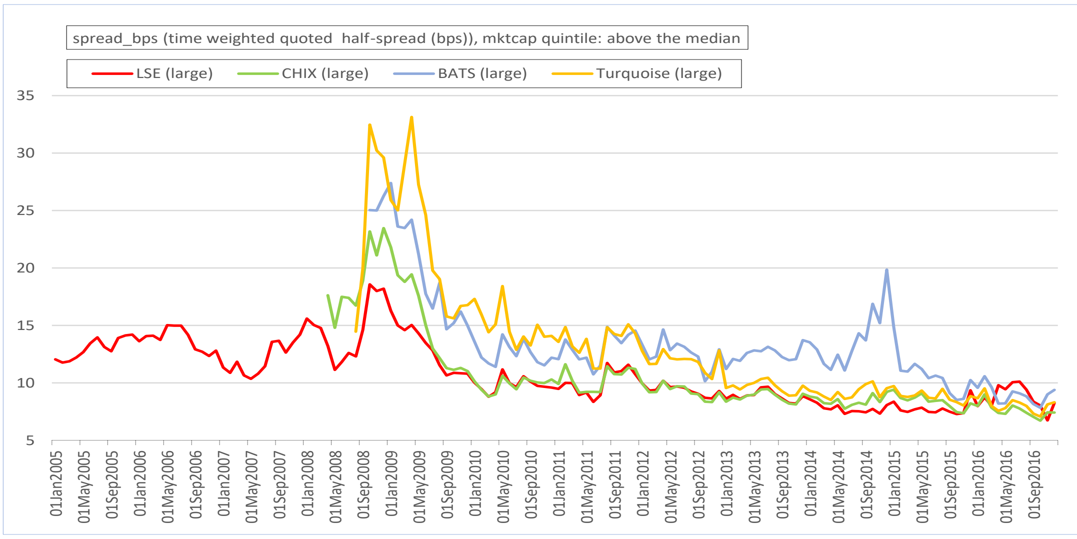


(b) Small stocks

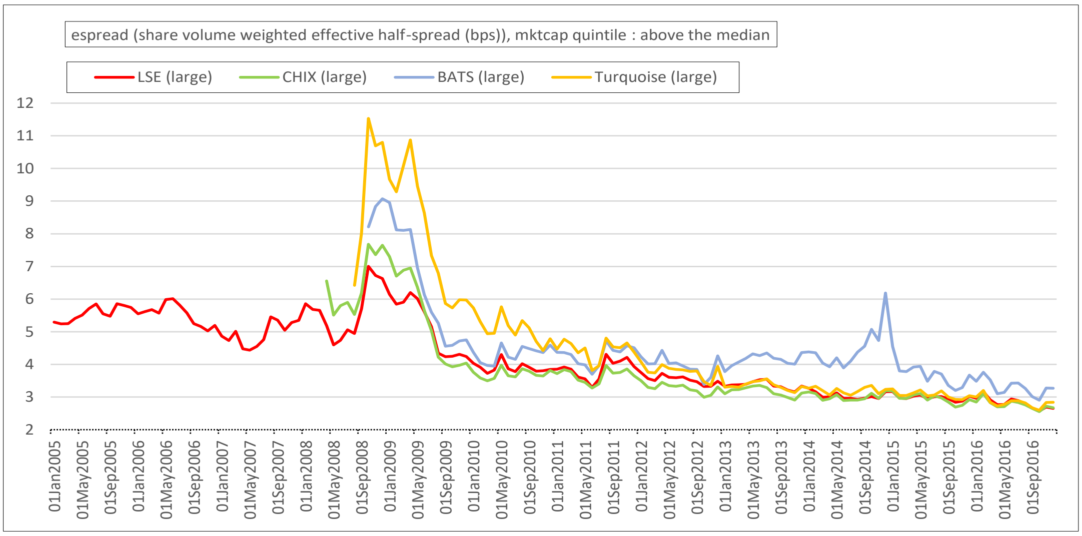
Fig. B.8: Trends in average trade sizes (number of shares)



(a) Trends in quotes update speed



(b) The average quoted spread (large stocks)



(c) The average effective half-spreads (large stocks)

Fig. B.9: Cross-market trends in speed competition and quoted spreads

References

Ait-Sahalia, Y. and Saglam, M. (2017). High frequency market making: Implications for liquidity.
URL: *http://dx.doi.org/10.2139/ssrn.2908438*

Aitken, M., Cumming, D. and Zhan, F. (2014). Trade size, high-frequency trading, and colocation around the world, *The European Journal of Finance* pp. 1–21.

Aitken, M. J., Harris, D. and Harris, F. H. D. (2015). Fragmentation and Algorithmic Trading: Joint impact on Market Quality.
URL: *https://dx.doi.org/10.2139/ssrn.2587314*

Aldridge, I. (2013). *High Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading System*, 2nd edn, John Wiley & Sons, Inc.

Biais, B. and Foucault, T. (2014). HFT and market quality, *Bankers, Markets and Investors* (128): 5–19.

Biais, B., Foucault, T. and Moinas, S. (2011). Equilibrium high frequency trading. AFA 2013 San Diego Meetings Paper; HEC Paris Research Paper No. 968/2013.
URL: *http://dx.doi.org/10.2139/ssrn.2024360*

Boehmer, E., Fong, K. Y. L. and Wu, J. (2015). International evidence on algorithmic trading. AFA 2013 San Diego Meetings Paper.
URL: *http://dx.doi.org/10.2139/ssrn.2022034*

Brogaard, J., Hagströmer, B., Nordén, L. and Riordan, R. (2015). Trading fast and slow: colocation and liquidity, *Review of Financial Studies* **28**(12): 3407–3443.

Brogaard, J., Hendershott, T., Hunt, S. and Ysusi, C. (2014). High-frequency trading and the execution costs of institutional investors, *Financial Review* **49**(2): 345–369.

Brogaard, J., Hendershott, T. and Riordan, R. (2014a). High-frequency trading and price discovery, *Review of Financial Studies* **27**(8): 2267–2306.

Brogaard, J., Hendershott, T. and Riordan, R. (2014b). Market integration and high frequency intermediation.
URL: *http://dx.doi.org/10.2139/ssrn.2908438*

Budish, E., Cramton, P. and Shim, J. (2015). The high-frequency trading arms race: Frequent batch auctions as a market design response, *The Quarterly Journal of Economics* **130**(4): 1547–1621.

Buti, S., Rindi, B. and Werner, I. M. (2011). Dividing into dark pools. Charles A. Dice Center Working Paper No. 2010-10; Fisher College of Business Working Paper No. 2010-03-010.
URL: *http://dx.doi.org/10.2139/ssrn.1630499*

Cantillon, E. and Yin, P. L. (2011). Competition between exchanges: A research agenda, *International Journal of Industrial Organization* **29**(3): 329–336.

Carrion, A. (2013). Very fast money: High-frequency trading on the NASDAQ, *Journal of Financial Markets* **16**(4): 680–711.

CFA Institute (2011). The structure, regulation, and transparency of European equity markets under MiFID, *Technical report*.
URL: *http://www.cfapubs.org/doi/pdf/10.2469/ccb.v2011.n3.1*

Conrad, J., Wahal, S. and Xiang, J. (2015). High-frequency quoting, trading, and the efficiency of prices, *Journal of Financial Economics* **116**(2): 271–291.

Degryse, H., De Jong, F. and Kervel, V. V. (2015). The impact of dark trading and visible fragmentation on market quality, *Review of Finance* **19**(4): 1587–1622.

European Securities and Markets Authority (2014). High-frequency trading activity in EU equity markets. ESMA economic report, Number 1.
URL: *https://www.esma.europa.eu/sites/default/files/library/2015/11/esma20141_-_hft_activity_in_eu_equity_markets.pdf*

Financial Markets Regulator (France) (2017). Study of the behaviour of the high frequency traders on Euronext, *Risk and Trends Mapping* (January).

Foucault, T. and Menkveld, A. J. (2008). Competition for order flow and Smart order routing systems, *The Journal of Finance* **63**(1): 119–158.

Foucault, T., Pagano, M. and Roell, A. (2013). *Market Liquidity: Theory, Evidence, and Policy*, Oxford University Press.

Friederich, S. and Payne, R. (2015). Order-to-trade ratios and market liquidity, *Journal of Banking & Finance* **50**: 214–223.

Frino, A., Mollica, V., Monaco, E. and Palumbo, R. (2017). The effect of algorithmic trading on market liquidity: Evidence around earnings announcements on Borsa Italiana, *Pacific Basin Finance Journal* **45**: 82–90.

Frino, A., Mollica, V. and Webb, R. I. (2014). The impact of co-location of security exchanges' and traders' computer servers on market liquidity, *The Journal of Futures Markets* **34**(1): 20–33.

Goettler, R. L., Parlour, C. A. and Rajan, U. (2009). Informed traders and limit order markets, *Journal of Financial Economics* **93**(1): 67–87.

Gomber, P., Arndt, B., Lutat, M. and Uhle, T. (2011). High-frequency Trading.
URL: *http://dx.doi.org/10.2139/ssrn.1858626*

Greene, W. (2003). *Econometric Analysis*, 5th edn, Prentice Hall.

Gresse, C. (2017). Effects of lit and dark market fragmentation on liquidity, *Journal of Financial Markets* **35**: 1–20.

Hagströmer, B. and Nordén, L. (2013). The diversity of high-frequency traders, *Journal of Financial Markets* **16**(4): 741–770.

Han, J., Khapko, M. and Kyle, A. S. (2014). Liquidity with high-frequency market making. Swedish House of Finance Research Paper No 14-06 Liquidity.
URL: *https://dx.doi.org/10.2139/ssrn.2416396*

Hasbrouck, J. and Saar, G. (2013). Low-latency trading, *Journal of Financial Markets* **16**(4): 646–679.

He, P. W., Jarnecic, E. and Liu, Y. (2015). The determinants of alternative trading venue market share: Global evidence from the introduction of chi-X, *Journal of Financial Markets* **22**(2015): 27–49.

Hendershott, T., Jones, C. M. and Menkveld, A. J. (2011). Does algorithmic trading improve liquidity?, *The Journal of Finance* **66**(1): 1–34.

Hendershott, T. and Moulton, P. C. (2011). Automation, speed, and stock market quality: The NYSE's Hybrid, *Journal of Financial Markets* **14**(4): 568–604.

Hendershott, T. and Riordan, R. (2013). Algorithmic trading and the market for liquidity, *Journal of Financial and Quantitative Analysis* **48**(4): 1001–1024.

Jarnecic, E. and Snape, M. (2014). The provision of liquidity by high-frequency participants, *The Financial Review* **49**: 371–394.

Jovanovic, B. and Menkveld, A. J. (2015). Middlemen in limit order markets.
URL: *https://dx.doi.org/10.2139/ssrn.1624329*

Kirilenko, A., Kyle, A. S., Samadi, M. and Tuzun, T. (2017). The flash crash: High-frequency trading in an electronic market, *Journal of Finance* **72**(3): 967–998.

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

Li, W. (2014). High frequency trading with speed hierarchies.
URL: *https://dx.doi.org/10.2139/ssrn.2365121*

Madhavan, A. (2012). Exchange-traded funds, market structure, and the flash crash, *Financial Analysts Journal* **68**(4): 20–35.

Menkveld, A. J. (2013). High frequency trading and the new market makers, *Journal of Financial Markets* **16**(4): 712–740.

Menkveld, A. J. (2014). High-frequency traders and market structure, *The Financial Review* **49**: 333–344.

Menkveld, A. J. (2016). The economics of high-frequency trading: Taking stock, *Annual Review of Financial Economics* **8**: 1–24.

Murray, H., Pham, T. P. and Singh, H. (2016). Latency reduction and market quality: The case of the Australian Stock Exchange, *International Review of Financial Analysis* **46**: 257–265.

O'Hara, M. (2015). High frequency market microstructure, *Journal of Financial Economics* **116**(2): 257–270.

O'Hara, M., Saar, G. and Zhong, Z. (2014). Relative tick size and the trading environment.
URL: *https://dx.doi.org/10.2139/ssrn.2463360*

O'Hara, M. and Ye, M. (2011). Is market fragmentation harming market quality?, *Journal of Financial Economics* **100**(3): 459–474.

Riordan, R. and Storkenmaier, A. (2012). Latency, liquidity and price discovery, *Journal of Financial Markets* **15**(4): 416–437.

Riordan, R., Storkenmaier, A. and Wagener, M. (2011). Do multilateral trading facilities contribute to market quality?
URL: *https://dx.doi.org/10.2139/ssrn.1852769*

The Netherlands Authority for the Financial Markets (2010). High frequency trading: The application of advanced trading technology in the European marketplace. AFM report on High Frequency Trading (HFT).
URL: *https://www.afm.nl/ /profmedia/files/rapporten/2010/hft-report-engels.pdf*

The Netherlands Authority for the Financial Markets (2016). A Case analysis of critiques on high frequency trading. Technical report.
URL: *https://www.afm.nl/en/nieuws/2016/jun/onderzoek-handelsstrategieen-hft*

Upson, J. and Van Ness, R. A. (2017). Multiple markets, algorithmic trading, and market liquidity, *Journal of Financial Markets* **32**: 49–68.

Zellner, A. and Theil, H. (1962). Three-stage least squares: Simultaneous estimation of simultaneous squations, *Econometrica* **30**(1): 54–78.