Tilburg University

Competition between stock exchanges and optimal trading

van Kervel, V.L.

Publication date:
2013

Competition between Stock Exchanges and Optimal Trading

Proefschrift ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit

op vrijdag 25 januari 2013 om 14.15 uur

door

Vincent Laurens van Kervel
geboren op 7 mei 1986 te Boxtel.

Promotores:

Prof. dr. Hans Degryse

Prof. dr. Frank de Jong

Overige leden van de Promotiecommissie:

Prof. dr. Albert Menkveld

Prof. dr. Bas Werker

Dr. Fabio Feriozzi

Dr. Peter Hoffmann

Acknowledgement

I am grateful to the many people that have supported me in writing this thesis. I thank Hans and Frank for the extensive discussions and comments that have greatly improved the quality of this work. I have enjoyed working with them very much, and hope to continue working with them in the future. In addition, I thank Fabio Feriozzi, Peter Hoffmann, Albert Menkveld and Bas Werker, the members of the committee, for the time they have spend on reading the thesis. Their detailed comments were critical but fair, and have greatly benefitted the work.

I also want to take this opportunity to thank my fellow PhD students for the good laughs we had in the past years. In particular, I thank Martijn for his patience in listening to many long stories on market microstructure in a confined office space for three years. Above all, I thank María José for her enthusiasm, empathy and loving support.

Contents

Introduction

The dissertation consists of three chapters that represent separate papers in the area of market microstructure. The first two chapters study competition between stock exchanges, and the third focuses on the optimal strategy to trade a large quantity of shares. Chapter 1 and 3 are joint work with Hans Degryse and Frank de Jong.

The first chapter is motivated by two recent changes in current equity markets. First, a large number of trading venues coexist with publicly displayed limit order books, whereas in the past only a single exchange existed. Second, a large fraction of trading takes place on dark markets without publicly displayed limit order books, such as crossing networks, dark pools, and over-the-counter. On dark markets the supply and demand of shares are not publicly disclosed, meaning investors cannot observe the trading interests of others. As a result, trading in the same stock is dispersed across many trading venues, both visible and dark, which causes a fragmented market.

We study the impact of dark trading and fragmentation in visible markets on liquidity. We find that fragmentation improves liquidity aggregated over all visible trading venues, which suggests that competition between liquidity suppliers has increased. Dark trading has a negative impact on liquidity, which confirms several recent theories but has not been shown empirically. In addition, liquidity available on the traditional market is lowered by fragmentation, meaning that the benefits of fragmentation are not enjoyed by market participants who resort only to the traditional market.

In the second chapter, I study competition between trading venues but focus on the interaction of liquidity demand and supply across these venues. The theoretical model shows that certain traders have an incentive to duplicate their liquidity supply across several venues. Then, after a trade on one venue (i.e., liquidity demand) they will quickly cancel their supply on others. The consequence is that a single transaction leads to a reduction of liquidity on all venues, such that the liquidity aggregated across venues in fact overstates the liquidity available in the market. The magnitude of the cancellations depends on the fraction of investors that may access several venues simultaneously, i.e., who use Smart Order Routing Technology (SORT). The reason is that market makers incur higher adverse selection costs when the investor trades at a competing venue first, because then the trade is larger and more informed on average. Consequently, a higher fraction of SORT investors reduces the incentive to duplicate liquidity supply across venues.

The empirical results shows that trades are followed by cancellations of limit orders on competing venues of more than 53% of the trade size. The magnitude of the cancellations varies according to the fraction of SORT traders active in the market, as predicted by the model.

The third chapter models the problem of a large trader who must trade a given amount of shares before a fixed deadline, and who wishes to minimize the total trading costs. The trading motive stems from liquidity or hedging needs rather than informational reasons, which means that she is a *liquidity trader* who does not have specific information about the fundamental value of the stock. This problem is relevant to many different traders, such as pension funds, insurance companies and certain hedge funds.

The model also considers an informed trader, who knows the fundamental value of the stock and trades to maximize profits, and noise traders, who trade random amounts each period. The liquidity trader, the informed trader and the noise traders submit their trades to the dealer (a market maker), who observes the aggregate trading volume and determines the price to clear the market. The market maker is risk neutral but cannot distinguish between trades from informed and uninformed investors. Therefore, the trades by the liquidity trader do affect the price, although they are liquidity motivated.

Our model shows that the liquidity trader optimally splits up the large trade over time to create predictability in the the order flow. The predictable order flow stems from the liquidity trader only, under the assumption that the informed traders have short-lived private information. In turn, the market maker knows that this predictable component stems from the liquidity trader and does not revise prices, such that these trades enjoy lower price impacts. In essence, we show that order splitting allows the market to distinguish between informed and uninformed trades, reducing the trading costs of the latter.

Chapter 1

The impact of dark trading and visible fragmentation on market quality

Abstract

Two important characteristics of current equity markets are the large number of trading venues with publicly displayed order books and the substantial fraction of dark trading, which takes place outside such visible order books. This paper evaluates the impact on liquidity of dark trading and fragmentation in visible order books. We find that fragmentation improves liquidity aggregated over all visible trading venues, whereas dark trading has a detrimental effect. In addition, liquidity available on the traditional market is lowered by fragmentation, meaning that the benefits of fragmentation are not enjoyed by market participants who resort only to the traditional market.

JEL Codes: G10; G14; G15.
Keywords: Market microstructure, Fragmentation, Dark trading, Liquidity

1.1 Introduction

Equity markets in the US, Europe, and Canada have seen a proliferation of new trading venues. The traditional stock exchanges are being challenged by a variety of trading systems, such as electronic communication networks (ECNs), broker-dealer crossing networks, dark pools, and over-the-counter markets (OTC). Consequently, trading has become dispersed over many trading venues—visible and dark—creating a fragmented marketplace. These changes in market structure follow recent changes in financial regulation, in particular the Regulation National Market System (Reg NMS) in the US and the Market in Financial Instruments Directive (MiFID) in Europe.

An important question is how market quality is affected by the many different competing venues. In this paper, we study the impact of market fragmentation on liquidity, which is an important aspect of market quality. We investigate the impact of different types of fragmentation by classifying trading venues, according to their pre-trade transparency, into visible and dark venues, i.e., with and without publicly displayed limit order books. According to this definition, US stocks have a dark market share of approximately 30% and European blue chips of 40%.[1] Recently, the SEC has been conducting a broad review of current equity markets, and it is particularly interested in the effect of dark trading on execution quality.[2]

The impact on equity markets of fragmentation in visible order books and of dark trading have long interested researchers, regulators, investors, and trading institutions. In a recent study, O'Hara and Ye (2011) find that fragmentation lowers transaction costs and increases execution speed for NYSE and Nasdaq

[1]Speech of SEC chairman Mary Schapiro, "Strengthening Our Equity Market Structure," US SEC New York, Sept 7, 2010, and Gomber and Pierron (2010) for Europe.

[2]See the speech of Schapiro, and the SEC concept release on equity market structure, February 2010, File No. S7-02-10.

stocks. They do not distinguish, however, between the differential impact on liquidity of fragmentation stemming from visible and dark trading venues. The main contribution of our paper is that we disentangle the liquidity effects of fragmentation in visible order books ("visible fragmentation") and dark trading. In addition, we address the regulatory issues of fair markets and retail investor protection. To this end, we distinguish between liquidity aggregated over all trading venues (global liquidity) and liquidity of the traditional market only (local liquidity). Global liquidity is available to investors using smart order routing technology (SORT), and local liquidity is accessible to investors who tap the traditional market only. We furthermore improve on previous research by employing a new data set that covers the relevant universe of trading platforms, provides a stronger identification of fragmentation, and allows for improved liquidity metrics.

Our main finding is that the effect of visible fragmentation on global liquidity is generally positive, while the effect of dark trading is negative. An increase in dark trading of one standard deviation lowers the global liquidity by 9%, which is consistent with most of the theoretical research but has not been documented empirically (see Section 1.2 for a literature review). The effect of visible fragmentation has an inverted U-shape, i.e., the marginal effect decreases as fragmentation increases. With our most conservative estimates, the optimal degree of visible fragmentation improves global liquidity by approximately 32% compared with a completely concentrated market. In addition, we find that the gains of visible fragmentation are strongest for liquidity close to the midpoint, i.e., at relatively good price levels, and are weaker for liquidity deeper in the order book, which improves by only 12%. This result suggests that newly entering trading venues with visible order books primarily improve liquidity close to the midpoint. Furthermore, trading in large stocks is more fragmented, and its liquidity benefits twice as much from fragmentation.

While global liquidity benefits from fragmentation, we find that the market quality at the traditional stock exchange is worse off: local liquidity close to the midpoint reduces by approximately 10%. Thus, investors without access to SORT are worse off in a fragmented market, especially for relatively small orders.

We address the impact of fragmentation on market liquidity by creating, for every firm, daily proxies of visible fragmentation, dark activity, and liquidity. Specifically, we use high-frequency data from all relevant trading venues from January 2006 (before fragmentation set in) to the end of 2009, when markets were quite fragmented. Similarly to Foucault and Menkveld (2008), we select all Dutch mid- and large-cap stocks, which are relatively large with an average market capitalization approximately twice that of the NYSE and Nasdaq stocks analyzed in O'Hara and Ye (2011). We measure the degree of visible fragmentation using the Herfindahl-Hirschman index (*HHI*, the sum of the squared market shares) based on the trading volumes of all visible venues. Dark trading is defined as the market share of trading volume on dark venues, which reflects OTC and internalization by broker dealers. Then, for each stock we construct a consolidated limit order book (i.e., the limit order books of all visible trading venues combined) to get a complete picture of the global liquidity available in the market. Based on the consolidated order book we analyze global liquidity at the best price levels and also deeper in the order book. The depth beyond the best price levels matters to investors because it reflects the quantity immediately available for trading and therefore the price of immediacy. In addition to global liquidity, we analyze local liquidity, which is available at the traditional exchange only.

Our panel dataset helps to identify the exogenous relationship between liquidity and fragmentation by means of firm-quarter fixed effects. The inclusion of firm-quarter dummies implies that the impact of fragmentation on liquidity stems from variation within a firm-quarter, making the analysis robust to various industry-specific shocks and time-varying firm-specific shocks. Furthermore, to address concerns about the endogeneity of visible fragmentation and dark trading, we use instrumental variables. Similarly to O'Hara and Ye (2011), we use as instruments for visible fragmentation the average order size of the visible competitors, and the number of limit orders to market orders on the visible competitors. Dark trading is instrumented by the size of the average dark order.

Our findings on liquidity are related to those of several recent studies. The positive effect of visible fragmentation is consistent with competition between liquidity suppliers, since the compensation for liquidity suppliers, the realized spread, reduces with fragmentation (e.g., Biais, Bisière, and Spatt (2010)). A similar argument is made by Foucault and Menkveld (2008), who use 2004 data to show that liquidity improves with competition between two traditional stock exchanges (the LSE and Euronext). The negative impact of dark trading is consistent with a "cream-skimming" effect, as the informativeness of trades, the price impact, increases strongly with dark activity (Zhu, 2011). Informed traders typically trade at the same side of the order book, such that they face low execution probabilities in dark pools and crossing networks. Consequently, dark markets attract relatively more uninformed traders, leaving the informed trades to visible markets. According to Hendershott and Mendelson (2000), the visible market might be used as a market of last resort which attracts mostly informed order flow.

In line with our results, Weaver (2011) shows that off-exchange reported trades, which mostly represent dark trades in his sample, negatively impact the market quality for US stocks. In contrast to our results, Buti, Rindi, and Werner (2010a) find that dark-pool activity is positively related to liquidity in the cross-section, but economically insignificant in the time series. They study pure dark pools however, whereas our measure of dark trading mostly reflects OTC and internalization.

In summary, our findings imply a deeper understanding of the more general conclusion of O'Hara and Ye (2011) that fragmentation does not harm market quality. We show that the composition of the fragmentation—visible versus dark—determines the total impact of fragmentation on market quality. Moreover, our conclusions relate to the issues raised by the SEC on the benefits and drawbacks of stock market fragmentation, showing that the benefits are not equally enjoyed by all stock market participants. This latter finding is particularly relevant to regulators who strive for fair markets and the protection of retail investors.

The remainder of this paper is structured as follows. Section 1.2 discusses the literature on competition between exchanges. The dataset and liquidity measures are described in Sections 1.3 and 1.4. Section 1.5 explains the methodology and main results, while Section 1.6 reports a series of robustness checks. Finally, Section 1.7 provides concluding remarks.

1.2 Literature on fragmentation and market quality

There is a trade-off between order flow fragmentation and competition. A single exchange has costs lower than those of a fragmented market structure. The latter costs are the fixed costs to set up a new trading venue; the fixed costs for clearing and settlement; the costs of monitoring several trading venues simultaneously; and the cost of advanced technological infrastructure to aggregate dispersed information in the market and to connect to several trading venues. Also, a single market that is already liquid will attract even more liquidity because of positive network externalities (e.g., Pagano (1989a), Pagano (1989b), and Admati and Pfleiderer (1991)). Each additional trader reduces the stock's execution risk for other potential traders, attracting more traders. This positive feedback should cause all trades to be executed at a single market, giving the highest degree of liquidity.

However, while network externalities are still relevant, nowadays they may be realized even when several trading venues coexist. This happens to the extent that the technological infrastructure seamlessly links the individual trading venues, creating effectively one market. From a broker's point of view, the market is then not fragmented, which alleviates the drawbacks of fragmentation (Stoll, 2006).[3] In addition, fragmentation might also enhance market quality, because increased competition among liquidity suppliers forces them to improve their prices, narrowing the bid-ask spreads (e.g., Biais, Martimort, and Rochet (2000) and Battalio (1997)). Confirming a competition effect, Conrad, Johnson, and Wahal (2003) find that alternative trading systems in general have lower execution costs compared with brokers on traditional exchanges. Furthermore, Biais, Bisière, and Spatt (2010) investigate the competition induced by ECN activity on Nasdaq stocks. They find that ECNs with smaller tick sizes tend to undercut the Nasdaq quotes and reduce the overall quoted spreads.

Differences between trading venues may arise to cater to the different needs of clientele. For example, investors differ in their

[3] Confirming a high level of market integration, Storkenmaier and Wagener (2011) find that at least two venues quote the best bid and offer 85% of the time for FTSE100 stocks in April/May 2010.

preferences for trading speed, order sizes, anonymity, and likelihood of execution (Harris (1993) and Petrella (2009)). In the US, Boehmer (2005) stresses the trade-off between speed of execution and execution costs on Nasdaq and NYSE, where Nasdaq is more expensive but also faster. To attract more investors, new trading venues may apply aggressive pricing schedules, such as make and take fees (Foucault, Kadan, and Kandel, 2009). The fact that some investors prefer a particular trading venue can also lead to varying degrees of informed trading at each exchange. For instance, the NYSE has been found to attract more informed order flow than the regional dealers (Easley, Kiefer, and O'Hara, 1996) and Nasdaq market makers (Bessembinder and Kaufman (1997) and Affleck-Graves, Hedge, and Miller (1994)). Furthermore, Barclay, Hendershott, and McCormick (2003) find that ECNs attract more informed order flow than Nasdaq market makers, because ECN trades have a larger price impact.

Stoll (2003) argues that competition fosters innovation and efficiency, but priority rules may not be maintained. Specifically, time priority is often violated in fragmented markets, and sometimes also price priority.[4] Foucault and Menkveld (2008) study the competition between an LSE order book (EuroSETS) and Euronext Amsterdam for Dutch firms in 2004, and find a trade-through rate of 73%. They call for a prohibition of trade-throughs since these events discourage liquidity provision. Possible explanations of trade-throughs are the high costs of monitoring multiple markets or the high variable and fixed trading fees and clearing and settlement costs. Gresse (2006) finds that trading activity on a crossing network improves the quoted

[4]Time priority is violated when two limit orders with the same price are placed on two venues and the later order is executed first. Price priority is violated, i.e., a trade-through, when an order gets executed against a price worse than the best quoted price in the market. In a partial trade-through only part of the order could have been executed against a better price.

spreads in the dealer market, especially when the dealers also trade on the crossing network.

In addition to competition between trading venues with visible liquidity, this paper is related to the literature on competition effects in dark markets, i.e., venues without publicly displayed order books. A few papers theoretically investigate the impact of dark trading on traditional markets. Hendershott and Mendelson (2000) model a crossing network that competes with a dealer market, and they find ambiguous effects on the dealer's spread. On the one hand, a crossing network may attract new liquidity traders and therefore lead to lower dealer spreads. On the other hand, when the dealer market is used as a market of last resort, the dealer's spread may increase. Also modeling the interaction between a crossing network and dealer market, Degryse, Van Achter, and Wuyts (2009) show that the order flow dynamics and welfare implications depend on the degree of transparency, but they do not endogenize the spread. Buti, Rindi, and Werner (2010b) model the competition between a dark pool and visible limit order book, and show that the initial level of liquidity determines the effect of the dark pool on the quoted spreads. That is, for liquid stocks both limit and market orders migrate to the dark pool, leaving the spread tight. For illiquid stocks the competition induced by the dark pool reduces the execution probability of limit orders, causing the spread to increase. In contrast, Zhu (2011) argues that dark pools "cream-skim" uninformed trades, leaving informed trades to the visible exchange. The intuition is that informed traders face a relatively low execution probability on the dark pool because they typically trade on the same side of the order book. Empirical evidence suggests that transparent markets allow for faster and cheaper access to information and are therefore more efficient (Hendershott and Jones, 2005, Boehmer, Saar, and Yu, 2005)

Finally, our paper is related to the literature on algorithmic trading, i.e., the use of computer programs to manage and execute trades in electronic limit order books. Algorithmic trading has strongly increased over time, and it has drastically affected the trading environment. In particular, it affects the level of market fragmentation analyzed in our sample, since computer programs and SORT allow investors to find the best liquidity in the market by comparing the order books of individual venues. Some algorithms are designed to split up trades over time to reduce implicit transaction costs (e.g., Huberman and Stanzl (2005)). Boehmer, Fong, and Wu (2012) find that an increase in algorithmic trading intensity improves liquidity, but also increases volatility. Programs are also used to identify deviations from the efficient stock price, by quickly trading on new information or price changes of other securities (Brogaard, Hendershott, and Riordan, 2012). Furthermore, programs may provide liquidity when the quoted spreads are large, e.g., when it is profitable to do so (Brogaard, Hendershott, and Riordan, 2012). Hasbrouck and Saar (2009) describe "fleeting orders," a relatively new phenomenon in the US and Europe, where limit orders are placed and canceled within two seconds if they are not executed. The authors argue that fleeting orders are part of an active search for liquidity and a consequence of improved technology, more hidden liquidity, and fragmented markets. Hasbrouck and Saar (2011) find that liquidity is positively affected by low-latency trading, i.e., extremely fast proprietary trading desks.

In summary, the literature suggests that the fragmentation of trading may improve liquidity, and it offers some empirical evidence for that. However, the empirical studies to date do not distinguish between fragmentation in visible and dark trading venues. This is precisely our contribution.

1.3 Market description, dataset and descriptive statistics

1.3.1 Market description

Our dataset contains 52 Dutch stocks forming the constituents of the so-called AEX Large and Mid cap indices. Over time, all these stocks are traded on several trading platforms, to a degree which is representative for the large European stocks analyzed by Gomber and Pierron (2010). In terms of size, the average market cap of our sample is approximately twice that of the NYSE and Nasdaq sample analyzed in O'Hara and Ye (2011). We can summarize the most important trading venues for these stocks into three groups as follows (Appendix A contains a more general description of current European financial markets).

First, there are regulated markets (RMs), such as NYSE Euronext, LSE and Deutsche Boerse. These markets have an opening and closing auction, and in between there is continuous and anonymous trading through the limit order book. Since Euronext merged with NYSE in April 2007, the order books in Amsterdam, Paris, Brussels and Lisbon act as a fully integrated and single market. For our sample, the LSE and Deutsche Boerse are not very important as they execute less than 1% of total order flow.

Second, there are the new ECNs (in European terminology Multilateral Trading Facilities) with visible liquidity, such as Chi-X, Bats Europe, Nasdaq OMX and Turquoise. Chi-X started trading AEX firms in April 2007, before the introduction of MiFID; Turquoise in August 2008 and Nasdaq OMX and Bats Europe in October 2008. Whether these MTFs will survive depends on the current level of liquidity, but also on the quality of the trading technology (e.g. the speed of execution), the number of securities traded, make and take fees and clearing and settlement costs. Nasdaq OMX closed down in May 2010, outside our sample period, as they did not meet their targeted market share.[5] A new trading venue in Europe typically starts with a test phase in which only a few liquid firms are traded, but will allow trading in all stocks of a certain index at once when it goes live.

The third group contains MTFs with completely hidden liquidity (e.g., dark pools), broker-dealer crossing networks, internalization and Over-The-Counter markets. This set of exchanges is waived from the pre-trade transparency rules set out by the MiFID, due to the nature of their business model. Most dark pools employ a limit order book with similar rules as those at Euronext for example. Crossing networks typically execute trades against the midpoint of the primary market, and do not contribute to price discovery. Gomber and Pierron (2010) report that the activity on dark pools, crossing networks and OTC has been fairly constant for European equities in 2008 - 2009, and they execute approximately 40% of total traded volume.

1.3.2 Dataset

Our dataset covers the AEX Large and Midcap constituents from 2006 to 2009, which currently have 25 and 23 stocks respectively. We remove stocks that are in the sample for less than six months or do not have observations in 2008 and 2009. Due to some leavers and joiners, our final sample has 52 stocks.

The data for the 52 stocks stem from the Thomson Reuters Tick History Data base. This data source covers the seven most relevant European trading venues for the sample stocks, which have executed more than 99% of visible order flow: Euronext, Chi-X, Deutsche Boerse, Turquoise, Bats Europe, Nasdaq OMX

[5]See "Nasdaq OMX to close pan-European equity MTF", www.thetradenews.com.

and SIX Swiss exchange (formerly known as Virt-X).[6] We employ data from all these venues but collect them only during the trading hours of the continuous auction of Euronext Amsterdam, i.e. between 09.00 to 17.30, Amsterdam time. Therefore, data of the opening and closing auctions at these venues are not included.[7]

Each stock-venue combination is reported in a separate file and represents a single order book. Every order book contains the ten best quotes at both sides of the market, i.e. the ten highest bid and lowest ask prices and their associated quantities, summing to 40 variables per observation.[8] All observations are time stamped to the millisecond. A new "state" of a limit order book is created when a limit order arrives, gets canceled or when a trade takes place. A trade is immediately reported and we observe its associated price and quantity, as well as an update of the order book. Price and time priority rules apply within each stock-venue order book, but not between venues. Furthermore, visible orders have time priority over hidden orders. Hidden orders are not directly observed in the dataset but are detected upon execution. Therefore, we have the same information set available to the market, i.e. the visible part of the order book on a continuous basis. We treat executions of hidden and 'iceberg' orders as visible, since these trades take place on predominantly visible trading venues.

[6]The visible order books of Dutch stocks on the LSE are discarded, as those stocks have different symbols, are denoted in pennies instead of Euros, and are in essence different assets. The remaining trading venues with visible liquidity attract extremely little order flow for the firms in our sample (e.g., NYSE, Milan stock exchange, PLUS group and some smaller exchanges).

[7]Unscheduled intra-day auctions are not identified in our dataset. These auctions, triggered by transactions that would cause extreme price movements, act as a safety measure and typically last for a few minutes. Given that we will work with daily averages of quote-by-quote liquidity measures, these auctions should not affect our results.

[8]Part of the sample only has the best five price levels: Euronext before January 2008. This affects only liquidity deep in the order book. As robustness, we execute the analysis separately for 2008 and 2009 in section 1.6.4; the results are unaffected.

Our dataset also provides information on "dark trades", i.e. trades at dark pools, broker-dealer crossing networks, internalized and Over The Counter (including trades executed over telephone). However, the fraction of trading volume in dark pools is very small in our sample period (<1%, according the FESE). These dark trades are reported in a separate file and are constructed by Markit Boat, a MiFID-compliant trade reporting company.[9] The file contains the price, quantity and time of the execution (time stamped to the millisecond). The file pools trades from all trading venues, but does not report the identity of the executing venue and does not contain any quotes. We complete the dark trades data by adding the OTC and internalized trades reported in separate files by Euronext, Xetra and Chi-X. The largest part of dark trading is internalization, but we do not know the exact decomposition.

1.3.3 Descriptive statistics

Figure 1.1 shows the evolution of the daily traded volume, aggregated over all AEX Large and Mid cap constituents. The graph shows a steady increase in total trading activity, which peaks around the beginning of 2008. Moreover, the dominance of Euronext over its competitors is strong, but slowly decreasing over time. This pattern is representative for all regulated markets trading European blue chip stocks, as analyzed by Gomber and Pierron (2010). Finally, while Chi-X started trading AEX firms in April 2007, the new competitors together started to attract significant order flow only as of August 2008 (4.5%). The slow start up shows that these venues need time to generate trading activity.

[9]There has been some discussion on issues with these dark data (e.g. double reporting). See the Federation of European Securities Exchanges (FESE) response to the MiFID consultation paper, February 2011. The market shares as reported in our data are consistent with those reported by FESE.

In Table A1 in the Appendix, the characteristics of the different stocks and some descriptive statistics are presented. There is considerable variation in firm size (market capitalization), price and trading volume. In the sample, 38 stocks have a market capitalization exceeding one billion Euro, while the 14 remaining stocks have market capitalization above 100 million Euro. The table also reports realized volatilities, computed by first dividing the trading day into 34 fifteen-minute periods and then calculating stock returns of each period, based on the spread midpoint at the beginning and end of that period. The standard deviation of these stock returns are daily estimates of realized volatility.[10] The table also shows the average market share of Euronext and dark trades, calculated as of November 2007 onwards, the period for which Markit Boat data have become available in the dataset.[11] According to our data, in 2009 37% of the total traded volume is dark; which can be split up into 38% for AEX large cap firms and 20% for mid cap firms.

1.4 Liquidity and fragmentation

1.4.1 The consolidated order book

The goal of this paper is to analyze the impact of equity market fragmentation on liquidity. We follow the approach of Gresse (2010) and distinguish between global traders and local traders. Global traders employ smart order routing technology (SORT) to access all trading venues simultaneously, while for local traders SORT is too expensive because of fixed trading charges and costs of adopting this trading technology. This distinction is empirically justified as SORT is not used by all investors (e.g. Foucault

[10] The use of realized volatility is well established, see e.g. Andersen, Bollerslev, Diebold, and Ebens (2001).

[11] The lack of Markit Boat data in 2006 and 2007 does not affect our results, as we execute the analysis separately for 2008 and 2009 only in section 1.6.4.

and Menkveld (2008) and Ende, Gomber, and Lutat (2009)). In our setting, Euronext Amsterdam is the local market and the consolidated order book of the different visible trading venues represents the global market.

To construct the consolidated order book, we follow the methodology of Foucault and Menkveld (2008) among others, based on snapshots of the limit order book. A snapshot contains the ten best bid and ask prices and associated quantities, for each stock-venue combination. Every minute we take snapshots of all venues and "sum" the liquidity to obtain a stock's consolidated order book. Therefore, each stock has 510 daily observations (8.5 hours times 60 minutes), containing the order books of the individual trading venues and the consolidated one.

1.4.2 Depth(X) liquidity measure

Our rich dataset allows to construct a liquidity measure that incorporates the limit orders beyond the best price levels; which we will refer to as the $Depth(X)$. The measure aggregates the Euro value of the number of shares offered within a fixed interval around the midpoint. Specifically, the midpoint is the average of the best bid and ask price of the consolidated order book and the interval is an amount $X = \{10, 20, ..., 50\}$ basis points relative to the midpoint.[12] The measure is expressed in Euros and calculated every minute. Equation 1.1 shows the calculation for the bid and ask side separately, which are summed to obtain $Depth(X)$. This measure is constructed for the global and local order book (i.e., Euronext Amsterdam) separately. Define price level $j = \{1, 2, ..., J\}$ on the pricing grid and the midpoint of the

[12] Foucault and Menkveld (2008) aggregate liquidity from one up to four ticks away from the best quotes. This approach is not appropriate in our setting, as tick sizes have changed over the course of our sample period. Furthermore, the tick size as a percentage of the share price is not constant.

consolidated order book as M, then

$$Depth\ Ask(X) = \sum_{j=1}^{J} P_j^{Ask} Q_j^{Ask} \mathbf{1}\{P_j^{Ask} < M(1+X)\}, \quad (1.1a)$$

$$Depth\ Bid(X) = \sum_{j=1}^{J} P_j^{Bid} Q_j^{Bid} \mathbf{1}\{P_j^{Bid} > M(1-X)\}, \quad (1.1b)$$

$$Depth(X) = Depth\ Bid(X) + Depth\ Ask(X). \quad (1.1c)$$

Figure 1.2 gives a graphical representation of the depth measure, where liquidity between the horizontal dashed lines is aggregated to obtain $Depth(20)$ and $Depth(40)$. The measure is averaged over the trading day, where $Depth(10)$ represents liquidity close to the midpoint and $Depth(50)$ also includes liquidity deeper in the order book. Comparing different price levels X reveals the shape of the order book. For example, if the depth measure increases rapidly in X, the order book is deep while if it increases only slowly, the order book is relatively thin.

The $Depth(X)$ measure is closely related to the Cost of Round-trip, $CRT(D)$ (e.g. Irvine, Benston, and Kandel (2000) and Barclay, Christie, Harris, Kandel, and Schultz (1999)), which also analyzes liquidity deeper in the order book.[13] More specifically, $CRT(D)$ fixes the quantity D of a potential trade, i.e. D equals €100.000, and analyzes the impact on price. In contrast, $Depth(X)$ fixes the price, i.e. X equals ten basis points around the midpoint, and analyzes the available quantity. Although both measures estimate the depth and slope of the order book, our approach solves two rather technical issues. First, the impact on price cannot be calculated when a stock's order book has insufficient liquidity to trade €100.000, such that the $CRT(D)$ does not exist. In contrast, if no additional shares are

[13] The $CRT(D)$ is also known as the Exchange Liquidity Measure, $XLM(V)$, (e.g. Gomber, Schweickert, and Theissen (2004)).

offered within the range of X and $X + \varepsilon$ basis points from the midpoint, then $Depth(X)$ has a zero increment and $Depth(X) = Depth(X+\varepsilon)$. Second, $CRT(D)$ may become negative when the consolidated spread is negative, i.e. when the best ask price of a venue is lower than the best bid price of another venue.[14] While negative transaction costs cannot be interpreted meaningfully, the midpoint and $Depth(X)$ are perfectly identified and reflect the available liquidity in a meaningful fashion.

An advantage of $Depth(X)$ over the traditional quoted depth and spread is that it is not sensitive to small, price improving orders. Such orders are often placed by algorithmic traders, whose activity has increased substantially over time. In addition, the quoted depth and spread are sensitive to changes in tick sizes.[15]

Figure 1.3 plots the 10, 50 and 90$^{\text{th}}$ percentile of the depth measure against the number of basis points around the midpoint. The vertical axis is plotted on a logarithmic scale, as we work with the logarithm of the depth measures in the regression analysis. Overall, the shape of the order book appears very linear. Also, there are large differences between firms, as the 90$^{\text{th}}$ percentile of $Depth(10)$ is €915.000, while the 10$^{\text{th}}$ percentile of $Depth(50)$ is €72.000. This is in line with high levels of skewness and kurtosis (not reported).

Table 1.1 contains the medians of the $Depth(X)$ measure for the global and local order book on a yearly basis, along with other liquidity measures discussed in the next section. As expected, the global and local depth measures vary substantially over time. However, some shocks affect liquidity close to the midpoint more than liquidity deep in the order book. That is, the ratio of $Depth(50)$ to $Depth(10)$ is not constant over time.

[14] Technically, a negative consolidated spread (or crossed quotes) is an arbitrage opportunity, which might not be exploited because of explicit trading costs for example.

[15] The effect of the tick size on quoted depth and spread have been subject of analysis in several papers, e.g. Goldstein and Kavajecz (2000), Huang and Stoll (2001).

1.4.3 Other liquidity measures

This section compares our $Depth(X)$ liquidity measure to the more traditional liquidity measures. These are the price impact, effective and realized spread, based on executed transactions, and the quoted spread and quoted depth, based on quotes in the local and global order books. The quoted depth sums the Euro amount of shares offered at the best bid and ask price, whereas the quoted spread looks at the associated prices. Appendix B contains a formal description of the measures.

The medians of the liquidity measures are reported in the upper panel of Table 1.1, based on daily observations and calculated yearly, for the global and local order book. The table shows several interesting results.

Depth close to the midpoint has reduced strongly over time, but liquidity deeper in the order book to a lesser extent. That is, the median of $Depth(10)$ has decreased by 35% from 2006 to 2009, while $Depth(50)$ by only 14%. In addition, the yearly standard deviations of the depth measures have decreased by approximately 50% over the years (not reported). While in 2006 and 2007 the local and global $Depth(X)$ are highly similar, in 2009 local $Depth(X)$ represents only about 50% of global depth.

Strikingly, between 2006 and 2009 the median quoted spread has improved by 9%, while the quoted depth (at the best quoted prices) has worsened by 68%. This is very likely due to the strong increase in very small orders. The $Depth(10)$ measure decreases by 35% over the same time period. This shows the shortcomings of the quoted depth and spread measures, because based on the quoted depth and spread alone, one cannot state whether an investor is better off in 2006 or 2009, as this depends on the traded quantity.

Turning to the liquidity measures based on executed trades, we observe that the median realized spread has reduced from

2.5 basis points in 2006 to 0 basis points in 2009. In this period, the price impact went up with 2.9 basis points while the effective spread reduced with 0.9 basis points. Because we show medians, the price impact and realized spread do not exactly add up to the effective spread.

Despite the reduction in $Depth(X)$, the local price impact, realized and effective spreads are almost identical to those of the global order book. This finding might be in line with "market tipping", where the local market switches between periods of relatively high liquidity, in which it attracts all trading, and periods of low liquidity, in which trading takes place at competing trading venues. As the price impact, effective and realized spread are based on trades, relatively liquid periods receive a larger weight in the calculation.

1.4.4 Equity market fragmentation

To proxy for the level of fragmentation in each stock, we construct a daily Herfindahl-Hirschman Index (HHI) based on the number of shares traded on each visible trading venue, similar to e.g. Bennett and Wei (2006) and Weston (2002). Formally, $HHI_{it} = \sum_{v=1}^{N} MS_{v,it}^2$, or the squared market share of venue v, summed over all N venues for firm i on day t. We then use $VisFrag = 1 - HHI$, short for visible fragmentation; such that a single dominant market has zero fragmentation whereas $VisFrag$ goes to $1 - 1/n$ in case of complete visible fragmentation. In addition, $Dark$ is our proxy for dark trading, calculated as the percentage of volume executed at dark pools, crossing networks, internalizers and OTC. We use the percentage of dark volume since we do not have information on fragmentation within the different dark venues. However, separating visible competition and dark trading is important, as they may affect liquidity in a different fashion. Our measure of fragmentation is more accurate than that of O'Hara and Ye (2011), where the origin of trades are classified as either Nasdaq, NYSE or external. The benefits of competition in their paper arise from the external venues, but the actual level of fragmentation, and whether they are dark or lit, is unclear.

Table 1.2 shows the yearly mean, quartiles and standard deviation of $VisFrag$ and $Dark$, based on the sample firms. In 2009, the sample average $VisFrag$ is 0.28, which is in line with other European stocks analyzed by Gomber and Pierron (2010). The US is more fragmented, as Nasdaq and NYSE combined have approximately 65% of market share in 2008 (O'Hara and Ye, 2011). As expected, fragmentation increases over time, since in 2006 and 2007 only few sample firms where traded on Virt-X and Deutsche Boerse. $Dark$ is fairly constant over time with on average 25% in 2009, but has a very high daily standard deviation of 17%.[16]

Figure 1.4 shows the 10, 50 and 90$^{\text{th}}$ percentile of $VisFrag$ over time, calculated on a monthly basis and covering all firms. The sharp increase in fragmentation refers to the period where Chi-X and Turquoise started to attract substantial order flow, September 2008. In the next section, we estimate the effect of fragmentation on various liquidity measures in a regression framework.

1.5 The impact of visible fragmentation and dark trading on global and local liquidity

This section first explains the methodology, and then presents the regression results of the base model, for the global and local order book.

[16] The dark share is calculated daily, and then averaged over all days and firms. When weighted by trading volume, 37% of all trading is dark in 2009, meaning that dark trades especially take place on high volume days.

1.5.1 Methodology

We employ multivariate panel regression analysis to study the impact of visible fragmentation and dark trading on liquidity. We have a panel dataset with 52 firms and 1022 days, from 2006 to 2009, which contains the liquidity and fragmentation measures discussed in section 1.4.

The panel approach allows for more flexibility compared to other papers investigating the impact of fragmentation on liquidity. For example, in contrast to the cross sectional regressions employed by O'Hara and Ye (2011), we add firm fixed effects to absorb unobservable firm characteristics, and also measure the time series variation in liquidity and fragmentation. By using a fragmentation measure based on the Herfindahl-Hirschman Index we improve on papers such as Foucault and Menkveld (2008), Chlistalla and Lutat (2011) and Hengelbrock and Theissen (2010), who study the introduction of a new trading venue (EuroSETS, Chi-X and Turquoise respectively). That is, these articles use a dummy variable that equals one after the introduction of the new venue, to estimate the effect of fragmentation on liquidity. Given the research question we are after, our approach has three advantages compared with the aforementioned papers. First, instead of a single trading venue we can analyze the effect of fragmentation on liquidity over many trading venues simultaneously. Second, we allow for cross sectional variation in fragmentation as some firms are more heavily traded on new venues than others. And third, we allow for variation in the time series and analyze a long time window. This approach takes into account that new trading venues might need time to grow, and allows the market as a whole to adjust to a new trading equilibrium.

In the regressions we include volatility, price, firm size and volume as control variables, which is common in this literature.[17] Descriptives of these control variables are presented in Table 1.1. In addition, we include a proxy for algorithmic activity, as this has been found to improve liquidity (e.g. Brogaard, Hendershott, and Riordan (2012)). We construct a measure similar to Hendershott, Jones, and Menkveld (2011). On average, algorithmic traders place and cancel many limit orders, so the daily number of electronic messages proxies for their activity, i.e. placement and cancelations of limit orders and market orders. This variable is divided by trading volume, as increasing volumes lead to more electronic messages even in the absence of algorithmic trading. Accordingly, $Algo_{it}$ is defined as the number of electronic messages divided by trading volume for firm i on day t.

The dependent variable in these regressions is one of the liquidity measures, and the independent variables are the level of fragmentation and dark trading, and several control variables. As the effect of fragmentation on liquidity might not be linear, we add a quadratic term. We employ $VisFrag_{it} = 1 - HHI_{it}$ and $VisFrag_{it}^2$ to measure fragmentation, where $VisFrag_{it} = 0$ if trading in a firm is completely concentrated. We add firm fixed effects to make sure the variation we pick up is due only to variability in fragmentation and dark trading relative to the firm's own average. We also add time effects to control for common, market wide fluctuations in all variables. We use quarterly time fixed effects, but the results are almost identical when using day or month dummies instead of quarter dummies. The regression equation thus becomes

$$Liq\,Measure_{it} = \alpha_i + \delta_{q(t)} + \beta_1 VisFrag_{it} + \beta_2 VisFrag_{it}^2 + \beta_3 Dark_{it} + \\ \beta_4 Ln(Volatility)_{it} + \beta_5 Ln(Price)_{it} + \beta_6 Ln(Size)_{it} +$$

[17] Weston (2000), Fink, Fink, and Weston (2006) and O'Hara and Ye (2011), among others, use similar controls.

$$\beta_7 Ln(Volume)_{it} + \beta_8 Algo_{it} + \varepsilon_{it}, \quad (1.2)$$

where α_i are the firm fixed effects and $\delta_{q(t)}$ are 16 quarterly dummies that take the value of one if day t is in quarter q, and zero otherwise. For the inference we use heteroscedasticity and autocorrelation robust standard errors (Newey-West for panel datasets), based on five lags.

1.5.2 Results: global liquidity

The regression results for the liquidity measures employing the global (consolidated) order book are reported in Table 1.3. The results of models (1) to (5) show that liquidity first strongly increases with visible fragmentation and then decreases, as the linear term $VisFrag$ has a positive coefficient and the quadratic term $VisFrag^2$ a negative one. The results are easier to interpret from Figure 1.5, which displays the implied results of the effect of visible fragmentation on liquidity for the five models. The figure clearly reveals an optimal level of visible fragmentation, where maximum liquidity is obtained at $VisFrag = 0.35$. This level of visible fragmentation is fairly close to the average level in 2009, which is 0.28. The pattern is similar for all depth levels, although liquidity levels close to the midpoint benefit somewhat more from visible fragmentation. The economic magnitudes of the variables are large, where the maximum effect on $ln(Depth(10))$ is 0.50, meaning that observations here have 65% more liquidity than observations in a completely concentrated market. For $Depth(50)$, liquidity improves by 50% at the maximum compared with $VisFrag = 0$. The standard deviation of visible fragmentation is 0.15 in the entire sample (Table 1.2), so variation in visible fragmentation has a large impact on liquidity throughout the entire order book.

We now investigate the impact of visible fragmentation on the other liquidity indicators, as reported in models (6) to (10) in Table 1.3. At the optimal degree of visible fragmentation, $VisFrag = 0.35$, the price impact and effective spread reduce by 6.3 and 6.8 basis points compared with a completely concentrated market, respectively. This is large, considering that the median effective spread in 2009 is 13.3 basis points (Table 1.7). The economic impact of the optimal degree of visible fragmentation on the effective spread in our analysis is larger than estimated in O'Hara and Ye (2011) for total fragmentation, where the benefit is approximately three basis points for NYSE and Nasdaq firms.[18] This difference can partly be explained by our inclusion of a separate dark trading variable, which has a positive effect on the effective spread and price impact. The effect of visible fragmentation on the realized spread is 0.5 at the optimal level, which is relevant given a median realized spread of virtually zero in 2009. The realized spread represents the reward of supplying liquidity, which reduces by the competition between liquidity suppliers in a fragmented market. The quoted spread in model (9) improves with eight basis point at $VisFrag = 0.37$, while the sample median is twelve basis points. In stark contrast, the results in model (10) show that quoted depth (at the best bid and ask quotes) reduces by 27% at $Frag = 0.37$. The results on the quoted depth point in the opposite direction of those of all other liquidity measures. Moreover, considering the low correlation between the quoted depth and $Depth(X)$ in Table 1.1, it appears that the quoted depth is not a suitable liquidity measure in the period we study. Possibly, this is a consequence of algorithmic traders who place many small and price improving orders.

We now turn to the effects of dark trading on liquidity. In

[18] O'Hara and Ye (2011) find a linear coefficient on "market share outside the primary markets" of 9 basis points, while the average level is 0.35, resulting in a benefit of approximately 3 basis points.

Table 1.3, the coefficients on *Dark* are strongly negative, with a coefficient of -0.91 for *Ln Depth*(10). As a result, a one standard deviation increase in the fraction of dark trading (0.18) reduces *Depth*(10) by 16%. In addition, the coefficient on the price impact of 4.1 suggests that dark trading leads to more adverse selection and informed trading on the visible markets. Both findings are consistent with the theoretical work of Hendershott and Mendelson (2000) and Zhu (2011), where dark markets are more attractive to uninformed traders, leaving the informed traders to the visible markets. The intuition is that informed traders typically trade at the same side of the order book, and therefore face relatively low execution probabilities in the dark pool or crossing network. As a result, the dark market "creamskims" uninformed order flow, worsening liquidity and adverse selection costs in the visible market. The reduction in depth at the visible exchanges is also consistent with the model of Buti, Rindi, and Werner (2010b), since limit orders migrate from the limit order book to the dark pool. Empirically, our results are consistent with Weaver (2011), who shows that off exchange reported trades, which mostly qualify as dark trades in his sample, negatively affect market quality for US stocks. Our results contrast Buti, Rindi, and Werner (2010a), who find that dark pool activity improves the quoted spread in the cross section. In time series regressions however, similar to ours, the authors find statistically marginally significant and economically insignificant results. In addition, the authors do not control for the degree of visible fragmentation, and for trades on crossing networks and OTC. Trading activity across such venues is likely to be correlated, implying an omitted variables bias. For example, dark pool activity is generally higher for larger firms, which also benefit more from higher levels of visible fragmentation in our sample.

The decision to trade in the dark might be endogenous as low levels of visible liquidity may induce an investor to trade in the dark, implying that they are substitutes. Alternatively, both markets can be considered complements, since a liquid OTC market forces limit order suppliers in the visible market to improve prices as well, and vice versa (e.g., Duffie, Garleanu, and Pedersen (2005)). We tackle such reverse causality issues with an instrumental variables regression in section 1.6.2, but our main results are robust.

Turning to the control variables of the regressions, we find that the economic magnitude of *Algo* is fairly small and negative. For example, a one standard deviation increase (0.36), lowers the $Depth(X)$ measures with 4%. However, as *Algo* might be indirectly related to fragmentation, we want to be careful in interpreting this result. The remaining control variables in the regressions have the expected signs. Larger firms tend to be more liquid, while the effect of price is marginally positive and economically small. As expected, increased trading volumes are related to better liquidity, but the causality might go either way. Finally, volatility has a negative impact on liquidity; especially for liquidity close to the midpoint. Not surprisingly, the price impact strongly increases in volatility, which proxies for the amount of information in the market.

1.5.3 Results: Local liquidity

We now turn to the impact of fragmentation available at the regulated market, which we call local liquidity. The estimates are reported in Table 1.4 and displayed in the lower panel of Figure 1.5. $Depth(10)$ first slightly improves with visible fragmentation, where the maximum lies at +10% at $VisFrag = 0.17$, but afterwards quickly reduces to −10% at $VisFrag = 0.4$. This reduction is in line with the theory of Foucault and Menkveld (2008), where the execution probability of the incumbent market diminishes as competing venues take away order flow. This side effect of competition makes the incumbent less attractive to liquidity providers, resulting in lower depth. The coefficients on $Dark$ are highly similar to those reported for the global order book.

Consequently, small investors, who mainly care for $Depth(10)$ and are limited to trading on Euronext only, are worse off. This result is in contrast to the empirical results of Weston (2002) for instance, who finds that the liquidity on Nasdaq improves when ECNs enter the market and compete for order flow. The difference is probably due to the market structure in the US, where Nasdaq market makers lost their oligopolistic rents after the entry of ECNs.

We now turn to the regressions of the remaining liquidity measures in Table 1.4, columns (6) to (10). In contrast to $Depth(10)$, these are not adversely affected by visible fragmentation. This is not surprising, as the $Depth(10)$ is a quote based liquidity measure, whereas the other measures are trade based. Some investors may time their trades and wait for liquid periods, and investors with access to all trading venues trade on Euronext mainly when that market is the most liquid. Therefore, the trade based measures reflect particularly liquid times. The quote based liquidity measures are time weighted, and reveal that quoted liquidity disappears in fragmented periods.

Finally, the quoted spread on Euronext improves with visible fragmentation, while the quoted depth reduces with 30% at $VisFrag = 0.35$. Possibly, in fragmented periods high-frequency traders place many small and price improving limit orders. Given the reduction in $Depth(10)$ however, the gains of improved prices are more than offset by the lower quantities offered.

1.6 Robustness checks

In this section we investigate the robustness of our main results. First, we control for potential endogeneity issues by introducing firm-quarter fixed effects. These control for the simultaneous interactions between market structure, the degree of fragmentation, liquidity and competition in the market. In addition, this approach controls for a specific reverse causality issue, where fragmentation tends to be higher for high volume and more liquid stocks (Cantillon and Yin, 2010). To tackle remaining endogeneity problems of the visible fragmentation and dark trading variables we use an instrumental variables estimator. The instruments are (i) the number of limit to market orders on the new competing venues, (ii) the logarithm of the average order size of the new competing venues and (iii) the logarithm of dark order size; and their respective squares. We conclude by analyzing large and small firms separately, along with some additional robustness checks.

1.6.1 Regression analysis: firm-time effects

In this section we add to (1.2) firm-quarter dummies. Instead of a single dummy for a period of four years, we add 16 quarterly dummies per firm. This approach is similar to Chaboud, Chiquoine, Hjalmarsson, and Vega (2009), who analyze the effect of algorithmic trading on volatility for currencies, and add separate quarter dummies for each currency pair. The procedure is aimed to solve the following issues.

First, the firm-quarter dummies make the analysis more robust to the impact of the financial crisis and industry specific shocks. For example, if the financial crisis specifically affects certain firms or industries (e.g., the financial sector), and affects both liquidity and fragmentation, then the previous analysis might suffer from an omitted variables problem, leading to a bias in the coefficients on fragmentation. The firm-quarter dummies capture industry shocks and time-varying firm specific shocks.

Second, the firm-quarter dummies can control for potential self-selection problems. For example, Cantillon and Yin (2010) raise the issue that competition might be higher for high volume and more liquid stocks; an effect that will be absorbed by the firm-quarter dummies as long as most variation in volume is at the quarterly level.

Third, the firm-quarter dummies can, at least partially, control for dynamic interactions between market structure, competition in the market, the degree of fragmentation and liquidity. Specifically, such interactions are dynamic as, for example, a change in the current market structure will affect the level of competition in the future, which, in turn, will affect the market structure and liquidity in the future. Our approach controls for the long-term interactions of such forces by only allowing for variation in liquidity and fragmentation within a firm-quarter. Accordingly, the dummy variables absorb the variation between quarters, which is likely to be more prone to endogeneity issues.

The results for global liquidity reveal a similar pattern as those presented in the base regressions, as shown in panel A of Table 1.5 and displayed in the upper part of Figure 1.6. For the sake of brevity, the table only reports the coefficients of $VisFrag$, $VisFrag^2$ and $Dark$ for the Depth(X) measures, as these are the main focus of the paper. Results of the control variables and other liquidity measures are in line with those reported in Tables 1.3 and 1.4, and available upon request.

In the first regression, we observe that $Depth(10)$ monotonically increases with visible fragmentation, as the maximum of the curve lies beyond the highest observed value of visible fragmentation. There appears to be no harmful effect of visible fragmentation on liquidity close to the midpoint. This is not the case for the other depth levels, as the maximum lies around $Frag = 0.40$, implying a trade-off in the benefits and drawbacks of fragmentation.

Two additional findings emerge from the figure. First, at $VisFrag = 0.40$, the effect of visible fragmentation on $Depth(10)$ improves to 0.28 and $Depth(50)$ to 0.10, compared with 0.50 and 0.40 in the base case regressions in Table 1.3. The effect of visible fragmentation on liquidity is smaller but still highly significant. This is easily explained as the firm-quarter dummies absorb long-term trends in visible fragmentation, while only the day-to-day fluctuations remain. From the regression results, it appears that removing the long-term variation dampens the estimated daily effects. Second, liquidity deeper in the order book benefits less from visible fragmentation than liquidity close to the midpoint does. This finding was also observed in Figure 1.5, but becomes more pronounced. The fact that $Depth(10)$ still improves strongly with visible fragmentation suggests that competition of new trading venues mainly takes place at liquidity close to the midpoint. The coefficients on $Dark$ show a similar pattern as those reported in Table 1.3, but are about 15% lower in magnitude. That is, the detrimental effect of dark activity on liquidity remains.

The impact of visible fragmentation on local liquidity, including firm-quarter effects, is shown in panel B of Table 1.5 and the lower part of Figure 1.6. The figure shows that the results for the local order book have become more negative, as all depth measures reduce by 8% at $VisFrag = 0.40$. In the base specification, this reduction of liquidity was only observed for $Depth(10)$.

1.6.2 An instrumental variables approach

In the instrumental variables regressions we aim to solve for more general reverse causality issues of fragmentation and dark trading. For example, $VisFrag$ might be high because a stock is very liquid on a particular day; or $Dark$ might be high when an investor substitutes the visible market for dark trading because the visible market is illiquid. In such cases $VisFrag$ and $Dark$ depend on liquidity, causing us to make incorrect interpretations of the regression coefficients.

We employ an instrumental variables specification to alleviate these problems. We instrument $VisFrag$, $VisFrag^2$ and $Dark$ with (i) the ratio of the number of limit orders to the number of market orders on the visible competitors (Bats Europe, Chi-X, Nasdaq OMX and Turquoise),[19] (ii) the logarithm of the visible competitors average order size and (iii) the logarithm of the average $Dark$ order size, for each stock and day. We also add the squares of these variables, summing to six instruments, because we have a linear and quadratic term for fragmentation. These instruments are specifically aimed to tackle the aforementioned reverse causality issues. The first instrument, the ratio of limit to market orders on the visible competitors, is negatively related to fragmentation. After the startup of a new venue, typically the number of transactions is very low, while the available liquidity can already be substantial. As the venue reaches critical mass, the number of transactions will increase sharply, lowering the ratio and boosting fragmentation. We argue that the instrument is exogenous, as higher levels of visible liquidity should not reduce the ratio of limit to market orders on the visible competitors. The second instrument, the logarithm of the visible competitors order size, positively relates to fragmentation as larger orders

[19]The number of limit orders represent placed, modified and canceled limit orders.

typically increase competitors market share.[20] Since the regression controls for total traded volume, a shift of volume from the primary market to the new competitors should not improve liquidity except via fragmentation. The third instrument, the logarithm of average dark order size, positively affects dark activity. In a similar fashion to the previous instrument, larger dark orders increase dark market share. The instrument seems exogenous, as lower visible liquidity should not increase the average dark order size.

Unreported first stage estimations reveal that all instruments are statistically and economically significant. As expected, especially the ratio of messages to transactions and the logarithm of average visible competitors order size are particularly useful instruments for $VisFrag$, with standardized coefficients of -0.15 and 0.23, respectively. The logarithm of the average $Dark$ order size is a very strong instrument for $Dark$, with a standardized coefficient of 0.4. The six instruments can strongly predict fragmentation and dark activity as the Kleibergen-Paap and Angrist-Pischke Wald tests for weak and under identification are strongly rejected in all regressions, reported in the bottom part of Table 1.5. Unreported tests also reject the redundancy of all individual instruments, meaning each instrument improves the estimators asymptotic efficiency.

The IV regressions include firm-quarter dummies, and we use the two stage GMM estimator which is efficient in the presence of heteroscedasticity (Stock and Yogo, 2002). The regression results are reported in panel C and D of Table 1.5 and displayed in Figure 1.7. First, we observe that the magnitudes of the coefficients on visible fragmentation have strongly increased and are highly significant. At $VisFrag = 0.35$, global $Depth(10)$ and $Depth(50)$ improve with 100% and 32% compared with a com-

[20] O'Hara and Ye (2011) also use the logarithm of average order size as an excluded instrument in their Heckman correction model.

pletely concentrated market. The standard errors have strongly increased, as the IV procedure reduces the accuracy with which the coefficients are estimated. Importantly, Figure 1.7 shows that the optimal level of visible fragmentation is similar to previous specifications, and we confirm again that $Depth(10)$ benefits most from visible fragmentation. The coefficients on $Dark$ have slightly increased in magnitude compared with those reported in panel A and B of Table 1.5 and are highly significant. Assuming exogenous instruments, in economical terms the initial estimates did not suffer from endogeneity issues.

Turning to the IV results for local liquidity, panel D of Table 1.5 and the lower panel in Figure 1.7, we observe the following. First, due to increased standard errors, only the coefficients of $Depth(10)$ and $Depth(50)$ are significantly different from zero. The standard errors have increased because the instruments need to generate variation in $VisFrag$ and $VisFrag^2$, which are very collinear. Accordingly, the plots do not reveal a clear pattern and we cannot confirm previous results. In contrast, the coefficients on $Dark$ are again highly significant and negative, similar to previous findings.

Finally, we test the requirement that the set of instruments are uncorrelated with the error term. The joint null hypothesis of the overidentifying restrictions test is that the instruments are valid, i.e., uncorrelated with the error term, and that the instruments are correctly excluded from the estimated equation. The Hansen J test statistics and p-values are reported in the bottom part of panel C and D of Table 1.5, and do not reject the overidentifying restrictions in eight out of ten regressions. Only for global $Depth(40)$ and $Depth(50)$ exogeneity of the instruments is questioned. A GMM distance test reveals that the logarithm of the visible competitors order size causes this rejection. In unreported regressions, using subsets of the instruments or treating $Dark$ as exogenous does not affect the main results.

However, we prefer the current setup, as it allows us to perform overidentifying restrictions tests.

1.6.3 Small versus large stocks

The benefits and drawbacks of fragmentation on liquidity might hinge on certain stock characteristics, such as firm size. We pursue the point in question by executing the base specification regressions for large stocks, with an average market cap exceeding ten billion Euro, and small stocks, with an average market cap below 100 million Euro. The results for the global and local order books of 15 large and 14 small sample stocks are reported in Table 1.6, panel A to D. The coefficients for the global order book are plotted in Figure 1.8, and show two interesting results. First, the benefits of visible fragmentation are higher for large stocks than for small stocks. For large firms, the $Depth(10)$ is 64% higher at $VisFrag = 0.35$, while for small firms the maximum, at $VisFrag = 0.18$, has 30% more liquidity compared with a completely concentrated market. Second, the figure shows that the benefit of visible fragmentation for large stocks is monotonically positive, meaning there are no harmful effects of fragmentation. By contrast, the liquidity of small stocks is negatively affected for levels of visible fragmentation exceeding 0.36. This suggests that the benefits of visible fragmentation strongly depend on firm size. The harmful effect of *Dark* activity on liquidity is similar for small and large stocks.

Turning to the regressions in panel C and D of Table 1.6, we find that the local liquidity of large stocks also increases with visible fragmentation, while that of small stocks strongly decreases. That is, at $Frag = 0.35$, $Depth(10)$ of large stocks improves by 12%, while that of small stocks reduces with 38%. Again, this confirms that the drawbacks of a fragmented market place mainly hold for relatively small stocks. The fact that large stocks benefit more from visible fragmentation is in line with their actual levels of fragmentation, which is 0.41 in 2009, while for small stocks only 0.21.

1.6.4 Additional robustness checks

To investigate the sensitivity of our results, we perform a number of robustness checks. First, we execute the regressions with firm-quarter dummies, but only use observations from 2008 and 2009. The results do not change (not reported), likely because fragmentation especially took place in 2008 and 2009. This provides an additional robustness to potential time effects (e.g. the financial crisis), as the coefficients on fragmentation are estimated within a smaller time window. In addition, this covers for the fact that our dataset contains the ten best price levels on Euronext Amsterdam as of January 2008, while before only the best five price levels (as mentioned in footnote 8). Finally, this solves the potential issue that the data by Markit Boat on dark trades is available only as of November 2007.

Second, we execute the regressions in first differences, i.e. use the daily changes instead of the daily levels. By analyzing the day-to-day changes, we remove the long-term trends in the data. The results are very similar to those using firm-quarter dummies (not reported).

Third, instead of using $VisFrag$ to measure visible fragmentation, we use the market share of the traditional market (Euronext Amsterdam), and the qualitative results do not change. Finally, we have plotted higher order polynomials of $VisFrag$, and the inverted U-shapes remain, indicating that the finding of an optimal level of visible fragmentation is robust.

1.7 Conclusion

Nowadays, stocks are simultaneously traded on a variety of different trading systems, creating a fragmented equity market. We show that the effect of fragmentation on liquidity crucially depends on the type of trading venue – visible versus dark. Our results reveal a key role for pre-trade transparency, which we define as having a publicly displayed limit order book. Liquidity seems to reap the gains of competition for order flow in case of visible fragmentation, whereas dark trading appears to have detrimental effects.

The positive effect of visible fragmentation stems from competition between liquidity suppliers, as evidenced by the reduction in the reward of supplying liquidity. The negative effect of dark trading is consistent with a "cream-skimming" effect, where the dark markets mostly attract uninformed order flow which in turn increases adverse selection costs on the visible markets. We relate this finding to pre-trade transparency, which has been shown to reduce adverse selection costs (e.g., Boehmer, Saar, and Yu (2005)). As such, we provide a deeper understanding of the current view that market fragmentation improves liquidity. More general, our results imply that the type of trading venue determines the overall costs and benefits of competition between trading venues.

Next to separating visible from dark fragmentation, we explicitly differentiate between global and local liquidity. Global liquidity takes all relevant trading venues into account while local liquidity only the traditional stock market. Although global liquidity improves with visible fragmentation, local liquidity does not. That is, limit orders migrate from the local exchange to the competing trading platforms, such that an investor with only access to the traditional market is worse off. The reduction in liquidity close to the midpoint, i.e. at relatively good prices, can be more than 10% compared to the case of no visible fragmentation. In addition, we find that competition between trading venues is fiercer for larger stocks, as these are more fragmented and have a higher marginal benefit of visible fragmentation. Also, large stocks do not face the drawbacks of visible fragmentation like small stocks do. This suggests that the benefits and drawbacks of fragmentation also depend on certain stock characteristics, size in particular.

In sum, our results add to the policy discussion on competition in financial markets, which is amplified by recent financial regulation (Reg NMS in the US and MiFID in Europe). A caveat is that we cannot observe the liquidity in the dark markets, yet, the result remains that investors without access to dark markets are worse off. This result should be seen in the light of fair markets and investor protection.

Appendix A: Background on European financial market

This section gives a brief discussion on the contents of the Markets in Financial Instruments Directive (MiFID), effective November 1, 2007. By implementing a single legislation for the European Economic Area, MiFID aims to create a level playing field for trading venues and investors, which would ultimately improve market quality. The regulation entails three major changes to achieve this goal.

First, competition between trading venues is introduced by abolishing the "concentration rule"[21] and allowing three types of trading systems to compete for order flow. These are reg-

[21] The "concentration rule", adopted by some EU members, obliges transactions to be executed at the primary market as opposed to internal settlement. This creates a single and fair market on which all investors post their trades, according to a time and price priority. The repeal of the rule however allows markets to become fragmented and increases competition between trading venues (Ferrarini and Recine, 2006).

ulated markets (RMs), Multilateral Trading Facilities (MTFs) and Systematic Internalisers (SIs). RMs are the traditional exchanges, matching buyers and sellers through an order book or through dealers. A firm chooses on which RM to list, and once listed, MTFs may decide to organize trading in that firm as well. MTFs, who closely resemble ECNs in the US, are similar to RMs in matching third party investors, but have different regulatory requirements and 'rules of the game'. For example, MTFs and RMs can decide upon the type of orders that can be placed, and the structure of fees, i.e. fixed fees, variable fees as well as make or take fees.[22] In order to survive, MTFs need to obtain a sufficient level of liquidity from order flow of their owners and outside investors. The largest MTFs with visible liquidity are Chi-X, Bats Europe, Nasdaq OMX and Turquoise. Lastly, SIs are organized by investment banks where customers trade against the inventory of the SI or with other clients, resembling market dealers.

MiFIDs second keystone refers to transparency which guarantees the flow of information in the market. As the number of trading venues increases, information about available prices and quantities in the order books becomes dispersed. Consequently, for investors to decide on the optimal venue and to evaluate order execution, a sufficient degree of pre-trade and post-trade transparency is necessary. Pre-trade transparency rules require trading venues to make (part of) their order books public and to continuously update this information. However, a number of waivers exist regarding pre-trade transparency. In particular, there is the "large-in-scale orders waiver", the "reference price waiver", the "negotiated-trade waiver", and the "order management facility waiver".[23] These waivers are used by MTFs such

[22]Make and take fees are costs charged to investors supplying and removing liquidity, respectively. Make fees can be negative, such that providers of liquidity receive a rebate for offering liquidity.

[23]See also Directive 2004/39/EC, article 29.

as dark pools and broker-dealer crossing networks who only have to report executed trades. Whether transparency has improved is a topic of current debate, which is complicated by increasingly fragmented markets, technological innovations and shortcomings in the quality of post-trade information.[24]

The third and final pillar of MiFID is the introduction of the best-execution rule, which obliges investment firms to execute orders against the best available conditions with respect to price, liquidity, transaction costs and likelihood and speed of execution. However, such a broad definition of best-execution policy allows investment firms to decide themselves where to route their orders to. For example, an investment firm may stipulate an execution policy of trading on one market only. In absence of a clear benchmark, it becomes difficult for investors to evaluate the quality of executed trades and the overall performance of an investment firm (Gomber and Gsell, 2006). This is the main difference between MiFID and its US counterpart, Reg NMS, which solely focusses on the price dimension.[25] For an extensive summary of the implementation process of MiFID we refer the interested reader to Ferrarini and Recine (2006).

Appendix B: liquidity measures

The liquidity measures other than $Depth(X)$ are explained in this section. We calculate the price impact and the effective and realized spreads based on trades and weighted over all trades per day. In contrast, $Depth(X)$, quoted spread and quoted depth are liquidity measures based on quotes offered in the limit order book and time weighted over the trading day. The effective spread measures direct execution costs while the realized spread

[24]CESR proposes changes to MiFID, July 29, 2010, ref. 10-926.

[25]In the U.S., the price of every trade is reported to the consolidated tape, such that the performance of a broker can clearly be evaluated.

takes the order's price impact into account. The realized spread is often considered to be the compensation for the liquidity supplier. Denote MQ_o as the quoted midpoint before an order takes place and MQ_{o+5} the quoted midpoint, but five minutes later and $D = \{+1, -1\}$ for a buy and a sell order respectively, then

$$Effective\, half\ spread = \frac{Price - MQ_o}{MQ_o} * D * 10.000, \quad (1.3)$$

$$Realized\, half\ spread = \frac{Price - MQ_{o+5}}{MQ_o} * D * 10.000, \quad (1.4)$$

$$Price\ impact = \frac{MQ_{o+5} - MQ_o}{MQ_o} * D * 10.000. \quad (1.5)$$

The price impact, realized and effective spread are first calculated per trade, based on the midpoint of that trading venue. Then, all calculations are averaged over the trading day, weighted by traded volume. Next, we average over trading venues, again weighted by trading venue. This approach gives the average spread in the whole market. Limited computer power is the reason we use the midpoint of the trading venue where the trade took place instead of the consolidated midpoint. That is, creating a consolidated midpoint quote-by-qoute, as is required for the effective and realized spreads, is computationally much more burdensome than creating a consolidated order book using one-minute snapshots.[26] The price impact and realized spread are calculated between 09.00 - 16.25, while the effective spread on 9.00 - 16.30. Therefore, $Effective\ spread \approx Realized\ spread + Price\ impact$. The global quoted spread is based on the best price in the consolidated order book (based on the one-minute snapshot data, see Section 1.4.1) and expressed in basis

[26]Our dataset also has a consolidated tape constructed by Thomson Reuters, containing best prices, quantities and all visible trades in the market. However, extensive checking shows that the time stamp of these trades may differ up to three seconds from the time stamp of the same trades in the original file.

points, while the local quoted spread is based on the order book of Euronext. In a similar fashion, the quoted depth aggregates the number of shares times their prices, expressed in Euros, or

$$Quoted\ spread = \frac{P^{ASK} - P^{BID}}{MQ_o} * 10.000, \qquad (1.6)$$

$$Quoted\ depth = P^{ASK} * Q^{ASK} + P^{BID} * P^{BID}. \qquad (1.7)$$

Note that the quoted depth on Euronext can be larger than that of the consolidated order book, for example when Chi-X offers a better price but with a lower quantity. The quoted spread of the consolidated order book is always equal or better than that of Euronext. Finally, the quoted depth is identical to $Depth(10)$ when the quoted spread equals 20 basis points.

Table (1.1) Descriptive statistics: time series.
The table shows the medians of the liquidity measures on a yearly basis for the global and local order book (Panel A), and additional descriptive statistics of the sample stocks (Panel B). The medians are based on 52 firms and 250 trading days per year (11.250 observations). *Depth*(X) is expressed in €1000s and represents the offered liquidity within X basis points around the midpoint. The effective spread, realized spread, price impact and quoted spread are measured in basis points. The price impact and realized spread are based on a 5 minute time window. The quoted depth is the amount of shares, in €1000s, offered at the best bid and ask price of the global and local order book. The descriptives show the natural logarithm of firm size, traded volume, realized return volatility (Ln SD) and algorithmic trading. Return volatility is defined as the daily standard deviation of 15 minute returns on the midpoint. Typically, this standard deviation is lower than one, so the natural logarithm becomes negative. *Algo* represents the number of electronic messages in the market divided by total traded volume (per €10.000). An electronic message occurs when a limit order in the order book is executed, changed or canceled.

Panel A: Liquidity measures								
	Global				*Local*			
	2006	2007	2008	2009	2006	2007	2008	2009
Depth(10)	102	134	50	66	101	127	39	36
Depth(20)	263	299	125	187	261	279	94	93
Depth(30)	367	404	183	291	359	366	141	155
Depth(40)	441	463	228	367	422	406	178	206
Depth(50)	488	505	258	420	463	426	205	244
Effective Spread	14.1	11.2	15.1	13.2	13.8	11.1	14.5	13.1
Realized Spread	2.5	1.1	-0.1	0.0	2.4	1.1	-0.2	0.1
Price Impact	10.4	9.4	14.3	13.3	10.4	9.5	14.2	13.5
Quoted Spread	13.3	10.9	14.5	12.0	13.5	11.5	16.8	14.7
Quoted Depth	101	82	41	32	102	85	40	30

Panel B: Descriptive statistics				
	2006	2007	2008	2009
Ln Size	14.7	15.0	14.7	14.4
Ln Volume	16.7	17.1	17.0	16.5
Algo	1.9	2.6	6.6	28.4
Ln SD	-6.2	-6.1	-5.5	-5.6

Table (1.2) Descriptive statistics of visible fragmentation and dark trading.
The yearly standard deviation, mean and quartiles of visible fragmentation and dark trading are reported. Visible fragmentation (VisFrag) is defined as $1 - HHI$, where HHI is based on the market shares of *visible* trading venues. Dark is the percentage of traded volume executed at dark pools, crossing networks and Over The Counter, available only as of November 2007. The statistics are based on daily observations per firm. As such, each observation is equally weighted; when weighing according to traded volume the average dark fraction is approximately 37%.

Year	Stdev	Mean	25^{th}	50^{th}	75^{th}
VisFrag					
2006	0.081	0.027	0.000	0.000	0.010
2007	0.066	0.026	0.000	0.000	0.017
2008	0.119	0.097	0.000	0.044	0.168
2009	0.153	0.275	0.143	0.291	0.403
Total	0.150	0.106	0.000	0.015	0.182
Dark					
2008	0.173	0.255	0.134	0.225	0.331
2009	0.169	0.250	0.131	0.221	0.327

Table (1.3) The effect of fragmentation on global liquidity.
The dependent variable in models (1) - (5) is the logarithm of the Depth(X) measure based on the consolidated order book. The Depth(X) is expressed in Euros and represents the offered liquidity within (X) basis points around the midpoint. The effective spread, realized spread, price impact and quoted spread, (6) - (9), are measured in basis points. Ln quoted depth is the logarithm of the quoted depth in Euros (10). VisFrag is the degree of visible market fragmentation, defined as $1 - HHI$. Dark is the percentage of order flow executed OTC, on crossing networks, dark pools and internalized. Algo represents the number of electronic messages divided by traded volume in the market (per €100); the other variables are explained in the descriptive statistics and Table 1.2. The regressions are based on 1022 trading days for 52 stocks, and have firm fixed effects and quarter dummies. T-stats are shown below the coefficients, calculated using Newey-West (HAC) standard errors (based on 5 day lags). ***, ** and * denote significance at the 1, 5 and 10 percent levels, respectively.

	(1) Ln Depth(10)	(2) Ln Depth(20)	(3) Ln Depth(30)	(4) Ln Depth(40)	(5) Ln Depth(50)	(6) Effective Spread	(7) Realized Spread	(8) Price Impact	(9) Quoted Spread	(10) Ln Quoted Depth
VisFrag	2.844***	2.080***	2.188***	2.334***	2.420***	-31.32***	1.047	-32.40***	-41.94***	-0.888***
	(15.9)	(21.0)	(26.4)	(28.5)	(29.7)	(-16.0)	(0.5)	(-17.2)	(-24.8)	(-11.6)
VisFrag^2	-4.069***	-2.875***	-3.081***	-3.381***	-3.616***	33.94***	-6.615**	40.61***	55.72***	0.305**
	(-13.4)	(-15.7)	(-19.2)	(-21.1)	(-22.7)	(9.8)	(-2.0)	(12.8)	(19.0)	(2.0)
Dark	-0.914***	-0.685***	-0.587***	-0.540***	-0.503***	2.960***	-1.147	4.101***	4.476***	-0.544***
	(-20.2)	(-23.7)	(-24.0)	(-22.9)	(-21.8)	(3.7)	(-1.4)	(7.8)	(9.8)	(-26.8)
Ln Size	1.008***	0.623***	0.491***	0.427***	0.387***	-6.996***	-3.220***	-3.779***	-4.906***	0.279***
	(24.2)	(24.8)	(24.6)	(22.5)	(20.9)	(-15.8)	(-7.6)	(-8.6)	(-9.4)	(17.7)
Ln Price	-0.012	0.062***	0.069***	0.069***	0.067***	-0.137	-0.207	0.0728	1.759***	-0.056***
	(-0.5)	(3.8)	(4.7)	(4.8)	(4.5)	(-0.4)	(-0.7)	(0.3)	(4.3)	(-4.2)
Ln Vol	0.576***	0.429***	0.385***	0.353***	0.327***	-2.304***	0.380**	-2.682***	-3.724***	0.233***
	(40.9)	(45.7)	(47.0)	(47.4)	(46.9)	(-11.3)	(2.0)	(-17.0)	(-29.4)	(43.2)
Ln SD	-0.619***	-0.537***	-0.466***	-0.420***	-0.384***	7.312***	-4.963***	12.28***	5.733***	-0.223***
	(-40.0)	(-52.2)	(-54.0)	(-52.5)	(-50.7)	(31.9)	(-23.6)	(53.0)	(33.5)	(-33.7)
Algo	-0.116***	-0.106***	-0.094***	-0.097***	-0.097***	4.565***	0.034	4.527***	4.514***	-0.007
	(-5.4)	(-7.1)	(-7.6)	(-8.4)	(-8.8)	(14.6)	(0.1)	(13.7)	(15.5)	(-0.8)
Obs	46879	46879	46879	46879	46879	46879	46879	46879	46879	46879
R^2	0.461	0.663	0.681	0.659	0.641	0.236	0.042	0.331	0.352	0.673

Table (1.4) The effect of fragmentation on local liquidity.
The dependent variable in models (1) - (5) is the logarithm of the Depth(X) measure based on the order book of Euronext Amsterdam. The Depth(X) is expressed in Euros and represents the offered liquidity within (X) basis points around the midpoint. The effective spread, realized spread, price impact and quoted spread, (6) - (9), are measured in basis points. Ln quoted depth is the logarithm of the quoted depth in Euros (10). VisFrag is the degree of visible market fragmentation, defined as $1 - HHI$. Dark is the percentage of order flow executed OTC, on crossing networks, dark pools and internalized. Algo represents the number of electronic messages divided by traded volume in the market (per €100); the other variables are explained in the descriptive statistics and Table 1.2. The regressions are based on 1022 trading days for 52 stocks, and have firm fixed effects and quarter dummies. T-stats are shown below the coefficients, calculated using robust Newey-West (HAC) standard errors (based on 5 day lags). ***, ** and * denote significance at the 1, 5 and 10 percent levels, respectively.

	(1) Ln Depth(10)	(2) Ln Depth(20)	(3) Ln Depth(30)	(4) Ln Depth(40)	(5) Ln Depth(50)	(6) Effective Spread	(7) Realized Spread	(8) Price Impact	(9) Quoted Spread	(10) Ln Quoted Depth
VisFrag	1.025***	0.006	0.162*	0.411***	0.589***	-31.07***	0.679	-31.79***	-35.21***	-0.416***
	(5.7)	(0.1)	(1.8)	(4.5)	(6.4)	(-14.7)	(0.3)	(-17.7)	(-18.3)	(-5.8)
$VisFrag^2$	-2.942***	-0.427**	-0.294	-0.624***	-0.960***	36.40***	-6.349*	42.82***	48.65***	-1.248***
	(-9.6)	(-2.2)	(-1.6)	(-3.3)	(-5.0)	(9.9)	(-1.7)	(14.2)	(17.0)	(-9.1)
Dark	-0.947***	-0.722***	-0.647***	-0.621***	-0.596***	2.348***	-1.784**	4.127***	4.043***	-0.541***
	(-20.8)	(-23.5)	(-23.3)	(-22.3)	(-21.6)	(2.9)	(-2.1)	(8.0)	(8.3)	(-27.5)
Ln Size	0.958***	0.542***	0.407***	0.344***	0.307***	-7.414***	-3.364***	-4.054***	-4.471**	0.224***
	(22.5)	(20.6)	(18.9)	(16.1)	(14.5)	(-16.5)	(-7.7)	(-9.3)	(-2.6)	(14.0)
Ln Price	-0.052**	0.044**	0.060***	0.060***	0.053***	0.069	-0.163	0.236	1.894***	-0.049***
	(-2.1)	(2.6)	(3.7)	(3.6)	(3.2)	(0.2)	(-0.5)	(0.8)	(4.7)	(-3.5)
Ln Vol	0.578***	0.426***	0.382***	0.351***	0.326***	-2.081***	0.598***	-2.676***	-4.066***	0.244***
	(40.1)	(43.1)	(43.1)	(42.6)	(41.7)	(-9.4)	(2.9)	(-17.4)	(-6.6)	(45.4)
Ln SD	-0.609***	-0.534***	-0.469***	-0.425***	-0.391***	7.312***	-5.057***	12.37***	8.337***	-0.223***
	(-38.4)	(-48.0)	(-47.7)	(-45.3)	(-43.2)	(30.8)	(-22.9)	(53.6)	(6.5)	(-34.4)
Algo	-0.128***	-0.187***	-0.206***	-0.216***	-0.214***	3.869***	0.108	3.756***	6.01***	0.056***
	(-5.9)	(-13.0)	(-16.2)	(-17.4)	(-17.5)	(12.4)	(0.4)	(11.9)	(18.4)	(6.7)
Obs	46879	46879	46879	46879	46879	46879	46879	46879	46858	46858
R^2	0.498	0.677	0.671	0.636	0.607	0.208	0.039	0.335	0.121	0.717

Table (1.5) The effect of fragmentation on liquidity: firm-quarter fixed effects and IV.
Panel A and B show the regression results for global and local depth respectively, where firm-quarter dummies are added. Panel C and D show the IV results, where VisFrag, $VisFrag^2$ and Dark are instrumented by (i) the number of electronic messages to transactions on the visible competitors, (ii) the logarithm of the visible competitors average order size, (iii) the logarithm of the average Dark order size; and their respective squares. The IV regressions also include firm-quarter dummies. The Hansen J statistic test the overidentifying restrictions (p-value reported below). The dependent variable is the logarithm of the Depth(X) measure based on the global and local order book. VisFrag is visible fragmentation and Dark is the percentage of order flow executed in dark markets. The control variables (not reported) are Ln size, Ln price, Ln volume, Ln volatility and Algo, as explained in Table 1.2. The regressions are based on 1022 trading days for 52 stocks. T-stats are shown below the coefficients, calculated using Newey-West (HAC) standard errors (based on 5 day lags). ***, ** and * denote significance at the 1, 5 and 10 percent levels, respectively.

	(1) Ln Depth(10)	(2) Ln Depth(20)	(3) Ln Depth(30)	(4) Ln Depth(40)	(5) Ln Depth(50)	(6) Ln Depth(10)	(7) Ln Depth(20)	(8) Ln Depth(30)	(9) Ln Depth(40)	(10) Ln Depth(50)
	Panel A: Global, Firm-Quarter dummies					Panel B: Local, Firm-Quarter dummies				
VisFrag	0.984***	0.756***	0.700***	0.659***	0.593***	0.259*	-0.0250	-0.0542	-0.0538	-0.0617
	(6.4)	(9.7)	(10.6)	(10.7)	(10.2)	(1.8)	(-0.4)	(-0.9)	(-0.9)	(-1.1)
$VisFrag^2$	-0.749***	-0.697***	-0.872***	-0.927***	-0.877***	-1.049***	-0.419***	-0.413***	-0.428***	-0.409***
	(-2.8)	(-4.7)	(-7.0)	(-7.9)	(-7.8)	(-4.1)	(-3.1)	(-3.4)	(-3.7)	(-3.6)
Dark	-0.750***	-0.532***	-0.480***	-0.443***	-0.417***	-0.723***	-0.535***	-0.491***	-0.458***	-0.430***
	(-20.2)	(-21.4)	(-26.5)	(-27.1)	(-26.8)	(-19.6)	(-21.5)	(-26.3)	(-26.7)	(-26.2)
	Panel C: Global, IV					Panel D: Local, IV				
VisFrag	8.146***	5.300***	3.933***	3.287***	2.773***	2.877**	0.653	-0.125	-0.255	-0.492
	(6.1)	(8.4)	(9.0)	(8.3)	(7.6)	(2.2)	(1.1)	(-0.3)	(-0.7)	(-1.4)
$VisFrag^2$	-17.63***	-11.27***	-8.164***	-6.844***	-5.668***	-7.307**	-1.659	0.545	0.850	1.491
	(-5.1)	(-6.7)	(-7.2)	(-6.7)	(-6.1)	(-2.1)	(-1.0)	(0.5)	(0.8)	(1.6)
Dark	-0.836***	-0.600***	-0.531***	-0.496***	-0.463***	-0.798***	-0.600***	-0.538***	-0.502***	-0.470***
	(-12.5)	(-14.2)	(-15.2)	(-15.4)	(-15.1)	(-12.6)	(-14.7)	(-15.0)	(-14.8)	(-14.2)
Hansen J	2.451	4.094	8.173	17.16	25.66	7.506	7.218	3.019	0.907	2.217
Hansen p	0.484	0.252	0.0426	0.001	0.000	0.057	0.065	0.389	0.824	0.529
	First stage results: Kleibergen-Paap weak ID F stat: 108. Angrist-Pischke weak ID F stat: 48 (Frag), 36 ($VisFrag^2$), 855 (Dark).									

Table (1.6) The effect of fragmentation on liquidity: large and small firms.
The base specification regressions are executed separately for the 15 smallest stocks (average market cap ¡ 100 million) and the 14 largest stocks (average market cap ¿ 10 billion); for the global and local order books. The dependent variable is the logarithm of the Depth(X) measure. The Depth(X) is expressed in Euros and represents the offered liquidity within (X) basis points around the midpoint. VisFrag is the degree of visible fragmentation, defined as $1 - HHI$. Dark is the percentage of order flow executed OTC, on crossing networks, dark pools and internalized. For the sake of brevity, the coefficients on the control variables are not reported, as they are very similar to those of Tables 1.3 and 1.4. The control variables are Ln size, Ln price, Ln volume, Ln volatility and Algo, as explained in Table 1.2. The regressions contain firm fixed effects and quarter dummies. T-stats are shown below the coefficients and calculated using Newey-West (HAC) standard errors (based on 5 day lags). ***, ** and * denote significance at the 1, 5 and 10 percent levels, respectively.

	(1) Ln Depth(10)	(2) Ln Depth(20)	(3) Ln Depth(30)	(4) Ln Depth(40)	(5) Ln Depth(50)	(6) Ln Depth(10)	(7) Ln Depth(20)	(8) Ln Depth(30)	(9) Ln Depth(40)	(10) Ln Depth(50)
	Panel A: Global, large firms					Panel B: Local, large firms				
VisFrag	1.458***	1.150***	1.072***	1.052***	1.058***	0.640***	0.478***	0.355**	0.352**	0.393**
	(10.1)	(9.1)	(8.1)	(7.4)	(7.1)	(5.1)	(3.9)	(2.6)	(2.3)	(2.4)
$VisFrag^2$	-0.555*	-0.463*	-0.334	-0.374	-0.438	-0.869***	-0.484*	0.0768	0.194	0.141
	(-2.0)	(-1.8)	(-1.2)	(-1.3)	(-1.4)	(-3.5)	(-1.9)	(0.3)	(0.6)	(0.4)
Dark	-0.833***	-0.598***	-0.497***	-0.461***	-0.443***	-0.813***	-0.671***	-0.579***	-0.540***	-0.520***
	(-23.0)	(-17.3)	(-13.8)	(-12.0)	(-11.3)	(-21.3)	(-16.6)	(-13.1)	(-11.4)	(-10.7)
	Panel C: Global, small firms					Panel B: Local, small firms				
VisFrag	2.992***	2.086***	1.805***	1.649***	1.443***	1.330**	0.380	0.173	0.139	0.0487
	-4.9	-6.8	-7.4	-7.5	-6.9	(2.1)	(1.2)	(0.6)	(0.5)	(0.2)
$VisFrag^2$	-8.300***	-5.373***	-4.706***	-4.290***	-3.825***	-6.230***	-3.045***	-2.406***	-2.144***	-1.844***
	(-6.9)	(-8.4)	(-8.9)	(-9.1)	(-8.7)	(-5.0)	(-4.5)	(-4.0)	(-3.7)	(-3.2)
Dark	-1.180***	-0.714***	-0.687***	-0.639***	-0.613***	-1.205***	-0.765***	-0.757***	-0.729***	-0.711***
	(-8.3)	(-9.6)	(-12.4)	(-13.1)	(-13.4)	(-8.3)	(-9.2)	(-10.9)	(-11.2)	(-11.3)

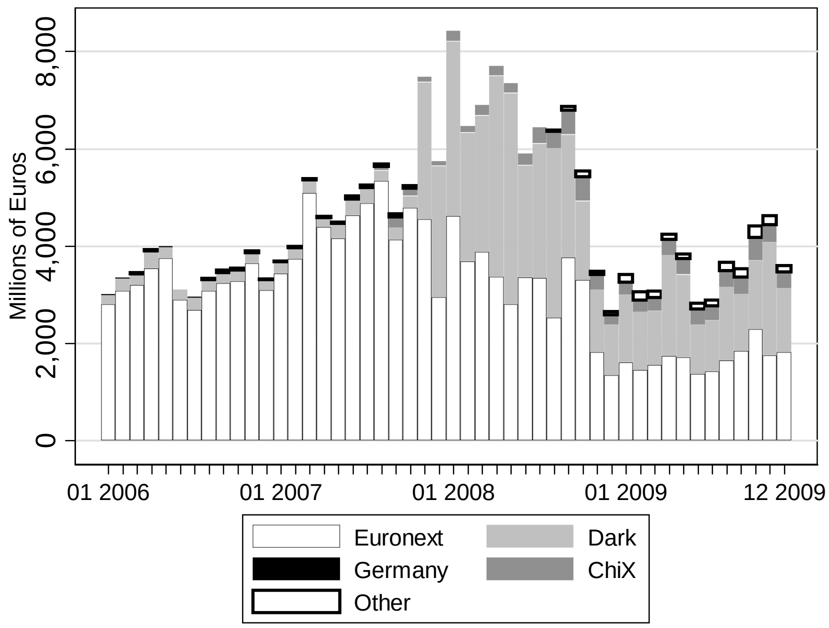

Figure (1.1) Traded Volume in millions of Euros.
The figure displays monthly averages of the daily traded volume in millions, aggregated over the 52 AEX Large and Mid cap constituents. Euronext consists of Amsterdam, Brussels, Paris and Lisbon. Germany combines all the German cities while Other represents Bats Europe, Nasdaq OMX Europe, Virt-x and Turquoise combined. Finally, Dark represents the orderflow executed Over The Counter, at crossing networks, dark pools and internalized; however, these numbers are not available prior to November 2007.

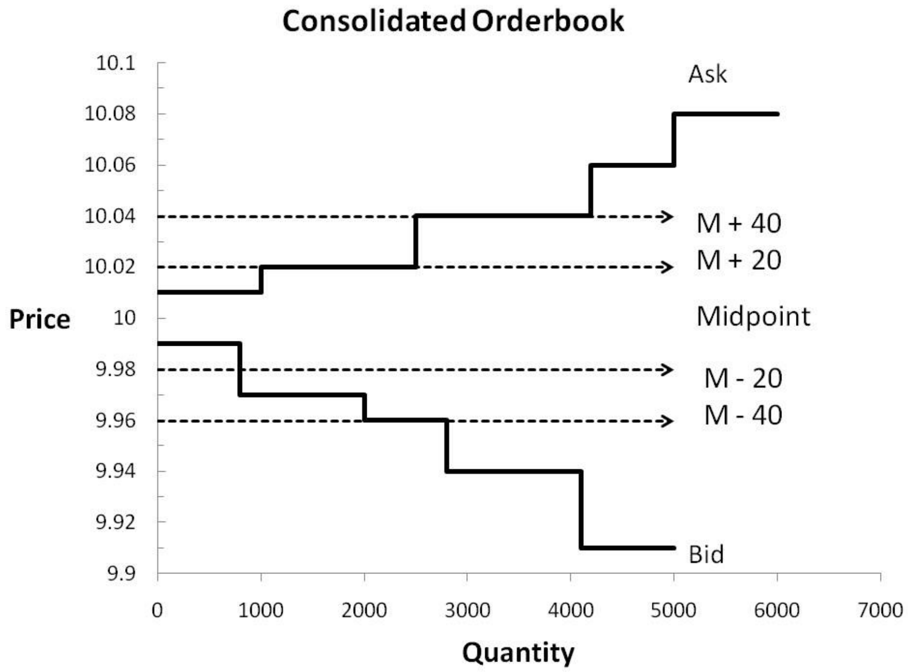

Figure (1.2) Snapshot of a hypothetical limit order book. Depth(20) aggregates liquidity offered within the interval of (M - 20bps, M + 20bps), which are 2500 shares on the ask side and 800 on the bid side. Depth(40) contains 4100 and 2800 shares on the ask and bid side respectively. The number of shares offered are converted to a Euro amount.

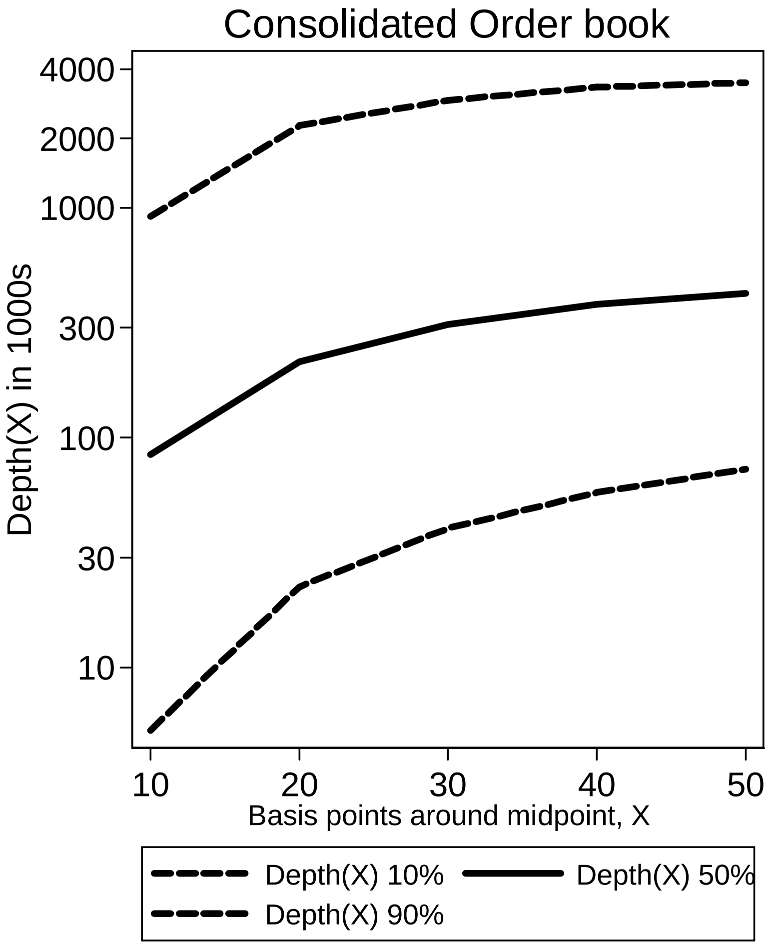

Figure (1.3) Depth in the consolidated order book.
The figure shows the 10, 50 and 90$^{\text{th}}$ percentiles of the Depth(X) measure, expressed on a logarithmic scale in €1000s. The measure aggregates the Euro value of shares offered within a fixed amount of basis points X around the midpoint, shown on the horizontal axes. The consolidated order book represents liquidity to a global investor, where the order books of Euronext Amsterdam, Deutsche Boerse, Chi-X, Virt-X, Turquoise, Nasdaq OMX Europe and Bats Europe are aggregated. The percentiles are based on the 52 AEX large and mid cap constituents between 2006 - 2009.

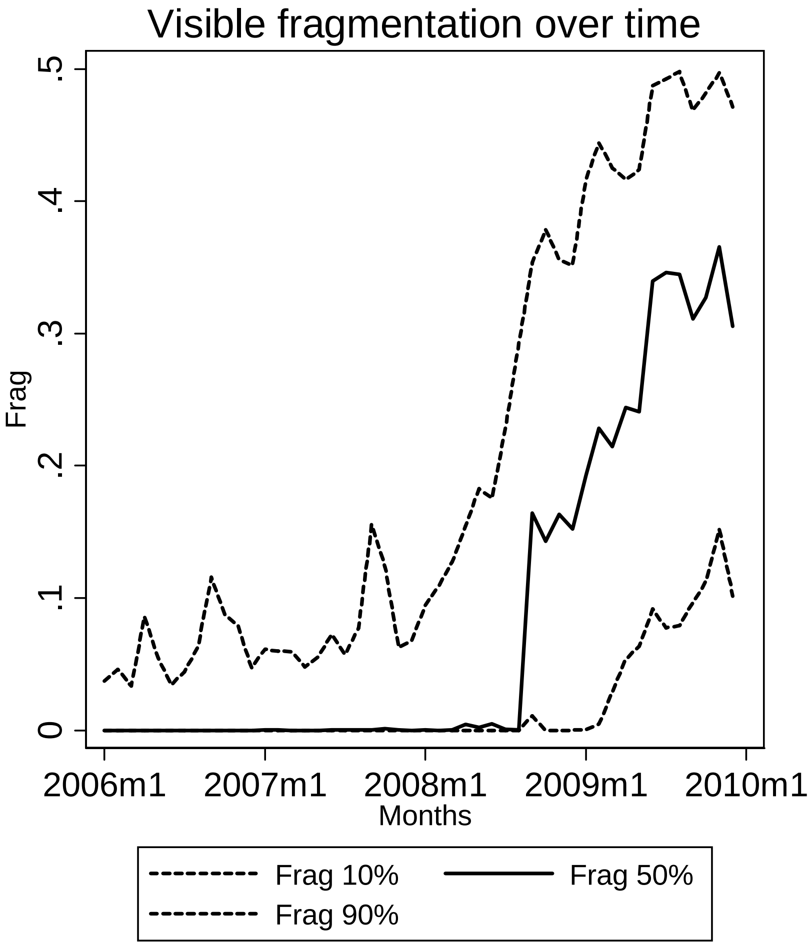

Figure (1.4) Visible fragmentation of AEX large and Mid cap firms.
The monthly 10, 50 and 90$^{\text{th}}$ percentiles of VisFrag are shown, for the 52 AEX large and mid cap stocks between 2006 - 2009. VisFrag equals 1 - HHI, based on the number of shares traded at the following trading venues: Euronext (Amsterdam, Brussels, Paris and Lisbon together), Deutsche Boerse, Chi-X, Virt-X, Turquoise, Nasdaq OMX Europe and Bats Europe. Trades executed OTC, on crossing networks, on dark pools or internalized are not taken into account, as we analyze the degree of market fragmentation of visible liquidity.

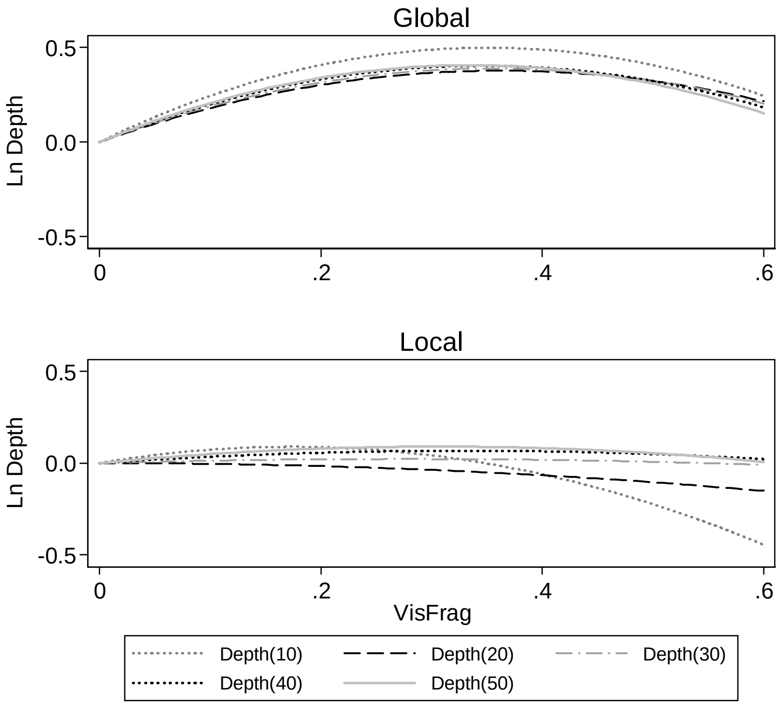

Figure (1.5) The effect of visible fragmentation on global and local liquidity.
The regression coefficients of visible fragmentation on liquidity are plotted, for the global order book (upper panel, model (1) - (5) of Table 1.3) and local order book (lower panel, model (1) - (5) of Table 1.4). The vertical axis displays the logarithm of the Depth(X), while the horizontal axis shows the level of visible fragmentation (Frag), defined as (1 - HHI).

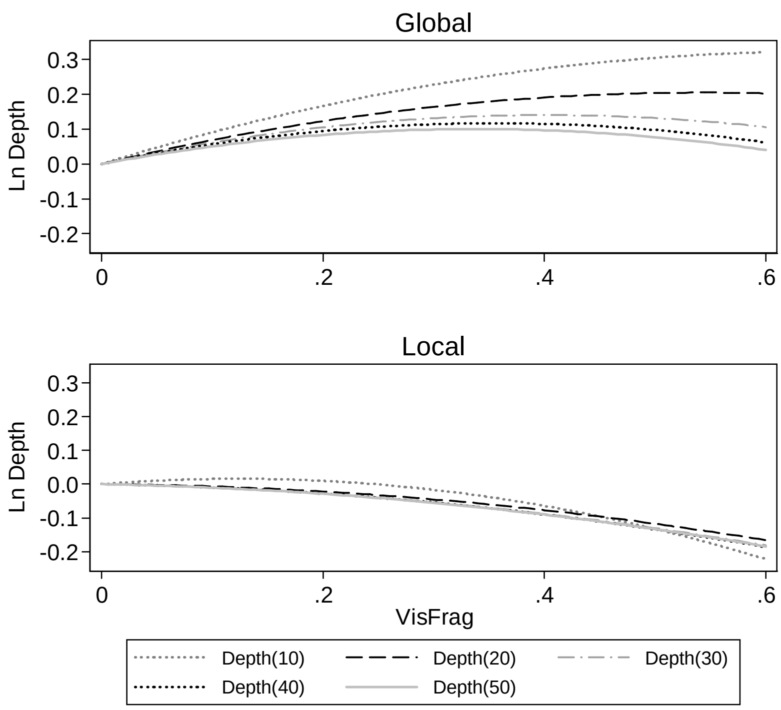

Figure (1.6) Visible fragmentation and liquidity: firm-quarter dummies.

The regression coefficients of visible fragmentation on liquidity of Table 1.5 are plotted, where the regressions have firm-quarter dummies. The upper panel shows the global order book and the lower panel the local order book. The vertical axis displays the logarithm of the Depth(X), while the horizontal axis shows the level of visible fragmentation, defined as (1 - HHI).

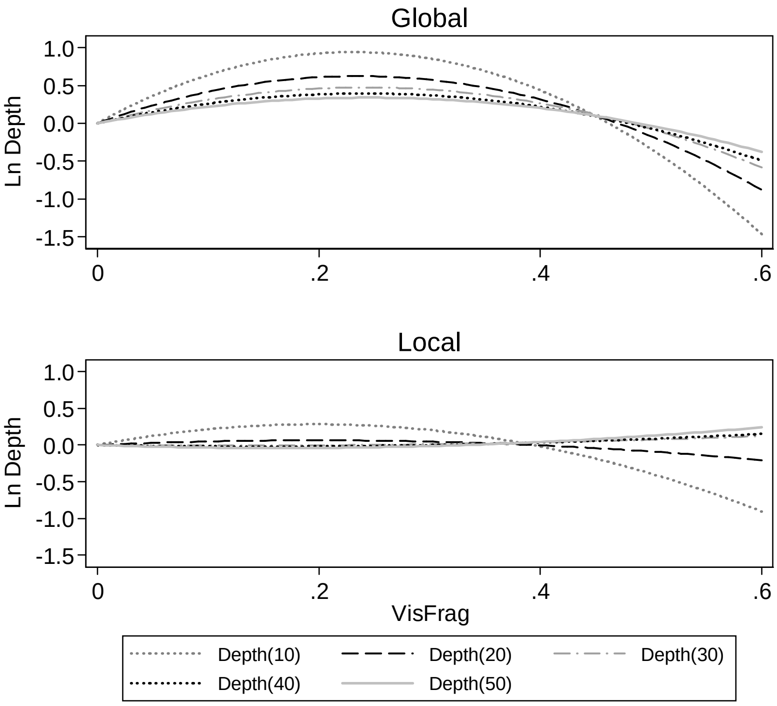

Figure (1.7) Visible fragmentation and liquidity: IV regressions. The IV regression coefficients of visible fragmentation on liquidity of Table 1.5 are plotted. The instruments are (i) the number of electronic messages to transactions on the visible competitors, (ii) the logarithm of the visible competitors average order size, (iii) the logarithm of the average Dark order size; and their respective squares. The regressions include firm-quarter dummies. The vertical axis displays the logarithm of the Depth(X), while the horizontal axis shows the level of visible fragmentation, defined as (1 - HHI).

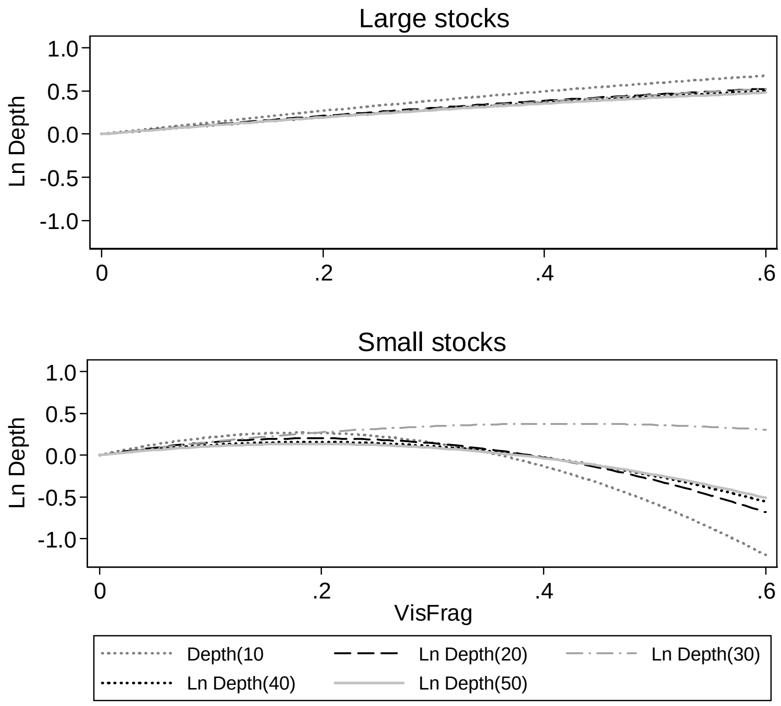

Figure (1.8) Visible fragmentation and global liquidity: small versus large stocks.
The regression coefficients of visible fragmentation on liquidity are plotted, for large and small stocks (regressions (1) - (5) in panel A and B, Table 1.6). Large stocks consist of the 14 stocks with an average market cap exceeding ten billion Euro, while the 15 small stocks have a market cap smaller than 100 million Euro. The regressions include firm-quarter dummies. The vertical axis displays the logarithm of the Depth(X), while the horizontal axis shows the level of visible fragmentation, defined as (1 - HHI).

Table (1.7) Appendix Table A1

Descriptive statistics of sample firms: cross section

The dataset covers daily observations for 52 AEX large and mid cap constituents, from 2006 to 2009. All variables in the table are averages. Firm size and traded volume are expressed in millions of Euros. Return volatility reflects the daily standard deviation of 15 minute returns on the midpoint and is multiplied by 100. Euronext represents the market share of executed trades on Euronext Amsterdam. Dark is the market share of Over The Counter trades, Systematic Internalisers and dark pools; this number is available as of November 2007.

Firm	Size	Price	Volume	Return Vol	Dark	Euronext
Aalberts	1.3	29.03	7.4	0.39	7.94	89.99
Adv. Metal. Group	0.6	21.24	8.6	0.78	17.95	78.88
Aegon	16.6	10.26	161.0	0.46	15.19	76.95
Ahold	11.4	8.52	120.0	0.28	18.59	74.96
Air France	5.6	19.95	63.9	0.40	15.03	78.07
Akzo nobel	12.3	44.93	147.0	0.30	19.55	73.47
Arcadis	0.9	31.40	3.3	0.41	10.78	87.52
Arcelor Mittal	44.1	35.69	388.0	0.50	24.15	70.16
Asm Int.	0.8	14.47	7.0	0.44	10.21	86.34
ASML	8.1	17.78	144.0	0.39	16.74	75.54
Bamn Group	1.7	18.84	14.8	0.41	11.81	83.47
Binckbank	0.6	10.60	4.2	0.36	10.50	88.54
Boskalis	2.2	39.50	13.5	0.42	12.73	83.66
Corio	3.5	51.12	26.9	0.36	14.96	79.07
Crucell	1.0	15.49	9.2	0.36	8.48	89.82
CSMN	1.5	20.93	8.0	0.29	11.99	86.04
Draka Hold.	0.5	13.01	3.6	0.55	16.21	77.53
DSM	6.1	32.02	73.0	0.29	16.86	76.89
Eurocomm. Prop	1.2	32.47	5.8	0.37	11.13	87.00
Fortis	34.5	22.83	437.0	0.38	13.37	83.51
Fugro	2.8	38.99	25	0.34	10.30	84.84
Hagemeyer	2.0	3.76	43.4	0.31	0.00	99.28
Heijmans	0.6	25.08	3.8	0.40	8.31	90.56
Heineken	16.7	34.06	100.0	0.28	18.85	74.16
Imtech	1.2	15.03	9.2	0.40	17.97	77.51
ING	50.1	22.76	904.0	0.44	14.22	81.26
Nutreco	1.5	42.41	15.1	0.28	12.34	85.30
Oce	0.9	9.95	8.5	0.40	10.52	86.85
Ordina	0.4	10.96	2.8	0.41	7.29	90.00
Philips	26.7	23.52	287.0	0.32	21.04	71.27
R. Dutch Shell	88.5	24.23	528.0	0.27	21.51	69.59
R. KPN	20.4	10.92	220.0	0.26	23.23	69.82
R. ten cate	0.5	26.69	2.3	0.40	10.09	87.67
R. Wessanen	0.6	8.49	4.3	0.32	9.07	87.53
Randstad	4.5	35.18	38.3	0.39	14.20	79.13
Reed Elsevier	8.1	11.31	74.1	0.27	20.47	72.01
SBM Offshore	2.8	26.39	35.6	0.36	13.17	80.13
Smit Int.	10.7	48.06	4.6	0.38	15.21	80.10
Sns Reaal	2.9	11.66	11.5	0.42	12.94	85.51
Tele Atlas	1.8	20.06	37.0	0.33	8.24	68.20
Tnt	10.7	24.96	99.3	0.33	19.63	73.61
Tomtom	3.0	25.33	47.7	0.54	10.43	83.15
Unibail Rodamco	11.9	143.62	172.0	0.36	34.65	57.61
Unilever	32.2	23.62	327.0	0.26	17.59	73.87
Usg People	0.6	9.27	8.5	0.71	19.01	68.01
Vastned	1.0	55.99	5.0	0.32	10.46	87.45
Vdr Moolen	0.2	4.74	2.2	0.35	2.12	97.45
Vedior	2.9	16.70	67.7	0.27	2.99	96.48
Vopak Int.	2.3	36.51	8.9	0.30	12.24	84.77
Wavin	0.7	7.94	5.8	0.49	11.80	86.92
Wereldhave	1.6	78.78	16.0	0.28	13.55	81.88
Wolters Kluwers	5.5	18.10	42.1	0.30	13.13	78.13

Chapter 2

Liquidity: What you see is what you get?

Abstract

In a model of competition between two limit order books I show that the aggregate liquidity may overstate the actual liquidity available to certain investors. The excess is caused by high-frequency traders operating as market makers, who have an incentive to duplicate their liquidity supply across markets. Then, after a trade on one venue they will quickly cancel outstanding limit orders on others. The magnitude depends on the fraction of traders that may access several venues simultaneously, i.e., those who use Smart Order Routing Technology (SORT). The empirical results strongly support the predictions of the model.

JEL Codes: G10; G14; G15;

Keywords: Market microstructure, Fragmentation, High Frequency Trading, Smart Order Routing

2.1 Introduction

Two important trends have drastically changed equity markets in recent years (SEC, 2010). First, technological innovations have led to high-frequency trading, a trading strategy where a computer algorithm analyzes market data and trades at extremely high speed. Second, competition between trading venues has caused a dispersion of trading volume and liquidity across venues, i.e., the market has become fragmented. These two changes in the structure of equity markets might strongly affect the optimal behavior of investors. Indeed, as high-frequency traders operate on several trading venues simultaneously, the order flow and liquidity of these venues become strongly interrelated.[1] This paper argues that the interrelation causes a substantial overestimation of liquidity when aggregated across trading venues.

I underpin this conjecture in a model of competition between two centralized limit order books. The model predicts that liquidity suppliers have an incentive to duplicate their limit order schedules on both venues, but cancel this additional liquidity after a trade on the competing venue. Due to the cancellations, a single trade reduces liquidity on all venues simultaneously such that the depth aggregated over both venues overstates true liquidity. The empirical results strongly support this hypothesis: within a second after a trade limit orders are cancelled on competing venues with a value of 58% of the trade size.

My first contribution to the literature is a theoretical model. The model is based on the framework of a pure limit order market with adverse selection of Glosten (1994), and extends the specification of Sandås (2001) to a two-venue setting. High-

[1]The high level of interaction between markets became apparent during the flash crash, i.e., between the E-mini S&P 500 futures and the individual stocks (SEC-CFTC "Findings regarding the market events of May 6, 2010", 2010).

frequency market makers supply liquidity with limit orders, whereas potentially informed traders demand liquidity with market orders. In this setting, only a fraction of the traders has the technological infrastructure to submit market orders to both venues simultaneously, i.e., use smart order routing technology (SORT). The model effectively creates market segmentation, because non-SORT traders cannot access the liquidity of both venues. In equilibrium, a limit order faces higher adverse selection costs when executed against a SORT trader. The reason is that conditional upon execution, the trader might have traded on the competing venue already and therefore the combined trades are larger and more informed on average. Consequently, a lower fraction of SORT traders reduces the adverse selection risk faced by the market makers, who in turn increase their equilibrium liquidity supply on both venues. However, the additional liquidity follows from market segmentation only, and will be cancelled after a trade on the competing venue; hence the term duplicate liquidity.

A general prediction of the model is that the information content revealed by a trade leads to cancellations of limit orders on the same side of the order books of competing venues. Cancellations would not occur in a single venue setting, which implies that I must test whether the observed cancellations are greater than zero.

In addition, the model offers three more specific predictions. First, a larger fraction of SORT traders reduces the amount of duplicate limit orders and therefore *decreases* consolidated liquidity. This result contrasts with Foucault and Menkveld (2008), who find that a larger fraction of SORT traders *increases* consolidated liquidity because of enhanced competition between market makers. While this channel is also present in my model, it gets dominated by the channel that a larger fraction of SORT traders reduces duplicate liquidity. However, when all investors use SORT competition between market makers strictly improves liquidity, as in Glosten (1998) and Foucault and Menkveld (2008). Second, a larger fraction of SORT traders reduces consolidated liquidity, and therefore also the *consolidated liquidity impact* of a trade, i.e., the change in consolidated liquidity due to a one unit trade. Third, a larger fraction of SORT traders active on the competing venue reduces the *cross-venue liquidity* impact of a trade on the current venue.

My second contribution to the literature is an empirical analysis. The dataset contains the entire limit order books of all relevant trading venues with publicly displayed data, for a sample of FTSE 100 stocks in November 2009.[2] These stocks are traded in a fairly fragmented environment, as the traditional market (the London Stock Exchange) executes 66% of lit trading volume, leaving 34% to four competing venues (Chi-X, Bats Europe, Turquoise and Nasdaq OMX Europe). I test the models' main predictions by investigating the short-term correlations between the supply and demand of liquidity across trading venues. Specifically, I regress the change in the liquidity supply at the bid or ask side of one venue on the buy and sell trading volumes of the five individual venues. The regressions include lagged trading volumes of up to ten seconds away to measure the impact of trades over time. I sample once every 100 milliseconds to analyze high-frequency trading strategies. To the best of my knowledge, this paper is the first to analyze the impact of high-frequency trading strategies on the liquidity supply across trading venues.

The empirical results strongly support the models' main prediction that trades are followed by limit order cancellations on competing venues. That is, within 100 milliseconds, transactions on the three most active trading venues are followed by cancel-

[2]The results are the same when I do the analysis for 10 stocks taken from the largest French index (the CAC40) in October 2009.

lations of limit orders on the *same side* of competing venues with a value of 38% to 85% of the transaction size. Further, these cancellations increase to 53% to 149% of the transaction size after one second (depending on the trading venue). As a result, the liquidity aggregated over all venues overstates liquidity available to non-SORT traders, since e.g. a 100 share trade reduces liquidity by more than 138 shares. This finding is particularly relevant to algorithms designed to split up large trades over time, both SORT and non-SORT, as the liquidity impact of each individual small trade is quite large. The fact that liquidity shocks immediately spill over to other venues is not captured by static liquidity measures, such as the quoted depth.

To test the remaining predictions of the model I construct a proxy for the fraction of SORT traders. In the model, the SORT traders submit two market orders to both venues simultaneously, such that the market makers are unable to revise or cancel their limit orders in between the two trades. Empirically, I state that two trades on different venues occur "simultaneously" when (i) the state of the limit order book of the venue executing the current trade has not changed since the previous trade, (ii) both trades occur within 100 milliseconds of each other and (iii) both trades are either purchases or sales. The first restriction directly follows from the model, and states that the current trade executes against limit orders which have not yet incorporated the information content of the previous trade. I group the strings of market orders send simultaneously to several venues into *trade sequences*, which overall represent 37% of LSE trading volume (45% of entrant volume). The percentage of trading volume part of trade sequences is a proxy for the fraction of SORT traders in the market. The details of this measure and some caveats are discussed in Section (2.3.3).

Per hour for each stock, I estimate the consolidated and cross-venue liquidity impact for trades on the traditional market and the combined new entrant venues. The average consolidated liquidity impact of LSE trades is -1.4 (-1.75 for the entrants), which implies that a 100 share trade is followed by cancellations of 40 shares, such that consolidated liquidity reduces by 140 shares. The cross-venue coefficients range between -0.58 and -0.75. Next, using one observation per hour of each stock, I regress the fraction of SORT traders on the consolidated and cross-venue liquidity impacts and on the level of consolidated liquidity.

I find strong empirical support for the specific predictions of the model. A one standard deviation increase in the fraction of SORT traders (0.09) reduces the level of consolidated liquidity by 7%. Furthermore, it increases the consolidated liquidity impact (i.e., reduces the magnitude) by 0.075 for LSE trades and 0.12 for entrant trades. Given the average of -1.4 and -1.75, these coefficients are economically large. The coefficients of the fraction of SORT traders on the cross-venue liquidity impacts range between 0.05 and 0.16 per standard deviation. The regression results are robust to controlling for trading volume, volatility, the average order size and a proxy for algorithmic trading.

The main policy implication of the model is that fair markets require investors to be able to split up trades *simultaneously* across several venues. When a trader leaves a millisecond delay between the split, high-frequency traders may observe the first part of the trade and will quickly cancel their limit orders on competing venues before the second part arrives.[3] This issue also applies to Intermarket Sweep Orders, which are specifically designed to demand liquidity across several venues. If all investors use SORT then trades would not be followed by excessive cancellations.

The model is a simple and stylized representation of an elec-

[3]For this reason, the Royal Bank of Canada introduced an order routing technology (THOR) that uses routing latency to synchronize the arrival time of trades at each venue.

tronic limit order book, and the purpose is to show (i) how segmentation arises when only a fraction of the traders uses SORT and (ii) how segmentation causes overestimation of consolidated liquidity. The model ignores potentially important features of real markets, such as the decision of traders to place limit or market orders, differences in speed between traders and between trading venues, and the inventory concerns of market makers, among others.

Most related to this work is the literature on competition between electronic limit order books. Pagano (1989b) predicts that all trading activity should divert to the trading system with the lowest transaction costs, and only unstable equilibria may exist when two venues have identical cost structures. In contrast, Glosten (1998) shows that two electronic limit order markets can coexist when tick sizes are discrete and time priority rules apply. Since time priority is absent *across* venues, competition between liquidity suppliers increases, which in turn raises aggregate liquidity. This point is further developed in Foucault and Menkveld (2008), who coin this channel the "queue-jumping" effect. Competition between exchanges also arises through differences in the tick size, where the venue with the smallest tick size becomes most liquid (Biais, Bisière, and Spatt (2010) and Buti, Rindi, Wen, and Werner (2011)). My model is consistent with the above findings, and adds the presence of duplicate limit orders when some investors cannot access all trading venues. Maglaras, Moallemi, and Zheng (2012) study competition between N venues, and show that the optimal order routing strategies cause an equilibrium in the trading rate and queue length of each venue, such that the N-dimensional problem collapses to a single dimensional problem. Their results hold mainly on lower frequencies because of the assumption that trading is smoothly and continuously, whereas my results hold mainly on high frequencies.

Empirical research on competition between exchanges typically focusses on its impact on aggregate liquidity and welfare. O'Hara and Ye (2011) find that competition between exchanges reduces the effective cost of trading. Degryse, de Jong, and van Kervel (2011) show that aggregate liquidity increases by competition between venues with publicly displayed limit order books, but worsens by competition of opaque markets. Instead, the current paper studies the effect of competing exchanges on the liquidity supply at each of the individual exchanges.

Market segmentation may also arise when investors have different trading speeds. Biais, Foucault, and Moinas (2011) show that high-frequency trading facilitates the search for trading opportunities, but increases adverse selection costs for slow traders. As a result, the equilibrium level of investment in high-frequency technology exceeds the welfare maximizing level. Pagnotta and Philippon (2012) obtain the same result in a model where competing exchanges invest in trading speed, rather than the investors, which reduces direct competition. Hoffmann (2012) shows that high-frequency traders face lower adverse selection costs on average, which in turn causes slow traders to post less aggressive limit orders. The result is an ambiguous effect on welfare, which depends on the stocks fundamental volatility. Cartea and Penalva (2011) show that high-frequency traders can extract rents from liquidity motivated traders when they operate as intermediaries. McInish and Upson (2011) provide empirical evidence that high-frequency traders pick off slower traders in the US, due to the regulation that effectively causes slow investors to trade against badly priced stale quotes. Hasbrouck and Saar (2011) argue that trading speeds affect the competition between liquidity suppliers in a single trading venue. This paper shows that market segmentation causes duplicate limit orders in fragmented markets.

The paper also relates to recent research on high-frequency traders who act as market makers (e.g., Jovanovic and Menkveld (2011), Menkveld (2011) and Guilbaud and Pham (2011)). Such market makers gain the bid-ask spread by offering liquidity at both sides of the limit order book, while simultaneously managing adverse selection costs (e.g., Glosten and Milgrom (1985)) and inventory risk (e.g., Ho and Stoll (1981)). This paper analyzes market making when trading is fragmented across electronic limit order books.

2.2 The model

In this section I first describe the duplicate limit order hypothesis. Then, I quantify this duplicate limit order effect by placing the problem into an adverse selection framework. In essence, the model introduces market segmentation into a combination of Glosten (1998), Sandås (2001) and Foucault and Menkveld (2008). I contribute to Foucault and Menkveld (2008) by allowing for adverse selection. Rather than analyzing the single exchange setting in Sandås (2001), I focus on two competing centralized limit order books. Compared to Glosten (1998), I introduce market segmentation by constraining some traders to have access to one trading venue only, which causes market makers to duplicate limit orders across markets. The model nests Glosten (1998) and Sandås (2001).

2.2.1 Duplicate limit order hypothesis

In a fragmented trading environment time priority is not enforced between trading venues, whereas price priority is enforced only when the trader has access to both venues.[4] Price priority implies that limit orders with a better price are executed before

[4]In the US, price priority across markets is enforced by law, Reg NMS. Time priority however, crucial for this hypothesis, is not enforced.

those with a worse price, while time priority entails that limit orders placed first are executed first.

Following from the absence of time priority across venues, I hypothesize that liquidity suppliers have an incentive to duplicate their limit order schedules across several venues. After a trade on one venue, they will cancel limit orders on other venues because the trade is informative about the fundamental value. By duplicating limit orders across venues the liquidity suppliers increase their execution probability and expected profits.[5] A tradeoff arises however, as there is a probability that limit orders on several venues will be executed simultaneously. In this case a liquidity supplier trades too much, which in the model is costly because of adverse selection risk. Therefore, liquidity suppliers have an incentive to become very fast to minimize the risk of simultaneous execution.

Who might pursue this duplicate limit order strategy? Arguably, high-frequency traders who operate as market makers can strongly benefit from the increased execution probability. At the same time, using state of the art technology they can monitor several venues simultaneously and may quickly cancel limit orders to reduce the risk of simultaneous execution. [6] In contrast, this strategy is likely not very attractive to "regular traders."For some traders, the technology to monitor continuously might be too expensive. Other traders might use algorithms to optimally split up large quantities over time, in which case they will not cancel limit orders since each child order is part of a large parent order.

[5]Note that this strategy does not work in a single exchange setting due to time priority.

[6]In this context, trades on two venues occur simultaneously when the liquidity supplier is not fast enough to adjust his outstanding limit orders after the first trade. Effectively, his quotes are stale when the second trade comes in.

2.2.2 Model setup

The goal is to study the equilibrium liquidity supply at two competing venues A and B in an adverse selection framework. I first provide a short description of the trading game, and leave the details about the strategies of the investors for the next subsections.

Consider two types of investors: market makers and traders. The market makers are risk neutral and supply liquidity by placing limit orders on one or both venues. They are profit maximizing, competitive and use high-frequency trading technology to quickly access all venues. The traders demand liquidity by placing market orders. Traders may have private information about the fundamental value or liquidity motives to trade, and therefore want to trade quickly using market orders.[7] The market orders are informed on average and therefore impose adverse selection risk to the market makers.

The timeline of the trading game is presented in Figure (2.1), which proceeds as follows. Trading occurs sequentially over periods indexed by t, and each period consists of three stages. First, the market makers arrive consecutively and place limit orders on one or both venues. They will do so until no market maker finds it optimal to place additional limit orders, i.e., until a competitive equilibrium is reached. Next a trader arrives who places market orders with total size x, and depending on her type executes limit orders on venue A, B or both. Finally, the executed trade reveals information about the fundamental value of the asset to all market makers, who update their expectation of the true value. Now, the game starts over and is repeated for every trade.

[7] Traders have both information and liquidity motives, and therefore the model does not need noise traders.

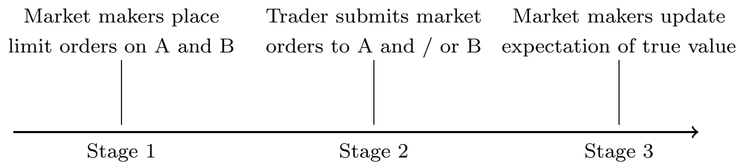

Figure (2.1) Timeline of the trading game
In stage 1 the competitive market makers place limit orders on both venues. In stage 2 a potentially informed trader submits market orders, with exogenous size x. The trader is one of three types: those with access to venue A only (fraction α), to B only (β), or to A and B ($\gamma = 1-\alpha-\beta$), in which case the trader uses smart order routing technology. In stage 3, all market makers observe the executed market orders and update their expectation of the fundamental value and limit order schedules accordingly. The trading game is repeated for every trade.

2.2.3 The trader

The trader is randomly drawn from a population of traders, which consists of three types. The first type only has access to venue A (fraction α), the second type only to venue B (fraction β), and the third type uses smart order routing technology (SORT) to access both venues simultaneously (fraction $\gamma = 1 - \alpha - \beta$). Simultaneously is defined here as sending market orders to both venues very quickly, such that the market makers are unable to update their limit order schedules between the arrival of the trades. When both venues offer the same best price, SORT traders are indifferent as to where to send their trades to. In this case, they simply use a tie-breaking rule, which posits that with probability π they first buy shares on venue A, and with the complementary probability they first buy shares on venue B. The parameters α, β and π are exogenous, as I focus on a high-frequency trading environment.

Four reasons motivate why some investors are not able to trade on both venues simultaneously. First, human traders with access to both venues might trade too slowly, creating a delay of several milliseconds when they split up a trade across two venues. In this case, high-frequency market makers have sufficient time to update their limit order schedules in between trades. This high-frequency trading strategy is known as latency arbitrage, and segments the market according to trading speed. As a result, human traders effectively have access to one exchange only.[8] For this reason, the Royal Bank of Canada introduced an order routing technology (named THOR) that takes each venues' routing latency into account to synchronize the arrival time of trades to all venues. In the US, inter-market sweep orders serve a similar purpose. Empirically, McInish and Upson (2011) show that slow traders are adversely selected because they often trade against stale quotes. A second source of millisecond delays stems from active searches for hidden liquidity. In this case a trader sends a (marketable) limit order to one venue, and must wait for the venues' response to confirm whether hidden liquidity is executed before choosing the quantity to trade at another venue. The delay possibly gives high-frequency traders sufficient time to update their limit orders in between the trades. Third, smart order routing technology might be too expensive for some traders, as it requires fixed costs for technological infrastructure, software, programmers and access and analysis of data feeds etc.[9] Fourth, fixed costs of splitting a trade across two venues could make it more economical to trade through one price but save on the transaction costs (such that fixed clearing and settlement costs are paid only once).

The trader is a buyer or seller with equal probability, and has

[8]In fact, when market makers can update in between the two trades, the slow trader will never prefer to split up his trade across two venues. The reason is that the market makers update their limit order schedules symmetrically across two venues. Therefore, if a trader finds it optimal to send the first part of the trade to venue A, then he will also send the second part there.

[9]A small investor could hire a SORT trader to execute his position for a small fee. However, this may lead to information leakage and dual trading (see e.g., Röell (1990)).

a reservation price $p_m > p_1$ at which she is not willing to buy.[10] The average order size is the same for all types, with mean ϕ and exponential density function

$$f(x) = \frac{1}{\phi} \exp(-\frac{x}{\phi}) \qquad \text{if } x > 0 \text{ (market buy).} \tag{2.1}$$

The cumulative distribution function is $F(\cdot)$. Assuming an exogenous order size simplifies the analysis, although in reality the size depends on several factors, such as the current state of the limit order book, the trader's expectation of the fundamental value and her current holdings. However, the interest of the model lies in the behavior of the market makers, described next.

2.2.4 The market makers

Market makers place limit orders on venue A and B, and know that trades are informative about the assets' fundamental value V_t. Market makers therefore face adverse selection risk, which will determine the equilibrium liquidity supply. Market makers monitor trades on both markets, and update their expectation of the fundamental value using a price impact function

$$E(V_{t+1}|x_t) = E(V_t) + \lambda x_t. \tag{2.2}$$

Following Sandås (2001), I assume a linear price impact function with coefficient $\lambda > 0$, since buy trades typically contain positive information with respect to the fundamental value (and similarly, sells contain negative information). Thus, larger orders cause more adverse selection costs and a greater price impact. In the remaining analysis I will omit the time subscript t.

Market makers place limit orders on a discrete pricing grid,

[10] This small assumption prevents the trade from walking up the limit order book too much in case of a thin order book, but does not affect the outcome of the model.

$\{p_1, p_2, ...p_k\}$ for the ask side. I focus on the ask side only, prices larger than p_0, as the bid side is analogous. Denote the total number of shares offered on venue $j \in \{A, B\}$ at each price level by $\{Q_{j1}, Q_{j2}, \ldots, Q_{jk}\}$.

Given the price impact function, we can calculate the expected profit of a limit order placed on any location q in the queue of limit orders on each venue. The expected profit depends on the expected value of the asset conditional upon execution of the limit order (the "upper-tail expectation" in Glosten (1994)). For a limit order on venue A, this value is $E(V|x > q)$ when the trader immediately goes to A (denoted $E_q(V)$ for brevity), and $E(V|x > q + Q_{B1})$ when the trader first buys all shares Q_{B1} on venue B and then goes to A. Denote the probability that the incoming order x is larger than q as $\overline{F}_q = 1 - F(q)$, then the expected profit of a limit order on price level 1 of venue A is

$$\Pi_{A,q} = (\alpha+\gamma\pi)\overline{F}_q(p_1-E_q(V))+\gamma(1-\pi)\overline{F}_{q+Q_{B1}}(p_1-E_{q+Q_{B1}}(V)). \tag{2.3}$$

In the first term, the limit order executes against traders going to venue A only (α) and against SORT traders who choose to trade on venue A first ($\gamma\pi$). Then, the expected profit is simply the price minus the expected value of the asset conditional on $x > q$. The second term represents SORT traders who first buy all the shares offered at price p_1 on venue B, and then buy shares on venue A ($\gamma(1 - \pi)$). As market makers only realize profits when their limit order is executed, the second term of the expected profit is lower since the incoming trade is larger and more informed, i.e., $E_{q+Q_{B1}}(V) > E_q(V)$.

Not surprisingly, we observe that the expected profit of a limit order on venue A depends on the number of shares offered on venue B. Therefore, to obtain the equilibrium liquidity supply we need to solve for the profit equation of limit orders on both venues simultaneously. The profitability of a limit order on price level 1 of venue B at location q is

$$\Pi_{B,q} = (\beta+\gamma(1-\pi))\overline{F}_q(p_1 - E_q(V)) + \gamma\pi\overline{F}_{Q_{A1}+q}(p_1 - E_{Q_{A1}+q}(V)). \tag{2.4}$$

2.2.5 Equilibrium

The model is in equilibrium when no market maker can profitably place an additional limit order on any price level (as in Glosten (1994), proposition 2). Therefore, the expected profit of the *marginal limit order*, i.e., the single share offered at the end of the queue of limit orders, must equal zero for all price levels on each venue. Following Sandås (2001), I substitute q by Q_{A1} (the marginal limit order) and integrate the profit equation over the distribution of the incoming order x, using equations (2.1) and (2.2)

$$\Pi_{A,Q_{A1}} = \int_{Q_{A1}}^{\infty} (\alpha + \gamma\pi)(p_1 - V - \lambda x)\frac{1}{\phi}\exp(-\frac{x}{\phi})\mathrm{d}x + \int_{Q_{A1}+Q_{B1}}^{\infty} \gamma(1-\pi)(p_1 - V - \lambda x)\frac{1}{\phi}\exp(-\frac{x}{\phi})\mathrm{d}x = 0.$$

The first integral goes to infinity, because the marginal limit order is executed for any trade larger or equal to Q_{A1}. For demonstrational purposes the previous equations contain γ, which I next substitute with $(1-\alpha-\beta)$ to calculate the solutions. Solving the integral equation yields

$$\Pi_{A,Q_{A1}} = (\alpha+\pi(1-\alpha-\beta))(p_1 - V - \lambda(\phi + Q_{A1}))\exp(-\frac{Q_{A1}}{\phi}) +$$

$$(1-\pi)(1-\alpha-\beta)(p_1-V-\lambda(\phi+Q_{A1}+Q_{B1}))\exp(-\frac{Q_{A1}+Q_{B1}}{\phi})=0.$$

The zero expected profit condition implies that the first line of the last equation is positive and the second line negative. In equilibrium, the market makers expect to lose to SORT traders that go to venue B first, and profit from traders that go to venue A first. Similarly, solving the integral for venue B yields

$$\Pi_{B,Q_{B1}}=\pi(\beta+(1-\pi)(1-\alpha-\beta))(p_1-V-\lambda(\phi+Q_{B1}))\exp(-\frac{Q_{B1}}{\phi})+$$

$$\pi(1-\alpha-\beta)(p_1-V-\lambda(\phi+Q_{A1}+Q_{B1}))\exp(-\frac{Q_{A1}+Q_{B1}}{\phi})=0.$$

The two equations with two unknowns can be solved implicitly

$$Q_{A1}=\frac{p_1-E(V)-\lambda\phi}{\lambda}-\frac{\gamma(1-\pi)Q_{B1}}{\gamma(1-\pi)+(\alpha+\gamma\pi)\exp(\frac{Q_{B1}}{\phi})}, \tag{2.5}$$

$$Q_{B1}=\frac{p_1-E(V)-\lambda\phi}{\lambda}-\frac{\gamma\pi Q_{A1}}{\gamma\pi+(1-(\alpha+\gamma\pi))\exp(\frac{Q_{A1}}{\phi})}. \tag{2.6}$$

The first terms in the above equations are identical, and equal the optimal quantity offered in a single venue setting (the solution obtained by Sandås (2001)).[11] However, we subtract a non-negative second term, implying that the offered quantities on the individual venues are weakly lower in a fragmented market. The second term reflects the adverse selection costs incurred by the market makers when a SORT trader buys at the competing venue first.

The zero expected profit condition holds for prices deeper in the order book too. Now, the expected profit equation consists

[11]Sandås also adds a fixed limit order execution cost, which is easily added but does not yield new insights in this model.

of three terms, as a limit order on price level 2 on venue A might get executed by traders of type α (who first buy Q_{A1} and then Q_{A2}), by type $\gamma\pi$ (who first buy Q_{A1}, Q_{B1} and then Q_{A2}), or by type $\gamma(1-\pi)$ (who first buy Q_{B1}, Q_{A1}, Q_{B2} and then Q_{A2}). I solve the system for prices deeper in the order book recursively, which gives an implicit solution similar to equation (2.6).

2.2.6 Testable implications

In this section I discuss a static prediction of the model, which holds at each point in time, and two dynamic predictions, which compare equilibrium liquidity before and after a trade.

The static prediction follows from the solutions for Q_{A1} and Q_{B1} in equation (2.6), and a simple proof is provided in the Appendix.

Proposition 1 *When the liquidity at both venues has reached the equilibrium, an increasing fraction of SORT traders γ strictly reduces consolidated depth. In addition, the consolidated depth in a fragmented market is strictly higher than the liquidity offered in a single exchange setting.*

A larger fraction of SORT traders implies that more traders have access to the liquidity of both venues simultaneously, such that the market is less segmented. Effectively, market makers face higher adverse selection costs when trading with a SORT trader, because there is a probability that she traded on the competing venue already. Therefore, a higher fraction of SORT traders increases expected adverse selection costs and reduces equilibrium liquidity supply.[12] This prediction is opposite to the result of Foucault and Menkveld (2008) that a larger fraction of SORT traders *increases* equilibrium liquidity supply. In

[12]The consolidated liquidity reduces by γ, because the derivatives of Q_{A1} and Q_{B1} with respect to γ are both negative.

their model, market makers may jump the queue of limit orders on one venue by placing limit orders on the other. Such "queue jumping" is more profitable when the fraction of SORT traders increases, which enhances competition between market makers and equilibrium liquidity supply. This channel is also present in the current model, but gets dominated by the channel that a larger fraction of SORT traders the increases adverse selection costs of market makers. Indeed, the queue jumping effect explains the second part of the proposition that the consolidated liquidity in a fragmented market is higher than that in a single exchange setting, even when all traders use SORT (which is also documented by Glosten (1998)).

The following two propositions relate to the impact of a trade on the equilibrium liquidity supply, i.e., the change in liquidity before and after a trade. These are the main contributions of the model.

Proposition 2 *Define the "consolidated liquidity impact" of a trade as the impact of a one unit trade on the equilibrium consolidated liquidity supply.*

1. *In the single exchange setting, the consolidated liquidity impact equals -1.*

2. *The consolidated liquidity impact decreases by γ. For $\gamma < 1$, the consolidated liquidity impact is strictly smaller than -1, such that the impact of a trade on consolidated liquidity is larger than the trade size.*

In part 1, a one unit trade reduces consolidated liquidity with one unit because the price impact of the trade equals the slope of the limit order book.[13] In part 2, a decrease in γ reduces the

[13]For a trade of size x (see equation (2.2)), the marginal impact on Q_{A1} is $\frac{\partial Q_{A1}}{\partial x} = \frac{\partial Q_{A1}}{\partial V}\frac{\partial V}{\partial x} = -\frac{1}{\lambda}\lambda = -1$.

market makers expected adverse selection costs, who in turn increase their liquidity supply (proposition 1). But given that the information content of a trade λ is held fixed, it must be that the increased liquidity is cancelled after the trade. Effectively, the private information of a trade is incorporated on the competing venue via cancellations of limit orders (and potentially, resubmissions of limit orders at higher prices). Duplicate liquidity arises in addition to the effect that for $\gamma > 0$, consolidated liquidity increases because of enhanced competition between liquidity suppliers.

In what follows, I define duplicate liquidity as the liquidity impact plus one, i.e., the impact of a trade on consolidated liquidity due to cancellations of limit orders.

The next proposition relates to the impact of a trade on the liquidity supply at the competing venue.

Proposition 3 *Define the "cross-venue liquidity impact" as the impact of a trade on the equilibrium liquidity supply of the competing exchange. The cross-venue liquidity impact is strictly negative, and the magnitude reduces by the fraction of SORT traders operating on the competing exchange.*

A trade reveals information about the fundamental value, which is incorporated on the competing venue via cancellations of limit orders. When a larger fraction of SORT traders operates on the competing venue, it becomes less liquid such that fewer limit orders have to be cancelled after a trade on the current venue. In practice, the fraction of SORT traders should be larger on an entrant venue than on the incumbent exchange, and therefore entrant trades will have a larger cross-venue liquidity impact.

2.2.7 Numerical example

In this section I substitute the parameters with realistic values and analyze the equilibrium. The main interest lays in the impact of the fraction of traders that might go to one venue only (α and β) on the equilibrium liquidity supply of both markets.

I choose the following parameter values. The average trade size is 1 unit ($\phi = 1$), the best ask price is £10.00 and the tick size is 0.5 cent, which is the relevant case for the sample stocks with a price of £10.00. The price impact of a one unit trade $\lambda = 20$ basis points, and the tie-breaking rule $\pi = 0.5$. The models' results come out clearest when I set the fundamental value V just above £9.99, such that the depth in the order book is constant at each price level for the case that all investors use SORT ($\gamma = 1$).[14]

The numerical outcomes for the four best price levels are shown in Table 2.1, where only α and β vary. The bottom panel shows the market shares and per-trade expected profits to the market makers of both venues.[15] The first row shows the single exchange setting, $\alpha = 1$, which is the Sandås (2001) solution. This is the benchmark case, and shows that 1.53 units are offered on the best price level, and 2.5 units on all subsequent price levels. In this case, the quantities offered beyond the best price are constant because of the tradeoff between improved prices and higher adverse selection costs, which equals the tick size divided by the price impact.

The second column shows the situation where all investors use

[14]Specifically, I set the fundamental value $V = 9.994943$. Small changes in V relative to the fixed pricing grid cause changes in offered liquidity at the best price, which interact with liquidity on the competing exchange and in turn with liquidity deeper in the order book.

[15]I obtain the market makers' expected profits as follows. For each type of trader I calculate the expected profits of all shares offered on each price level by integrating the marginal profit equation (e.g. equation (2.5)) over the number of shares offered at that price level. Then, the total expected profit is the sum of the profits at each price level, weighted by the fraction of traders of each type.

smart order routing technology ($\gamma = 1, \alpha = \beta = 0$). Compared to the benchmark case, consolidated liquidity is 63% higher on price p_1 (2.5 versus 1.53), and identical on all subsequent levels. This corresponds to the second part of proposition (1) that liquidity in a fragmented market is strictly higher than in a single venue setting. The bottom panel confirms that the increased liquidity comes at the expense of the market makers, whose expected profits decline from from 2.24 to 1.50 basis points per trade.

From columns three to six the fraction of SORT traders gradually decreases, which increases consolidated liquidity (part 1 of proposition (1)). When we move towards the full duplicate limit orders case ($\gamma = 0$, column (6)), consolidated liquidity is twice that of the single exchange setting. In each case, the table reports the market shares and the per trade expected profits to markets makers of both venues.

The main prediction of the model is that a higher fraction of SORT traders reduces the amount of duplicate limit orders (proposition (2)). To illustrate this point, we analyze the impact of a 2.5 share trade on the liquidity of both venues. The price impact is $2.5\lambda = 1$ tick, meaning that all quantities shift up exactly one price level after market makers have revised their limit orders. When all investors use SORT (column (2)), the trade consumes the 1.25 units of Q_{A1} and 1.25 of Q_{B1}, and then the limit order books are immediately in equilibrium (market makers will not need to revise their limit orders). In effect, there are no duplicate limit orders and the consolidated liquidity impact of a trade is exactly one (this is not generally the case). In contrast, when no investors use SORT (column (6)), the trade will consume 2.5, but then also 2.5 will be cancelled on the competing venue because of the price impact of the trade. In effect, after the revisions by the market makers, the 2.5 share trade reduces liquidity on both venues by 5 shares, meaning that

100% of the order size is cancelled.

Proposition (2) and (3) are confirmed in column (7), where 50% of the investors have access to venue A only and 50% use SORT. When more SORT traders operate on venue B, the cross-venue impact of a trade on A is lower. As above, afer executing 1.25 on venue A and B simultaneously all quantities shift up one price level, such that 0.17 is cancelled on Q_{A1} (from 1.42-1.25 to 0) and 0.95 on Q_{A2} (from 2.37 to 1.42), whereas no cancellations occur on Q_{B1} and Q_{B2}. Indeed, column (7) shows the realistic setting in Europe, where a large fraction of investors is able to trade only on the traditional venue A (50%), and the remaining investors use SORT to access both venues.

2.3 Empirical analysis

This section first presents a brief overview of the FTSE 100 stocks' trading environment, followed by a data description. Next I explain the DepthAsk(X) and DepthBid(X) liquidity measures, and an empirical proxy for the fraction of SORT investors. Then I test the propositions of the model and discuss the results.

2.3.1 Background and Data

Market structure FTSE100 stocks

The FTSE100 stocks are primarily listed on the London Stock Exchange (LSE). In November 2009 the LSE executes 61% of trading volume (excluding dark pool and Over-The-Counter volume).[16] These stocks are traded on an electronic limit order market which integrates liquidity provision by market makers. Note that the market makers in the model are regular investors,

[16] As reported by Fidessa, see http://fragmentation.fidessa.com.

who use high-frequency technology to operate like traditional market makers. Continuous trading occurs between 08:00 and 16:30, local time.

Once stocks are listed on the LSE, alternative trading venues may decide to organize trading in them as well.[17] Four important entrants have emerged which also employ publicly displayed limit order books: Chi-X, Bats, Turquoise and Nasdaq OMX Europe. These venues are regulated as Multi-lateral Trading Facilities (MTFs), the European equivalent to ECNs. While these entrants in effect have the same market model as the LSE, they differ with respect to trading technology (speed in particular), fixed and variable trading fees, and some of the types of orders that may be placed (e.g., pegging a limit order price to the midpoint, such that it always equals the midpoint $+n$ ticks). Investors can demand ("take") liquidity by issuing a market order or supply ("make") liquidity by issuing limit orders at any moment in time. All markets allow for visible, partially hidden (iceberg) and fully hidden limit orders.

Chi-X started trading in April 2007 and is the most successful entrant in terms of market share with 24% of trading volume in November 2009. Turquoise and Nasdaq OMX started trading FTSE 100 firms as of September 2008, and Bats two months later. Their market shares are substantially lower, with 5.5%, 1.8% and 7.6%, respectively. In May 2010 Nasdaq OMX closed down, as they did not meet their targeted market shares.[18] As of July 2009, the five trading venues use identical tick sizes, which depend on the stock price. Compared to the US, the tick sizes in the UK are very small with for example £0.001 for a stock price between £5.00 and £10.00, and £0.005 for a price between £10.00 and £50.00. All the new competitors employ

[17]This feature makes the current study inherently different from literature on cross-listings, where firms themselves may choose to list on several exchanges.

[18]See "Nasdaq OMX to close pan-European equity MTF", www.thetradenews.com.

a maker - taker pricing schedule, where executed limit orders receive a rebate of 0.18 to 0.20 basis points, while market orders are charged 0.28 to 0.30 basis points of traded value. These make-take fees are very small compared to the tick size of around 5 basis points on average, and small compared to those of US stocks.

The trading venues with publicly displayed limit order books execute approximately 60% of total volume, while the remaining 40% is executed on dark pools, broker-dealer crossing networks, internalized and OTC.

Data

The analysis is based on a subsample of ten FTSE100 stocks, each randomly selected from one market cap decile of the 100 constituents (i.e., a size stratified sample).[19] The sample period consists of 10 trading days (November 2 - 13, 2009), and high-frequency data are taken from the Thomson Reuters Tick History database. For each stock, the data contain separate limit order books for the five trading venues. As a robustness, I also do the analysis for 10 stocks taken from the largest French index (the CAC40) during October 5 - 16, 2009, and the results are similar (available upon request).

For each transaction, I observe the price, traded quantity and execution time to the millisecond,[20] while for each limit order placement, modification or cancellation, the data set reports the timestamp and the ten best prevailing bid and offer prices and their associated quantities.

While the time stamp is per millisecond, I take snapshots of the limit order books at the end of every 100^{th} millisecond,

[19] I choose ten stocks during ten trading days as computational limitations prevent me from using the full sample of stocks or more trading days. Table (2.2) shows the list of stocks.

[20] If a single market order is executed against several outstanding limit orders, separate messages are generated for each limit order.

resulting in approximately 30 million observations. Higher frequencies are not useful when comparing multiple trading venues, as it may lead to inaccuracies because of latency issues, i.e. millisecond reporting delays.[21] Per snapshot, I observe the outstanding liquidity and the trading volumes on the buy and sell sides of every venue. The advantage of taking snapshots is that the data become evenly spaced, such that lagged variables in the regressions become easily interpretable.

I do not directly observe hidden and iceberg limit orders. However, from trades I can construct the executed hidden quantity, based on the state of the order book directly before and after the trade and the traded quantity. As such, I do observe hidden liquidity that gets 'hit' by a market order. These data are identical to those offered by several information vendors, meaning I use the information set available to the market.

The upper panel of Table 2.2 presents summary statistics for the sample stocks. As I select stocks from each size decile, there is a large variation in market cap: the mean is £21 billion with a £37 billion standard deviation. Accordingly, also trading volume (in shares and pounds) and realized volatility vary substantially. In contrast, the market shares of the five trading venues are fairly stable between firms, and highly representative for the entire FTSE100 index.

The lower panel of Table 2.2 presents summary statistics on the average number of limit order book modifications and transactions per minute. While the LSE's market share is largest by far, the number of transactions lay much closer together (i.e., trade sizes are smaller on the entrant venues). In fact, the number of limit order modifications on Chi-X greatly exceeds that of the LSEs, on average 218 versus 160 per minute.

The limit order books of entrant venues are highly active, de-

[21] The reporting delays are smaller than five milliseconds according to industry professionals.

spite the relatively small market shares. The ratio of limit order modifications to trades is 31:1 for the LSE, 51:1 for Chi-X and increases as a venues market share goes down to 123:1 for Nasdaq. This is mostly due to high-frequency traders placing and cancelling many limit orders. Chi-X and Bats are the most successful new competing venues in terms of market share, number of transactions and limit order book activity.

2.3.2 The DepthAsk(X) and DepthBid(X) liquidity measures

This subsection explains the DepthAsk(X) and DepthBid(X) measures, also used in Degryse, de Jong, and van Kervel (2011).

The DepthAsk(X) aggregates all shares offered at prices between the midpoint and the midpoint plus X basis points. Similarly, the DepthBid(X) sums the shares offered within the midpoint and the midpoint minus X basis points. The midpoint is the average of the best bid and ask price available in the market (the NBBO), and I choose $X = 10$ basis points relative to the midpoint. The price constraint X guarantees that only liquidity at prices close to the midpoint is aggregated, i.e. at good price levels. This is important, as liquidity offered deeper in the order book is less likely to be executed, and therefore less relevant to investors. The number of shares in the interval are then converted to the value in GBPs.

Formally, define price level $j = 1, 2, ..., J$ on the pricing grid and the midpoint M, then for venue v,

$$DepthAsk(X)_v = \sum_{j=1}^{J} P_{j,v}^{Ask} Q_{j,v}^{Ask} \mathbf{1}\left(P_{j,v}^{Ask} < M(1+X)\right), \quad (2.7a)$$

$$DepthBid(X)_v = \sum_{j=1}^{J} P_{j,v}^{Bid} Q_{j,v}^{Bid} \mathbf{1}\left(P_{j,v}^{Bid} > M(1-X)\right). \quad (2.7b)$$

The measures are calculated at the end of every 100 millisecond interval and represent liquidity offered at the bid and ask side, per trading venue. When taking larger values for X, liquidity deeper in the order book is also incorporated. Then, comparing for different price levels X reveals the shape of the order book. For example, if the depth measure increases rapidly in X then the order book is relatively deep. The order book is asymmetric when the absolute difference between DepthAsk(X) and DepthBid(X) is large.

The depth measure has several features which make it highly suitable for the empirical approach. First, the measure is calculated per venue, for the bid and ask sides, which allows us to analyze the correlations between the demand and supply of liquidity across venues, sides and over time. Second, the depth measure is easily related to trading volume as both are measured in currency (GBPs). Third, the measure incorporates limit orders beyond the best price levels, making it robust to small, price improving limit orders. Such orders are often placed by high-frequency traders, who mostly drive the dynamics in the model. Fourth, by choosing a fixed interval the measure is independent of the tick size, which varies across stocks. For a detailed discussion of this measure and a comparison with related liquidity measures such as the Cost of Roundtrip ($CRT(D)$) and Exchange Liquidity Measure ($XLM(V)$), I refer the interested reader to Degryse, de Jong, and van Kervel (2011).

Table 2.3 contains summary statistics on the Depth(10) and Depth(50) measures for the bid and ask side, reported in GBPs and calculated per exchange. The statistics are based on single observations per tenth of a second per stock, equal weighted over all stocks. Depending on the tick sizes, the Depth(10) aggregates liquidity of two to five price levels on the bid and ask side. The Depth(50) often represents the entire limit order book. First and most strikingly, we observe that the Depth(10) offered on

Chi-X is 86% of the LSE, while they execute only a third of the LSE volume. The liquidity at Bats is roughly 40% of the LSE Depth(10), while Turquoise and Bats have approximately 20% each. On average, the ask and bid sides are symmetrical.

The regression analysis works with *changes* in DepthAsk(10) and DepthBid(10), i.e., the value of the current minus the previous observation. As equation (2.7) shows, these changes depend on the activity in the limit order book and on the level of the midpoint. The model predicts changes in the depth measures due to limit order book activity, i.e., the placement, cancellation, modification and execution of limit orders. Therefore, I define $Chg\_DepthAsk(X)_t$ as the difference in DepthAsk between each period, holding the midpoint constant (similarly for the $Chg\_DepthBid(X)_t$)

$$Chg\_DepthAsk(X)_t = DepthAsk(X_t, M_{t-1}) - DepthAsk(X_{t-1}, M_{t-1}). \tag{2.8}$$

The measure simply shows how much liquidity in GBPs is added or removed from the limit order book in a period.

2.3.3 Trade sequences by SORT traders

In the model, the SORT traders submit two market orders simultaneously to both venues, but in reality this is not perfectly possible. However, the crucial element is that market makers are unable to revise their limit order schedules in between the two trades. Essentially, each market order executes against limit orders which have not yet incorporated the information content of the other market order. Empirically, I use this feature to identify market orders that are sent "simultaneously" to several venues, which I call *trade sequences*.

I use the definition that two trades occur simultaneously when

(i) the state of the limit order book of the venue executing the current trade has not changed since the previous trade, (ii) both trades occur within 100 milliseconds after each other,[22] and (iii) both trades are either purchases or sales. The first restriction immediately follows from the model and states that the trade executes against stale limit orders.

In more detail, I use the following approach. Per stock I aggregate the trades of the five venues in a single file,[23] chronologically sorted and timestamped to the millisecond. I add the state of the limit order books of the five venues just before each trade takes place. Then, I specify that the current and previous trade occur simultaneously if (i), (ii) and (iii) above are satisfied. Next, I group the simultaneous trades into a trade sequence, i.e., a string of market orders that are send to several venues simultaneously. The number of venues that are accessed simultaneously is the length of a sequence, which ranges from one to five. A sequence with a length of one is an individual trade on one venue, whereas a sequence with a length larger than one is assumed to stem from a SORT trader. The fraction of trading volume of sequences with a length larger than one is the empirical proxy for the fraction of SORT traders.

The top panel of Table 2.4 shows the percentage of trades that are part of a sequence of a certain length, for each trading venue. The first row shows that the LSE executes about 250.000 trades, of which 69.8% has a sequence length of one, 17.2% a length of two, and the remaining 13% is part of a sequence with a length between three and five. According to my definition, 30.2% of the LSE trades stem from SORT traders. On Chi-X and Bats 39%

[22]Admittedly, 100 milliseconds is arbitrarily chosen, but it is consistent with the remainder of the analysis. The mean and standard deviation of the time between two trades classified as simultaneously is 23 and 22 milliseconds respectively.

[23]In the data, a single market order may execute against several limit orders, generating multiple observations. These partial executions have the same time stamp (at the millisecond level), which I pool to obtain a single trade.

and 51% of the trades have a sequence length larger than one, which means that a higher fraction of the traders on these venues use SORT. Similarly, the fraction of SORT traders on Turquoise and Nasdaq OMX is 46.4% and 52.6%. There are 652,000 trades in the sample, of which 252,000 (38.7%) are part of the same sequence, i.e., occur simultaneously. An unreported calculation shows that the first restriction is binding frequently, as without it the number of simultaneous trades would increase to 335,000 (51%). While trades often occur on several venues in the same 100 millisecond interval, in many cases the market makers have already revised their limit order schedules in between the trades.

The middle panel of Table 2.4 shows the average size of trades part of a sequence of a certain length, per trading venue. The average trade size ranges between £15,000 and £19,000 for the LSE, and between £4,000 and £9,000 for the new entrants. The average order sizes are fairly constant across trades part of different sequence lengths. Therefore, the total *sequence* volume increases almost linearly in the number of venues accessed. This finding is perhaps not surprising as order splitting is more attractive when trading larger amounts.

In the analysis, I use the fraction of trading volume executed by sequences with a length exceeding one to proxy for SORT trading activity. The bottom panel of Table 2.4 shows that 38.6% of the overall trading volume stems from SORT traders. On the LSE, 34.5% of trading volume stems from SORT traders. This is approach differs from Foucault and Menkveld (2008), who estimate the fraction of SORT traders by looking at trade-throughs (violations of price priority). However, their approach is not suitable when five trading venues coexist, as traders may have access to one or any combination of several venues. The current approach correctly reflects the adverse selection risk faced by market makers to trade against traders who accessed another venue already, which in the model is a function of the fraction of SORT traders and their order routing preferences.

The trade sequences do have some caveats. First, it is possible that a SORT trader may execute his entire trade on a single venue, in which case I would classify that trade as stemming from a non-SORT trader. Second, it is possible that an arbitrageur (or intermediary) trades first with a limit order on one venue and then quickly closes the position with a market order on a competing venue to benefit from make-take fees or crossed quotes for example. In this case, both trades would be incorrectly classified as stemming from a SORT trader. Note that absent trader identities, I am not certain whether the market orders in a trade sequence indeed stem from the same SORT trader. However, this is not an issue because the adverse selection risk in the model is driven by the size of the sequence, and not by the identity of the trader. That is, the market makers face the same picking off risk when two simultaneous market orders stem from the same or different traders.

2.3.4 The cross-venue liquidity impact

The regression analysis consists of two parts. In this subsection, I establish the general prediction of the model that a trade on one venue is followed by cancellations of limit orders at competing venues. This analysis shows the cross-venue liquidity impact of trades on each individual trading venue. In the next subsection I test the three propositions of the model that relate the fraction of SORT traders to the level of consolidated liquidity, the consolidated liquidity impact and the cross-venue liquidity impacts.

Methodology

Cancellations of limit orders occur when the $Chg\_DepthAsk(10)^V$ is negative for venue V, after controlling for trading volume on that venue. To estimate cross-venue cancellations, I regress the $Chg\_DepthAsk(10)^V$ on contemporaneous and lagged buy and sell volumes of the individual venues. I add lags up to ten seconds (100 periods) to allow all markets to incorporate the information content of trades, which should be sufficiently long in a high-frequency trading environment. Instead of estimating 100 individual lagged coefficients, I add five variables that average trading volume of 1, 2-4, 5-10, 11-20 and 21-100 periods away, per venue for buy and sell volumes. Section 2.5.2 in the appendix explains in more detail how I obtain the cumulative impact of a transaction over time, and the corresponding standard errors.

A trade is classified as a Buy or $Sell$, and define the trading venue $v = 1, ..., 5$, and the current and five lagged groups $l = 0, ..., 5$ for stock i and time t. Then, for each venue V, I regress

$$Chg\_DepthAsk(10)^V_{it} = c_i + \sum_{v=1}^{5}\sum_{l=0}^{5}\left(\beta^{Buy}_{l,v} \times Buy^v_{it-l} + \beta^{Sell}_{l,v} \times Sell^v_{it-l}\right) + \sum_{v=1}^{5}\left(\beta^{Buy}_{v} \times BuyHid^v_{it} + \beta^{Sell}_{v} \times SellHid^v_{it}\right) + \varepsilon_{it}. \quad (2.9)$$

The term after the firm fixed effects represents the buy and sell volumes (in GBPs) for the five venues covering the contemporaneous and five lagged groups. The term on the second line controls for contemporaneous hidden buy and sell liquidity (observed when executed), which is added for the following reason. The effect of a buy trade on $Chg\_DepthAsk(10)^V$ of venue V should mechanically be -1 : a one pound trade reduces the depth with exactly one pound. However, a trade executed against a hidden limit order does not reduce visible liquidity, and therefore I control for executed hidden liquidity.

This regression is executed ten times: for the bid and ask sides of the five trading venues. The result shows how many pounds close to the midpoint are submitted or cancelled after a one pound buy or sell trade on any venue. Note that the effects are permanent and do not die out over time, as for example a buy trade might contain positive price information such that some limit orders will permanently be cancelled on the ask side.

Results

The regression results are reported in Table 2.5, with the change in DepthAsk and DepthBid of all venues as dependent variables. Each column represents one regression, showing separate coefficients for buy and sell volumes, per trading venue. Within each venue, the displayed coefficients represent the cumulative effect over time (the running sum). I only show the contemporaneous effect (within the 100 millisecond interval), after 1 second and after 10 seconds. Intermediate lagged values are estimated to improve the model fit, but for brevity not reported. Standard errors are omitted in the Table (available on request), but a single asterisk indicates significance at the 1% level. Next I will discuss the findings for the DepthAsk only, as the results for the DepthBid are symmetric.

The first column shows that the immediate effect of a one pound LSE purchase on LSE DepthAsk is -0.83 pounds. This implies that while the trade removed 1 pound, either 17 cents is immediately replenished, or, first a new limit order is placed which immediately provokes the trade. The first explanation is consistent with iceberg orders that reveal an additional hidden component after execution, and the latter with the findings of Hasbrouck and Saar (2009).

Consistent with the main prediction of the model, a one pound buy trade at Chi-X is immediately followed by cancellations on the LSE of -0.21 pounds (-21%). After ten seconds, the effect is -0.61, meaning that more than half of the Chi-X trade size is cancelled on the LSE. The coefficients for Bats are similar, -0.27 immediately and -0.54 after ten seconds (all significant at the 1% level). This effect is economically very large, and implies that trades on entrant venues are immediately followed by cancellations on the traditional market. Note that this effect cannot be explained by investors who simultaneously trade on several venues, since the regression controls for trades on all venues. The effect of Nasdaq and Turquoise trades on LSE liquidity are negative, but fairly small.

The LSE responds more strongly to Chi-X and Bats trades than vice versa, which is consistent with proposition (3) since more SORT traders are active on the entrant venues. In column (2) and (3), LSE trades reduce liquidity on Chi-X and Bats with -0.18 and -0.05 after ten seconds (which is significantly different from the -0.61 and -0.54 above). The one-second impact of an LSE trade on the change in liquidity at competitors is 0.53 (the row sum). An alternative explanation why the new entrant venues respond less to trades on the LSE might be that these new entrants are inactive at times. When they do not offer liquidity at the best price, cancellations should not occur.

A Chi-X buy trade reduces Chi-X DepthAsk with -1.31 (column (2)), implying that beyond the reduction of 1 pound, an additional and significant 31 cents is cancelled. A likely explanation is that Chi-X allows for "pegged to midpoint" limit orders, which are automatically repriced when the midpoint changes. If a trade moves the midpoint, these limit orders are essentially cancelled and resubmitted at a higher price. Displayed pegged to midpoint orders are allowed on Bats too (which has a coefficient of -1.26), but not on the LSE in our sample period.

Cancellations and trades are strongly interlinked between Bats and Chi-X, which have cross-coefficients of -0.58 and -0.18 (column (2-3)). These are the most successful entrants in terms of market share. In contrast, Turquoise and Nasdaq seem more independent, as their liquidity does not respond much to trades on the LSE, Chi-X and Bats and vice versa.

The immediate effect of any venues' sales on the LSE DepthAsk(10) is economically large and positive, with coefficient ranging from 0.10 to 0.30 (column (1), bottom panel). This result is consistent with an information effect: the sell trade conveys negative information about the stock, such that market makers improve prices and quantities on the ask side.

The results for sell trades on the change in liquidity on the bid side, the right part of the table, are highly symmetrical. As a robustness check, I execute the analysis for 10 stocks taken from the largest French index (the CAC40) during October 5 - 16, 2009. Using a different sample and time period (one month earlier), the estimated coefficients are similar (available upon request). This is not surprising however, given that the four new trading venues who compete in FTSE100 stocks also trade the French CAC40 stocks (along with most of the other European indices).

2.3.5 Testing the propositions: The effect of SORT traders

In this section I test the three propositions of the model. I relate the fraction of SORT traders to the level of consolidated liquidity, the consolidated liquidity impact and the aggregate cross-venue liquidity impact. These variables are estimated at the hourly frequency, and I use the time series variation within a stock to identify the relation.

Methodology

The first proposition states that market makers place fewer duplicate limit orders when the fraction of SORT traders increases. Therefore, the level of consolidated liquidity $Ln(Depth(10))^{Cons}$

should decrease in $\%SORT$, defined as the fraction of trading volume part of sequences with a length exceeding one. I run the following regression, with and without a set of control variables X_{ih} that are known to affect the consolidated depth.

$$Ln(Depth(10))_{it}^{Cons} = c_i + \delta_{(h)} + \beta_1 \%SORT + \beta_2 X_{ih} + \varepsilon_{ih}. \quad (2.10)$$

The c_i and $\delta_{(h)}$ are firm and daily dummy variables. The vector X_{ih} consists of the logarithm of turnover, realized volatility (calculated hourly as the sum of the 12 squared 5-minute returns), the logarithm of the average order size and a proxy for algorithmic trading. These variables are important in the model: turnover combined with volatility may proxy for asymmetric information,[24] and the average order size is directly used in the model. The proxy for algorithmic trading, $Algo\_trad$, may capture the cross-market activity of market makers, and is defined as the negative of trading volume in hundreds of pounds divided by the number of electronic messages (the placement, execution and cancellation of limit orders), following Hendershott, Jones, and Menkveld (2011).

Proposition two and three relate the fraction of SORT traders to the consolidated and cross-venue liquidity impact, which I estimate as follows. I run regression 2.9 for hour t of stock i (360,000 observations), but use as dependent variable the change in *consolidated Depth(10)* on the ask or bid side. These regressions represent the impact of a trade at venue v on market wide depth. In a new dataset I store the estimated coefficients $Cons\_coef_{v,it}$ accumulated over one second after the trade, for the effect of buy trades on the ask side liquidity and sell trades on the bid side liquidity.

In addition, I run regression 2.9 for each venue v of hour t

[24] For example, holding volatility constant, more turnover should imply less informative trades and lower price impacts.

and stock i using as dependent variable the change in liquidity of the current venue, $Chg\_DepthAsk(10)_{it}^{v}$. Again, I save the estimated coefficients accumulated over one second $Own\_coef_{v,it}$, which represents the impact of a trade at venue v on the change in liquidity of that venue. Then, I define the cross-venue liquidity impact as the difference between the impact of a trade on consolidated liquidity and its own liquidity, $Cross\_coef_{v,it} = Cons\_coef_{v,it} - Own\_coef_{v,it}$. I calculate these separately for the effect of purchases on the DepthAsk(10) and sales on the DepthBid(10).

The sample size is extended from 10 to all 21 trading days in November 2009 for the same 10 stocks used in the previous analysis, which results in 1890 observations. The estimated coefficients have substantial noise when only a few trades occur in a particular stock-hour, which is frequently the case for some of the lesser active entrants. Therefore, per observation I average the estimates across the four new entrant venues, weighted by trading volume. Important structural differences exist between the new entrants and the traditional exchange. For example, the entrant venues all have very fast trading systems, similar order types, small transaction sizes, low transaction costs and a similar make/take fee breakdown (see section (2.3.1)). An advantage is that studying the LSE and the combined entrants corresponds better to the two-exchanges model.

To reduce the impact of noisily estimated coefficients, I apply weighted least squares below using the number of trades on the LSE or the new entrants in the particular stock-hour as weight variable. In addition, I winsorize the cross-venue and consolidated liquidity impacts at the 1% and 99% level to reduce the impact of outliers.

Based on the new data set, I test proposition two and three with the following regressions, for $V = \{LSE,\ New\ Entrants\}$

$$\begin{aligned} Cons\_coef_{V,ih} &= c_i + \delta_{(h)} + \beta_1 \%SORT + \beta_2 X_{ih} + \varepsilon_{ih}. \quad (2.11) \\ Cross\_coef_{V,ih} &= c_i + \delta_{(h)} + \beta_1 \%SORT^{Entrants} + \beta_2 \%SORT^{LSE} + \\ & \quad \beta_3 X_{ih} + \varepsilon_{ih} \quad (2.12) \end{aligned}$$

I add the following independent variables. As before, $\%SORT$ is the fraction of total trading volume submitted by SORT traders, and $\%SORT^{Entrants}$ is the fraction of trading volume submitted by SORT traders on the entrant venues (similarly defined for $\%SORT^{LSE}$). The breakdown follows from proposition 3, which states that the fraction of SORT traders on the competing venue determines the cross-venue impact of a trade on the current venue. The vector X_{ih} is defined above. Note that using estimated coefficients as dependent variables in a second step regression does not bias the coefficients of the second step.[25]

Summary statistics of the regression variables are presented in Table 2.6. The average trading volume by SORT traders is 40% (with a standard deviation of 0.09). The SORT volume on entrants is slightly higher than on the LSE (45% versus 37%). The average level of $Ln(Depth(10))^{Cons}_{it}$ is 12.4 (£250,000), and has a standard deviation of 1. The proxy for algorithmic trading shows that on average, £193 of trading volume is executed per electronic message, with a standard deviation of £162.[26]

The middle and bottom panels show the consolidated and cross venue liquidity impacts for the LSE and the new entrants. The average consolidated impact of buys on the ask side is -1.39 for the LSE and -1.76 for the entrants, which have large standard deviations (0.77 and 0.98 respectively). The standard errors may reflect measurement error or variation in the cross-

[25]The measurement error from the first step only increases the standard errors of the coefficients in the second step.

[26]Hendershott, Jones, and Menkveld (2011) find \$1,100 per message in 2005 for the large cap quintile of NYSE stocks.

market activity of market makers and SORT and non-SORT traders. The cross-venue impact ranges between -0.58 to -0.75 (-0.58 is mentioned in the introduction).

Results

Table 2.7 confirms proposition 1 that a larger fraction of trading volume by SORT traders reduces consolidated depth. Column 1 shows that a one-standard deviation increase in $\%SORT$ (0.09) reduces consolidated liquidity by %7, which is significant at the 99% level. Algorithmic trading appears to have no effect. The coefficients remain similar when adding control variables (column 2). As expected, turnover and the average order size are positively related to consolidated depth, although the causality may go both ways. Volatility has a strong negative sign, which is common in this literature.

The upper panel of Table 2.8 tests proposition (2) and the lower panel proposition (3), for the ask side (Table 2.9 shows the results for the bid side). Consistent with the theory, $\%SORT$ has a positive coefficient in all specifications, meaning that the consolidated liquidity impact becomes less negative (i.e., smaller in magnitude) when $\%SORT$ increases. For the LSE, when including the control variables (column 3), a one standard deviation change in $\%SORT$ increases the consolidated liquidity impact by 0.075, which is large given a mean of -1.39. The impact for the entrants is larger with 0.12 per standard deviation.

The control variables have the following signs. Algorithmic trading increases the magnitude of the cancellation impact in all specifications, i.e., increases duplicate liquidity. The proxy for algorithmic trading depends on the number of electronic messages, which may represent the cross-market activity of the market makers in the model. Volatility diminishes the cancellation impact, perhaps because the risk of simultaneous executions becomes larger in volatile times.

The bottom panel of the Table 2.8 shows the results for the cross-venue liquidity impact. Consistent with the model, a larger fraction of SORT traders on the competing venues reduces the magnitude of the cross-venue liquidity impact. The coefficient of $\%SORT^{Entrants}$ on the LSE cross-venue impact is 1.61, such that a one standard deviation change (0.1) increases the average LSE cross-venue impact from -0.58 to -0.42 (i.e., reduces the magnitude). The coefficient of $\%SORT^{LSE}$ on the cross-venue impact of the entrants is 0.45, and statistically significant at the 1% level when adding the control variables. The control variables have the same sign as in the regressions with the consolidated liquidity impact as dependent variable.

Table 2.9 shows the regression results for the impact of sales on the consolidated and cross venue liquidity on the bid side. The signs and magnitudes of the coefficients are similar as before, which suggests the results are robust.

2.4 Conclusion

In a fragmented equity market, I show that liquidity suppliers have an incentive to duplicate their limit order schedules across venues. They will cancel the duplicate orders after a trade on the competing venue, such that liquidity shocks between venues become strongly correlated. An important determinant of the amount of duplicate liquidity is the fraction of traders using smart order routing technology.

The model focusses on two trading venues only, but can already predict a substantial fraction of duplicate limit orders. Therefore, the relevance of the model is only strengthened by the fact that most European stocks trade on more than four trading venues (with publicly displayed limit order books) and some US stocks on so much as twelve trading venues. A larger number of trading venues encourages market makers to duplicate their limit order schedules.

The main policy implication of the model is that fair markets require traders to split up trades *simultaneously* across venues. When a trader leaves a millisecond delay between the split, high-frequency traders can observe the first part of the trade and quickly cancel duplicate limit orders on other venues before the second part of the trade arrives. When all traders would use SORT, duplicate liquidity disappears and trades are not followed by excessive cancellations of limit orders.

2.5 Appendix

2.5.1 Theory

This section offers the proof of Proposition (1): The consolidated liquidity in a fragmented market is strictly larger than liquidity in a single exchange setting.

Proof. Denote the liquidity in a single exchange setting as Q_1, which is the solution from Sandås (2001). I repeat equation (2.6), and rewrite as

$$Q_{A1} = \frac{p_1 - V_t - \lambda\phi}{\lambda} - \frac{\gamma(1-\pi)Q_{B1}}{\gamma(1-\pi) + (\alpha+\gamma\pi)\exp(\frac{Q_{B1}}{\phi})} \equiv Q_1 - c_{B1}Q_{B1} \quad \text{A1(a)}$$

$$Q_{B1} = \frac{p_1 - V_t - \lambda\phi}{\lambda} - \frac{\gamma\pi Q_{A1}}{\gamma\pi + (1-(\alpha+\gamma\pi))\exp(\frac{Q_{A1}}{\phi})} \equiv Q_1 - c_{A1}Q_{A1}. \quad \text{A1(b)}$$

I need to show that $Q_{A1} + Q_{B1} > Q_1$, which is equivalent to $c_{B1}Q_{B1} + c_{A1}Q_{A1} < Q_1$.

$$c_{B1}Q_{B1} + c_{A1}Q_{A1} = c_{B1}Q_{B1} + c_{A1}Q_1 - c_{A1}c_{B1}Q_{B1}$$

$$= c_{A1}Q_1 + (1 - c_{A1})c_{B1}Q_{B1}$$
$$< c_{A1}Q_1 + (1 - c_{A1})Q_1 = Q_1.$$

In the first equality I simply replace Q_{A1} with the first line of equation (A1(a)); which I then rewrite in the second line. The inequality holds because $c_{B1}Q_{B1} < Q_1$, since $c_{B1} < 1$ and $Q_{B1} \leq Q_1$. ■

2.5.2 Empirical

This section shows that the regression methodology in section 2.3.4 measures cumulative effects over time. That is, in regression (2.9) I add contemporaneous terms and lagged values of 100 periods ago (i.e., ten seconds). Instead of estimating 100 coefficients, I create six variables representing the averaged lagged volumes of the current, 1, 2-4, 5-10, 11-20 and 21-100 periods away, per venue for buy and sell volumes. Define t as the current period, i as the start and j as the end of the intervals (e.g., $i = 2$ and $j = 4$). Then

$$Vol_{i,j} = \frac{1}{j - i + 1} \sum_{n=i}^{j} Vol_{t-n}.$$

The periods of lagged values are chosen to maximize the model fit. An example of the data is shown in the table below, where a £1.00 trade occurs on some venue at time $t = 1$. The first four columns show the values of the contemporaneous and three lagged groups. The fifth column shows the cumulative effect of the regression coefficients over time, calculated as a running sum of the individually estimated coefficients. By constructing the variables as averages, the long-term effect of a trade is simply the sum of the estimated coefficients. The standard errors are also calculated based on this sum.

Data example of a trade at time $t = 1$.
Each β_t represents the estimated coefficient of lagged average volumes t periods away, as described in regression (2.9).

Time	Vol_0	Vol_1	$Vol_{2,4}$	$Vol_{5,10}$	Cumulative effect
0	0	0	0	0	0
1	1.00	0	0	0	β_0
2	0	1.00	0	0	$\beta_0 + \beta_1$
3	0	0	0.33	0	$\beta_0 + \beta_1 + 0.33\beta_{2,4}$
4	0	0	0.33	0	$\beta_0 + \beta_1 + 0.66\beta_{2,4}$
5	0	0	0.33	0	$\beta_0 + \beta_1 + \beta_{2,4}$
6	0	0	0	0.20	$\beta_0 + \beta_1 + \beta_{2,4} + 0.2\beta_{5,10}$
7	0	0	0	0.20	$\beta_0 + \beta_1 + \beta_{2,4} + 0.4\beta_{5,10}$

Table (2.1) Numerical Example.
The model is solved for the equilibrium quantities offered on the four best ask price levels of venue A and B. Each column shows a different combination of the fraction of investors with access to venue A only (α) and venue B only (β), such that the fraction of investors with smart order routing technology ($\gamma = 1 - \alpha - \beta$) varies accordingly. The lower panels show the market shares and the per trade expected profits to market makers of both venues. The expected profits are expressed in basis points relative to the midpoint. The remaining model parameters are held fixed. The average trade size is 1, with a per unit price impact of 20 basis points. The best ask price is £10.00, the tick size is 0.5 cents, and I specifically set the fundamental value to £9.994943, such that offered liquidity is constant at each price level when all investors use SORT (column 2). When both venues offer the best price, SORT traders are equally likely to go to venue A or B first ($\pi = 0.5$).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
α	1	0	0.1	0.1	0.4	0.5	0.5
β	0	0	0	0.1	0.4	0.5	0
γ	0	1	0.9	0.8	0.2	0	0.5
Venue A							
Q_{A1}	1.53	1.25	1.29	1.30	1.46	1.53	1.42
Q_{A2}	2.50	1.25	1.78	1.87	2.47	2.50	2.37
Q_{A3}	2.50	1.25	2.49	2.86	2.58	2.50	2.55
Q_{A4}	2.50	1.25	2.64	2.94	2.52	2.50	2.50
Venue B							
Q_{B1}	0	1.25	1.25	1.30	1.46	1.53	1.25
Q_{B2}	0	1.25	1.23	1.87	2.47	2.50	1.15
Q_{B3}	0	1.25	0.78	2.86	2.58	2.50	0.15
Q_{B4}	0	1.25	0.03	2.94	2.52	2.50	0.00
Market shares							
A	1.00	0.50	0.58	0.50	0.50	0.50	0.78
B	0.00	0.50	0.42	0.50	0.50	0.50	0.22
Market maker profits in basis points							
A	2.24	0.75	0.93	0.84	1.06	1.12	1.55
B	0.00	0.75	0.66	0.84	1.06	1.12	0.35

Table (2.2) Summary statistics of sample firms.
The table presents summary statistics for the 10 FTSE100 sample stocks using data of November 2009. The upper panel shows market cap (in millions), price, average daily traded volume (in millions of shares), turnover (in millions of GBP), realized volatility and the market shares of the five trading venues. The trading venues are the London Stock Exchange (LSE), Chi-X, Bats Trading, Turquoise and Nasdaq OMX Europe. The lower panel shows order book data, where statistics are equally weighted based on one observation per 100 milliseconds, per stock. For each venue, statistics on the average number of transactions and limit order book modifications per minute are shown, per venue. The stocks in our sample are Aviva, Hsbc Holdings Plc, Itv Plc, Kingfisher, Lonmin, National Grid, Pearson, Sage group, Vedanta Resources and Xstrata.

	Mean	Stdev	Max
Stock characteristics			
Market Cap	21,062	37,827	125,930
Price	8.06	7.13	23.3
Volume (shares)	527	554	1,883
Turnover (GBP)	3,104	4,188	13,650
Realized Volatility	3.92	1.38	6.10
Share LSE	66.2	5.24	73.8
Share Chi-X	20.5	3.52	25.8
Share Bats	6.44	1.68	9.49
Share Turquoise	5.17	1.07	6.29
Share Nasdaq	1.77	0.65	2.73
Order book data			
Trades LSE	5.16	9.70	347
Trades Chi-X	4.32	7.46	283
Trades Bats	1.80	3.59	84
Trades Turquoise	1.19	2.41	51
Trades Nasdaq	0.61	1.54	56
Limits LSE	159.84	216.31	6,999
Limits Chi-X	218.46	373.08	11,934
Limits Bats	123.92	231.60	7,308
Limits Turquoise	98.25	146.95	3,296
Limits Nasdaq	75.63	154.66	7,081

Table (2.3) Summary statistics Depth(X) measure.
The table presents summary statistics of the Depth(X) measure for the sample stocks, based on limit order book data with a sample frequency of once per 100 milliseconds. The statistics are equally weighted over all observations. The mean, standard deviation and maximum of Depth(10) and Depth(50) on the ask and bid side are shown. The Depth(10) on the ask side reflects the available liquidity, in GBP, offered with prices in the interval of Midpoint and Midpoint + 10bps. Similarly, the Depth(10) on the bid side reflects the liquidity offered with prices between Midpoint - 10bps and Midpoint. Depth(50) sums liquidity within 50 basis points from the midpoint.

Ask side	Mean	Stdev	Max
Depth(10) LSE	66,620	117,961	8,940,000
Depth(10) Chi-X	57,693	78,948	1,100,000
Depth(10) Bats	26,311	39,358	527,947
Depth(10) Turquoise	14,915	18,817	411,974
Depth(10) Nasdaq	12,824	21,345	213,733
Depth(50) LSE	446,687	371,316	9,040,000
Depth(50) Chi-X	268,467	210,066	1,650,000
Depth(50) Bats	88,363	94,112	800,103
Depth(50) Turquoise	71,495	47,170	570,524
Depth(50) Nasdaq	58,096	44,992	332,529
Bid side			
Depth(10) LSE	63,244	90,327	4,420,000
Depth(10) Chi-X	55,907	77,345	1,020,000
Depth(10) Bats	25,404	38,187	553,743
Depth(10) Turquoise	14,273	19,277	525,517
Depth(10) Nasdaq	13,241	22,891	538,609
Depth(50) LSE	431,658	339,853	5,100,000
Depth(50) Chi-X	270,657	213,624	1,790,000
Depth(50) Bats	85,748	92,792	1,170,000
Depth(50) Turquoise	74,360	51,850	638,623
Depth(50) Nasdaq	61,156	51,443	664,181

Table (2.4) The length and volume of trade sequences.
For each trading venue the table shows the percentage of trades that are part of a sequence of a certain length. Sequences are strings of market orders submitted to several venues simultaneously. The current and previous trade are part of the same sequence when (i) the state of the limit order book of the venue executing the current trade has not changed since the previous trade, (ii) the current and previous trade occur within 100 milliseconds after each other, and (iii) both trades are either purchases or sales. Each sequence has a length ranging from one to five, which represents the number of venues accessed simultaneously. The column Total represents the total number of trades at each venue in the top panel. The middle panel shows the average size of a trades that are part of a sequence of a given length. The bottom panel shows the total trading volume by non-SORT and SORT investors (in millions of GBP), where SORT volume is defined as trading volume part of sequences with a length exceeding one.

	Sequence Length					
	1	2	3	4	5	
% of trades						Total
LSE	69.8	17.3	8.1	3.4	1.5	251,381
Chi-X	61.0	23.1	10.1	4.1	1.6	215,876
Bats	48.9	26.4	14.7	7.0	3.0	90,200
Turquoise	53.6	24.2	12.6	6.7	2.9	64,224
Nasdaq OMX	47.4	25.5	14.5	7.9	4.7	30,103
Average trade size						Avg
LSE	14,884	18,095	18,196	17,261	19,076	17,503
Chi-X	7,973	8,012	8,586	8,628	9,009	8,442
Bats	6,301	6,434	5,852	5,815	5,904	6,061
Turquoise	6,219	7,577	7,017	6,265	6,968	6,809
Nasdaq OMX	4,236	5,103	5,278	5,737	5,963	5,263

Volume	Non-SORT	SORT	%SORT
LSE	2610	1374.2	34.5
Chi-X	1050	696	39.9
Bats	278	283.4	50.5
Turquoise	214	214.6	50.1
Nasdaq OMX	60.5	84.2	58.2
Overall	4212.5	2652.4	38.6

Table (2.5) The cumulative impact of turnover on Depth(10).
Each column represents one regression, showing the cumulative effect over time of buy and sell turnover on changes in DepthAsk(10) and DepthBid(10) of one venue. The cumulative effect over time (i.e., the running sum) of contemporaneous trades, and trades one and ten seconds ago are displayed. Changes in the DepthAsk(10) reflect changes in liquidity offered with prices in the interval of Midpoint and Midpoint + 10bps. These changes stem from limit order book activity (placement, cancellations and execution of limit orders). The data consist of one observation per 100 milliseconds, for all stocks. The independent variables are contemporaneous and lagged buy and sell trading volumes of the five venues, denominated in GBP. Accordingly, each panel shows the immediate and short term effects of one venues' transactions on another venues' liquidity. The regressions also contain executed hidden volume as control variables (not reported for brevity). Standard errors are clustered per firm - halfhour, a single asterisk indicates significance at the 1% level.

		Ask Side					Bid Side				
£ *Buys*	Sec	LSE	Chi-X	Bats	Turq.	Nasdaq	LSE	Chi-X	Bats	Turq.	Nasdaq
LSE	0	-0.83*	-0.25*	-0.09*	-0.02*	-0.02*	0.28*	0.24*	0.09*	0.02*	0.01*
LSE	1	-0.80*	-0.30*	-0.14*	-0.04*	-0.05*	0.35*	0.31*	0.15*	0.05*	0.05*
LSE	10	-0.67*	-0.18*	-0.05*	-0.03*	-0.04*	0.33*	0.23*	0.09*	0.04*	0.04*
Chi-X	0	-0.21*	-1.31*	-0.18*	-0.02*	-0.03*	0.26*	0.68*	0.18*	0.03*	0.03*
Chi-X	1	-0.52*	-1.47*	-0.46*	-0.09*	-0.13*	0.50*	1.00*	0.45*	0.13*	0.12*
Chi-X	10	-0.61*	-1.29*	-0.37*	-0.09*	-0.13*	0.67*	1.11*	0.46*	0.15*	0.14*
Bats	0	-0.27*	-0.58*	-1.26*	-0.04*	-0.10*	0.37*	0.60*	0.51*	0.05*	0.08*
Bats	1	-0.46*	-0.79*	-1.21*	-0.07*	-0.17*	0.52*	0.88*	0.69*	0.09*	0.16*
Bats	10	-0.54*	-0.83*	-1.01*	-0.08*	-0.15*	0.87*	1.16*	0.81*	0.14*	0.16*
Turq	0	-0.04	-0.04*	-0.05*	-0.70*	-0.03*	0.13*	0.11*	0.08*	0.14*	0.04*
Turq	1	-0.11*	-0.08*	-0.02	-0.69*	-0.04*	0.22*	0.15*	0.06*	0.19*	0.08*
Turq	10	-0.13	-0.06	0.04	-0.68*	-0.02	0.17	0.13	0.01	0.18*	0.07*
Nasdaq	0	-0.03	0.03	-0.01	0.01	-0.75*	-0.05	-0.08	0.04	0.07*	0.14*
Nasdaq	1	-0.08	0.08	0.04	0.04	-0.63*	0.00	-0.18	-0.08	0.02	0.12*
Nasdaq	10	-0.24	0.02	0.07	0.06	-0.62*	0.43	-0.19	-0.11	0.06	0.19*
£ *Sells*											
LSE	0	0.30*	0.27*	0.10*	0.02*	0.02*	-0.78*	-0.29*	-0.10*	-0.02*	-0.02*
LSE	1	0.38*	0.35*	0.17*	0.06*	0.06*	-0.75*	-0.35*	-0.16*	-0.05*	-0.06*
LSE	10	0.39*	0.29*	0.12*	0.05*	0.05*	-0.65*	-0.23*	-0.06*	-0.04*	-0.04*
Chi-X	0	0.27*	0.70*	0.19*	0.02*	0.02*	-0.24*	-1.28*	-0.20*	-0.02*	-0.03*
Chi-X	1	0.46*	0.99*	0.43*	0.12*	0.11*	-0.51*	-1.40*	-0.43*	-0.08*	-0.13*
Chi-X	10	0.57*	1.08*	0.44*	0.13*	0.12*	-0.53*	-1.15*	-0.31*	-0.08*	-0.13*
Bats	0	0.29*	0.56*	0.43*	0.04*	0.08*	-0.21*	-0.58*	-1.15*	-0.05*	-0.10*
Bats	1	0.37*	0.81*	0.62*	0.07*	0.15*	-0.38*	-0.77*	-1.11*	-0.09*	-0.16*
Bats	10	0.53*	1.00*	0.70*	0.09*	0.16*	-0.41*	-0.69*	-0.88*	-0.08*	-0.15*
Turq	0	0.10*	0.08*	0.08*	0.13*	0.04*	-0.06*	-0.03	-0.07*	-0.63*	-0.05*
Turq	1	0.17*	0.12*	0.06*	0.18*	0.08*	-0.08	-0.03	-0.06*	-0.63*	-0.03*
Turq	10	0.33*	0.11	-0.02	0.20*	0.09*	-0.17	-0.02	0.00	-0.59*	-0.04
Nasdaq	0	0.21*	0.04	0.12*	0.11*	0.16*	-0.24*	-0.04	-0.05	-0.03	-0.82*
Nasdaq	1	0.22	-0.01	0.07	0.11*	0.20*	-0.43*	-0.04	0.00	-0.03	-0.76*
Nasdaq	10	0.35	-0.14	-0.05	0.09*	0.23*	-0.84*	-0.18	-0.00	-0.03	-0.68*
R-squared		.123	.115	.069	.025	.026	.104	.113	.071	.022	.027

Table (2.6) **Consolidated liquidity impact: Descriptive statistics.**
The table shows the descriptive statistics of the dependent and independent variables of regression equation 2.10, 2.11 and 2.12. The variables are constructed per hour, of 10 stocks for 21 days ($9 \times 10 \times 21 = 1890$ observations). The top panel shows the fraction of trading volume by SORT traders, which is defined as the volume from trade sequences with length exceeding one (see Table 2.4). %SORT LSE and %SORT Entrants represent the fraction of SORT volume on the LSE and the new entrant venues. Ln(Depth(10) Cons) represents the logarithm of consolidated liquidity, i.e., the sum of the DepthAsk(10) and the DepthBid(10) of the five venues. Algo_trad is a proxy for algorithmic trading, defined as the negative of trading volume (in hundreds of pounds) divided by the number of electronic messages. Ln(realized volatility) is defined as the sum of squared 5 minute returns measured hour-by-hour, which is often negative because of the logarithm. Ln(Turnover) is the logarithm of turnover summed over all venues, and Ln(Order size) represents the logarithm of the average transaction size. The middle and bottom panel shows the impact, accumulated over one second, of a buy (sell) trade on the change in consolidated and cross-venue ask (bid) liquidity, for the LSE and the new entrant venues. This impact is the estimated by regression 2.9 per hour of each stock. The estimated variables are winsorized at the 1% and 99% level.

	Mean	Stdev	Max
%SORT	0.40	0.09	0.68
%SORT LSE	0.37	0.12	1.00
%SORT Entrants	0.45	0.10	1.00
Ln(Depth(10) Cons)	12.43	1.01	16.09
Algo_Trad	-1.93	1.63	-0.03
Ln(Realized volatility)	-10.23	2.01	0.66
Ln(Turnover)	14.99	1.38	19.04
Ln(Order size LSE)	9.10	1.33	12.13
Ln(Order size Entrants)	8.54	0.54	9.81
Consolidated liquidity impact	Mean	Stdev	Min
Buy-Ask LSE	-1.39	0.77	-3.94
Sell-Bid LSE	-1.42	0.81	-4.05
Buy-Ask Entrants	-1.76	0.98	-8.41
Sell-Bid Entrants	-1.74	0.92	-8.05
Cross venue liquidity impact	Mean	Stdev	Min
Buy-Ask LSE	-0.58	0.53	-2.50
Sell-Bid LSE	-0.60	0.54	-2.44
Buy-Ask Entrants	-0.75	0.73	-5.96
Sell-Bid Entrants	-0.74	0.68	-6.07
Observations	1890		

Table (2.7) The impact of SORT traders on consolidated liquidity.
The regression shows the impact of the fraction of SORT traders on the consolidated depth, defined as the sum of the DepthAsk(10) and the DepthBid(10) of the five venues. The independent variables are described in Table (2.6), and consist of the fraction of volume executed by SORT-traders, a proxy for algorithmic trading, the logarithm of market wide turnover, the logarithm of realized volatility based on 5 minute squared returns and the logarithm of the average transaction size. The variables are observed once per hour of each stock. I add firm and day fixed effects. For inference I apply robust Newey-West standard errors (HAC) with 9 lags; the standard errors are shown below the coefficients.

	Consolidated Depth(10)	
	(1)	(2)
% Sort	-0.860***	-0.768***
	0.267	0.262
Algo Trading	-0.006	0.024
	0.017	0.022
Ln Turnover		0.249***
		0.030
Ln Volatility		-0.129***
		0.018
Ln Order Size		0.172**
		0.085
Observations	1,859	1,859
R-squared	0.017	0.249

Table (2.8) **Determinants of the consolidated and cross venue liquidity impact on the ask side.**

I show how market characteristics affect the consolidated liquidity impact and the cross venue impact, which are estimated using Equation 2.9, per hour of each stock. The consolidated liquidity impact is the impact accumulated over one second of a venues buy-trades on the consolidated Depth(10) of the ask side (top panel), and the cross venue impact is that on the Ask Depth(10) of the competing venues (bottom panel). The impacts are estimated for trades on each of the five venues, but I average the impacts over the four new entrant venues (volume weighted). These impacts are estimated and used as the dependent variables in the weighted linear regressions below. As weight variable, I use the number of transactions per stock-hour for the LSE and the new entrants. The independent variables are described in Table (2.6), and consist of the fraction of volume executed by SORT-traders, a proxy for algorithmic trading, log market wide turnover, log realized volatility based on 5 minute squared returns and the log average order size. Table 2.9 shows the results for sales on the bid liquidity. I add firm and day fixed effects. For inference I apply robust Newey-West standard errors (HAC) with 9 lags; the standard errors are shown below the coefficients.

	Buy trades on Consolidated Ask liquidity			
	LSE	Entrants	LSE	Entrants
%Sort	0.459	1.280***	0.749**	1.177***
	0.337	0.344	0.351	0.339
Algo Trading	-0.104***	-0.214***	-0.052**	-0.207***
	0.0239	0.0403	0.026	0.042
Ln Turnover			-0.081	-0.332***
			0.056	0.061
Ln Volatility			0.072***	0.101***
			0.013	0.018
Ln Order Size LSE			0.444***	
			0.142	
Ln Order Size Entrants				0.865***
				0.154
Observations	1,857	1,857	1,857	1,857
R-squared	0.057	0.140	0.100	0.195
Weight variable	LSE trades	Entrant trades	LSE trades	Entrant trades
	Buy trades on Cross-venue Ask liquidity			
	LSE	Entrants	LSE	Entrants
%Sort Entrants	1.609***		1.223***	
	0.212		0.191	
%Sort LSE		0.216		0.447***
		0.203		0.170
Algo Trading			-0.039**	-0.133***
			0.016	0.028
Ln Turnover			-0.065*	-0.242***
			0.037	0.042
Ln Volatility			0.049***	0.06***
			0.007	0.012
Ln Order Size LSE			0.266***	
			0.091	
Ln Order Size Entrants				0.578***
				0.104
Observations	1,857	1,857	1,857	1,857
R-squared	0.064	0.001	0.154	0.155
Weight variable	LSE trades	Entrants trades	LSE trades	Entrants trades

Table (2.9) Determinants of the consolidated liquidity impact on the bid side.

This Table continues from 2.8, but shows the results for the consolidated and cross-venue liquidity impact of sales on the bid side.

	Buy trades on consolidated Bid liquidity			
	LSE	Entrants	LSE	Entrants
%Sort	0.038	1.889***	0.430	1.802***
	0.347	0.488	0.368	0.455
Algo Trading	-0.165***	-0.158***	-0.093***	-0.146***
	0.021	0.035	0.02	0.036
Ln Turnover			-0.076*	-0.287***
			0.042	0.056
Ln Volatility			0.082***	0.097***
			0.011	0.015
Ln Order Size LSE			0.602***	
			0.095	
Ln Order Size Entrants				0.834***
				0.154
Observations	1,857	1,857	1,857	1,857
R-squared	0.132	0.097	0.193	0.151
Weight variable	LSE trades	Entrant trades	LSE trades	Entrant trades
	Buy trades on cross-venues Bid liquidity			
	LSE	Entrants	LSE	Entrants
%Sort Entrants	1.655***		1.135***	
	0.188		0.173	
%Sort LSE		0.488**		0.625***
		0.206		0.214
Algo Trading			-0.055***	-0.088***
			0.015	0.027
Ln Turnover			-0.068**	-0.208***
			0.03	0.038
Ln Volatility			0.046***	0.055***
			0.007	0.011
Ln Order Size LSE			0.414***	
			0.063	
Ln Order Size Entrants				0.512***
				0.105
Observations	1,857	1,857	1,857	1,857
R-squared	0.062	0.005	0.217	0.103
Weight variable	LSE trades	Entrants trades	LSE trades	Entrants trades

Chapter 3

Does order splitting signal uninformed order flow?

Abstract

We study the problem of a large liquidity trader who must trade a fixed amount before a deadline and wishes to minimize the expected cost of trading. We add this trader to the Kyle (1985) framework to endogenize the price impact of trading. Under the assumption that the informed traders have short-lived private information, we show that the autocorrelation in the order flow stems only from the trades of the liquidity trader. In turn, the market maker perceives this autocorrelated component as uninformed and does not revise prices, such that the liquidity trader enjoys lower price impacts. We thus show that order splitting is a noisy form of preannouncing trades, i.e., sunshine trading (Admati and Pfleiderer, 1991). The model also offers a novel explanation for resiliency, i.e., why liquidity replenishes after a trade. If the market beliefs a certain trade belongs to a series of liquidity motivated trades, then the trade should not affect prices and liquidity will be replenished very quickly.

JEL Codes: G10; G11; G14;

Keywords: Market microstructure, Kyle model, Order splitting, Algorithmic trading, Optimal execution problem

3.1 Introduction

In the last decade equity turnover has increased sevenfold, while average order sizes have declined tenfold (Chordia, Roll, and Subrahmanyam, 2011).[1] Among other reasons, this is a consequence of algorithmic trading and the practice of order splitting, where a large trade is split up into many small packages which are traded over time. Order splitting, or "working the order", is a standard practice in the investment management industry.

This paper studies the optimal execution problem of an institutional investor who must trade a given quantity before a deadline, and wishes to minimize the expected execution costs. The investor optimally chooses the quantity to trade in each period and trades for liquidity motives. We endogenize the price impact parameter by placing this problem in the Kyle (1985) framework, and show that the price impact (Kyle's lambda) is strongly affected by the strategy of the liquidity trader. Moreover, we find that order splitting is a noisy form of preannouncing trades, i.e., sunshine trading (Admati and Pfleiderer (1991)). In the model, the autocorrelated component in the order flow stems from the liquidity trader only, which is a noisy signal of her uninformed trading interest. This mechanism is an additional explanation of the increasing popularity of order splitting algorithms.

The model builds on the multiperiod discrete-time model of Kyle (1985), where we add a discretionary liquidity trader who optimally splits up trades over time. The liquidity trader, the informed trader and the noise traders submit market orders to the market maker, who observes the aggregate order flow and determines the price to clear the market. The market maker

[1] Average trade sizes have decreased from \$80,000 to \$7,000, and monthly turnover increased from 6% to 40% of market cap for CRSP stocks in the period 1993 - 2008 (Chordia, Roll, and Subrahmanyam, 2011).

is risk neutral but cannot distinguish between trades from informed and uninformed investors. Therefore, the trades by the discretionary liquidity trader do affect the price, and because the market is anonymous she cannot simply preannounce her trading interest.[2]

The model's key assumption is short-lived private information, meaning that in each trading round a new informed trader arrives who may only trade in that round. Also, in each period a new innovation in the fundamental value of the asset occurs, such that the degree of informed trading becomes constant over time. While the assumptions seem strong, we motivate them by the following example of the different types players. The informed traders use high-frequency trading strategies with short trading horizons, such as arbitrage, structural or directional strategies (as defined by the SEC (2010)).[3] The discretionary liquidity trader is an institutional trader with a long trading horizon of one (or several) days, which is typical according to Campbell, Ramadorai, and Schwartz (2009). The market maker in the model is represented by a group of high-frequency traders with a passive market making strategy, who predominantly use limit orders to earn the bid-ask spread and liquidity rebates. Brogaard, Hendershott, and Riordan (2012) confirms that high-frequency traders who demand liquidity are informed traders, i.e., their trades push the price towards the fundamental value. In addition, the high-frequency traders who supply liquidity are market makers, i.e., their trades go in opposite direction to permanent price changes and are adversely selected. The informed traders in the theory of Foucault, Hombert, and Rosu (2012)

[2]While an investor could say that he will buy shares in some future trading round, there is no mechanism that forces him to actually to do so, i.e., preannouncement is a non-credible commitment.

[3]The SEC defines four broad high-frequency trading strategies: passive market making, arbitrage, structural (e.g., trading on latency and the use of flash orders) and directional (trading on fundamentals, momentum and order anticipation). See the SEC concept release on equity market structure, February 2010, File No. S7-02-10.

also have short-lived private information, and are motivated as high-frequency traders who observe news an instance before the rest of the market.

Our main contribution is that in this setting, order splitting by the discretionary liquidity trader is a noisy form of trade preannouncement, i.e., sunshine trading. This follows from a general result of the Kyle model that to the market maker, the trades of the informed trader have zero autocorrelation. Then, only the trades of the discretionary liquidity trader cause autocorrelation in the order flow, and these trades are unrelated to the fundamental value of the asset. Therefore, the market maker attaches a zero price impact to the predictable component of the order flow. Here, the assumption of short-lived private information creates the separating equilibrium, in the sense that the informed traders cannot mimic the strategy of the liquidity trader.[4] Effectively, the discretionary liquidity trader sends a credible signal of his uninformed trading interest, and is rewarded by lower price impacts.

The second contribution is that we endogenize the price impact parameter in the optimal execution problem of a large liquidity trader. The liquidity trader must trade a fixed amount before a deadline, and has a U-shaped optimal execution strategy: the first and last trades are large, and intermediate trades are small. The optimal trade size in each period depends on the following tradeoff. On the one hand, a larger trade increases the expected order flow in all future rounds, which then receive a zero price impact. On the other hand, a larger trade increases the price of the current and all future rounds, because prices only slowly revert to the fundamental value via informed trading. For the initial trade the first effect is large whereas for the last trade the second effect is small, which generates the

[4]If the informed traders would also split up trades across periods, then the autocorrelation in the order flow is not strictly uninformed and the equilibrium breaks down.

U-shape. The price impact parameter (illiquidity) comoves negatively with the quantity traded by the liquidity trader, simply because uninformed trading reduces the price impact parameter (as in Kyle (1985)).

The third contribution of the model is that the predictable component of the order flow offers a novel explanation for why resiliency exists in electronic limit order markets, i.e., why liquidity replenishes after a trade. If the market perceives that a certain trade belongs to a series of liquidity motivated trades, then the liquidity consumed by the trade should be replenished quickly. This mechanism is an alternative explanation as to why the slope in the limit order book is very high (i.e., illiquid) compared to the actual price impact of trades over time (as discovered by Sandås (2001), see also Hasbrouck (2007), Chapter 13). This empirical fact implies that a substantial part of the liquidity available in the market is not offered in the limit order book. According to our model, the realized price impact is lower than the instantaneous price impact because the anticipated component of the order flow stems from uninformed investors. Similarly, the presence of the predictable component in the order flow may also explain why liquidity on the bid and ask side is asymmetrical at times.[5]

The predictable component of the order flow does not affect prices, but is empirically unobservable. However, a direct consequence is that the immediate price impact (based on the liquidity in the limit order book) is larger than the price impact of trading volume in the short run (e.g., within minutes), which in turn is larger than the price impact of trades in the long run (e.g., over hours or days). The intuition is that at lower frequencies, a larger fraction of the cumulative order flow is expected

[5]Relatedly, Van Achter (2008) shows that asymmetric liquidity may result from heterogenous trading horizons of investors, which affects the decision to place limit or market orders.

which reduces the price impact. The economic force behind this prediction is that liquidity traders have a longer trading horizon than the informed traders.

We solve the problem numerically, as a closed form solution is not available. The problem has many state variables because the market maker must learn not only about the fundamental value, but also about the trading interest of the liquidity trader. In addition, the optimization problem of the liquidity trader is constraint, as she must trade exactly the exogenously given quantity.

The model explains several empirical findings. Griffin, Harris, and Topaloglu (2003) estimate a VAR model with five-minute returns, institutional order imbalance and retail order imbalance, and find that the positive autocorrelation of the institutional order imbalance over the preceding 30 minutes does not affect current returns (i.e., is uninformed). In addition, current returns positively predict future institutional order imbalance; which coincides with our theory as mainly the *unexpected* component of the order flow affects current returns and signals future liquidity trading interest and order imbalance. The predictions of our model also confirm several empirical results of Chordia, Roll, and Subrahmanyam (2002, 2004). In particular, they find that the daily order imbalance is strongly autocorrelated whereas returns have virtually zero autocorrelation, which suggests that predictable order flow is uninformative at the daily level. Almgren and Lorenz (2006) state that the deadline of institutional traders is typically the end of the trading day, and in this case our model matches the empirically observed U-shaped patterns of liquidity and trading volume. Heston, Korajczyk, and Sadka (2010) argue that predictable patterns in volume, returns and order imbalance are caused by systematic trading patterns of institutional investors.

Our model supports the theoretical results of Obizhaeva and

Wang (2005) and Alfonsi, Fruth, and Schied (2010). They also find an optimal U-shaped trading pattern, as a large initial trade creates a price pressure that attracts many new limit orders to the limit order book. Essentially, this mechanism is also present in our model as the informed trades reduce the price pressure caused by the strategic liquidity trader. Novel in our paper is the channel that expected uninformed order flow has zero impact on prices, and that the resiliency is explicitly modeled in an adverse selection framework. Chordia, Roll, and Subrahmanyam (2004) study a two-period model with a discretionary liquidity trader. They obtain a closed-form solution, but the discretionary trader is limited to trade either in one period only, or to split up his trades equally across both periods. Because trading is restricted to two rounds, they do not obtain the U-shape trading pattern and the liquidity trader cannot update his strategy over time. While their model focusses on the relation between order imbalance and stock returns across days, our model focusses on the optimal intraday trading strategy.

This paper contributes to the following three strands of literature, which are discussed in detail in the literature section. Above all, the paper relates to studies on the optimal execution problem of a large liquidity trader, who must trade a given quantity before a deadline and aims to minimize execution costs. In addition, we relate to the extensions of Kyle (1985) that focus on the problem of strategic liquidity traders. In general, the paper relates to dynamic models that study optimal investor behavior over time.

3.2 Literature review

This paper is particularly related to three strands of literature: the optimal execution problem of a liquidity trader, extensions to Kyle (1985) that focus on strategic liquidity traders, and dynamic models that study strategic investor behavior over time.

Our model contributes to the literature on the optimal execution problem, which is the problem of a large liquidity trader who must trade a given quantity before a deadline who aims to minimize execution costs (Bertsimas and Lo, 1998).[6] Almgren and Chriss (1999, 2000), Engle, Ferstenberg, and Russell (2012) find the optimal execution strategy in a mean-variance framework. Obizhaeva and Wang (2005) and Alfonsi, Fruth, and Schied (2010) study this problem in a limit order book market, and show that the optimal strategy strongly depends on the resiliency of the book, i.e., the speed with which the limit order book recovers after a trade. Huberman and Stanzl (2005) add transaction costs, which is important as in the continuous time limit the execution cost of the problem of Bertsimas and Lo becomes in fact independent of the actual strategy.[7] They also allow for a time-varying price impact function, which can be obtained in our framework easily as well by changing the variance of the noise trades and the fundamental innovations over time. Easley, Lopez de Prado, and O'Hara (2012) show that the order imbalance affects the endogenously determined trading horizon and the price impact function. These papers are partial equilibrium models in the sense that the price dynamics are exogenously determined. We endogenize the price impact function and show how it is affected by order splitting.

Extensions to Kyle (1985) with strategic liquidity traders, i.e., discretionary traders, are most closely related to the current paper. Admati and Pfleiderer (1988) model a group of discretionary traders who may decide in which period to submit there entire trade, and find that in equilibrium the informed and discretionary traders will trade in the same period. In contrast, we

[6] See also Kissell, Glantz, and Malamut (2003), Chapter 15 in Hasbrouck (2007).

[7] In continuous time, there are infinitely many trading rounds before the fixed deadline, which becomes meaningless.

analyze the behavior of a single discretionary liquidity trader, and find that she trades smoothly over time. The equilibrium of Admati and Pfleiderer has a coordination problem however, as in an anonymous market liquidity traders do not know when other traders will trade. Subrahmanyam (1995) analyzes circuit breakers, and extends the Admati and Pfleiderer model to a two-period version, where the discretionary trader is limited to trade either in one period only, or to split up his trades equally across both periods.While in equilibrium the discretionary trader indeed splits up across periods, the informed traders are restricted to trade in the first period only.[8] In our model a new informed trader arrives every period, which introduces an important dynamic aspect as the price pressure from uninformed trades in the current period affects the informed order flow in the next.

Back and Pedersen (1998) extend the Admati and Pfleiderer model by allowing for long-lived private information, and find that market depth and volatility are constant over time. Spiegel and Subrahmanyam (1992) extend the Kyle model by replacing the noise traders with risk averse and price sensitive liquidity traders who trade for hedging motives.[9] Massoud and Bernhardt (1999) extend the previous model to a two-period version, and find that some results of Kyle get reversed. For example, the price impact becomes steeper over time, because the liquidity traders wish to trade in earlier periods as to avoid the pricing risk in later periods. In Spiegel and Subrahmanyam (1995) risk averse discretionary liquidity traders also trade for hedging motives, and will trade either at the beginning of the day, or later in the same direction as the market makers—effectively providing liquidity. Mendelson and Tunca (2004) find that the risk aversion of liquidity traders generally reduces informational ef-

[8] Chordia, Roll, and Subrahmanyam (2004) also make the assumptions of equal order splitting and a single period with informed trading.

[9] The liquidity traders have hedging motives in Vayanos (1999) too.

ficiency, and that insider trading may improve the welfare of risk averse liquidity traders because they reduce the volatility of prices.

This paper also relates to the preannouncement of trading interest (Admati and Pfleiderer, 1991), as order splitting generates a predictable and uninformed component in the order flow. Huddart, Hughes, and Williams (2010) analyze the case where the insider must preannounce his trades, but also has liquidity motives to trade (e.g., risk sharing). Suboptimal risk sharing follows, because even though an insider might trade for liquidity reasons, the market makers revise prices because the trade may reflects private information. Huddart, Hughes, and Levine (2001) analyze the Kyle model where informed traders must announce their trades after submission, like employees of a corporation need to. In this case, the insider adds some noise to his strategy to jam the signal of the market maker.

Several dynamic models also study the strategic investor behavior over time. Foucault (1999) models the decision to place limit or market orders, where limit orders face adverse selection costs and non-execution risk. Hoffmann (2012) extends this model by allowing traders to invest in speed, and shows that fast traders face lower adverse selection costs because they can cancel limit orders quickly after news arrives. In turn, this affects the tradeoffs between limit and market orders and the overall gains from trade. Foucault, Kadan, and Kandel (2005) do not model asymmetric information costs, and show that the choice between limit and market orders depends on the tradeoff between the cost of waiting (impatience) and better prices. Rosu (2009) also studies waiting costs, and shows that the bid and ask prices comove because investors revise limit orders after a liquidity shock. Goettler, Parlour, and Rajan (2005) allow traders to also choose the size of the order and to place limit orders deeper in the book, and demonstrate a numerical procedure to find the equilibrium. Goettler, Parlour, and Rajan (2009) add strategically informed traders into the dynamic limit order market, who may purchase information on the assets fundamental value. Rosu (2010) allows investors to revise limit orders continuously, and models the optimal behavior of informed traders and the resulting price impacts. We contribute to these papers by analyzing the problem of a liquidity trader who may trade repeatedly within a certain horizon.

3.3 Model setup

Consider a Kyle (1985) framework, where trading occurs sequentially in a number of auctions or trading rounds. Each auction is organized as an anonymous batch market where investors submit market orders. Trading begins at time 0 and ends at time 1, and takes place during $n = 1, ..., N$ periods, each of length $1/N$. Time 0 represents the beginning of the trading day for example, and time 1 the end. Three players exist in the Kyle model. There is a risk neutral informed trader who observes the fundamental value of the asset and trades to maximize profits. In addition, a group of noise traders trade random amounts every period. Then, a competitive market maker first observes the total order flow in each period, and next chooses the price and his quantity that clears the market.

We deviate from the standard multi-period Kyle model in two important ways. First, we introduce a strategic liquidity trader who must trade a given amount before the deadline at period N. She chooses the optimal quantity to trade in each round, and is a "discretionary" liquidity trader following the terminology of Admati and Pfleiderer (1988).[10] The total quantity is drawn

[10]The discretionary trader in our model submits an optimal fraction each period, whereas the liquidity trader in Admati and Pfleiderer may only choose a single period to submit the entire trade.

from a normal distribution before trading starts, and does not change afterwards.

Second, we assume that in *every* trading round a new informed trader arrives, who observes the fundamental value and may trade only once. Also, in every period an innovation in the fundamental value of the asset occurs. In essence, we assume that private information is short lived, which resembles Foucault, Hombert, and Rosu (2012) where high-frequency traders (HFTs) respond to news (i.e., short-term information) extremely quickly. This setup generates a constant level of informed trading, which is realistic as it may follow from news revealed by trades in correlated assets for example.

The traders are all risk neutral, and submit market orders. Denote by $S \sim N(0, \sigma_S^2)$ the total quantity that the strategic liquidity trader must trade before the deadline, where a positive value represents a purchase while a negative value a sell. This quantity is exogenously determined and she cannot access other trading venues. In each periods n she submits a fraction $f_n S$, where she chooses the vector $f_1, ..., f_N$, subject to $\sum_{n=1}^{N} f_n = 1$. Denote by x_n the strategically chosen order flow of the informed investor, and $u_n \sim N(0, \sigma_u^2)$ the randomly determined uninformed order flow of non-discretionary noise traders. Then, the total order flow each period is

$$y_n = x_n + f_n S + u_n. \tag{3.1}$$

The process of the fundamental value is given by

$$v_n = v_0 + \sum_{j=1}^{n} \varepsilon_j, \tag{3.2}$$

where $\varepsilon_j \sim N(0, \sigma_\varepsilon^2)$ and IID.

An important element in the model is that a trade by the liquidity trader affects the current price, which in turn affects the strategy of informed traders in future periods. In fact, price pressures due to trades in the current period are beneficial to informed traders in future periods, as the pricing error gets larger in magnitude. This mechanism does not exist in the models of Bertsimas and Lo (1998), Almgren and Chriss (2000) and Huberman and Stanzl (2005) for example.

Like in Kyle's model, we restrict attention to the recursive linear equilibrium, such that the market maker's pricing function and the strategies of the traders are linear in their information set. The market maker determines the price P_n after observing the current and past order flow, which are summarized in a vector $I_n = \{y_1, ..., y_n\}$ that represents his information set. The information set contains the sequence of past order flow, but not past prices. Since prices depend linearly on the order flow, the order flow contains all information. We conjecture (and verify in equation (3.6)) that the pricing schedule is

$$P_n = P_{n-1} + \lambda_n(y_n - E(y_n|I_{n-1})), \tag{3.3}$$

where $E(y_n|I_{n-1})$ represents the expected order flow. Intuitively, the price set by the market maker is unaffected by the expected uninformed order flow. Given short lived private information, $E(y_n|I_{n-1}) = f_n E(S|I_{n-1})$, i.e., the expected order flow depends only on the market makers expectation of the quantity traded by the strategic liquidity trader. The informed order flow is unpredictable to the market maker, which is a result obtained by Kyle. Intuitively, the autocorrelated component of the informed order flow is concealed by the camouflage of noise trading. A proof by contradiction is that if informed trades were autocorrelated, the market maker could immediately set a price that reflects this information which eliminates the autocorrelation. In the model, the market maker behaves competitively and incorporates all information available to him.

3.3.1 Problem of the informed trader

The problem of the informed trader is identical to the one period version of Kyle (1985). The informed trader observes v_n (a perfect signal) and I_{n-1}, and knows the pricing function (3.3). He chooses x_n to maximize his expected profits

$$\max_{x_n} E[x_n(v_n - P_n)|v_n, I_{n-1}]$$
$$= \max_{x_n} \left(x_n \left(v_n - P_{n-1} - \lambda_n(E(y_n|v_n, I_{n-1}) - E(y_n|I_{n-1})\right)\right)\right). \tag{3.4}$$

Given $y_n = x_n + u_n + f_n S$ we have $E(y_n|v_n, I_{n-1}) - E(y_n|I_{n-1}) = x_n$. We set the first order condition to zero with respect to x_n,

$$0 = E[(v_n - P_{n-1} - 2\lambda_n x_n)]$$
$$x_n = \beta_n(v_n - P_{n-1}), \qquad \text{with } \beta_n = \frac{1}{2\lambda_n}, \tag{3.5}$$

where β_n represents the aggressiveness of the informed trader in period n. Importantly, the market maker cannot predict informed order flow, as $E(x_n|I_{n-1}) = 0$ because to the market maker $E(v_n|I_{n-1}) = E(v_{n-1}|I_{n-1}) = P_{n-1}$.

Note that x_n does not depend on the number of trading rounds like in the multi-period version of Kyle, because each insider trades only once. Importantly, the insider only observes v_n, but not the history of v; which would otherwise enable him to reconstruct the history of x and therefore predict S much clearer than the market maker can. The equilibrium would then break down, because the insider has an incentive to front-run the strategic liquidity trader which in turn would affect the liquidity traders' strategy. However, a new informed trader arrives every period and therefore they cannot observe the history of x.

3.3.2 Problem of the market maker

The strategic liquidity trader submits a fraction of her total quantity S each period, of which a part is expected and a part unexpected, i.e., reveals new information about her trading interest. In this setup, the market maker needs to learn about the fundamental value of the asset v_n and about the quantity S. Learning about S improves the prediction of uninformed order flow, such that the informed component of the order flow provides a clearer signal about the true value. The market maker observes the order flow y_n, and learns about S and v_n by updating the conditional expectations and variances.

With respect to the fundamental value, market efficiency states that the market maker chooses the price such that $P_n = E(v_n|I_n)$. We repeat that $E(y_n|I_{n-1}) = f_n E(S|I_{n-1})$, because to the market maker the informed trades have zero autocorrelation. This is a result obtained in Kyle's model, as noise trading provides camouflage which conceals the autocorrelated component of the informed order flow.

Using the linear projection theorem we can write the conditional expectation of the fundamental value and quantity S recursively (the derivation is in the Appendix):

$$\begin{aligned} P_n &= P_{n-1} + \lambda_n \left(y_n - f_n E(S|I_{n-1})\right), \\ \lambda_n &= \frac{\beta_n Var(v_n|I_{n-1})}{\beta_n^2 Var(v_n|I_{n-1}) + \sigma_u^2 + f_n^2 Var(S|I_{n-1})}, \end{aligned} \tag{3.6}$$

and

$$\begin{aligned} E(S|I_n) &= E(S|I_{n-1}) + \varphi_n(y_n - f_n E(S|I_{n-1})), \\ \varphi_n &= \frac{f_n Var(S|I_{n-1})}{\beta_n^2 Var(v_n|I_{n-1}) + \sigma_u^2 + f_n^2 Var(S|I_{n-1})}. \end{aligned} \tag{3.7}$$

The conditional variance of S is $Var(S|I_{n-1})$, and of v_n is $Var(v_n-$

$P_{n-1}|I_{n-1}) = Var(v_n|I_{n-1})$. Comparing the solution of P_n to the pricing rule (equation (3.3)) confirms that the informed trader trades linearly on his signal, and provides the definition of λ_n. The informativeness of the order flow about quantity S is represented by $f_n\varphi_n$ (the signal strength). The φ_n and λ_n depend on the variances of S and v_n conditional on the information set I_{n-1}. These are obtained recursively using the general rule for conditional variances[11] (see Appendix):

$$Var(v_{n+1}|y_n, I_{n-1}) = Var(v_n|I_{n-1}) + \sigma_\varepsilon^2 - \tag{3.8}$$

$$\frac{\beta_n^2 Var(v_n|I_{n-1})^2}{\beta_n^2 Var(v_n|I_{n-1}) + \sigma_u^2 + f_n^2 Var(S|I_{n-1})}. \tag{3.9}$$

The recursion of the conditional variance of the quantity S depends on past realizations of y but not on past prices. Using the definition of φ_n above, the conditional variance of S is

$$Var(S|I_n) = Var(S|I_{n-1})(1 - \varphi_n f_n). \tag{3.10}$$

The recursions for $Var(S|I_n)$ and $Var(v_{n+1}|I_n)$ are forward recursions, and start with initial values σ_S^2 and σ_ε^2. Now, we have five equations for $Var(v_{n+1}|I_n)$, λ_n, β_n, φ_n and $Var(S|I_{n-1})$, which we can solve in terms of exogenous parameters, recursive parameters and the yet unknown f_n. The only assumption we have made so far is the linear pricing rule in (3.3):

$$Var(v_{n+1}|I_n) = 1/2 Var(v_n|I_{n-1}) + \sigma_\varepsilon^2,$$

$$\lambda_n = \frac{\sqrt{Var(v_n|I_{n-1})}}{2\sqrt{\sigma_u^2 + f_n^2 Var(S|I_{n-1})}},$$

[11] See for example Chapter 4 of Hamilton (1994) for the general rule of conditional expectations and variances.

$$
\begin{aligned}
\beta_n &= \frac{1}{2\lambda_n},\\
\varphi_n &= \frac{f_n Var(S|I_{n-1})}{2\left(\sigma_u^2 + f_n^2 Var(S|I_{n-1})\right)},\\
E(S|I_n) &= (1 - f_n\varphi_n)E(S|I_{n-1}) + \varphi_n y_n,\\
Var(S|I_n) &= (1 - f_n\varphi_n)Var(S|I_{n-1}). \qquad (3.11)
\end{aligned}
$$

Important to note is that the parameters $Var(v_{n+1}|I_n), \lambda_n, \beta_n, \varphi_n$ and $Var(S|I_{n-1})$ do not depend on the realizations of the quantity S, noise trades and innovations in the asset value. In fact, for a given f_n they are static and determined before the trading day begins. Only the expectation of the size of the liquidity trade $E(S|I_n)$ and the price P_n depend on the order flow (which in turn is a function of the realization of S, noise trades and the innovations in the asset value).

3.3.3 Problem of the strategic liquidity trader

The goal of the strategic liquidity trader is to find an optimal strategy $f_1^*, ..., f_N^*$ that minimizes the expected total execution costs. Furthermore, this strategy must be a rational expectations equilibrium, in the sense that the market maker must correctly anticipate the strategy to determine his response, and given the response of the market maker the strategy must indeed be optimal for the liquidity trader.

Crucial for the optimal strategy is the realization of the quantity S, which in fact contains two important sources of information. First, although the liquidity traders' trades are uninformed, they do affect prices and therefore create a pricing error. Because this pricing error will be exploited by the informed traders in future rounds, the liquidity trader can predict future order flow and prices. This information depends on her own trades, and is static in the sense that it is known before trading starts.

Second, after observing any period's order flow the liquidity trader can filter out her own trades, such that the remaining component reveals a more precise signal about the fundamental value (as compared to the market maker). Effectively, she may construct an improved λ_n (equation (3.11)) because the conditional variance of S is zero to her. This information is dynamic and affects the optimal trading strategy depending on market conditions, i.e., the realizations of noise trades and asset innovations.

We proceed by analyzing the dynamic problem, where the liquidity trader may learn from the order flow and each period recalculate the optimal trading strategy of all the remaining periods. The static solution is simply the dynamic solution before any trades have occurred, i.e., based on the liquidity traders information set at time 0. We find numerical solutions for the static and dynamic problems, but only the static solution is also a rational expectations equilibrium. The dynamic strategy is not a rational expectations equilibrium, because the market maker cannot foresee the changes in the liquidity traders strategy.[12] However, Monte Carlo simulations presented in the results section reveal that the deviations in the strategy are relatively small, which suggest that the main results are robust.

Dynamic problem

In the dynamic problem the strategic liquidity trader updates her strategy according to the realizations of the order flow. The liquidity trader can form a better expectation (and variance) of the fundamental value than the market maker. She constructs these expectations in a similar fashion to the market maker, but conditions on a bigger information set.

[12]Intuitively, the changes in the liquidity traders' strategy are based on information only available to her.

The liquidity trader uses the expectation of the fundamental value to predict the informed order flow and the market makers' expectation of the quantity S in future periods. These elements combined with the liquidity traders own strategy determine the expected pricing schedule in each period. The expected prices of each period are then a function of the strategy of the liquidity trader, who minimizes the total expected execution costs.

The strategic liquidity trader extracts a signal about the fundamental value from the order flow. To denote the information set of the liquidity trader, we add a superscript S, for example I^S_{n-1}. After observing y_{n-1}, the expectation of v_n becomes (see Appendix)

$$E(v_n|y_{n-1}, I^S_{n-2}) = E(v_{n-1}|I^S_{n-2}) + \big(y_{n-1} - E(y_{n-1}|I^S_{n-2})\big) \times \frac{\beta_{n-1} Var(v_{n-1}|I^S_{n-2})}{\beta^2_{n-1} Var(v_{n-1}|I^S_{n-2}) + \sigma^2_u}, \tag{3.12}$$

where

$$\begin{aligned} E(y_{n-1}|I^S_{n-2}) &= E(\beta_{n-1}(v_{n-1} - P_{n-2}) + f_{n-1}S + u_{n-1}|I^S_{n-2}), \\ &= \beta_{n-1}\big(E(v_{n-1}|I^S_{n-2}) - P_{n-2}\big) + f_{n-1}S, \end{aligned}$$

and

$$Var(v_{n-1}|I^S_{n-2}) = \sigma^2_\varepsilon + Var(v_{n-2}|I^S_{n-3}) - \frac{\beta^2_{n-2} Var(v_{n-2}|I^S_{n-3})^2}{\beta^2_{n-2} Var(v_{n-2}|I^S_{n-3}) + \sigma^2_u}. \tag{3.13}$$

In the next section we show that the liquidity trader will use the current expectation of the fundamental value to predict the entire future paths of all the recursive variables in the model. Based on these predicted values she then minimizes the expected total execution costs.

3.4 Model Solution

To solve for the optimal strategy of the strategic liquidity trader, we proceed by backward induction. At time 0, the problem of the discretionary liquidity trader is

$$\min_{f_1,\dots,f_N} E_0^S \left[\sum\nolimits_{n=1}^{N} P_n f_n S\right],$$
$$s.t. \quad \sum\nolimits_{n=1}^{N} f_n = 1.$$

Define the value function V_n as the total expected expenditures on trades in rounds $n, ..., N$, assuming that the liquidity trader makes the best possible decision in each period. Because trading finishes after period N, we have that $V_{N+1} = 0$. At any period $n < N$, V_n will depend on the quantity left to trade and the set of state variables $State_{n-1} = P_{n-1}$, $Var(v_n|I_{n-1})$, $E(S|I_{n-1})$, $Var(S|I_{n-1})$, $E(v_n|I_{n-1}^S)$, $Var(v_n|I_{n-1}^S)$. The recursive solutions of these state variables are defined earlier, and are a sufficient statistic for the history of order flow until period n. The state variables allow the liquidity trader to make the optimal decision in each period by submitting a fraction f_n of the total quantity S. Denote by F_n the fraction of the total order she still needs to trade (before the start of period n), then

$$F_n = F_{n-1} - f_{n-1}, \tag{3.14}$$

such that $F_1 = 1$ and $F_{N+1} = 0$. In period N the trader has no choice but to submit the remaining quantity $f_N S = F_N S$. Then, using the pricing equation (3.3), the expected costs are

$$\begin{aligned} V_N(F_N, State_{N-1}) &= E_{N-1}^S[P_N f_N S] \\ &= (P_{N-1} + \lambda_N f_N (S - E(S|I_{N-1}))) f_N S. \end{aligned} \tag{3.15}$$

In period $N-1$, she trades $f_{N-1}S$ leaving the remaining $(F_{N-1}-f_{N-1})S$ for period N. Thus, in period $N-1$ she minimizes

$$V_{N-1}(F_{N-1}, State_{N-2}) = \min_{f_{N-1}} E^S_{N-2}[P_{N-1}f_{N-1}S + V_N(F_N, State_{N-1})]. \tag{3.16}$$

In this equation we must use the recursive solutions for the parameters (3.11), the rule for the fraction left to trade (3.14), and the solutions for $E(v_n|I^S_{n-1})$ and $Var(v_n|I^S_{n-1})$. The typical procedure is to fill in the recursive solutions and take expectations, such that the value function depends only on variables known to the liquidity trader at time $N-2$, and then solve the first order condition with respect to f_{N-1} (see e.g., Bertsimas and Lo (1998)). However, we cannot obtain a closed form solution of this problem due to the complexity introduced by the many state variables, i.e., the value equation contains high-degree polynomials. Therefore, we use numerical methods to find the solution.

3.4.1 Numerical approach

We solve the model using the following numerical procedure. We start by taking numerical values for the exogenous parameters of the model, $N, \sigma^2_u, \sigma^2_\varepsilon$ and σ^2_S, which are the number of trading rounds, and the variances of the noise trade, innovations in the fundamental value and the quantity of the strategic liquidity trader. Based on these parameters, we can solve for $\lambda_n, \beta_n, Var(v_n|I_{n-1}), Var(S|I_{n-1})$ and $Var(v_n|I^S_{n-1})$ as a function of $f_1, ..., f_N$ only. Next, we take the value function described in Equation (3.16) and iterate it backwards to period 1, using the recursive solutions of equations (3.11), (3.14) and (3.12). Note that at each iteration n we add P_nf_nS, such that each step represents the total expected costs to be paid in periods $n, ..., N$. Then, the value function in period 1 is a single equation that contains the total expected costs, as a function of $f_1, ..., f_N$ and the realizations of $u_1, ..., u_N$ which are set to zero in expectation (the ε_n are incorporated via equation (3.12)).

We next iteratively find a rational expectations equilibrium, where the market maker correctly anticipates $f_1, ..., f_N$ and determines his response (i.e., chooses λ_n and φ_n), and given the response of the market maker the sequence $f_1, ..., f_N$ minimizes the expected trading costs of the liquidity trader. The iteration starts by the assumption that the market maker believes $f_n = 1/N$ for all n, and then we minimize the value function yielding the liquidity traders optimal strategy $f_1, ..., f_N$. We substitute this solution into the market makers belief, and again we calculate the liquidity traders optimal strategy. We iterate this process until convergence, i.e., until the optimal strategy is arbitrarily close to the market makers expectation of the strategy.[13] Now we have obtained the static solution, which generates fixed values for $\lambda_n, \beta_n, \varphi_n, Var(v_{n+1}|I_n)$ and $Var(S|I_n)$.

Note that the optimal strategy is independent of the realization of S, i.e., the fraction submitted by the liquidity trader each period is independent of the total quantity she must trade. This is a necessary requirement as the market maker cannot observe S, and otherwise would not be able to form a rational expectation of the sequence $f_1, ..., f_N$.

The dynamic solution continues from the static solution. In round 1, the strategic liquidity trader submits f_1 from the static solution (which now becomes realized), and afterwards observes y_1. Next, we again start from the value function V_{N-1} from equation (3.16), but iterate backwards to period 2 (as period 1 has realized). As before, at each step n we add $P_n f_n S$, such that in period 2 we obtain a single equation that contains the total expected costs of period $2, ..., N$. We substitute the realization

[13]In the results section, we assume convergence is reached when $\sum_{i=1}^{N} |f_i^* - f_i| < 0.00001$, where f_i^* is the optimized sequence and f_i the market makers' expectation.

of y_1 and set the expectation of the realizations of $u_2, ..., u_N$ to zero. We find the sequence $f_2, ..., f_n$ that minimizes this equation given the static beliefs of the market maker. Typically, this solution differs from the static solution. In round 2, the insider submits f_2 (which now becomes realized), and afterwards observes y_2. This process is iterated until we reach period N. Essentially, this procedure repeats the static solution for every time period. In each period we find the global solution of the sequence f_n conditional upon the information set, because we minimize the entire problem in a single equation.

The dynamic solution typically differs from the static solution. The reason is that in the dynamic solution, the liquidity trader learns from the previous order flow and updates her entire (expected) future trading strategy. She may form a more precise estimate of the fundamental value than the market maker, because she can filter out her own (uninformed) trades from the aggregate order flow to obtain a clearer estimate of the informed order flow. Using this additional information updates, she decides to speed up or delay her trades.

3.5 Results

We first analyze the optimal strategy of the liquidity trader before trading begins, but after the realization of $S, N, \sigma_u^2, \sigma_\varepsilon^2$ and σ_S^2. The liquidity trader incorporates in her strategy the impact of her trades on the predictable order flow and prices, which in turn affects future periods informed order flow, predictable order flow and prices.

3.5.1 Static solution

We solve the model numerically for the base case first, and then change the value of one parameter at a time to analyze the impact of that parameter on the equilibrium outcome. For the base case we use the following parameter values. The number of trading periods $N = 5$, the strategic liquidity trader must trade the quantity $S = 3$, which has an unconditional variance of $\sigma_S = 5$. The volatility of noise trading $\sigma_u = 1$ and innovations in the fundamental value $\sigma_\varepsilon = 1$ per trading round.

Figure 3.1 plots the parameters of interest over time for the base case. The upper panel shows the U-shaped trading strategy of f_n, which ranges between $16.2\% - 26.4\%$ and the closely corresponding inverted U-shape for the price impact parameter λ_n. The variables comove, since a larger amount of uninformed trading in a certain period increases the liquidity of that period (as in Kyle (1985)). The bottom panel shows the total order flow y_n, informed order flow x_n and the market makers expectation of the uninformed order flow $E(S|I_n)$, which are expectations formed d by the liquidity trader. After round 1, the informed trades are in opposite direction to the trades by the strategic liquidity trader, such that the aggregate order flow lies relatively close to zero.

The results for the other parameters and the deviations to the base case are presented in Table 3.1. The variance of the pricing error is constant over time and equals $2\sigma_\varepsilon^2$, because in each period the variance of the new innovation in the fundamental value equals the information revealed by the informed trader. Most of the learning on S (column 5) occurs in the first trading round because f_1 is large, and then $Var(S|I_1)$ reduces from 5 to $4, 35$. The expected order flow y_1 is 0.793, which is large and generates a substantial price impact (the price increases from 10 to 10.48). Therefore, in period 2 the informed trader sells 0.363 and pushes the price partly back towards the fundamental value (in fact, he halves the initial pricing error). In period 2 the liquidity trader buys 0.519, such that the expected net order flow y_2 is 0.156. The first cell in the last column shows the expected execution cost of a benchmark model where the market maker does not learn about the trading interest of the liquidity trader. Then she trades an optimal fraction $1/N = 0.2$ each period, and the model is reduced to a repeated single period Kyle (1985) model. In the benchmark case the expected execution costs are 31.874, and with optimal order splitting 31.685. The gains of optimal order splitting and learning are 0.19, on a total of 1.685 (note that the fundamental value of the three shares remains 30.00).

The main result of the model is that the market maker learns about the trading interest of the uninformed liquidity trader, which he incorporates in setting the price. In particular, the initial large liquidity trade increases the market makers expectation of S to $E(S|I_1) = 0.389$, such that in period 2 the market maker subtracts the impact of the expected order flow $\lambda_2 f_2 E(S|I_1) = 0.044$ from the price. The price reduction lasts for all future periods as well, such that the remaining 74% of the order size receives a discount. After round 2, an additional $\lambda_3 f_3 E(S|I_2) = 0.045$ is subtracted from P_3 and the future prices, and similar amounts in the remaining periods. Summed over all periods, the predictable component of the order flow reduces P_5 by 0.20, which is large relative to the total price impact of trading of 0.66 $(P_5 - v_0)$. The predictable component in the order flow reduces prices as it effectively signals the liquidity traders uninformed trading interest, i.e., order splitting is a noisy form of sunshine trading (Admati and Pfleiderer (1991)).

The market makers' expectation of the uninformed order flow $E(S)$ does not exceed 0.44 in any period and therefore seems small. In fact, signalling is quite costly to the liquidity trader for two reasons. First, the market maker cannot distinguish informed from uninformed order flow, which impairs learning about the interest of the liquidity trader. In fact, equation (3.11) shows that the impact of an additional unit of order flow on $E(S|I_n)$ is strictly smaller than 0.5, i.e., $f_n \varphi_n < 0.5$. Second, a larger purchase by the liquidity trader in the current period causes a price pressure which will be followed by sells of informed traders in future periods. These sells reduce the market maker's expectation of S.

The parameter values explain why the optimal trading strategy is U-shaped, which follows from a tradeoff. On the one hand a larger current trade provides a stronger signal about future trades of the liquidity trader, which then enjoy a lower price impact. On the other hand, a larger current trade increases both the current and future prices, as trades by informed traders only slowly push the price back towards the fundamental value. The positive effect is particularly important in the first trading period, as a large trade generates a clear signal to the market maker about future uninformed trading interest, which is beneficial for all remaining trading periods. In the last trading period, the negative effect is small because the impact of the last trade on future prices is irrelevant to the liquidity trader.

An immediate consequence of the predictable component in the order flow is that the price impact is relatively large in each individual period, but gets smaller when measured across several periods. For example, the combined order flow in period 1 and 2 is 0.95 and the price change 0.54, which gives a price impact of 0.56 (as compared to $\lambda_1 = 0.61$ and $\lambda_2 = 0.66$). The price impact measured over all the five periods is 0.48.[14] This is an important testable prediction of the model.

The impact of expected uninformed order flow on prices offers a novel explanation for resiliency in electronic limit order markets, i.e., why liquidity replenishes after a trade. If the market perceives that a certain trade belongs to a series of trades and is likely to be uninformed, then the market attaches a low price impact to that trade such that liquidity will be replenished

[14]The price impact over the five periods is $P_5 - P_0 = 0.665$ divided by the cumulative order flow 1.377.

quickly. In this paper we model a batch auction where the resiliency occurs instantaneously, i.e., that $\lambda_2 f_2 E(S|I_1)$ is simply subtracted from P_2, but in a limit order market traders needs some time to respond. Note that the expected order flow affects the price level, but not the slope of the limit order book λ_n. Empirically, the expected order flow is unobservable however.

The mechanism of expected uninformed order flow may also explain why liquidity on the bid and ask side of the order book can be asymmetrical at times. If the market expects substantial order flow from buyers, the ask side should be more liquid. A trade executed on the bid side is more likely to stem from informed traders, such that the bid side is less liquid.

The second panel of Table 3.1 shows the case where the strategic liquidity trader trades only a small quantity, $S = 1$. A different realization of S affects the realized and expected order flow, and in turn the price and future periods' informed order flow (but not the other parameters). The third panel shows the case with relatively large innovations in the fundamental value, which increases the illiquidity parameter λ_n and therefore expected prices and trading costs. In panel four $N = 3$, so that the liquidity trader must trade large quantities each period, which in turn strongly increases the expected price and execution costs. This price increase gets partly mitigated by the lower values of λ_n, which follows from the relatively large amount of liquidity trading each period. In the bottom panel the unconditional variance of S is low, $\sigma_S = 1$. In this case, the market maker expects to learn very little about S, such that the expected order flow remains small, i.e., the liquidity trader cannot signal her trading interest very well. For this reason, there is less curvature in the U-shape of the sequence of f_n.

3.5.2 Dynamic solution

In the dynamic version of the model, in each period the strategic liquidity trader forms expectations of current and all future parameter values, conditional upon her information set. After a trading round she observes the order flow, and updates her expectations and optimal strategy accordingly. Therefore, for each trading round n she constructs N estimates, which gives an $N \times N$ matrix for every parameter.

The dynamic results of the model are shown in Table 3.2 and 3.3. Parameter values in bold font type have realized, whereas the other values are expectations of the strategic liquidity trader. Table 3.2 shows the parameters in an environment where the fundamental value rises, as we set $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 1, \varepsilon_4 = \varepsilon_5 = 0$ and all noise trades to zero. Table 3.3 studies an environment where only the noise traders purchase, and set $u_1 = u_2 = u_3 = 1, u_4 = u_5 = 0$ and all innovations in the fundamental value to zero.

In the first panels we observe that the realizations of the order flow cause only small dynamic adjustments to the optimal strategy of the strategic liquidity trader. The largest change in the left panel is for example from $E(f_3|I_0^S) = 0.162$ to $E(f_3|I_2^S) = 0.156$, which in the right panel becomes $E(f_3|I_2^S) = 0.158$. Due to the buying pressure of the informed traders (Table 3.2) and noise traders (Table 3.3), the strategic liquidity trader beliefs the stock is overpriced and therefore postpones her own purchases a little.[15] By postponing her trades, she anticipates that informed traders in future rounds will reduce the pricing error—making it cheaper to buy then. The second panel shows the expected and realized prices. The expected prices adjust slowly to the order flow, because the liquidity trader is uncertain whether order

[15]The bottom panel of the left block shows that $E(v_3|I_2^S) = 11.303$, and $E(P_3|I_2^S) = 11.706$, i.e., the liquidity trader indeed believes the stock is overpriced by 0.40.

flow stems from the informed or noise trader. The expected and realized order flow in the next block shows more clearly that the order flow is difficult to predict. If current purchases stem from noise traders, then next periods' informed trader will sell (to correct the mispricing), whereas if current purchases stem from the informed trader, then next periods' informed trader will also purchase (because informed traders reveal only half their signal). As a result, predicting future order flow is difficult for the liquidity trader.

The fifth block shows the liquidity traders expectation of the market makers expectation of S. The market maker beliefs that the buying pressure of the informed traders (Table 3.2) and noise traders (Table 3.3) might stem from the liquidity trader, and therefore the expectation of $E(S|I_n)$ increases over time. The last block shows the realizations of the remaining parameters. While these also change over time, the liquidity trader's expectations of the future values equals that of the current value, and therefore we report them in a single column.

3.5.3 Model shortcoming

In this version of the model, only the static version is a rational expectations equilibrium. The dynamic model is not, as the liquidity trader updates her strategy over time which cannot be predicted by the market maker. The static model is appropriate if the liquidity trader cannot modify her strategy over time, or if she does not filter out her own trades from the order flow to learn about the fundamental value.[16] In practise however there is no mechanism that prevents her from dynamically updating. Absence of a rational expectations equilibrium also implies that the market maker does not earn zero expected profits.

Therefore, we execute monte carlo simulations in the next sec-

[16] In the latter case she has no incentive to change her strategy.

tion, and show that the dynamic changes are relatively small. Thus, the main result of the model that the optimal trading strategy is U-shaped should hold. Also, the loss to the market maker because of dynamic updating appears small, and is not statistically significantly different from zero for 10.000 simulations. Still, this issue must be solved in a future version of the paper.

3.5.4 Monte Carlo simulations

We do two simulations of 10.000 trading games. In the first we draw random values for the realizations of S and all ε_n and u_n (with variances $5, 1$ and 1 respectively), whereas in the second we set $S = 3$ as in the base case and draw all ε_n and u_n. The overall profits to the traders are reported in Table 3.4.

The upper panel shows that for random S, the market maker earns a loss on average (-0.085), which is small compared to the standard deviation of 6.6 per game. This loss is not statistically significant with a t-statistic of $0.085/(6.6/\sqrt{10000}) = -1.2$, which implies that the test has either insufficient power (and we must increase the number of simulations), or the market maker is not worsely affected by the dynamic updates of the liquidity trader. In fact, by dynamically updating her strategy, the liquidity trader trades on fundamental information and effectively competes with the informed traders. This might be beneficial to the market maker.

The average informed traders' profit is 4.25, and the loss to the liquidity trader and noise traders is 0.88 and 3.28 respectively. The fifth column shows the loss of the liquidity trader in a model where the market maker does not learn about the trading interest of the liquidity trader. By signalling, the strategic liquidity trader reduces expected losses from 0.98 to 0.88.

The middle panel of the table shows the profits for the base case where $S = 3$. Here, the market maker earns a positive profit (0.21), because the realization of $S = 3$ is a relatively large draw compared to $Var(S) = 5$, and more uninformed trading is profitable to the market maker. Although $E(S) = 0$, the trading strategy and profits are symmetric for positive and negative values of S. Therefore, the expected magnitude of the liquidity trader is the absolute value of S, which is $E(|S|) = \sigma_S^2\sqrt{2/\pi} = 1.784$ for $S \sim N(0,5)$. Then, for $|S| > 1.784$, the market maker earns a positive expected profit. Compared to the upper panel, the liquidity trader earns a greater loss when $S = 3$ (-1.59 compared to -0.88), which is beneficial to the informed traders and the market maker.

The bottom panel shows the static and average dynamic strategy of the liquidity trader (the sequence of f_n) for the simulations of random S. The absolute differences are small (less than 0.006), and for 10.000 simulations only f_2 is significantly smaller in the dynamic strategy. This suggests that the main results are robust; i.e., the U-shaped pattern of the optimal strategy. In fact, it appears that the dynamic strategy has a slightly stronger U-shape than the static solution.

3.5.5 Extant literature

In this section we relate our modeling assumptions and results to the literature.

The assumption of short-lived private information is strong and deserves further explanation. In the model, the assumption creates a separating equilibrium between the informed traders and the strategic liquidity trader, such that the predictable component in the order flow is strictly uninformed. Without the assumption the informed traders could mimic the strategy of the liquidity trader, and the equilibrium would break down. Note that short-lived private information is not necessary for the result that to the market maker the order flow of the informed trader is not autocorrelated.[17]

Our motivation of short-lived private information is the difference in the length of the trading horizons of informed traders and the strategic liquidity trader. The strategic liquidity trader might be an institutional trader with a trading horizon of one (or several) days, which is common according to Campbell, Ramadorai, and Schwartz (2009). The informed traders in the model might be considered high-frequency traders with a very short trading horizon and a structural or directional trading strategy (see the HFT strategies as defined by the SEC (2010)). The market maker too might be a high-frequency trader with a market making strategy, which is a common but very different strategy from structural or directional trading. Brogaard, Hendershott, and Riordan (2012) show empirically that HFT trades that demand liquidity are indeed informed, i.e., push prices towards the fundamental value, whereas HFT trades that supply liquidity are adversely selected, like the market maker in our model. The informed traders in the theory of Foucault, Hombert, and Rosu (2012) are high-frequency traders who observe news an instance before the market, and respond very quickly and aggressively. A similar result may also be obtained when there are multiple informed traders who compete with each other.[18] Our model only requires that the trading horizon of the liquidity trader is longer than that of the informed trader, such that the predictable trading at the low frequency is uninformed.

The optimal trading strategy of our liquidity trader is U-shaped, similar to the limit order book models of Obizhaeva and Wang (2005) and Alfonsi, Fruth, and Schied (2010). In their papers, the initial transaction is large and pushes the price above

[17]Also in the multi-period model of Kyle (1985), to the market maker the informed order flow has zero autocorrelation.

[18]With multiple informed traders who observe correlated signals, prices quickly become fully revealing as in Holden and Subrahmanyam (1992).

the fundamental value to attract new limit orders.[19] Then, the future trades are of equal size and exactly consume the new limit orders that are placed. The last trade is larger again, and pushes up the price after which the price dynamics are not important anymore. Our model has this feature too, because the informed traders push back the price towards the fundamental value each period, such that they essentially provide liquidity to the market. In addition, our model has the mechanism that the market maker learns about the trading interest of the uninformed liquidity trader which receives a zero price impact. This second channel increases the curvature of the U-shape pattern. In unreported results, when we set $\sigma_S^2 = 0$ the market maker does not learn about S,[20] and we obtain an optimal strategy that is symmetrically U-shaped similar to Obizhaeva and Wang (2005).

In models like Almgren and Chriss (2000), the optimal strategy is a constant fraction each period (if the price of risk is zero), because this strategy minimizes the quadratic trading costs. In these models, liquidity and prices do not recover from the trading shocks and signalling of uninformed order flow is absent. When the liquidity trader is risk averse, the optimal trade sizes decline over time to reduce the exposure to future price swings (which is also obtained by Huberman and Stanzl (2005)). Bertsimas and Lo (1998) add an AR(1) news process, and find that trading a constant fraction each period is no longer optimal. Compared to Bertsimas and Lo (1998) and Almgren and Chriss (2000), our model can easily accommodate time variation in the price impact by allowing the variance of innovations and noise trade to vary over time. This is important, as intraday data

[19]In Alfonsi, Fruth, and Schied (2010) the trade increases the quoted spread, which attracts new limit orders.

[20]Technically, setting $\sigma_S^2 = 0$ implies that $S = 0$ because $S \sim (0, \sigma_S^2)$. Then, by choosing $S > 0$, we force that the market maker does not learn about S, which leads to a symmetric optimal strategy.

typically reveal U-shape patterns in trading volume and liquidity over the trading day (see also Almgren and Lorenz (2006), where investors learn about the trading interests of others).

3.6 Conclusion

We study trading of a stock in a market with asymmetric information. It is shown that a liquidity motivated trader with a long trading horizon can credibly signal her uninformed trading interest by splitting up her order over time. Then, the predictable component in the order flow stems only from her trades, such that the market can distinguish this component from the total order flow. Accordingly, the predictable component of the order flow does not affect prices, such that the liquidity trader enjoys a lower price impact.

This result is driven by the main assumption that informed traders only have short-lived private information. While this might not be realistic in practise, for our main result to hold it should be sufficient that the liquidity trader has a longer trading horizon than the informed traders. In this case, the predictability in the order flow at frequencies lower than the horizon of the informed traders represents the trading interest of the liquidity trader. This assumption seems very reasonable, given that institutional investors have trading horizons of sometimes several days, while high-frequency traders often have trading horizons of less then a few minutes.

An interesting extension would be to analyze our problem with multiple long-term liquidity traders, and for a portfolio of stocks with correlated liquidity trading.

3.7 Appendix

Conditional Expectation: Equation (3.6) uses the linear projection theorem to calculate the conditional expectation of the fundamental value and the quantity S. In our setting, all variances and covariances are scalars. The derivation is $\begin{bmatrix} P_n \\ E(S|I_n) \end{bmatrix} =$

$$
\begin{aligned}
&= E\begin{bmatrix} v_n|I_n \\ S|I_n \end{bmatrix} \\
&= E\begin{bmatrix} v_n|I_{n-1} \\ S|I_{n-1} \end{bmatrix} + [y_n - E(y_n|I_{n-1})]\,[\sigma_{yy}|I_{n-1}]^{-1}\begin{bmatrix} \sigma_{vy}|I_{n-1} \\ \sigma_{Sy}|I_{n-1} \end{bmatrix}' \\
&= \begin{bmatrix} P_{n-1} \\ E\left(S|I_{n-1}\right) \end{bmatrix} + (y_n - f_n E(S|I_{n-1}))\,Var(y_n|I_{n-1})^{-1}\times \\
&\quad \begin{bmatrix} Cov\left(v_n, y_n|I_{n-1}\right) \\ Cov\left(S, y_n|I_{n-1}\right) \end{bmatrix}' \qquad (3.17) \\
&= \begin{bmatrix} P_{n-1} \\ E\left(S|I_{n-1}\right) \end{bmatrix} + (y_n - f_n E(S|I_{n-1}))\times \\
&\quad \left(\beta_n^2 Var(v_n|I_{n-1}) + \sigma_u^2 + f_n^2 Var(S|I_{n-1})\right)^{-1}\begin{bmatrix} \beta_n Var(v_n|I_{n-1}) \\ f_n Var(S|I_{n-1}) \end{bmatrix}'. \qquad (3.18)
\end{aligned}
$$

Line four uses that $y_n = x_n + u_n + f_n S$ and $x_n = \beta_n(v_n - P_{n-1})$, and that $Var(y_n|I_{n-1}) = \beta_n^2 Var(v_n|I_{n-1}) + \sigma_u^2 + f_n^2 Var(S|I_{n-1})$ and $Cov\left(v_n, y_n|I_{n-1}\right) = Cov(v_n, \beta_n(v_n - P_{n-1}) + u_n + f_n S|I_{n-1}) = \beta_n Var(v_n|I_{n-1})$. The last line consists of two equations that provide the recursive solutions for P_n and $E(S|I_n)$.

The expectation of the fundamental value conditional upon the information set of the strategic liquidity trader I_{n-1}^S is

$$
\begin{aligned}
E(v_n|y_{n-1}, I_{n-2}^S) = E(v_n|I_{n-2}^S) + \left(y_{n-1} - E(y_{n-1}|I_{n-2}^S)\right)\times \\
\frac{Cov(y_{n-1}, v_n|I_{n-2}^S)}{Var(y_{n-1}|I_{n-2}^S)},
\end{aligned}
$$

$$= E(v_{n-1}|I^S_{n-2}) + \big(y_{n-1} - E(y_{n-1}|I^S_{n-2})\big) \times \frac{\beta_{n-1} Var(v_{n-1}|I^S_{n-2})}{\beta^2_{n-1} Var(v_{n-1}|I^S_{n-2}) + \sigma^2_u}, \tag{3.19}$$

where

$$\begin{aligned} E(y_{n-1}|I^S_{n-2}) &= E(\beta_{n-1}(v_{n-1} - P_{n-2}) + f_{n-1}S + u_{n-1}|I^S_{n-2}), \\ &= \beta_{n-1}\big(E(v_{n-1}|I^S_{n-2}) - P_{n-2}\big) + f_{n-1}S. \end{aligned}$$

Conditional Variance: The general rule for conditional variances provides a recursive solution

$$\begin{aligned} Var(Y|X_1, X_2) =& Var(Y|X_2) - Cov(Y, X_1|X_2) \times \\ & Var(X_1|X_2)^{-1} Cov(X_1, Y|X_2). \end{aligned}$$

Note that information set I_n contains y_n, I_{n-1}, then the conditional variance of v_{n+1}, equation (3.9), is

$$\begin{aligned} Var(v_{n+1}|I_n) &= Var(v_{n+1}|y_n, I_{n-1}) \\ &= Var(v_{n+1}|I_{n-1}) - \frac{Cov(v_{n+1}, y_n|I_{n-1})^2}{Var(y_n|I_{n-1})}, \\ &= Var(v_n|I_{n-1}) + Var(\varepsilon_{n+1}) - \\ &\quad \frac{Cov(\varepsilon_{n+1} + v_n, \beta_n(v_n - P_{n-1})|I_{n-1})^2}{Var(\beta_n(v_n - P_{n-1}) + u_n + f_n S|I_{n-1})}, \\ &= Var(v_n|I_{n-1}) + \sigma^2_\varepsilon - \\ &\quad \frac{\beta^2_n Var(v_n|I_{n-1})^2}{\beta^2_n Var(v_n|I_{n-1}) + \sigma^2_u + f^2_n Var(S|I_{n-1})}. \end{aligned} \tag{3.20}$$

The conditional variance of S, equation (3.10), is

$$\begin{aligned} Var(S|I_n) &= Var(S|y_n, I_{n-1}) \\ &= Var(S|I_{n-1}) - \frac{Cov(S, y_n|I_{n-1})^2}{Var(y_n|I_{n-1})} \end{aligned}$$

$$= Var(S|I_{n-1}) - \frac{f_n^2 Var(S|I_{n-1})^2}{\beta_n^2 Var(v_n|I_{n-1}) + \sigma_u^2 + f_n^2 Var(S|I_{n-1})}$$
$$= Var(S|I_{n-1})(1 - \varphi_n f_n). \tag{3.21}$$

The variance of the fundamental value conditional upon the information set of the strategic liquidity trader, equation (3.13), is

$$\begin{aligned} Var(v_{n+1}|I_n^S) &= Var(v_{n+1}|y_n, I_{n-1}^S), \\ &= Var(v_{n+1}|I_{n-1}^S) - \frac{Cov(v_{n+1}, y_n|I_{n-1}^S)^2}{Var(y_n|I_{n-1}^S)}, \\ &= Var(v_n|I_{n-1}^S) + \sigma_\varepsilon^2 - \frac{\beta_n^2 Var(v_n|I_{n-1}^S)^2}{\beta_n^2 Var(v_n|I_{n-1}^S) + \sigma_u^2}. \end{aligned} \tag{3.22}$$

Table (3.1) Numerical example of the static model.
We solve the model numerically, and show the expected parameter values conditional upon the information set of the strategic liquidity trader. We use the following parameter values for the base case (upper panel). The number of trading periods $N = 5$, the strategic liquidity trader must trade the quantity $S = 3$, which has variance $\sigma_S^2 = 5$. The volatility of noise trading $\sigma_u = 1$ and innovations in the fundamental value $\sigma_\varepsilon = 1$. The first cell of column $\Sigma cost$ represents the total costs if the market maker would not learn about the the trading interest of the liquidity trader, i.e., absent signalling. The last cell shows the execution costs when the market maker does learn about the trading interest of the liquidity trader. Each panel shows the effect of a change in one of the parameters values. The solution are calculated before any trade has taken place, i.e., before any realizations of u and ε. Note that the submitted fractions f_n, and the anticipated and realized order flow $E(S|I_n)$ and y_n depend on the realizations of u and ε. The dynamic solution of the model is shown in Table 3.2.

				Base Case: $N=5, S=3, \sigma_S^2=5, \sigma_\varepsilon=1, \sigma_u=1$.						
N	f_n	λ_n	$V_n(v_n)$	$V_n(S)$	$E_{n-1}(S)$	y_n	x_n	P_n	$cost_n$	$\Sigma cost$
1	0.264	0.609	2	4.352	0.000	0.793	0.000	10.483	8.318	31.874
2	0.173	0.665	2	4.102	0.389	0.155	-0.363	10.541	5.461	
3	0.162	0.672	2	3.903	0.418	0.083	-0.403	10.552	5.123	
4	0.174	0.669	2	3.697	0.422	0.109	-0.412	10.575	5.515	
5	0.227	0.648	2	3.401	0.433	0.237	-0.444	10.665	7.267	31.685
				Small quantity liquidity trader: $S=1$						
N	f_n	λ_n	$V_n(v_n)$	$V_n(S)$	$E_{n-1}(S)$	y_n	x_n	P_n	$cost_n$	$\Sigma cost$
1	0.264	0.609	2	4.352	0.000	0.265	0.000	10.161	2.688	10.208
2	0.173	0.665	2	4.102	0.130	0.051	-0.121	10.180	1.756	
3	0.162	0.672	2	3.903	0.139	0.028	-0.134	10.184	1.648	
4	0.174	0.669	2	3.697	0.141	0.037	-0.137	10.192	1.773	
5	0.227	0.648	2	3.401	0.144	0.079	-0.148	10.222	2.322	10.187
				Volatile innovations: $\sigma_\varepsilon=3$						
N	f_n	λ_n	$V_n(v_n)$	$V_n(S)$	$E_{n-1}(S)$	y_n	x_n	P_n	$cost_n$	$\Sigma cost$
1	0.265	1.054	6	4.352	0.000	0.794	0.000	10.836	8.599	33.245
2	0.172	1.152	6	4.102	0.389	0.155	-0.363	10.938	5.666	
3	0.162	1.164	6	3.903	0.418	0.083	-0.403	10.955	5.319	
4	0.174	1.158	6	3.697	0.422	0.109	-0.412	10.997	5.735	
5	0.227	1.122	6	3.401	0.433	0.237	-0.444	11.153	7.599	32.919
				Close deadline: $N=3$						
N	f_n	λ_n	$V_n(v_n)$	$V_n(S)$	$E_{n-1}(S)$	y_n	x_n	P_n	$cost_n$	$\Sigma cost$
1	0.357	0.553	2	4.026	0.000	1.072	0.000	10.592	11.350	32.410
2	0.291	0.610	2	3.513	0.584	0.389	-0.485	10.726	9.376	
3	0.351	0.591	2	2.982	0.680	0.440	-0.614	10.844	11.433	32.159
				Low volatility of quantity liquidity trader: $\sigma_S^2=1$						
N	f_n	λ_n	$V_n(v_n)$	$V_n(S)$	$E_{n-1}(S)$	y_n	x_n	P_n	$cost_n$	$\Sigma cost$
1	0.243	0.687	2	0.972	0.000	0.729	0.000	10.501	7.659	32.013
2	0.178	0.696	2	0.958	0.084	0.175	-0.360	10.612	5.670	
3	0.168	0.698	2	0.945	0.097	0.064	-0.439	10.646	5.351	
4	0.179	0.697	2	0.931	0.101	0.073	-0.463	10.684	5.732	
5	0.232	0.690	2	0.909	0.105	0.202	-0.496	10.806	7.534	31.946

Table (3.2) **Numerical example dynamic model: Rise in value**

We solve the model numerically, and show the expected parameter values conditional upon the information set of the strategic liquidity trader. We use the parameter values of the base case: $N = 5$, demand of strategic liquidity trader $S = 3$, with variance $\sigma_S^2 = 5$. The volatility of noise trading $\sigma_u = 1$ and innovations in the fundamental value $\sigma_\varepsilon = 1$. The current Table shows the case “rise in value”, where the fundamental value increases by 1 unit in periods 1, 2 and 3, and there is zero noise trade. Table 3.3 shows the case “Buying noise traders”, where noise traders buy 1 unit in period 1, 2 and 3, and the innovations in the fundamental value are zero. Each panel shows how the liquidity traders expectations of a particular variable get updated as trading periods passes. Each column shows the information set of the liquidity trader, i.e., the order flow of the periods he has observed. Values in bold are realizations of that variable, rather than expectations. Therefore, the panels show how the strategic traders' expectations are updated according to realizations of the order flow, which depends on ε_n and u_n. The bottom panel shows the realizations of the remaining model parameters. These are reported in columns, as the expectation in future periods simply equals that of the current period.

	Rise in value $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 1, \varepsilon_4 = \varepsilon_5 = 0$				
N	I_0^S	I_1^S	I_2^S	I_3^S	I_4^S
			f_n		
1	**0.264**	**0.264**	**0.264**	**0.264**	**0.264**
2	0.173	**0.168**	**0.168**	**0.168**	**0.168**
3	0.162	0.160	**0.156**	**0.156**	**0.156**
4	0.174	0.172	0.172	**0.168**	**0.168**
5	0.227	0.234	0.239	0.244	**0.244**
			P_n		
1	10.483	**10.983**	**10.983**	**10.983**	**10.983**
2	10.541	11.024	**11.737**	**11.737**	**11.737**
3	10.552	11.034	11.706	**12.555**	**12.555**
4	10.575	11.055	11.716	12.523	**12.936**
5	10.665	11.149	11.804	12.599	11.690
			y_n		
1	0.793	**1.615**	**1.615**	**1.615**	**1.615**
2	0.155	0.198	**1.270**	**1.270**	**1.270**
3	0.083	0.146	0.144	**1.407**	**1.407**
4	0.109	0.174	0.215	0.220	**0.837**
5	0.237	0.332	0.399	0.461	-1.535
			x_n		
1	0.000	**0.822**	**0.822**	**0.822**	**0.822**
2	-0.363	-0.307	**0.764**	**0.764**	**0.764**
3	-0.403	-0.334	-0.323	**0.940**	**0.940**
4	-0.412	-0.343	-0.301	-0.285	**0.333**
5	-0.444	-0.371	-0.319	-0.269	-2.266
			$E_{n-1}(S)$		
1	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
2	0.389	**0.791**	**0.791**	**0.791**	**0.791**
3	0.418	0.812	**1.168**	**1.168**	**1.168**
4	0.422	0.816	1.154	**1.533**	**1.533**
5	0.433	0.826	1.159	1.519	**1.706**
	v_n	$E_{n-1}^S(v_n)$	$V_{n-1}^S(v_n)$	$V_n(S)$	$V_{n-1}(v_n)$
1	11	10.000	2.000	4.352	2
2	12	10.574	1.851	4.102	2
3	13	11.303	1.905	3.903	2
4	13	12.174	1.927	3.697	2
5	13	12.602	1.928	3.401	2

Table (3.3) **Numerical example dynamic model: Buying noise traders.** This Table continues from Table 3.2, but shows the case "Buying noise traders", where noise traders buy 1 unit in period 1, 2 and 3, and the innovations in the fundamental value are zero. Table 3.2 shows the case "rise in value", where the fundamental value increases by 1 unit in periods 1, 2 and 3, and there is zero noise trade.

	Buying noise traders $u_1 = u_2 = u_3 = 1, u_4 = u_5 = 0$				
	I_0^S	I_1^S	I_2^S	I_3^S	I_4^S
N			f_n		
1	**0.264**	**0.264**	**0.264**	**0.264**	**0.264**
2	0.173	**0.168**	**0.168**	**0.168**	**0.168**
3	0.162	0.160	**0.158**	**0.158**	**0.158**
4	0.174	0.172	0.172	**0.171**	**0.171**
5	0.227	0.235	0.237	0.238	**0.238**
			P_n		
1	10.483	**11.092**	**11.092**	**11.092**	**11.092**
2	10.541	11.130	**11.446**	**11.446**	**11.446**
3	10.552	11.139	11.437	**11.598**	**11.598**
4	10.575	11.160	11.453	11.606	**11.012**
5	10.665	11.253	11.543	11.694	10.843
			y_n		
1	0.793	**1.793**	**1.793**	**1.793**	**1.793**
2	0.155	0.209	**0.684**	**0.684**	**0.684**
3	0.083	0.159	0.158	**0.398**	**0.398**
4	0.109	0.188	0.206	0.207	**-0.681**
5	0.237	0.350	0.380	0.391	-0.067
			x_n		
1	0.000	**0.000**	**0.000**	**0.000**	**0.000**
2	-0.363	-0.295	**-0.820**	**-0.820**	**-0.820**
3	-0.403	-0.320	-0.315	**-1.076**	**-1.076**
4	-0.412	-0.329	-0.310	-0.307	**-1.195**
5	-0.444	-0.356	-0.332	-0.323	-0.781
			$E_{n-1}(S)$		
1	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
2	0.389	**0.879**	**0.879**	**0.879**	**0.879**
3	0.418	0.898	**1.056**	**1.056**	**1.056**
4	0.422	0.902	1.052	**1.124**	**1.124**
5	0.433	0.911	1.059	1.127	**0.858**
	v_n	$E_{n-1}^S(v_n)$	$V_{n-1}^S(v_n)$	$V_n(S)$	$V_{n-1}(v_n)$
1	10	10.000	2.000	4.352	2
2	10	10.699	1.851	4.102	2
3	10	11.022	1.905	3.903	2
4	10	11.187	1.927	3.697	2
5	10	10.572	1.928	3.401	2

Table (3.4) **Monte Carlo simulations**

We simulate the model 10.000 times, for random realizations of the innovations in the fundamental value and quantity trader by noise traders. The upper panel shows the simulations for random values for the quantity trader by the strategic liquidity trader S, and the middle panel for the base case where $S = 3$. We show the mean and standard deviation of the profits to each of the trader types summed over the five periods (the informed traders, the market maker (Mm), the noise traders and the liquidity trader (LT)). We also show the liquidity traders profits in the case where the market makers would not learn about the trading interest of the liquidity trader, i.e., absent signalling. The last column shows the absolute quantity traded by the liquidity trader (in number of shares). The bottom panel shows for random S the static and average dynamic strategy of the strategic liquidity trader, i.e., the fractions she submits each period. Since we simulate 10.000 times, the standard deviation of the mean of the variables is simply the standard deviation divided by 100. The parameter values are as described in Table 3.2: $N = 5$, variance of S is $\sigma_S^2 = 5$, the volatility of noise trading $\sigma_u = 1$ and innovations in the fundamental value $\sigma_\varepsilon = 1$.

	Overall profits to the players: *Informed*	*Mm*	*Noise*	*LT Strategic*	*LT Naive*	Diff	Abs(S)
				Random S			
Mean	4.246	-0.085	-3.281	-0.880	-0.982	0.102	1.789
St dev.	7.507	6.605	4.256	3.603	4.707	-1.104	1.343
				$S = 3$			
Mean	4.632	0.211	-3.252	-1.591	-1.748	0.157	3
St dev.	7.842	6.859	4.271	4.530	6.082	-1.553	0

	Strategy f-sequence			
	Period	Static	Dynamic	Diff
	f_1	0.265	0.265	0.000
	f_2	0.173	0.168	-0.005
	f_3	0.162	0.160	-0.002
	f_4	0.174	0.180	0.006
	f_5	0.227	0.228	0.001

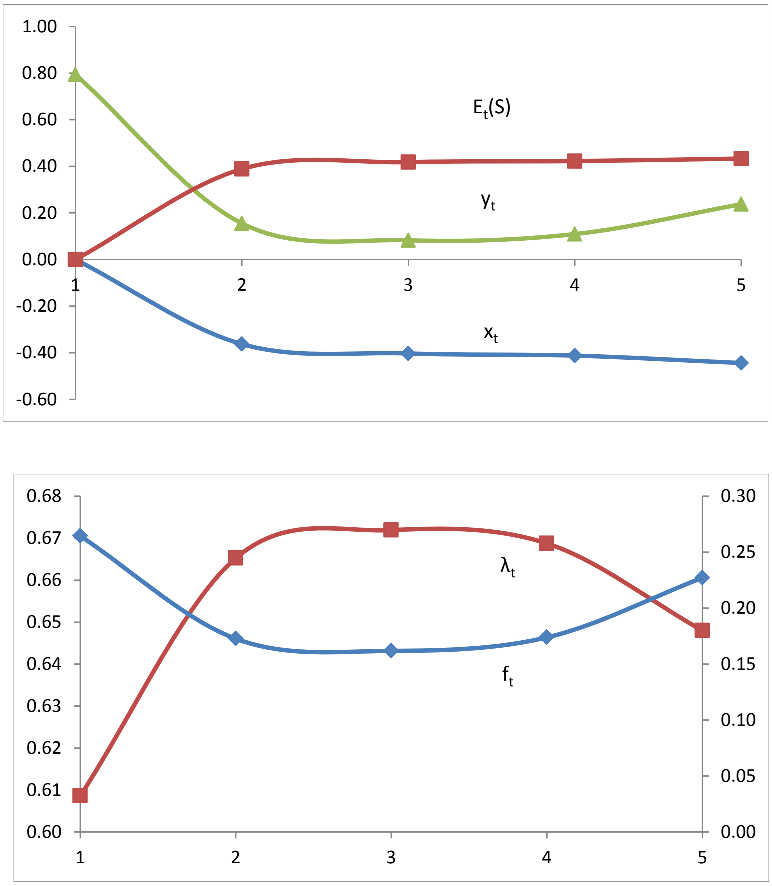

Figure (3.1) The evolution of the parameters over time. The upper figure shows the price impact parameter (λ_t, left axis) and the liquidity traders optimal strategy over time (f_t, right axis). The bottom figure shows the net trading volume (y_t), the volume by the informed trader (x_t) and the market makers expectation of the uninformed trading volume ($E_t(S)$). These values are expectations of the strategic liquidity trader before trading starts, based on numerical results of the base case.

Bibliography

Admati, A., and P. Pfleiderer, 1988, "A Theory of Intraday Patterns: Volume and Price Variability," *The Review of Financial Studies*, 1(1), 3–40.

——— , 1991, "Sunshine trading and financial market equilibrium," *The Review of Financial Studies*, 4(3), 443–481.

Affleck-Graves, J., S. Hedge, and R. Miller, 1994, "Trading Mechanisms and the Components of the Bid-Ask Spread," *Journal of Finance*, 49(4), 1471–1488.

Alfonsi, A., A. Fruth, and A. Schied, 2010, "Optimal execution strategies in limit order books with general shape functions," *Quantitative Finance*, 10(2), 143 – 157.

Almgren, R., and N. Chriss, 1999, "Value Under Liquidation," *Risk*, pp. 61–63.

——— , 2000, "Optimal Execution of Portfolio Transactions," *Journal of Risk*, pp. 5–39.

Almgren, R., and J. Lorenz, 2006, "Bayesian Adaptive Trading with a Daily Cycle," *Working Paper ETH Zurich.*

Andersen, T., T. Bollerslev, F. Diebold, and H. Ebens, 2001, "The distribution of realized stockreturn volatility," *Journal of Financial Economics*, 61, 43–76.

Back, K., and H. Pedersen, 1998, "Long-lived information and intraday patterns," *Journal of Financial Markets*, 1(34), 385 – 402.

Barclay, M., W. Christie, J. Harris, E. Kandel, and P. Schultz, 1999, "Effects of Market Reform on the Trading Costs and Depths of NASDAQ Stocks," *The Journal of Finance*, 54(1), 1–34.

Barclay, M., T. Hendershott, and D. McCormick, 2003, "Competition among trading venues: Information and trading on electronic communications networks," *Journal of Finance*, 58, 2637 – 2666.

Battalio, R., 1997, "Third Market Broker-Dealers: Cost Competitors or Cream Skimmers?," *Journal of Finance*, 52(1), 341–352.

Bennett, P., and L. Wei, 2006, "Market structure, fragmentation, and market quality," *Journal of Financial Markets*, 9(1), 49 – 78.

Bertsimas, D., and A. Lo, 1998, "Optimal control of execution costs," *Journal of Financial Markets*, 1, 1–50.

Bessembinder, H., and H. Kaufman, 1997, "A Comparison of Trade Execution Costs for NYSE and NASDAQ-Listed Stocks," *The Journal of Financial and Quantitative Analysis*, 32, 287–310.

Biais, B., C. Bisière, and C. Spatt, 2010, "Imperfect Competition in Financial Markets: An Empirical Study of Island and Nasdaq," *Management Science*, 56(12), 2237–2250.

Biais, B., T. Foucault, and S. Moinas, 2011, "Equilibrium High Frequency Trading," *Working Paper Toulouse School of Economics.*

Biais, B., D. Martimort, and J. Rochet, 2000, "Competing Mechanisms in a Common Value Environment," *Econometrica*, 68(4), 799–837.

Boehmer, E., 2005, "Dimensions of execution quality: Recent evidence for US equity markets," *Journal of Financial Economics*, 78, 553–582.

Boehmer, E., K. Fong, and J. Wu, 2012, "International Evidence on Algorithmic Trading," *Working Paper EDHEC Business School.*

Boehmer, E., G. Saar, and L. Yu, 2005, "Lifting the Veil: An Analysis of Pre-Trade Transparency at the NYSE," *Journal of Finance*, 60, 783–815.

Brogaard, J., T. Hendershott, and R. Riordan, 2012, "High Frequency Trading and Price Discovery," *Working Paper University of Washington.*

Buti, S., B. Rindi, Y. Wen, and I. Werner, 2011, "Tick Size Regulation, Intermarket Competition and Sub-Penny Trading," *Working Paper Fisher College of Business, Ohio State University.*

Buti, S., B. Rindi, and I. Werner, 2010a, "Diving into Dark pools," *Working Paper Fisher College of Business, Ohio State University.*

——— , 2010b, "Dynamic Dark Pool Trading Strategies in Limit Order Markets," *Working Paper Fisher College of Business, Ohio State University.*

Campbell, J. Y., T. Ramadorai, and A. Schwartz, 2009, "Caught on tape: Institutional trading, stock returns, and earnings announcements," *Journal of Financial Economics*, 92(1), 66–91.

Cantillon, E., and P. Yin, 2010, "Competition between Exchanges: A Research Agenda," *International Journal of Industrial Organization*, 29(3), 329–336.

Cartea, l., and J. Penalva, 2011, "Where is the Value in High Frequency Trading?," *Working Paper Universidad Carlos III de Madrid.*

Chaboud, A., B. Chiquoine, E. Hjalmarsson, and C. Vega, 2009, "Rise of the machines: Algorithmic trading in the foreign exchange market," *Discussion paper Federal Reserve System.*

Chlistalla, M., and M. Lutat, 2011, "The impact of new execution venues on European equity markets liquidity: The case of Chi-X," *Financial Markets and Portfolio Management, forthcoming.*

Chordia, T., R. Roll, and A. Subrahmanyam, 2002, "Order imbalance, liquidity, and market returns," *Journal of Financial Economics*, 65(1), 111 – 130.

———, 2004, "Order imbalance and individual stock returns: Theory and evidence," *Journal of Financial Economics*, 72(1), 485–518.

———, 2011, "Recent trends in trading activity and market quality," *Working Paper UCLA*, 101(2), 243–263.

Conrad, J., K. Johnson, and S. Wahal, 2003, "Institutional Trading and Alternative Trading Systems," *Journal of Financial Economics*, 70, 99–134.

Degryse, H., F. de Jong, and V. van Kervel, 2011, "The Impact of Dark trading and Visible Fragmentation on Market Quality," *Working paper Tilburg University.*

Degryse, H., M. Van Achter, and G. Wuyts, 2009, "Dynamic order submission strategies with competition between a dealer market and a crossing network," *Journal of Financial Economics*, 91, 319–338.

Duffie, D., N. Garleanu, and L. Pedersen, 2005, "Over-The-Counter Markets," *Econometrica*, 73(6), 1815–1847.

Easley, D., N. Kiefer, and M. O'Hara, 1996, "Cream-Skimming or Profit-Sharing? The Curious Role of Purchased Order Flow," *The Journal of Finance*, 51, 811–833.

Easley, D., M. Lopez de Prado, and M. O'Hara, 2012, "Optimal Execution Horizon," *Working Paper Cornell University*.

Ende, B., P. Gomber, and M. Lutat, 2009, "Smart order routing technology in the New European Equity Trading Landscape," *Software Services for E-Business and E-Society: 9th IFIP WG 6.1 Conference*.

Engle, R., R. Ferstenberg, and J. Russell, 2012, "Measuring and Modeling Execution Cost and Risk," *The Journal of Portfolio Management Science*, 38(2), 14–28.

Ferrarini, G. A., and F. Recine, 2006, *Investor Protection In Europe: Corporate Law Making, the MiFID and Beyond*. Oxford University Press.

Fink, J., K. Fink, and J. Weston, 2006, "Competition on the Nasdaq and the growth of electronic communication networks," *Journal of Banking & Finance*, 30, 2537–2559.

Foucault, T., 1999, "Order flow composition and trading costs in a dynamic limit order market," *Journal of Financial Markets*, 2, 99–134.

Foucault, T., J. Hombert, and I. Rosu, 2012, "News Trading and Speed," *Working Paper HEC Paris*.

Foucault, T., O. Kadan, and E. Kandel, 2005, "Limit Order Book as a Market for Liquidity," *Review of Financial Studies*, 18(4), 1171–1217.

——— , 2009, "Liquidity Cycles and Make/Take Fees in Electronic Markets," *EFA 2009 Bergen Meetings Paper*.

Foucault, T., and A. Menkveld, 2008, “Competition for order flow and smart order routing systems,” *Journal of Finance*, 63(1), 119–158.

Glosten, L., 1994, “Is the Electronic Open Limit Order Book Inevitable?,” *Journal of Finance*, 49(4), 1127–1161.

——— , 1998, “Competition, design of exchanges and welfare,” *Working paper Columbia University.*

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

Goettler, R., C. Parlour, and U. Rajan, 2005, “Equilibrium in a Dynamic Limit Order Market,” *Journal of Finance*, 60(5), 2149–2192.

——— , 2009, “Informed traders and limit order markets,” *Journal of Financial Economics*, 93, 67–87.

Goldstein, M., and K. Kavajecz, 2000, “Eights, sixteenths, and market depth: changes in the tick size and liquidity provision on the NYSE,” *Journal of Financial Economics*, 56, 125–149.

Gomber, P., and M. Gsell, 2006, “Catching Up with Technology - The Impact of Regulatory Changes on ECNs/MTFs and the Trading Venue Landscape in Europe,” *Competition and Regulation in Network Industries,*, 1(4), 535–557.

Gomber, P., and A. Pierron, 2010, “MiFID: Spirit and Reality of a European Financial Markets Directive,” *Report Published by Celent.*

Gomber, P., U. Schweickert, and E. Theissen, 2004, “Zooming in on Liquidity,” *Working Paper Series.*

Gresse, C., 2006, "The Effect of Crossing-Network Trading on Dealer Market's Bid-Ask Spread," *European Financial Management*, 12(2), 143–160.

——— , 2010, "Multi-Market Trading and Market Quality," *Working Paper Paris-Dauphine.*

Griffin, J. M., J. H. Harris, and S. Topaloglu, 2003, "The Dynamics of Institutional and Individual Trading," *The Journal of Finance*, 58(6), 2285–2320.

Guilbaud, F., and H. Pham, 2011, "Optimal high frequency trading with limit and market orders," *Working paper Universit Paris.*

Hamilton, J. D., 1994, *Time Series Analysis.* Princeton University Press.

Harris, L., 1993, "Consolidation, Fragmentation, Segmentation and Regulation.," *Financial Markets, Institutions & Instruments*, 5, 1–28.

Hasbrouck, J., 2007, *Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading.* Oxford University Press.

Hasbrouck, J., and G. Saar, 2009, "Technology and Liquidity Provision: The Blurring of Traditional Definitions," *Journal of Financial Markets*, 12, 143–172.

——— , 2011, "Low-latency trading," *Working paper Stern school of business.*

Hendershott, T., and C. Jones, 2005, "Island Goes Dark: Transparency, Fragmentation, and Regulation," *The Review of Financial Studies*, 18(3), 743–793.

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

Hendershott, T., and H. Mendelson, 2000, "Crossing networks and dealer markets: competition and performance," *Journal of Finance*, 55(5), 20712115.

Hengelbrock, J., and E. Theissen, 2010, "Fourteen at one blow: The market entry of Turquoise," *Working Paper University of Bonn, University of Mannheim.*

Heston, S., R. Korajczyk, and R. Sadka, 2010, "Intraday Patterns in the Cross-section of Stock Returns," *The Journal of Finance*, 65(4), 1369–1407.

Ho, T., and H. R. Stoll, 1981, "Optimal dealer pricing under transactions and return uncertainty," *Journal of Financial Economics*, 9(1), 47–73.

Hoffmann, P., 2012, "A dynamic limit order market with fast and slow traders," *Working Paper European Central Bank.*

Holden, C. W., and A. Subrahmanyam, 1992, "Long-Lived Private Information and Imperfect Competition," *The Journal of Finance*, 47(1), pp. 247–270.

Huang, R., and H. Stoll, 2001, "Tick Size, Bid-Ask Spreads and Market Structure," *Journal of Financial and Quantitative Analysis*, 36, 503–522.

Huberman, G., and W. Stanzl, 2005, "Optimal Liquidity Trading," *Review of Finance*, 9(2), 165–200.

Huddart, S., J. S. Hughes, and C. B. Levine, 2001, "Public Disclosure and Dissimulation of Insider Trades," *Econometrica*, 69(3), 665–681.

Huddart, S., J. S. Hughes, and M. Williams, 2010, "PRE-ANNOUNCEMENT OF INSIDERS TRADES," *Working Paper Pennsylvania State University.*

Irvine, P., G. Benston, and E. Kandel, 2000, "Liquidity beyond the inside spread: measuring and Using Information in the Limit Order Book," *Working Paper, Emory University and Hebrew University.*

Jovanovic, B., and A. Menkveld, 2011, "Middlemen in Limit Order Markets," *Working paper New York University.*

Kissell, R., M. Glantz, and R. Malamut, 2003, *Optimal Trading Strategies: Quantitative Approaches for Managing Market Impact and Trading Risk.* American Management Association.

Kyle, A., 1985, "Continuous Auctions and Insider Trading," *Econometrica*, 53(6), 1315–1335.

Maglaras, C., C. Moallemi, and H. Zheng, 2012, "Optimal Order Routing in a Fragmented Market," *Working Paper Columbia University.*

Massoud, N., and D. Bernhardt, 1999, "Stock market dynamics with rational liquidity traders," *Journal of Financial Markets*, 2(4), 359 – 389.

McInish, T., and J. Upson, 2011, "Strategic liquidity supply in a market with fast and slow traders," *Working Paper University of Texas, El Paso.*

Mendelson, H., and T. I. Tunca, 2004, "Strategic Trading, Liquidity, and Information Acquisition," *Review of Financial Studies*, 17(2), 295–337.

Menkveld, A., 2011, "High Frequency Trading and The New-Market Makers," *Working paper Free University Amsterdam.*

Obizhaeva, A., and J. Wang, 2005, "Optimal Trading Strategy and Supply/Demand Dynamics," *NBER Working Paper.*

O'Hara, M., and M. Ye, 2011, "Is Market Fragmentation Harming Market Quality?," *Journal of Financial Economics*, 100(3), 459–474.

Pagano, M., 1989a, "Endogenous Market Thinness and Stock Price Volatility," *The Review of Economic Studies*, 56, 269–287.

———, 1989b, "Trading volume and asset liquidity," *Quarterly Journal of Economics*, 104, 25–74.

Pagnotta, E., and T. Philippon, 2012, "Competing on Speed," *Working Paper New York University.*

Petrella, G., 2009, "MiFID, Reg NMS and Competition Across Trading Venues in Europe and United States," *Working Paper University of Milan.*

Röell, A., 1990, "Dual-capacity trading and the quality of the market," *Journal of Financial Intermediation*, 1(2), 105 – 124.

Rosu, I., 2009, "A Dynamic Model of the Limit Order Book," *The Review of Financial Studies*, 22(11), 4601–4641.

———, 2010, "Liquidity and Information in Order Driven Markets," *Working Paper University of Chicago.*

Sandås, P., 2001, "Adverse Selection and Competitive Market Making: Empirical Evidence from a Limit Order Market," *The Review of Financial Studies*, 14(3), 705–734.

SEC, 2010, "Concept release on equity market structure," *Release No. 34-61358; File No. S7-02-10.*

Spiegel, M., and A. Subrahmanyam, 1992, "Informed speculation and hedging in a noncompetitive securities market," *Review of Financial Studies*, 5(2), 307–329.

Spiegel, M., and A. Subrahmanyam, 1995, "On Intraday Risk Premia," *The Journal of Finance*, 50(1), pp. 319–339.

Stock, J., and M. Yogo, 2002, "Testing for Weak Instruments in Linear IV Regression," *NBER Working Paper No. T0284*.

Stoll, H., 2003, *Market Microstructure*. Handbook of the Economics of Finance.

——— , 2006, "Electronic Trading in Stock Markets," *Journal of Economic Perspectives*, 20(1), 153–174.

Storkenmaier, A., and M. Wagener, 2011, "European Market Integrity: Regulating Equity Trading in Fragmented Markets," *Working paper Karlsruhe Institute of Technology*.

Subrahmanyam, A., 1995, "Circuit Breakers and Market Volatility: A Theoretical Perspective," *Journal of Finance*, 49(1), 237–254.

Van Achter, M., 2008, "A Dynamic Limit Order Market with Diversity in Trading Horizons," *Working Paper Rotterdam School of Management*.

Vayanos, D., 1999, "Strategic Trading and Welfare in a Dynamic Market," *The Review of Economic Studies*, 66(2), 219–254.

Weaver, D., 2011, "Off-Exchange Reporting and Market Quality in a Fragmented Market Structure," *Working Paper Rutgers Business School*.

Weston, J., 2000, "Competition on the Nasdaq and the Impact of Recent Market Reforms," *Journal of Finance*, 55(6), 2565–2598.

——— , 2002, "Electronic Communication Networks and Liquidity on the Nasdaq," *Journal of Financial Services Research*, 22, 125–139.

Zhu, H., 2011, "Do Dark Pools Harm Price Discovery?," *Working Paper Stanford University.*