ESSAYS ON INTRADAY LIQUIDITY AND PRICE MOVEMENTS

By
PAOLO MAZZA

SUBMITTED TO
UNIVERSITE CATHOLIQUE DE LOUVAIN
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTEUR EN SCIENCES ECONOMIQUES ET DE GESTION
AT LOUVAIN SCHOOL OF MANAGEMENT

DOCTORAL COMMITTEE:
Prof. Claude PECHEUX (Université catholique de Louvain, Belgium), chairman
Prof. Mikael PETITJEAN (Université catholique de Louvain, Belgium), supervisor
Prof. Catherine D'HONDT (Université catholique de Louvain, Belgium)
Prof. Gunther WUYTS (KULeuven, Belgium)
Prof. Carole GRESSE (Université Paris-Dauphine, France)
Prof. Ian MARSH (City University of London, United Kingdom)

MONS, BELGIUM
OCTOBER 2013

CONTENTS

Acknowledgements **13**

1 Introduction **17**

1.1 Where it all began... 17

1.2 Market microstructure and liquidity 21

1.3 A four-essay doctoral research 26

1.3.1 Essay I 27

1.3.2 Essay II 30

1.3.3 Essay III 31

1.3.4 Essay IV 33

2 Price Dynamics and Market Liquidity: An Intraday Event study on Euronext **36**

2.1 Introduction 36

2.2 Literature review 39

2.2.1 Price dynamics 39

2.2.2 Price dynamics and liquidity 42

2.3 Data 46

2.3.1 Data and Sample 46

2.3.2 Liquidity measures 49

2.4 Methodology 53

2.5 Results 55

2.5.1 Overall picture 55

2.5.2 Details 56

2.6 Detecting the Presence of Informed Trading 63

2.7 Granger causality 65

2.8 Robustness checks 69

2.8.1 Reversal configurations 72

2.8.2 Changing the time interval 74

2.9 Conclusion 76

2.10 Appendix 80

3 Intraday liquidity, price dynamics and uncertainty in cap-based portfolios 81

3.1 Introduction 81

3.2 Data and methodology 86

3.2.1 Liquidity proxies 86

3.2.2 Price movement variables 89

3.2.3 Model specification 94

3.3 Empirical analysis 95

3.3.1 General panel regressions 99

3.3.2 Controlling for volatility 106

3.3.3 Robust and median regressions 110

3.3.4 Addressing causality and endogeneity issues 115

3.3.5 Time intervals 122

3.4 Conclusion 125

4 Rethinking zero returns in the liquidity puzzle of a limit order market 128

4.1 Introduction 128

4.2 Zero returns and (il)liquidity 131

4.2.1 The zero return measure 131

4.2.2 Zero returns, informed trading and liquidity 133

4.3 Data and liquidity proxies 137

4.3.1 Data 137

4.3.2 Liquidity proxies 139

4.4 Event study 144

4.4.1 Methodology 144

4.4.2 Results 145

4.5 Logit regressions 149

4.5.1 Methodology 149

4.5.2 Results 151

4.5.3 Robustness checks 153

4.5.4 Daily zero returns 155

4.6 Conclusion 156

5 Do Japanese candlesticks help solve the trader's dilemma? 159

5.1 Introduction 159

5.2 Japanese candlesticks and liquidity 162

5.3 Data 167

5.3.1 Sample 167

5.3.2 Sequences of orders 169

5.3.3 Transaction costs measures 171

5.4 Panel regressions 173

5.4.1 Methodology 173

5.4.2 Results 177

5.5 Order processing simulation 182

5.5.1 Methodology 182

5.5.2 Results 184

5.6 Conclusion 185

6 Conclusion **187**

LIST OF FIGURES

2.1 HLOC dynamics in charts 40

2.2 Doji 41

2.3 Consensus configurations by interval before filters 48

2.4 Abnormal liquidity around consensus configurations 1/2 60

2.5 Abnormal liquidity around consensus configurations 2/2 61

2.6 Abnormal number of trades around consensus configurations 62

2.7 Hammer-like structures 70

2.8 Consensus and reversal structures 80

3.1 High, low, opening and closing (HLOC) price dynamics 83

3.2 HLOC price movements and relative spread 99

4.1 Zero returns by interval 139

4.2 Abnormal liquidity around zero returns 146

4.3 Abnormal trading activity around zero returns 148

5.1 Candlestick formation process 162

5.2 Doji 164

5.3 Hammer-like structures 165

5.4 Doji by interval 168

5.5 Candlestick-based strategy 182

5.6 Random strategy 183

LIST OF TABLES

2.1 Consensus configurations 42

2.2 Events count 49

2.3 Liquidity proxies and trading activity measures 50

2.4 Non-parametric tests on the relationship between PIN and the number of consensus configurations per day. 64

2.5 Granger causality tests 68

2.6 Reversal configurations 71

2.7 Liquidity dynamics around reversal configurations 73

2.8 Events count for 30 and 60-minutes price series 75

3.1 Descriptive statistics 90

3.2 Correlation Matrix - levels 95

3.3 Correlation Matrix - first differences 95

3.4 A first look at liquidity and HLOC price movements 98

3.5 Regressions with clustered standard errors 101

3.6 Controlling for volatility 108

3.7 Robust regressions with realized volatility 111

3.8 Median regressions with realized volatility 112

3.9 Multivariate VARX(4,2) model: Small cap portfolio 118

3.10 Multivariate VARX(4,2) model: Mid cap portfolio 119

3.11 Multivariate VARX(4,2) model: Large cap portfolio 120

3.12 Robustness check : 10-minute interval regressions with clustered standard errors 123

3.13 Robustness check : 20-minute interval regressions with clustered standard errors 124

4.1 Descriptive statistics 142

4.2 Correlation matrix 150

4.3 Conditional logit regressions 152

4.4 Conditional logit regressions: Robustness checks 154

4.5 Conditional logit regressions: Daily data 155

5.1 Number of signals 169

5.2 Sequences - Descriptive statistics 170

5.3 Market impact and market timing - Hammer-like configurations 178

5.4 Market timing - Doji 179

5.5 Market timing - Dragonfly and Gravestone Doji 179

5.6 Market impact - Doji 180

5.7 Market impact - Dragonfly and Gravestone Doji 181

5.8 Simulation 184

ACKNOWLEDGEMENTS

This doctoral thesis was conducted from September 2010 to September 2013 at the Center for Studies in Asset Management (CESAM) of the Louvain School of Management. A doctoral research is such a long process that it necessarily involves many people. When a soccer player scores, his name is the only one that appears on the board even if the goal results from the hard work done by a whole team. The following paragraphs express all my gratitude to that "team".

I first would like to express my deep gratitude to Mikael Petitjean, my supervisor, who offered me the unique opportunity to conduct this research when I was employed by an asset management company. From the beginning of the doctoral process to its end, he has always provided me with good advice, comments and a huge motivation. His availability, his research experience and econometrics knowledge have always been extremely precious to me.

I would like to thank all the members of my dissertation committee. I am thankful to Catherine D'Hondt for her availability, for the mountains of comments she made about the papers of this thesis, and for the numerous autographs. I also would like to thank Gunther Wuyts for his numerous comments and his thorough knowledge of the microstructure of financial markets. I am so thankful to them for having accepted to be members of my supervisory panel. I am also very much indebted to Carole Gresse and Ian Marsh who agreed to

be involved in this committee.

I greatly appreciate the help of the Finance Department of the Louvain School of Management. I would like to thank Rudy De Winne for his huge research expertise, his SAS and LaTEX knowledge, his sound and wise advice as well as his availability. I am much indebted to Jean-Yves Filbien and Christophe Majois for always dedicating time to the reading of my papers, despite a busy schedule, as well as Alexis Cellier for his SAS macros. I also would like to thank all my colleagues involved in the doctoral process: Anthony Bellofatto, Saifeddine Ben Hadj, Benoit Detollenaere and Thibaut Moyaert. Our discussions have always been very constructive and improved our respective works. I also would like to thank Matthieu Duvinage for his help and most particularly when I was trying to tame the Matlab software.

I would like to express all my gratitude to all the people who made some comments about my work, during conference, discussions or by email: Jan Annaert, Sophie Béreau, Benoît Detollenaere, Laurent Deville, Catherine D'Hondt, Rudy De Winne, Jean-Yves Filbien, Thierry Foucault, Carole Gresse, Xing Han, Bruce Lehmann, Wael Louhichi, Christophe Majois, Thibaut Moyaert, Marco Pagano, Mikael Petitjean, Florian Resch, Patrick Roger, Marko Savor, Elvira Sojli, Frederick Van Gysegem, Kathleen Walsh, Gunther Wuyts as well as four anonymous referees.

I would like to thank NYSE Euronext for providing the data used in this study and all the people that worked on the database before.

I also would like to thank the institution and all the people who work on the Mons Campus of the Louvain School of Management for the smiles, the jokes and everything that contributes to the unique atmosphere of this campus. I also would like to say my special thanks to Vincianne Claus and Sandrine Delhaye who do a great job in the management of the doctoral process. They have always been available for any question regarding the doctoral program.

This research has been funded by the “Action de recherche Concertées” subsidies from the Communauté française de Belgique granted by the Académie universitaire “Louvain”. I am very much indebted to this funding that allowed me to do research and to present my results in conferences.

I would like to thank all my friends and all the people who have shown some interest for this work. I hope they now know it is not always appropriate to ask a PhD student about the completion of his doctoral thesis.

I also would like to express all my gratitude to all my family and most specially to my parents for giving me the opportunity to go to university and supporting me in every decision I made. I hope that they now understand what I have been working on during the last three years and that it is not just about “selling Japanese candlesticks”. I think I will have to translate this paragraph. Maybe is it a sufficient motivation for them to learn the language of Shakespeare!

Last but not least, I would like to thank Sarah Brohé who has always been present for the best and worst moments. Both of us being on a doctoral process was an inexhaustible source of mutual motivation. You are the person to whom I am the most indebted. Thank you for reading all my productions, your understanding, your tolerance, and your everlasting support. Now, your turn!

CHAPTER

ONE

INTRODUCTION

"All truths are easy to understand once they are discovered; the point is to discover them"
Galileo Galilei

1.1 Where it all began...

Financial markets have become of paramount importance in today's economic environment. We can hardly imagine living without them even though their origin cannot be precisely determined. The idea of a barter economy in which our old ancestors were trading stones against food, has become quite inconceivable across time. Barter was slowly but surely replaced by money which in turn led to the creation of financial markets.

As outlined by Ferguson (2008), "*Money, it is conventional to argue, is a medium of exchange, which has the advantage of eliminating inefficiencies of barter; a unit of account,*

which facilitates valuation and calculation; and a store of value, which allows economic transactions to be conducted over long periods as well as geographical distances. To perform all these functions optimally, money has to be available, affordable, durable, fungible, portable and reliable".[1] Money appeared as a store of value and a better valuation were required. The Incas for instance did not need any kind of money, since they were not willing to convert precious metals into storing value and labor was their unique unit of value. They also appreciated the aesthetic characteristics of rare metals. Spanish conquistadors however had an insatiable lust for silver and gold which they got their fill of in the Cerro Rico mountain, the "rich hill", in the city of Potosí in Peru.[2] Conquistadors' lust for silver ore cost the life of many Peruvian miners that were forced to work at horrendous conditions. Spaniards rapidly needed more workforce and found it through the importation of African slaves. Earlier, the Aztecs used cacao beans for money, even if, in this case, cacao was rather a way to value goods than a purchase mean, as outlined by Weatherford (2009). More interestingly, he argues that salt was the commodity money of China, North Africa and the Mediterranean, leading to the word *salario* from which the English word *salary* is derived.

The mountain of silver extracted from Potosí served to create the world's first global currency: The Spanish piece of eight. After the creation of this conventional mean of exchange, people began to understand its limits: money is only worth what someone else agrees to pay. However, Spaniards believed that by extracting more and more silver ore, they were becoming richer. They thought it was a blessing to find such a huge quantity of precious metals but it was actually a curse: the value of precious metals went down. They were in fact experiencing the first inflation crisis.

Nevertheless, money has been borrowed and lent for ages. Historians found traces of the first moneylenders back to around 3000 BC in ancient Mesopotamia. Mesopotamians used clay tablets where they indicated their promises to give goods to the holder of the clay tablets. These clay tablets were intrinsically worth nothing but are among the first traces of

[1] Ferguson (2008), pp. 23-24.
[2] The expression "*Valer un Potosí*" has remained in Spain and is translated as "to be worth a fortune".

debt and credit that allowed payment to be transferable. Afterwards, religions were entrusted with a dramatic power that enable them to infer changes in individuals' way of living. For instance, lending at interest was a sin liable to excommunication for Christians. In such an environment, we can hardly imagine how banks would have been created. Nevertheless, Shakespeare's play *The Merchant of Venice* presents an anecdote where a young man visits a Jewish moneylender. Jews were only forbidden to lend upon usury to their brothers, following a verse from the Old Testament. They were in fact able to provide commercial credit to merchants in the city of Venice, where they lived in ghettos.

Near the end of the Middle Ages, finance in Italy was at its apogee, most particularly in Venice, Florence and Pisa. The Medici contributed to this financial development. They circumvented the financial rules introduced by the Church that stipulated that no interests may be earned on any money lending. They were simple foreign exchange traders who remunerate the moneylenders by giving participation to the profits of the firm, which are not considered as interests. At that time, there were many different systems of coinage and trading was difficult. The Medici profitted from the situation by converting the different currencies with each other. Since Venice was an international port, they rapidly made a fortune. Both Jews and the Medici have contributed to the birth of banks by creating a solid banking system that Northern countries rapidly took in example.

The first stock exchange was created at about the same moment, during the fifteenth century, in Antwerp, even though companies, i.e. an association of people who pool welfare together to achieve a given purpose, have been existing for a long time. In 1460, the first business center was created in the port of Antwerp. The first objective was to bring together merchants from all parts of the world to trade goods, contracts, as well as both government and individual debt. The first company that was financed by the selling of shares to the public was the Dutch East India Company in 1602. Shares were easily transferable since they were materialized by papers. During the first part of the seventeenth century, another type of product knew an important development: tulip bulbs. The prices rapidly rose as many

investors wanted to profit from the opportunity to make money. However, the prices dropped afterwards, creating the first crash.

Today's stock exchanges began to appear at the end of the eighteenth century, the London Stock Exchange (LSE) in 1773 and the New York Stock Exchange, 19 years later. Since then, many stock exchanges have been created with several types of securities bringing a new complex organization: Different instruments, different marketplaces, several ways to organize them, etc. Mergers, acquisitions and crises have constituted turning points in the history of financial markets. For instance, October 19, 1987's crisis has tremendously affected the way academic research understands the markets. The *Black Monday*, an indirect reference to the *Black Thursday* of the 1929 crisis, occurred when a steep rise in the treasury bonds' interest rates generated a sharp drop in stocks prices. On that Monday, the Dow Jones Industrial Average felt by more than 22%. The causes have not been clearly established but program trading, i.e. the first trading algorithms, is often cited to be one of the main drivers. Just after October 1987, market microstructure has raised a huge interest for research. For instance, Blume et al. (1989) analyze the behavior of stock prices during the financial crisis of October 1987 and identify a strong relation between these movements and order imbalances. With the advent of high frequency trading as well as algorithmic trading, more and more researchers have been working on microstructure to understand what happens when algorithmic traders are active in the market. As an example, the Flash Crash of May 6, 2010 has been a very hot topic and many papers have been discussing the effect of high frequency trading on the microstructure of financial markets. Among them, Easley et al. (2010) address the impact of the order toxicity that resulted from the Flash Crash on liquidity using their VPIN measure. Hendershott et al. (2011) also investigate whether or not algorithmic trading improves market quality. They find that liquidity is positively influenced by the presence of algorithmic traders and that the quotes are much more informative since their apparition. Among many other examples, the Flash Crash clearly stresses the dire need to further investigate the microstructure of financial markets.

1.2 Market microstructure and liquidity

Market microstructure is the branch of finance that investigates how markets are organized and how trades occur.[1] Depending on the type of instruments, this organization may be dramatically different. For instance, the microstructure of the foreign exchange is totally distinctive from the microstructure of the stock market. The structure of a market influences the way traders behave and information is conveyed, which in turn leads to price formation and trade determination that imply transaction costs and market frictions. Since it has to deal with a huge number of players, whose behavior is assumed to be rational, understanding the functioning of markets and the trading process is not an easy task.

A market is a place, physical or electronic, where buyers and sellers meet to trade instruments. As traders do not arrive at the same moment in the market, the market must deal with this asynchronous arrival. Some particular types of traders, i.e. dealers and brokers, help organize traders' meetings. Brokers find counterparty for their clients. They do not personally trade; they only help their clients fill their requirements. Dealers offer counterparty for their clients and propose both selling and buying prices. The dealers then try to make profit by taking the opposite trade on the market. As such, they are most interested in selling high and buying low. The difference between the prices they provide for buying and selling, respectively named bid and ask prices, constitutes their profit, i.e. the bid-ask spread. Dealer's biggest fear is to trade with a better informed trader who knows the fundamental value of the stock, towards which prices tend to move. The rapid evolution of prices towards their fundamental value prevents dealers to earn the spread. Dealers compensate this risk by offering worse prices to their uninformed clients. They do so in widening the bid-ask spread. This is not only applicable to dealers: every trader fears to trade with informed traders. The information asymmetries are of utmost importance in market microstructure research and several papers thoroughly investigate this issue. This asymmetry is often characterized by the fact that domestic investors have better information than foreign investors, what is called

[1] Several books present this discipline of finance. Among them, Harris (2003) is the most accessible for both practitioners and academics.

home bias (Kang and Stulz, 1997; Brennan and Cao, 1997; Grinblatt and Keloharju, 2001; Chan et al., 2008).

Each trader has a different motivation to trade. Some of them are just gambling as an entertainment, or as a way to burn the household's money. Investors trade to move some welfare from the present to the future while borrowers do the opposite. Hedgers trade to protect themselves from a risk that they run. Harris (2003) distinguishes three main styles of traders. Profit-motivated traders rationally expect to make profit with their trades, e.g. dealers. Utilitarian traders wish to increase their utility, in addition of the potential profits they make from trading. Gamblers, investors, borrowers, hedgers enter this category as their primary purpose is not the profit. Finally, futile traders await to make profit but trade too irrationally to do so. They do not own a particular information set that would allow them to make the best choices.

Another way to classify traders relates to the information they own. Informed traders are profit-motivated traders who own sufficient information to precisely know what is the fundamental value of the stock. The fundamental value is the price which all traders would agree upon, should they own and understand all information regarding the instrument. They are motivated to trade when the current trading price is different from the fundamental value that they estimate. Their estimations vary depending on the private information they are disposing on and the publicly available information. By trading, they force the market price to move towards its fundamental value. Uninformed traders are not in possession of particular information and trade for different purposes. Uninformed traders' activity makes prices less efficient and moves them away from their fundamental value. All uninformed traders are therefore noise traders.

The existence of both informed and uninformed traders generates transactions. As outlined by Harris (2003), symmetrically informed traders do not trade with each other, since they would trade in the same direction. A transaction is the result of a mutual search: sell-

ers are looking for buyers and conversely, buyers for sellers. This search is based on three interdependent criteria: price, quantity and patience. Sellers want to get the highest price for the security while buyers prefer to trade at a low price. They are not only seeking for the best price, they also want their desired quantity to be effectively transacted. If the size is too large, their order will require several traders to be filled. The transaction price also depends on the trader's patience. A patient trader will trade at better prices, since he is willing to wait for better conditions that facilitate trading. An impatient trader needs direct trade completion and is therefore willing to pay to get her orders executed most rapidly. Traders who submit limit orders offer the other traders the ability to trade inexpensively. In doing so, they give other traders free options to meet their needs. They therefore provide liquidity to the market. Those traders are often patient traders that are not willing to pay the spread to trade immediately. Impatient traders submit market orders that directly pick pending limit orders, using the liquidity usually provided by patient traders. Limit orders, however, imply a non-execution risk, if nobody is willing to use this option. They may also bear the risk of a possible adverse price evolution which triggers trades that are not profitable anymore. They still can cancel their order and resubmit it at better conditions. Since trading is a zero-sum game, from a static point of view, the profits made by informed traders are the losses that uninformed traders incur.

Limit and market orders submission is related to liquidity which is, along with information asymmetry, one of the most discussed issues in market microstructure research. As outlined by Harris (2003), liquidity denotes "*the ability to trade large size quickly, at low cost, when you want to trade*".[1] This definition suggests that liquidity is multidimensional: immediacy ("*quickly*"), width ("*low cost*"), and depth ("*large size*").[2] He also identifies another dimension: resiliency, which is related to the time to recovery from a shock.

Liquidity measurement is far from being simple and many research studies have addressed the different market structures and their impact on liquidity supply, which affects

[1] Harris (2003), pp. 394.

[2] Depth refers to the quantities that are pending in the order book.

the execution process. Two main types of markets coexist: quote-driven and order-driven markets. In quote-driven markets, also called dealer(ship) markets, the dealer is the counterparty in every trade. Traders may not directly trade with each other. For instance, foreign exchange and bond markets are mostly dealer markets, as well as the NASDAQ and the London Stock Exchange. In dealership markets, the liquidity is provided by the market maker. The dealer proposes liquidity when the probability to trade with an uninformed trader is high. When trading with an informed trader, the dealer loses on average and therefore enlarges the spread to cover this risk. The inventory issue is also very important for the dealers, since they need capital to finance it.

In order-driven markets, also called limit order markets, there is no determined dealer so that buyers and sellers trade with each other without intermediation. However, some dealers can actually trade on the stock. Trading is organized by a plethora of rules, e.g. order precedence rules which determine the orders with priority. These rules are of paramount importance, since they tremendously affect the way trading occurs and have a direct incidence on market liquidity. In these markets, limit orders supply liquidity, whatever the trader who places the order. Marketable orders, i.e. market orders or limit orders that generate a trade, consume liquidity that may be displayed or hidden. Hidden orders, also called iceberg orders, propose quantities that do not appear in the limit order book but which are available to trade. Order-driven markets are therefore very flexible.

In an attempt to describe the functioning of both dealership and limit order markets, academics have set up different theoretical models. Kyle (1985) is among the most cited studies in market microstructure and proposes a model of price adjustment with asymmetric information. Glosten and Milgrom (1985) also model dealers' behavior when they face privately informed and uninformed traders' demands. For limit order markets, best known models include Parlour (1998), Foucault (1999), and Foucault et al. (2005). Modeling limit order markets is more difficult than modeling dealer markets, as it is done in Kyle (1985) and Glosten and Milgrom (1985), since agents are very different and have many possible

reactions.

Given the multidimensionality of liquidity that we previously mentioned, an accurate estimation of liquidity has always been a particular challenge. It is however of particular interest since the state of liquidity determines transaction costs. Lower liquidity implies higher implicit transaction costs. Implicit transactions costs are the part of the costs that a trader may not accurately estimate before trading. It comprises the relative spread, the market impact, the price appreciation, the timing risk and other opportunity costs, including missed trade opportunity costs. An exhaustive definition of each transaction costs' component is beyond the purpose of this research.[1] Since there are practical consequences of poorly estimating liquidity, research has recently focused on efficient liquidity proxies. For instance, Goyenko et al. (2009) and Aitken and Comerton-Forde (2003) investigate whether existing liquidity measures are appropriate. The analysis is not straightforward, since there is a huge number of proxies that belong to different categories and each new paper proposes refinement to existing measures. Liquidity proxies broadly fall into two main categories. On the one hand, trade-based proxies are computed with trade data. They are typically related to trading volumes, market returns, trade imbalances, trade frequency, or a combination of these variables, such as in the case of Amihud (2002)'s illiquidity ratio. One of the main advantages is that data which they are based on is widely available even for emerging markets or for old historical datasets. On the other hand, order book-based proxies are computed with order data. They are often calculated with the order book and include spreads, depth, height, dispersion, slope, order imbalance, etc. These proxies are more directly related to the liquidity dimensions and are therefore more in line with transaction costs' expectations. It is however not easy to build such a reliable dataset.

[1] The interested reader should refer to Chan and Lakonishok (1995), Keim and Madhavan (1997), or Kissell et al. (2004).

1.3 A four-essay doctoral research

This four-essay doctoral research study contributes to the literature on liquidity measurement by investigating the relationships between price dynamics and liquidity, measured in the order book. This relationship has been documented in the existing literature but price dynamics are characterized by variations in the closing prices only. For instance, Blume et al. (1989) identify strong and positive correlation between order imbalances and price movements. They enlighten a cascade effect of order imbalance on stock prices during the stock market crisis of October 19 and 20, 1987 and show that the crisis was a direct consequence of the inability of the current market structure to cope with large selling orders. Using U.S. market data, Chordia et al. (2001) empirically find that liquidity and trading activity are influenced by both market returns and volatility. Their results demonstrate that effective and quoted spreads increase dramatically in down markets. Chordia et al. (2002) further highlight a significant impact of daily order imbalance and state that it is unwise to trade when the order book is highly imbalanced and waiting costs are low. Chordia and Subrahmanyam (2004) present a theoretical framework for Chordia et al. (2001) and obtain the same results at the individual stock level. Using the TORQ database, Harris and Panchapagesan (2005) reveal that the limit order book is informative about future price movements and that NYSE specialists use this information to submit orders. Chan (2005) further demonstrate that positive returns lead traders to more (less) aggressive buying (selling) behavior and conversely. Chordia et al. (2008) also show that the predictability of future short-term returns is negatively affected by liquidity, i.e. the lower the spread, the lower the predictability. Using Australian Stock Exchange data, Cao et al. (2009) conclude that the ASX order books are associated with future returns. The relationship between liquidity and price jumps has been thoroughly examined in the literature, showing that a significant proportion of jumps are the consequences of illiquid states of the book which cannot be explained by the occurrence of news (Boudt and Petitjean, 2013). This doctoral thesis contributes to the existing literature on the accuracy of low-frequency liquidity proxies, on price discovery, and on information-based trading in the limit order book.

As previously outlined, the level of data accuracy is essential in market microstructure. In this thesis, we use Euronext market data that have been provided by Euronext Paris. This unique and rich dataset covers a period extending from February 1, 2006 to April 30, 2006 for 701 stocks. It contains all records for orders and trades, upon which the order books have been reconstructed.[1] It also includes undisclosed information, such as hidden orders and market members' ID that we use to disentangle buyer and seller-initiated trades without any error margin. By using this database, we have been able to rebuild intraday prices series that we use to characterize price movements. Since the implementation phase of MiFID starting in November 2007, volumes have been shifting from national exchanges to Multilateral Trading Facilities (MTF). the key advantage of this dataset is to avoid that phenomenon. In order to still be representative of market activity, more recent datasets must include sufficient information from MTF and market data, which has become extremely difficult in today's decentralized trading environment.

The doctoral thesis work is divided in four chapters, each one corresponding to one essay.[2]

1.3.1 Essay I

The first essay analyzes the relationship between liquidity and price movements. This relation has already been discussed in the literature. However, all studies consider closing prices to be the only driver of price movements. In this essay, we extend these analyzes by examining the information content of High-Low-Open-Close (HLOC) dynamics that are easily represented by Japanese candlesticks or bar charts. The information that is gathered in these charts is much more complete than the traditional return. Japanese candlesticks have been used for centuries in eastern countries to trade on markets as old as Osaka and Sakata's

[1] We would like to express all our gratitude to all the people who worked on this database and reconstructed the order books.

[2] All these essays are aimed to be read separately. As a result, we were not able to avoid some repetitions in the definitions of concepts.

rice futures markets. The legend says that a local lord in the region of Sakata, Munehisa Homma, was the first to use past prices to detect future price changes using this method. Steve Nison was the first to bring this way of representing prices on charts to Occident in 1991. His book *Japanese Candlestick Charting Techniques : A Contemporary Guide to the Ancient Investment Technique of the Far East* is one of the most popular technical trading books.[1]

In this essay, we outline a link between 15-minute HLOC price dynamics and liquidity as measured in the limit order book.[2] There are several reasons for the existence of such a relationship. First of all, Kavajecz and Odders-White (2004) find a strong relationship between liquidity and technical analysis, i.e. the study of past prices and their potential predictive power of future prices and returns. Technical analysis has been historically dismissed by academics as it is inconsistent with the efficient market hypothesis proposed by Fama (1970). The goal of this thesis is not to study the predictability of stock returns. It instead focuses on the relationship between price movements and liquidity.

Kavajecz and Odders-White (2004) propose to investigate an unexplored feature of technical analysis by examining how it is related to the liquidity provision. They use moving averages as well as support and resistance to characterize price movements. On the one hand, they relate moving averages to the relative position of depth in the order book. On the other hand, they associate support and resistance levels to peaks in depth in the order book. They consider four specific liquidity measures that are associated to the depth dimension. They also conduct Granger causality tests that show that these technical indicators help locate depth that is already present in the order book.

More recently, Wang et al. (2012) have also outlined a link between technical analysis

[1]For more details on the legend, historical information or candlesticks, please refer to Nison (1991) or Nison (1994).

[2]In this doctoral thesis, we use 15-minute data in the core analyzes of the essays. This interval length is the best trade-off between noise and information and is usually used in papers dealing with HLOC dynamics, e.g. Fiess and MacDonald (2002).

and order submission, which is directly related to the state of liquidity in the order book. They show that institutional investors, and most particularly foreign investors, submit buy (sell) orders after buy (sell) signals.

In this doctoral thesis, we find that some HLOC patterns are related to particular price dynamics that result from special market conditions. For example, the so-called Doji is a consensus structure which occurs when opening and closing prices are close from each other and represents a temporary agreement on the value of the instrument that stops bulls, i.e. traders that buy and push the price up, or bears, i.e. traders that sell and push the price down, to move prices. This temporary agreement is consistent with the findings of Bloomfield et al. (2005) who show that informed traders change their way of trading when the fundamental value is inside the bid-ask spread. Informed traders typically trade on their information set that helps them accurately estimate the fundamental value of the instrument. This activity forces the prices to quickly evolve towards this fundamental price. However, when the difference between the current price and the fundamental value is small, informed traders are not encouraged to enter the market. As a result, they behave as dealers, supplying liquidity, willing to earn the spread that other traders are willing to pay for immediacy. They are motivated in this way as they own information that other traders do not and therefore can submit liquidity inside the existing quotes. They also do not bear any adverse selection cost since they are informed. Bloomfield et al. (2005) also demonstrate that liquidity and noise traders change their behavior as their client's requirements have not been met, forcing them to become more aggressive and taking the liquidity provided by the new informed traders-dealers.

A fundamental value inside the quotes is similar to the idea of the price agreement that leads in the occurrence of this type of consensus configuration. We therefore expect these structures to denote moments where informed traders become dealers. We investigate intraday liquidity behavior around the occurrence of these consensus HLOC patterns through the use of an event study methodology. We use different liquidity proxies from the limit

order book: relative spread, displayed and hidden depth, order imbalances, dispersion and slope. We also analyze trading activity and find that the explanation on the trade dynamics provided by Bloomfield et al. (2005) is consistent with the behavior of liquidity. We further study the relationship on Euronext's blue chips by investigating Granger causality as well as the presence of informed trading that we measure with the PIN indicator, presented in Easley et al. (1996). We check the robustness of the results by testing the same relationship on other time frames. We conduct the analysis on 30-minute and 60-minute samples and find less significant results that indicate that the relationship is rather short-lived. We also address other types of HLOC patterns that show less interesting results. All in all, these outcomes confirm that the main driver of the previously outlined relation emanates from the price agreement that occurs when these configurations appear.

1.3.2 Essay II

The second essay further investigates the relation between HLOC price dynamics and liquidity behavior by quantifying the price movements rather than relating them to common patterns.[1] We conduct panel regressions, where stocks' fixed-effects are controlled by using the clustering approach that is discussed in Petersen (2009). This regression framework addresses the within correlation issues that affect panel datasets when stocks' fixed effects do exist. The sample is split in three portfolios based on the market capitalization: Small, mid and large caps. Each portfolio contains the hundred largest market capitalizations of each category. We include price variables that summarize all possible HLOC dynamics: Open-Close and High-Low ranges, the interaction term between these two variables, as well as dummies for the occurrence of zero returns, price windows and price evolution. Liquidity behavior is investigated by analyzing both order book-based and trade-based liquidity proxies, such as the depth, the relative spread, the dispersion, the slope, the number of trades, the average trade size, trading volume and Amihud (2002)'s illiquidity ratio.

[1] This paper is a joint work with Mikael Petitjean.

Our results corroborate the previous findings and suggest that liquidity is higher when opening and closing prices are close to each other, which is similar to a consensus structure. They indicate that liquidity is negatively related to both OC and HL ranges and positively related to the interaction between these two variables. The results are also consistent with previous literature on the relationship between liquidity and volatility, as the High-Low range is negatively related to liquidity. We also confirm the findings of Boudt and Petitjean (2013) by demonstrating that price gaps occur in illiquid states of the market.

Since price movements may easily be attributed to volatility, we use the same specification with the inclusion of an additional variable that controls for the realized volatility which has proven to be a reliable indicator of volatility (De Vilder and Visser, 2008). The results of these panel regressions show that the inclusion of this variable does not influence the significance of the price movement variables, suggesting that the effect of HLOC price dynamics on liquidity is not captured by volatility. This outcome is particularly interesting and contributes to a large literature on the negative liquidity-volatility relationship by indicating that price ranges contain a liquidity effect beyond the one incorporated in the realized volatility.

We further address this relationship by running robust and median regressions that control for the presence of outliers. The results are not affected and remain significant for all portfolios. As a robustness check, we confirm these outcomes by using 10-minute and 20-minute samples.

1.3.3 Essay III

The third essay focuses on zero returns which are often used in the literature to characterize liquidity. This chapter more specifically examines the illiquidity measure proposed by Lesmond et al. (1999) who demonstrate that the occurrence of zero returns is related to illiquid states of the book. Their reasoning grounds on the fact that transaction costs constitute a

threshold that the value of the new information must exceed to encourage informed traders to trade. If informed traders do not evaluate the upcoming information sufficiently profitable, they decide not to trade, which in turn leads to a zero return, as the effect of other traders does not force prices to move. Following this statement, zero returns should be positively related to transactions costs and illiquidity. Different research papers use this proxy to characterize liquidity when order book or trade data are not available. Bekaert et al. (2007) study the impact of liquidity on asset pricing on a sample of emerging market's stocks using zero returns as a proxy for illiquidity. Levine and Schmukler (2006) also addresses cross-listing and estimate liquidity with the frequency of zero returns. However, several elements in the literature suggest that this proxy is inappropriate to measure illiquidity. First, Goyenko et al. (2009) test this proxy and find that it performs poorly in capturing changes in the effective spread. Second, Bloomfield et al. (2005) show that informed traders become dealers when they estimate that the fundamental value is inside the quotes. Therefore, they provide liquidity, even if the return is zero, which reduces transaction costs. Lesmond et al. (1999)'s measure is also opposed to the relationship we outline in our two first chapters. A zero return is a particular case of consensus where the closing and opening prices are exactly alike. As we have outlined a strong correlation between consensus structures and liquidity, we also expect that this relationship still holds if we restrict these configurations to the zero return. Another justification is provided by Amihud (2002)'s illiquidity measure that propose a liquidity proxy where the return is negatively related to liquidity, i.e. the lower the return, the higher the liquidity.

In this essay, we apply an event study methodology, where the incidence of a zero return is the event. By analyzing liquidity behavior around the event, our findings clearly demonstrate that zero returns are related to a liquid state of the order book. We split the sample using the market capitalization criterium and define three groups of stocks: Small, mid and large caps. We also run conditional logit regression to further address this issue. Our results are in line with the outcomes of the event study. These outcomes are corroborated by the analysis of lower frequency intervals: 20, 30 and 60-minute intervals. In addition, we consider daily zero returns in our sample and find that their occurrence is more related

to liquidity than to illiquidity. In this chapter, we extensively discuss the role of informed traders and we confirm that informed traders are not trading the way we expect them to trade, i.e. aggressive trades that move prices to the fundamental value. We however argue that the market structure has a strong influence on the effectiveness of the occurrence of zero returns as a liquidity proxy which seems to be most appropriate in limit order markets. In a dealer market, the detection of informed traders, or of their willingness to trade, by the market makers leads them to enlarge the spread which in turn impacts the return that informed traders would have made without the dealer's intervention. They are therefore not willing to trade anymore. The market becomes more illiquid and the probability of a zero return is higher. In a limit order market, Bloomfield et al. (2005) argue that informed traders behave as dealers when the fundamental value is inside the quotes. The resulting return is zero and liquidity is higher since they do not move prices by actively trading and they supply liquidity that they usually take away from the limit order book.

1.3.4 Essay IV

The fourth essay is devoted to the analysis of the implication of the previous findings for transaction costs estimation.[1] This is the next logical step after having assessed that liquidity is higher when particular HLOC price dynamics occur. Transaction cost management has always been a particular concern for trade execution. Transactions costs are divided into several components. Explicit costs are the first category and may be accurately estimated before trading occurs. They include brokerage commissions, market fees, clearing and settlement costs as well as taxes. When we refer to transaction costs, the second category is much more important. Implicit costs consists in all costs that may not be accurately evaluated in advance. They include the bid-ask spread, market impact and several opportunity costs. The bid-ask spread is the cost that bear liquidity demanders and the return that receive liquidity suppliers. The market impact is the cost incurred when a large order successively

[1] This essay is a joint work with Benoit Detollenaere and has been accepted for publication in the *Journal of Banking and Finance*.

dries out best quantities at the opposite side. Opportunity costs are costs that traders bear before their trade decision has been implemented. Market timing and missed trade opportunity costs belong to this category. All these components vary in transparency and associated cost. For instance, explicit costs are much more visible and less influential in the total trading costs. Interestingly, it is not possible to reduce all costs components simultaneously. As a consequence, a trader must decide which type of component best suits her execution strategy and reduce it. The most tricky issue is related to the splitting of orders. Splitting orders may significantly reduce market impact but this action bears the risk of an adverse future price variation that may hinder the global execution. Traders have to decide which splitting they want to implement, if any, to reduce market impact or market timing cost. This decision is called the trader's dilemma.

In this paper, we address the following question: Do HLOC price dynamics and Japanese candlesticks help solve the trader's dilemma? Based on our previous findings showing that some HLOC price dynamics are related to a higher liquidity in the order book, we examine whether this relationship may help traders decide to split or not their trades. The existence of a possible relationship is also motivated by the findings of Kavajecz and Odders-White (2004) that are presented in the previous paragraphs. Wang et al. (2012) also show that order submission depends on the occurrence of technical signals in the Taiwan Stock Exchange. In this research study, we investigate sequences of orders submitted by the same market member and concentrate the core of the paper on the Doji and the Hammer-like configurations identified on a 15-minute time frame. We conduct fixed-effect panel regressions, where we control for within correlation among firms using the clustering approach that is described in Petersen (2009). The dependent variables are market timing or market impact costs and the set of explanatory variables includes the most common cost drivers, e.g. the number of orders in the sequence and its duration, depth and spread liquidity proxies, as well as candlestick identification dummies. We investigate both contemporaneous and lagged signals to check whether an execution strategy may profit from HLOC configurations that have just occurred.

Our results indicate that market timing may not be improved when these HLOC configurations occur. This is verified for both Doji and Hammer-like structures and is consistent with previous literature which shows that candlesticks fail to predict future price evolution. However, market impact costs are significantly lower when or after a Doji has occurred. The outcomes further show that traders may profit from a window ranging from 15 to 30 minutes to place aggressive orders at lower cost. Some Hammer-like configurations also display statistically lower market impact costs but the effect is only valid for buy orders, is related to much less significance and implies shorter-lived effects.

In order to further check these results, we also conduct order processing simulations where we compare two submission strategies. The first one is an original Doji-based strategy where orders are submitted just after a Doji has occurred. The other one is a random strategy that consists in placing n orders each day, where n is the number of orders that are submitted in the original strategy, randomly during the day. This latter strategy is replicated five hundred times to control for lucky draws in the choice of time intervals where orders are submitted. The results are consistent with the panel regressions since the Doji-based strategy presents significantly lower market impact costs. All in all, this essay suggests that liquidity demanders may benefit from candlesticks to improve their execution strategy.

The last chapter is devoted to a general conclusion and proposes different avenues for future research.

CHAPTER

TWO

PRICE DYNAMICS AND MARKET LIQUIDITY: AN INTRADAY EVENT STUDY ON EURONEXT

2.1 Introduction

Liquidity has recently become of utmost importance in finance. The liquidity black holes observed during the Subprime crisis, the recent emergence of liquidity dark pools, and the surge in high frequency trading have drawn the attention of an increasing number of researchers and practitioners. The flash crash of May 6th, 2010 has shown in particular that liquidity can be very unstable through time. In such a trading environment, the ability to find and estimate intraday liquidity in a fast and accurate way is extremely precious but also very challenging. Liquidity is indeed not unidimensional. Harris (2003) defines liquidity as "*the ability to trade large size quickly, at low cost, when you want to trade*"[1] and attributes three main dimensions to liquidity (i.e. immediacy, width and, depth), before identifying

[1] Harris (2003), pp. 394.

resiliency as the fourth dimension. In this paper, we study intraday price dynamics observed on Euronext in order to determine whether they help better characterize market liquidity in real time.

The seminal paper in this area of research is Kavajecz and Odders-White (2004) who first reveal some unexpected features of price dynamics with respect to liquidity on the NYSE. These authors study the information content of technical analysis for liquidity provision. They do not investigate the return predictability power of technical indicators but examine similarities between support and resistance levels and the level of depth in the limit order book. They also study moving average indicators and assess their information content in the order book. Their results show that signals in price charts are significantly related to the state of liquidity in the order book. They find that support and resistance levels and moving averages are strongly correlated with liquidity. The authors also conduct Granger causality tests which reveal that technical analysis helps discover depth available in the book. However, their study has several limitations. Firstly, the four liquidity measures they define are customized proxies that are not easily extracted from the limit order book. Secondly, they focus on 30-minute intervals only, without trying to extend their results to longer and shorter time frames. Thirdly, as support and resistance levels are often related to round price limits, the relationship they outline between technical analysis and liquidity may just be due to the fact that limit orders submissions occur at these round prices, implying a higher depth for round price limits.[1] Last but not least, price dynamics are captured by changes in closing prices only.

To respond to each of these limitations, we focus on High-Low-Open-Close (HLOC) price dynamics. Widely available and easily depicted in charts, intraday HLOC prices provide market participants with a quick snapshot of buying and selling pressures, as well as turning points. As such, they might help estimate intraday liquidity in a fast and accurate way. Using market data on a sample of European stocks, we study the relationship between

[1]There is an extensive literature on round prices that we do not discuss in the paper. Ball et al. (1985), Harris (1991) and Kandel et al. (2001) are some examples.

liquidity and price movements by applying an event study methodology on 15-minute intervals for several HLOC patterns. As outlined by Kavajecz and Odders-White (2004), price dynamics are expected to be related to modifications in the state of the limit order book and to the supply of liquidity. For this purpose, we analyze various standard liquidity proxies: relative spread, one-sided displayed and hidden depth (at the best bid and offer and at the five best limits), order imbalance, dispersion, and slope. We also analyze the following trading activity measures: number of buyer and seller-initiated trades, as well as trade imbalances.

In the empirical analysis, HLOC price movements that characterize a consensus on the value of a security are shown to contain important information regarding liquidity. Informed traders supply more liquidity when they evaluate that the fundamental value of the security is inside the best quotes, as suggested by Bloomfield et al. (2005) among others. For all liquidity proxies, our results suggest that liquidity is higher when such a consensus HLOC price configuration appears on screen. There is nevertheless less trading activity at that particular time. This reinforces the hypothesis of price agreement for the security and is consistent with limit order book models and empirical evidence, as indicated by Chakravarty and Holden (1995) and Foucault (1999) among others. This consensus implies a narrower spread, higher depth and less dispersion due to higher competition that liquidity suppliers face. The duration of the liquidity changes depends on the proxy, but the patterns are typically short-lived. The position of opening and closing prices on the high-low range is also related to different changes on bid and ask sides: prices close to the highest point of the interval are linked to changes at the ask while prices near the lowest point are related to changes at the bid.

We investigate whether these results are related to the probability of informed trading. Based on the PIN measure proposed by Easley et al. (1996), our analysis shows that consensus HLOC price configurations characterize moments at which the PIN is lower, indicating that informed traders are less likely to generate trade imbalances. We also conduct Granger causality tests in order to address the direction of the relationship between HLOC price movements and liquidity. Our results indicate that changes in liquidity proxies 'Granger

cause' changes in the Close-Open range and that price movements are good indicators for liquidity. As robustness checks, we extend our analysis to other types of configurations. We also check the significance of our results when we consider 30-minute and 60-minute intervals. We find similar but less significant results indicating that the relationship is rather short-lived.

All in all, our analysis indicates that HLOC dynamics may be used to quickly characterize the four dimensions of liquidity in a single chart. Since information displayed in the limit order book is often difficult to interpret in real time, the quick estimation of liquidity through HLOC price dynamics may bring significant improvements in the execution of buy and sell decisions taken by investors, portfolio managers and brokers.

The remainder of the paper is organized as follows. Section 2.2 provides a brief review of the literature on price dynamics and liquidity. We give theoretical evidence on the links between liquidity and the consensus HLOC price configurations. Section 2.3 describes the dataset and the different liquidity measures that are used. Section 2.4 presents the methodology that we apply. Section 2.5 reports the findings of the event study, Section 2.6 presents the results of the PIN analysis and, Section 2.7 is devoted to a causality analysis. Section 2.8 contains the robustness checks that are performed. The final section concludes.

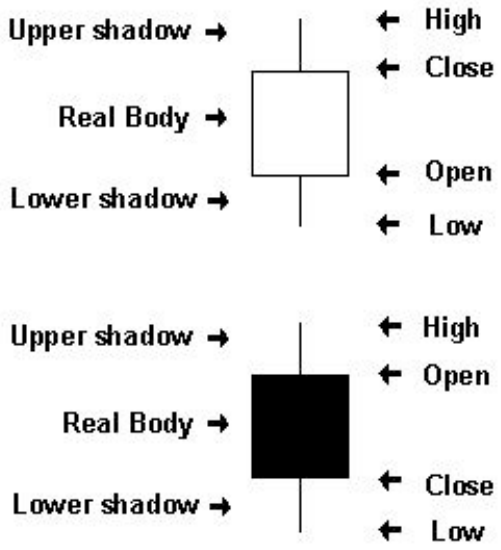
2.2 Literature review

2.2.1 Price dynamics

HLOC price dynamics may easily be represented in charts where each interval is related to a configuration similar to the one depicted in Figure 2.1. These types of charts include bar charts and Japanese candlesticks. Japanese candlesticks are a technical analysis charting

technique based on High-Low-Open-Close prices.[1] Depending on the length of the shadows and the size and color of the bodies, many structures may be identified, from one to five candles. These candlesticks emphasize what happened in the market at that particular moment. Each configuration can be translated into traders' behaviors through price dynamics implied by buying and selling pressures.

Figure 2.1: HLOC dynamics in charts



This type of representation is interesting because it summarizes a lot of information in one single chart: the closing price, the opening price as well as the lowest and highest prices. With the rising interest in high frequency trading and the narrowing of trading intervals, they have been increasingly used to capture short term price dynamics and are available in almost all trading softwares. To our knowledge, this paper is the first research study that investigates the information content of HLOC price movements for intraday liquidity.

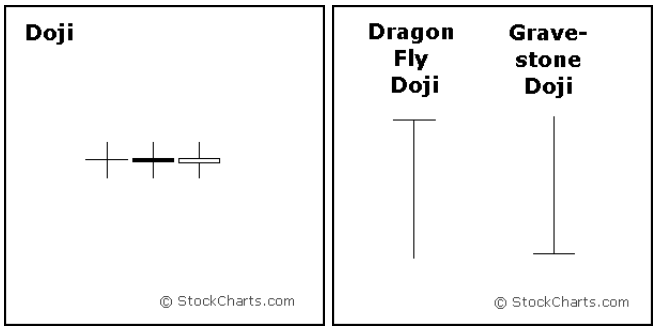
Previous research has outlined strong relationships between price movements and trade

[1] Steve Nison was the first to bring this method to the west in the nineties, even if Japanese candlesticks have been used for centuries in eastern countries. See Nison (1991), Nison (1994), Morris (1995) and Bigalow (2001) for more information on candlestick charting.

measures (Blume et al., 1989). We also expect a relationship between the occurrence of particular HLOC configurations and trading activity measures. Fiess and MacDonald (2002) also argue in favor of that point. A relation between these patterns and liquidity measures is also expected as trading activity measures are linked to the state of the order book, i.e. a trade occurs when supply meets demand and this matching is realized through the limit order book. This is the case of the consensus structures, also called Doji.

The Doji is one of the core structures of the literature on Japanese candlesticks.[1] A Doji appears when the closing price is (almost) equal to the opening price and denotes moments where there is no particular pressure on the price, resulting in a consensus on the fundamental value of the security. The security is said to be “on a rest”. We observe different types of Doji. The most frequent Doji is a "plus", i.e. no real body and almost equal shadows. If both closing and opening prices are also equal to the highest price of the interval, the Doji becomes a Dragonfly Doji (top consensus) which appears when a strong buying pressure directly follows a strong selling pressure implying an upper shadow almost equal to zero. By contrast, it becomes a Gravestone Doji (bottom consensus) when both closing and opening prices are equal to the lowest price of the interval which indicates that the buyers have dominated the first part of the session and the sellers, the second one. In essence, the Doji is not an indicator of price reversal: it only helps detect the end of the current trend. These configurations are presented in Figure 2.2.

Figure 2.2: Doji



[1] Nison (1991), Nison (1994) and Morris (1995).

Table 2.1 presents the different types of consensus and the rules that they need to satisfy.

Table 2.1: Consensus configurations

Consensus	HLOC	Representation
Consensus - Doji	Open $\approx$ Close	+
Top consensus - Dragonfly Doji	Open $\approx$ Close $\approx$ High	$\top$
Bottom consensus - Gravestone Doji	Open $\approx$ Close $\approx$ Low	$\perp$

This table describes the different consensus configurations that may appear on HLOC charts. The second column indicates the rules that the structure must satisfy. The third column presents the graphical representation of this type of consensus structure. More details are available in appendix.

2.2.2 Price dynamics and liquidity

The relationship between liquidity and price movements is well documented in the literature. For instance, Blume et al. (1989) study the interactions between liquidity and price dynamics by investigating the impact of order imbalances on stock price movements during the stock market crisis of October 19 and 20, 1987. The authors identify strong and positive correlation between order imbalances and price movements. They also enlighten a cascade effect of order imbalance on stock prices. The authors conclude that their results are consistent with the hypothesis that the stock decline was due to the incapacity of the market structure to absorb large selling orders. Using data from the Paris Bourse, Biais et al. (1995) investigate the supply and demand of liquidity as well as traders' aggressiveness. They find that traders place orders inside the quotes when the spread or the depth (at the best quotes) is large. Chordia et al. (2001) empirically find that liquidity and trading activity are influenced by market returns and volatility. They also find that effective and quoted spreads increase dramatically in down markets. This effect is asymmetric because spreads do not decrease much in up markets. Chordia et al. (2002) reach the same results and identify a significant

impact of daily order imbalance on both market returns and volatility. They also state that it is unwise to trade when the order book is highly imbalanced if waiting costs are low. Chordia and Subrahmanyam (2004) present a theoretical framework for Chordia et al. (2001) and obtain the same results at the individual stock level. Harris and Panchapagesan (2005) also outline a relationship between the limit order book and future price movements using the TORQ database. Chan (2005) find evidence that order placement strategies depend on previous returns, i.e. traders are more aggressive in buying and place fewer sell orders after positive returns and conversely, a decrease in price cause sellers to be more aggressive by placing more orders at the best quotes, larger orders, or by reducing the spread. More recently, Chordia et al. (2008) find that short-term return predictability is lower when liquidity (as measured by the bid-ask spread) is higher. They also point out that prices have been more efficient after the change to decimal tick size, supporting the positive relationship between liquidity and market efficiency. Cao et al. (2009) analyze the predictive power of limit order book information for 100 stocks quoted on the Australian Stock Exchange. They find evidence that it facilitates price discovery and is associated with future short-term returns. Boudt and Petitjean (2013) also consider the dynamics of liquidity around price jumps and the information content of window formation in intraday price charts with an event study. They find that liquidity drops sharply and is particularly low at the time of formation as well as in the following thirty minutes of the jump. All these studies investigate the returns based on the closing prices only, without taking into account HLOC dynamics which are directly related to trading activity and volatility. This paper fills that gap by thoroughly investigating liquidity around particular HLOC price patterns, among which consensus structures.

Beyond the fact that consensus structures are related to less volatility, hence higher liquidity,[1] they denote moments where there exists an agreement between buyers and sellers on the fundamental value of the stock. In this situation, informed traders are likely to situate the value of the stock inside the spread. After a price discovery process, the current price

[1]The negative relationship between liquidity and volatility has been extensively discussed in the literature, e.g. Benston and Hagerman (1974), Stoll (1978a), Stoll (1978b), Amihud and Mendelson (1980), Ho and Stoll (1981), Copeland and Galai (1983), Admati and Pfleiderer (1988), Pagano (1989), Foster and Viswanathan (1990), Stoll (2003), Pastor and Stambaugh (2003), Harris (2003) and, Beaupain et al. (2010).

is efficient and informed traders do not trade to benefit from their information set since its value is small. They therefore submit limit orders inside the quotes, leading to competition among traders which reduces the spread and the dispersion while increasing the slope and best quotes quantities. This phenomenon has been outlined by Chakravarty and Holden (1995) who extend Glosten and Milgrom (1985)'s framework and show that informed traders use a combination of limit and market orders when the fundamental value is inside the spread. Bloomfield et al. (2005) further demonstrate that informed traders evolve in their way of trading when the current price is becoming more and more efficient. They are likely to submit limit orders after a rally that made the price informative. Their motivation in doing so resides in their ability to trade without adverse selection costs to earn the spread that the liquidity traders are willing to pay to meet their client's requirements. This is particularly true as Bloomfield et al. (2005) also show that liquidity traders are becoming more aggressive when their target is not inexpensively met by submitting limit orders, which frequently happens when they participate to a rally led by informed traders. Kaniel and Liu (2006) argue in favor of that point and also suggest that the horizon of private information plays an important role in the choice to submit limit orders or not. This reasoning is also valid in the framework of the limit order book models that have been developed in Parlour (1998), Foucault (1999), Foucault et al. (2005), Goettler et al. (2005) and more recently, Goettler et al. (2009). Roşu (2009) even more specifically address the competition and patience of informed traders in his theoretical dynamic model.

In order to understand the functioning of the order book when a consensus HLOC price configuration on the fundamental value arises, as it is the case when a consensus occurs, let us suppose that all informed traders own the same private information set that allows them to accurately find the fundamental value, at time t for a given security, v_t that strictly belongs to $]PB1_t, PA1_t[$, where $PB1_t$ and $PA1_{it}$ are the best available prices at time t on the bid and ask sides, respectively. Let us denote the bid-ask spread $s_t = PA1_t - PB1_t$ and the midquote $m_t = \frac{PA1_t + PB1_t}{2}$. The order book at time t is displayed as follows:

$$
\begin{array}{cc}
 & m_t \\
PB1_t & PA1_t \\
PB2_t & PA2_t \\
PB3_t & PA3_t \\
PB4_t & PA4_t \\
PB5_t & PA5_t
\end{array}
$$

At time t, v_t is included in $]PB1_t, PA1_t[$ and may be different from the midpoint m_t. Informed traders change their way of trading at that moment and become dealers to still profit from their informational advantage. To earn the spread s_t, an informed trader must submit orders on both sides. This also prevents him to have an unbalanced inventory that would confront him to the risk of a possible opposite price evolution. In general, a dealer has no interest in speculating on future price evolution. This is consistent with the literature on inventory models (Stoll, 1978b; Amihud and Mendelson, 1980). As a consequence, he will reduce the spread by submitting a limit order at the bid with a price $PB1_{t+1}$ and a limit order at the ask with a price $PA1_{t+1}$ inside the existing quotes.[1]

Then, a competition effect between the informed traders occurs as the book carries information that traders see. Orders that are far from the best quotes experience a lower execution probability and informed traders resubmit them at better quotes. They compete to submit the best limit orders that are most likely to be picked off by other traders. As shown in Kyle (1985), this Bertrand competition drives informed traders-dealers' profits to zero. Harris (2003) also outlines that informed traders who own the same information do not trade with each other. On that account, only liquidity or noise traders are willing to trade to consume this abundant liquidity from the book. The situation evolves when a new information comes up that changes the fundamental value of the stock, v_{t+n}, or when some traders have an

[1]The determination of the optimal price of these orders is not important in our study, since informed traders will enter a Bertrand competition to submit better and better price limits. In Foucault (1999), there is a discussion on this optimal level but, in that framework, traders choose between limit and market orders, depending on the probability of execution and the expected price change that the order will imply.

imperfect estimation of the new stock's value.[1]

The logical consequences of this reasoning are a higher liquidity and a lower trading activity. Since these consensus structures characterize moments where there is an agrement on the price, we may clearly situate their occurrences after the price discovery process and it should be associated with the change of behavior of informed traders, which in turn leads to higher liquidity. Further evidence of this is given by the drivers of the consensus configuration which will most likely situate it at the end of a rally but not necessarily implying a price reversal. We expect our results to be short-lived and the dynamics to appear in the window close to the signal. We also expect different outcomes between top and bottom consensus configurations as they come from different succession of price pressures: the bottom (top) consensus is made with a previous bullish (bearish) rally and ends up with a strong selling (buying) pressure. These expectations are opposed to the findings of Lesmond et al. (1999) who characterize zero-return days as illiquid days with higher transaction costs. The behavior of informed traders outlined by Bloomfield et al. (2005), which we build our reasoning on, however stands against the proposition of Lesmond et al. (1999).

2.3 Data

2.3.1 Data and Sample

We use Euronext market data on 81 stocks belonging to three national indexes: BEL20, AEX and CAC40. We have tick-by-tick data for 61 trading days from February 1, 2006 to April 30, 2006. The database include trades, orders and order book data, which allows

[1] As the order book is informative, if some informed traders have different beliefs on the initial fundamental value, i.e. $v_t + \tilde{u}$ with $\tilde{u} \sim \mathcal{N}(0, 1)$, they will update their beliefs and include the informational content of the book. These traders will most likely not actively participate the Bertrand competition as they do not precisely estimate the fundamental value.

the most comprehensive analysis. The key advantage of this dataset is to avoid the shift of volumes from national exchanges to Multilateral Trading Facilities (MTF) that has been occurring since the implementation phase of the MiFID directive. More recent datasets must instead include sufficient information from market data and MTF to be representative of market activity. Needless to say, it has become extremely difficult to build such a reliable dataset in today's decentralized trading environment. In addition, our dataset includes market members' ID that we use to disentangle buyer-initiated and seller-initiated trades, without any error margin. Traditionally, in market microstructure studies, the Lee and Ready (1991)'s algorithm is used to categorize buyer and seller-initiated trades. This algorithm has proved to be sufficiently efficient but misclassification may still occur. In our case, we are able to precisely assess which order initiates the transaction. Finally, we are also provided with undisclosed data on hidden orders.[1]

We have rebuilt High-Low-Open-Close prices from this database for the 81 stocks over the whole sample period. As tick data are not adapted for this analysis, we build 15-minute-intervals which leads to 34 intervals a day, from 9:00 AM to 5:30 PM (CET). This interval length is the best trade-off which allows to include intraday trends and to avoid noisy patterns resulting from non-trading intervals.[2] We use the HLOC prices calculated above in order to identify the most widespread price configurations as defined in the TA-Lib library.[3] We obtain a total of 167068 records (81 firms, 61 days, 34 intervals/day).

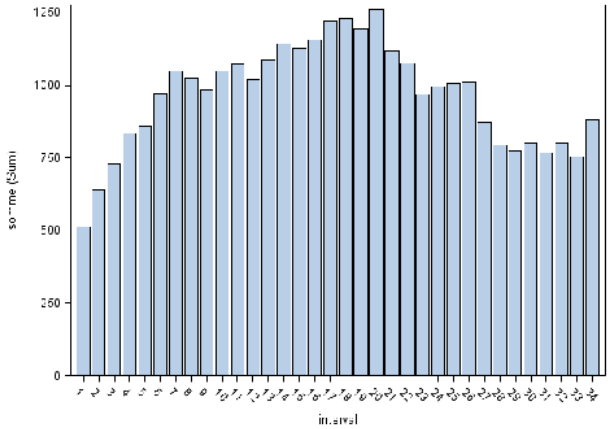
[1] Hidden orders are orders that gradually display part of their total amount. For instance, a hidden order of 500 may appear on the book with a quantity of 100 and will automatically be refilled when 100 shares have been consumed.

[2] At the 5-minute interval, we observe a lot of non-trading intervals, i.e. Doji structures with a zero volume. In this analysis, only Doji configurations with positive volume are investigated since we want to relate them to trader's behavior. In the robustness checks section, we evaluate the sensitivity of the results to a change in the interval length by analyzing 30-minute and 60-minute intervals.

[3] For each type of configuration and for each record, The TA-lib library returns "1" if the bullish part of the structure is identified, "-1" for the bearish part and "0" otherwise. As the structures are bullish, bearish or both, for each event type, the values that may appear are [0 ; 1], [-1 ; 0] or [-1 ; 0 ; 1]. Each structure is identified by observing the relation between the four prices. The TA-lib allows some flexibility in the recognition of the configurations. As it is an open source library, we have been able to check the parametrization of the structures. Events are recognized according to the standard flexibility rules presented in Nison (1991) and Morris (1995). The TA-lib contains 61 pre-programmed structures. We however dropped 17 of them due to a lack of intraday significance. These 17 structures are meaningful when longer time periods are considered. However, they are too frequent on intraday data and do not give any signal. The list of these configurations is available upon request.

We look at the occurrences of the consensus structures and check whether they appear at a particular moment during the day. Figure 2.3 shows that the distribution of these configurations is roughly uniform with the most significant peaks occurring during lunch time and maybe resulting from non-trading. They also seem to not occur frequently during the first two intervals of the day. This may be explained by the strong unidirectional movement that appears at that moment, as trends are at their beginning.

Figure 2.3: Consensus configurations by interval before filters



This figure displays the number of consensus structures in each time interval. The 34 intervals correspond to 15-minute interval starting at 9:00 AM until 5:30 PM.

We filter the sample as follows. First, we remove non-trading patterns as they result in 'four price consensus'.[1] Second, we only keep the events, i.e. the occurrence of a consensus structure, for which we do not observe any other structure in the previous and next three intervals in order to avoid any contagion between events. If there are many events in a single window, it is difficult to assess which event is the main driver of the observed pattern. Therefore, there is only one event by analyzed window. As we study a moving window containing three 15-minute intervals before and after the signal, we do not consider events in the first and last three intervals for each day to avoid constructing event windows over

[1]This configuration occurs when all the prices are equal. When they occur in daily, weekly or monthly charts, they are a strong clue of a potential reversal, as outlined by Nison (1991). However, in intraday price charts, they often represent non-trading intervals. We only remove four-price Doji structures that are associated with a volume equal to zero.

different days. After the removal of all the possible contagious data, we have a total of 2959 HLOC consensus structures, among which 653 are top consensus and 614 are bottom consensus. Table 2.2 shows event occurrences.

Table 2.2: Events count

Name	Bull/Bear	Count
Consensus	1	2959
Top Consensus	1	653
Bottom Consensus	1	614

2.3.2 Liquidity measures

We measure liquidity at the end of each trading interval, which enables us to measure the direct relation between the event and liquidity. We first calculate the traditional liquidity proxies such as relative spread, depths (displayed and hidden, in number of shares) and order imbalances (at different order book levels). We also use dispersion and slope measures which are respectively presented in Kang and Yeo (2008) and Næs and Skjeltorp (2006). Then, we compute trading activity measures as buyer and seller-initiated trades and imbalances, in number of trades.[1] Finally, we evaluate volatility using the High-Low measure for each time interval. Table 2.3 presents the different measures.

[1]These measures are computed with the sum over each interval.

Table 2.3: Liquidity proxies and trading activity measures

Name	Median	Name	Median
Quantities at the Best Bid	1978	5 Best Limits Hidden Imbalance	0.04
Quantities at the Best Ask	1953	Relative Spread	0.07
Quantities First Limits (Bid+ask)	4541.5	Dispersion Bid	0.02
Hidden Quantities at the Best Bid	0	Dispersion Ask	0.02
Hidden Quantities at the Best Ask	0	Dispersion	0.02
Hidden Quantities First Limits (Bid+ask)	453	Bid Slope	4245.62
Displayed Quantities 5 Best Bid	13409	Ask Slope	4269.39
Displayed Quantities 5 Best Ask	13612	Slope	4302.42
Quantities 5 First Limits (Bid+ask)	27908	Number of buyer-initiated trades	28
Hidden Quantities 5 Best Bid	1559	Number of seller-initiated trades	31
Hidden Quantities 5 Best Ask	1800	Imbalance Number of Trades	-0.05
Hidden Quantities 5 First Limits (Bid+ask)	8410.5	High - Low	0
First Limits Imbalance	0	Buy-side Aggressiveness	-0.42
First limits Hidden Imbalance	-0.04	Sell-side Aggressiveness	-0.37
5 Best Limits Imbalance	0.01		

"Quantities at the best bid (ask)" denotes the amount of shares displayed at the best bid (ask) limit. "Hidden" indicates the quantities that are not displayed. "Displayed Quantities 5 Best Bid (Ask)" represents the total amount displayed at the five best bid (ask) limits. "Hidden Quantities 5 Best Bid" denotes the total hidden amount at the five best bid (ask) limits. "Quantities First Limits (Bid+ask)" is the sum of displayed best bid and offer quantities. "Hidden Quantities First Limits (Bid+ask) only takes into account hidden quantities. Quantities 5 First Limits (Bid+ask) and Hidden Quantities 5 First Limits (Bid+ask) are computed across the five best price limits while Total Quantities (Bid+ask)" and "Hidden Total Quantities (Bid+ask)" consider the whole book. "First limits Imbalance" is the best limits displayed imbalance ($Imbalance_{i,t} = \frac{QuantitiesBid_{i,t} - QuantitiesAsk_{i,t}}{QuantitiesBid_{i,t} + QuantitiesAsk_{i,t}}$, where i denotes a given security and t a given interval.). "First limits Hidden Imbalance" only considers hidden quantities. The same measures are computed for the five best limits ("5 Best Limits Imbalance" and "5 Best Limits Hidden Imbalance"). "Relative Spread" denotes the relative spread. The other measures are extensively detailed in the text (Section 2.3.2). The High-Low has been computed relatively to the midpoint.

Dispersion Kang and Yeo (2008) present two measures to quantify the density of the limit order book, i.e. how limits are far from each other or from the quoted midpoint. One of these two measures is the dispersion:

$$Dispersion_{i,t} = \frac{1}{2}\left(\frac{\sum_{j=1}^{n} w_{i,j,t}^{Bid} Dst_{i,j,t}^{Bid}}{\sum_{j=1}^{n} w_{i,j,t}^{Bid}} + \frac{\sum_{j=1}^{n} w_{i,j,t}^{Ask} Dst_{i,j,t}^{Ask}}{\sum_{j=1}^{n} w_{i,j,t}^{Ask}}\right), \tag{2.3.1}$$

where, for security i and interval t, $w_{i,j,t}$ are the weights which are equal to quantities,

offer and bid sizes, at the j^{th} price limit normalized by the total depth of the five best limits, $Dst^{Bid}_{i,j,t} = (Price^{Bid}_{i,j-1,t} - Price^{Bid}_{i,j,t})$ and, $Dst^{Ask}_{i,j,t} = (Price^{Ask}_{i,j,t} - Price^{Ask}_{i,j-1,t})$. The midquote is used for the distance of the first best limits.

As Kang and Yeo (2008) outline, dispersion is small under fierce competition as each trader wants to gain price priority.

Slope The slope is computed by averaging the price elasticity of quantities over the five best quotes. A steep slope represents an order book where volumes are concentrated at a given limit (low elasticity) while a gentle slope denotes an order book where volumes are not aggregated at a given limit (high elasticity). A steep slope also means that traders agree about the value of the security while a more gentle slope indicates that traders have different estimations of the fair price of the security.

We calculate the slope of the book following Næs and Skjeltorp (2006), that is:

$$SLOPE_{i,t} = \frac{DE_{i,t} + SE_{i,t}}{2}, \tag{2.3.2}$$

where $DE_{i,t}$ and $SE_{i,t}$ are the demand and supply elasticities respectively and are computed as:

$$DE_{i,t} = \frac{1}{5}\left(\frac{v^B_1}{|p^B_1/p_0 - 1|} + \sum_{\tau=1}^{4} \frac{v^B_{\tau+1}/v^B_\tau - 1}{|p^B_{\tau+1}/p^B_\tau - 1|}\right), \tag{2.3.3}$$

$$SE_{i,t} = \frac{1}{5}\left(\frac{v^A_1}{p^A_1/p_0 - 1} + \sum_{\tau=1}^{4} \frac{v^A_{\tau+1}/v^A_\tau - 1}{p^A_{\tau+1}/p^A_\tau - 1}\right). \tag{2.3.4}$$

p_τ^B and p_τ^A are the prices, respectively at the bid and at the ask, appearing at the quote τ. p_0 denotes the quoted midpoint. Finally, v_τ^B and v_τ^A are the natural logarithm of accumulated total share volume at the limit τ respectively for the bid and the ask.[1]

Trade Imbalance In order to calculate trade imbalance, we first have to sign transactions. Most empirical studies use Lee and Ready (1991)'s algorithm, which categorizes buyer and seller-initiated trades based on the position of the transaction price relative to the bid-ask spread. With our database, we are able to match for each transaction the orders that generate the trade. Then, the sign of the transaction is found by comparing the submission time of the orders, i.e. the last order being the determinant of the transaction. After that, we compute the total number of trades and quantities respectively for buyer and seller-initiated trades. With these variables, we compute the imbalance as follows:

$$ImbalanceN_{i,t} = \frac{NTrades_{i,t}^{Buy} - NTrades_{i,t}^{Sell}}{NTrades_{i,t}^{Buy} + NTrades_{i,t}^{Sell}} \tag{2.3.5}$$

where $NTrades_{i,t}^{Buy}$ is the number of buyer-initiated trades occurring at the t^{th} interval for stock i and $NTrades_{i,t}^{Sell}$ is the number of seller-initiated trades occurring at the t^{th} interval for stock i.

This measure is computed separately for each interval and for each security.

Aggressiveness We compute our aggressiveness measure separately for bid and ask sides. Our measure captures the number of aggressive orders, i.e. orders that consume liquidity and generate a trade, compared to the total number of orders that occurs during a given time interval:

[1] By accumulated, we mean the sum of the quantities outstanding at that limit and the sum of all quantities outstanding at each better quote.

$$AggressivenessB_{i,t} = \frac{NbAggressive_{i,t}^{Buy} - NbPassive_{i,t}^{Buy}}{NbAggressive_{i,t}^{Buy} + NbPassive_{i,t}^{Buy}} \tag{2.3.6}$$

where $NbAggressive_{i,t}^{Buy}$ is the total number of marketable buy orders occurring at the t^{th} interval for stock i and $NbPassive_{i,t}^{Buy}$ is the total number of non-marketable buy orders occurring at the t^{th} interval for stock i. A similar process is applied to sell orders.

2.4 Methodology

With our extended dataset containing HLOC prices, HLOC patterns identification variables and liquidity proxies, we perform intraday event studies of liquidity behavior around HLOC dynamics. This original median-based event study methodology has been proposed by Boudt and Petitjean (2013). Our event is the occurrence of a consensus structure. The null hypothesis of this event study is: The occurrence of a consensus configuration is not related to a change in the liquidity proxy. The alternative hypothesis postulates that: The occurrence of a consensus configuration is related to a positive or negative change in the liquidity proxy. We focus on an event window of [-3,+3] containing seven observations: Three observations before the signal, the time of the signal and three observations after the signal. This leads us to consider liquidity behavior 45 minutes before and after the apparition of the event. As we have different types of measures, we compute our abnormal measures depending upon the nature of the variables, as suggested by Boudt and Petitjean (2013).

For imbalances and aggressiveness measures, which may display negative values and are already standardized, we compute the abnormality as follows:

$$Abnormal_{i,t,p} = Proxy_{i,t,p} - Median_{i,t,p}^{NE}, \tag{2.4.1}$$

where $Proxy_{i,t,p}$ is the analyzed liquidity proxy p for stock i for the time interval t and $Median^{NE}_{i,t,p}$ is the median of the proxy p for stock i across all non-events[1] occurring during the time interval t.

For the other measures, the calculation method is:

$$Abnormal_{i,t,p} = \frac{Proxy_{i,t,p} - Median^{NE}_{i,t,p}}{Median^{NE}_{i,t,p}}, \tag{2.4.2}$$

where $Proxy_{i,t,p}$ is the analyzed liquidity proxy for stock i for the time interval t and $Median^{NE}_{i,t,p}$ is the median of the proxy p for stock i across all non-events occurring during the time interval t.

We need to separate these two categories of variables as imbalances and aggressiveness measures may be negative and belong to [-1,1] by construction. We apply these processes separately for each liquidity measure. We then aggregate the results obtained for each stock to form median patterns of liquidity behavior around each event and for each proxy. We choose the median as it is a more robust measure of central tendency, compared to the mean, as the distributions of liquidity proxies are heavily skewed. We analyze these patterns to check whether the abnormality is significantly different from zero. For this purpose, we use a standard non parametric sign test that does not need any assumptions about the shape of the distribution. Our null hypothesis postulates that the median of the abnormal measure equals zero. The alternative hypothesis postulates the opposite. The statistic M is computed as follows:

$$M = \frac{N_+ - N_-}{2}, \tag{2.4.3}$$

where M follows a binomial distribution, N_+ is the number of positive values and N_- is the

[1]The 17 dropped pre-programmed structures have been included in the non-event sample as they are too frequent on intraday data and are not linked to any particular signal.

number of negative values. Values equal to zero are discarded.

We then analyze the p-values of each time interval of the window and check whether the differences are significant or not. If the p-value at the signal interval is significant, the structure is associated to a particular state of (il)liquidity. This signal may thus lead to a given configuration of the limit order book. If p-values are significant before the signal, the signal may be a response to a particular state of liquidity. If p-values are significant after the signal, a change in liquidity may correspond to a response to the signal.

2.5 Results

2.5.1 Overall picture

The general conclusion of the event study is that liquidity is higher when a consensus configuration appears, consistent with our expectations based on informed traders' change in behavior. We also observe that there are fewer trades which also confirms the rationale we provide in Section 2.2. A consensus structure is an indication of the presence of liquidity on each side of the book. The strategy of placing liquidity-taking orders at that moment is likely to cost less in terms of implicit costs. Liquidity providers are also numerous and, as a result, face more competition to supply liquidity. Interestingly, these results do not hold when jumps occur. As pointed out by Boudt and Petitjean (2013), jumps are linked to illiquid states of the book. As a consequence, HLOC movements that occur inside previous interval's levels would be related to liquidity while those with jumps would be related to illiquidity.[1]

[1]We test this hypothesis by comparing consensus configurations that occur inside previous interval's High-Low range to consensus configurations that appear outside this range. We confirm that consensus structures that gap up or down are linked to less liquid states of the book than other types of consensus structures. The drop in liquidity is effective in terms of spread, depth, dispersion and slope. These results are not reported here but are available upon request.

The identification of the different types of consensus also brings interesting information about the thickness of each side of the book. The top consensus seems to be associated with more significant changes at the ask while the bottom consensus seems to rather affect the bid. These movements occur about 15 minutes before the occurrence of these events, implying that the consensus itself is the consequence of depth dynamics. This result is confirmed by imbalances which show that the order book is imbalanced before these structures and comes back to equilibrium after their occurrence. These findings show that price dynamics are related to the state of the limit order book and confirm previous evidence of such a relationship, as in Kavajecz and Odders-White (2004) or Chordia et al. (2002).

We also observe disparities in trading activity. While buyer-initiated trades are less frequent and smaller in size when there is a bottom consensus, seller-initiated trades are less frequent and smaller in size when a top consensus occurs. Buy orders are also more aggressive after a top consensus while sell orders are more aggressive when a bottom consensus appears.

Our analysis finally shows that liquidity is negatively correlated with volatility as the High-Low is significantly lower while liquidity seems to be higher when a consensus configuration appears, consistent with the existing literature on the liquidity-volatility relationship.

All in all, these findings corroborate the theoretical evidence and limit order book dynamics that we outline in Section 2.2 and are in line with Bloomfield et al. (2005)'s outcomes.

2.5.2 Details

Figure 2.4 presents the depth results. Panels a to d refer to displayed depth while Panels e and f show hidden depth results. The quantities at the best bid are significantly higher, at a 1% confidence level, at the time of the signal. Quantities at the best bid are also significantly

higher just before the apparition of a bottom consensus. The pattern at the ask is more noisy but indicates that the top consensus presents a significantly higher best ask depth just before its occurrence. These structures seem to be the consequences of a particular state of liquidity in the limit order book as the pattern starts in $t-1$. The bottom consensus is likely to occur just after an accumulation of depth at the bid that does not translate into more trades. The top consensus seems to appear when depth at the ask is higher and when there are less trades. The results of the bid side are confirmed if we consider the five best quotes. However it is not the case for the ask side. If we look at the sum of both sides, we also observe a peaking depth for the first quote and for the five best quotes. This is interesting as the reduction of the spread does not take place at the cost of a lower depth. Liquidity is higher over these two dimensions. The bottom consensus presents a peak in hidden depth at the bid just before its apparition while depth at the ask peaks just before the top consensus. These findings indicate that the position of the Close-Open range (near the highest or lowest price of the time interval) on the High-Low range has a one-sided impact on depth, hidden or not.

Figure 2.5 shows that the spread significantly drops when a consensus configuration appears. The recovery is fast after the trough in all cases. As outlined by Bloomfield et al. (2005), informed traders seem to agree on the fundamental value of the stock and situate it inside the spread. They are acting as market makers and increase the competition to submit the best limit to increase the trading probability. We observe opposite patterns for the top and bottom consensus structures. Order book imbalance is significantly higher just before and when a bottom consensus appears while it is significantly lower before and when a top consensus occurs. The traditional consensus configuration does not exhibit any particular pattern, however.

The dispersion drops when a consensus configuration appears, meaning that the competition in the book is higher at that particular moment and that the limits are closer from each other. This is consistent with the idea of consensus. Informed traders are competing to gain price priority implying a narrower spread and a smaller dispersion. If we disentangle ask and

bid sides, we observe that dispersion is significantly lower at the bid for the top consensus and at the ask for the bottom consensus while it remains unchanged on the other side. These findings are also coherent with the philosophy behind these two structures: a strong buying pressure creates the top consensus while the bottom consensus appears after a selling rally. This is also in line with the one-sided impact that occurs for depth. An incomplete price discovery at the beginning of the consensus may also be a possible cause of the one-sided effect.

The slope significantly peaks at the moment of the event, on both supply and demand sides. Still consistent with Bloomfield et al. (2005) and the reasoning we propose in Section 2.2, these outcomes are also in line with the agreement on the price and with dispersion measures which show an increase in order book density when these signals appear.

Figure 2.6 indicates that the number of trades is lower when the signal appears, whatever the direction of the trade. After the signal, there is a quicker return to normal values for buy trades than for sell trades which remain low 15 minutes after the signal. Traders seem to delay buying and selling activities when these signals occur, even if liquidity is higher. As depth does not fall and even increase, there are more pending orders in the book. If we investigate the pattern for the bottom and the top consensus configurations and keeping in mind the philosophy behind these structures, we observe that buy trades are less numerous and less big for the bottom consensus compared to sell trades while we observe the opposite pattern for the top consensus. Since, the bottom consensus appears when a selling rally follows a strong buying pressure, our results suggest that the selling rally does not come from an increased sell volume but from a decrease in buy volume. The opposite interpretation may be done for the top consensus. Trade imbalances confirm these findings with a sharp drop for the bottom consensus and a significant peak for the top consensus.

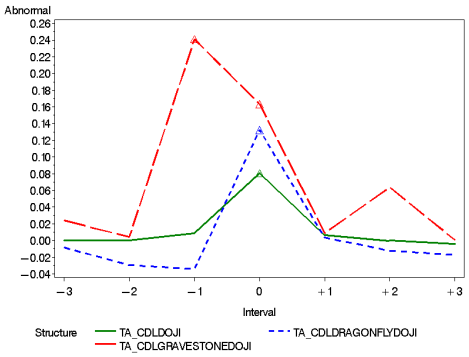
Sell orders are more aggressive when a consensus structure appears while buy orders are more aggressive only 15 minutes after the apparition of the consensus. Traders seem to

place more marketable orders than they usually do, even if trading activity is lower. This suggest that buy traders become aggressive after the apparition of the structure. Sellers react quicker than buyers do. The results also suggest that sellers are even more aggressive when there is a bottom consensus and buyers seem to be even more aggressive in the case of a top consensus. These results are consistent with Bloomfield et al. (2005) who show that liquidity and noise traders are likely to become more aggressive when their clients' requirements are not met. They may also decide to trade more aggressively when they estimate liquidity to be sufficiently abundant to trade at a reasonable cost.

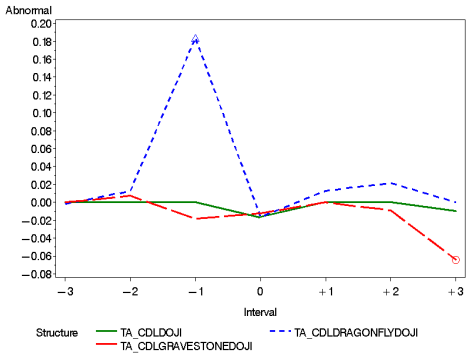
Whatever the consensus structure, the High-Low volatility measure sharply falls when the signal appears. The level is significant at 1%. These configurations are thus linked to lower volatility. This also corroborates the findings on liquidity, since the negative relationship between liquidity and volatility has been clearly established in the literature, e.g. in Pagano (1989).

Figure 2.4: Abnormal liquidity around consensus configurations 1/2

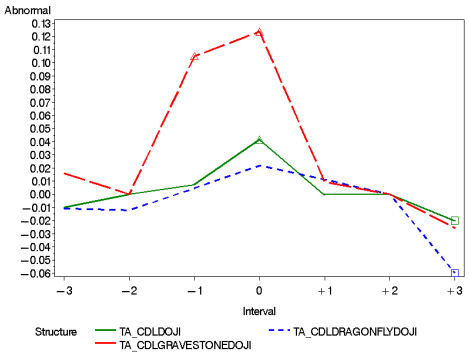
(a) Best Bid



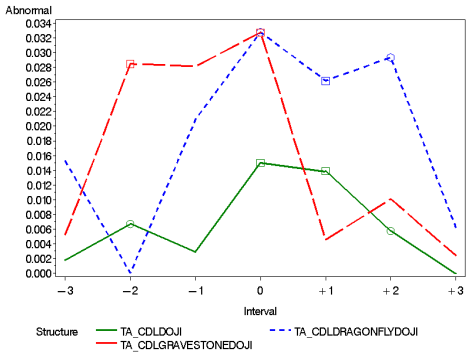
(b) Best Ask



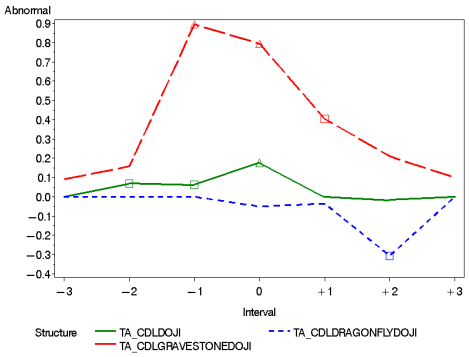
(c) First quote (Both sides)



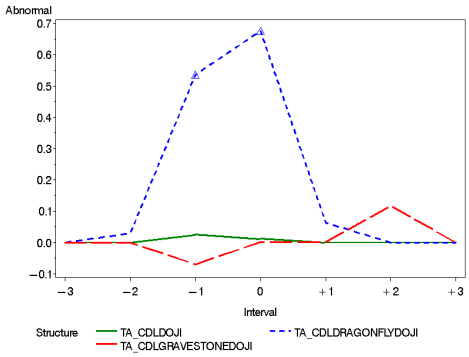
(d) Five quotes (Both sides)



(e) Hidden: 5 Bids

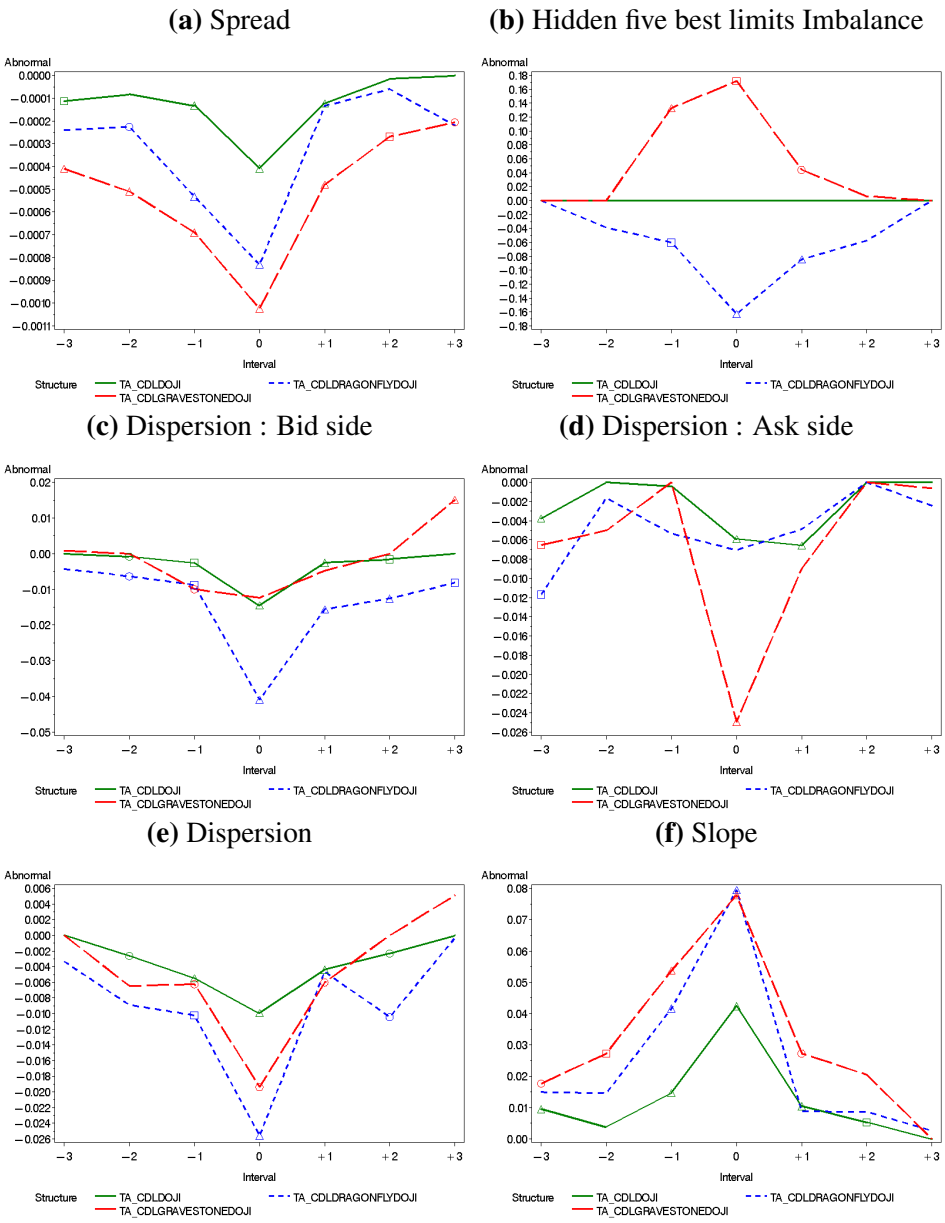


(f) Hidden: 5 Asks



Full, dotted and dashed lines represent the intra-window median pattern for the abnormal spread respectively for the consensus (TACDLDOJI), the top consensus (TACDLDRAGONFLYDOJI) and, the bottom consensus (TACDLGRAVESTONEDOJI). Triangles (△), squares (□), and circles (○) indicate a rejection of the null hypothesis respectively at the 99%, 95% and 90% confidence levels.

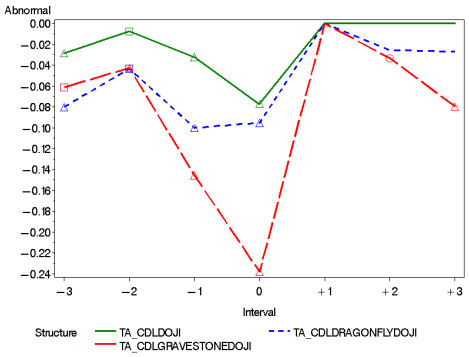
Figure 2.5: Abnormal liquidity around consensus configurations 2/2



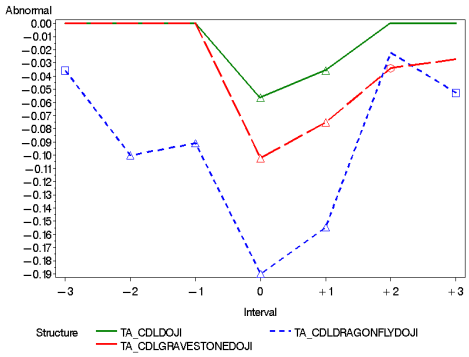
Full, dotted and dashed lines represent the intra-window median pattern for each of the abnormal dispersion respectively for the consensus (TACDLDOJI), the top consensus (TACDLDRAGONFLYDOJI) and, the bottom consensus (TACDLGRAVESTONEDOJI). Triangles (△), squares (□), and circles (○) indicate a rejection of the null hypothesis respectively at the 99%, 95% and 90% confidence levels.

Figure 2.6: Abnormal number of trades around consensus configurations

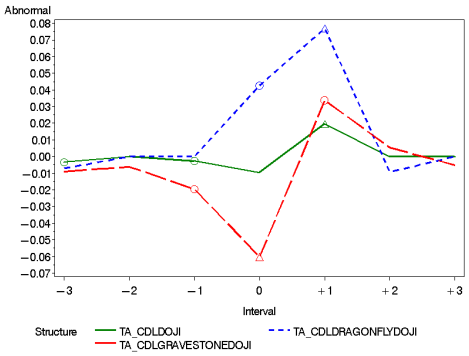
(a) Number of buy trades



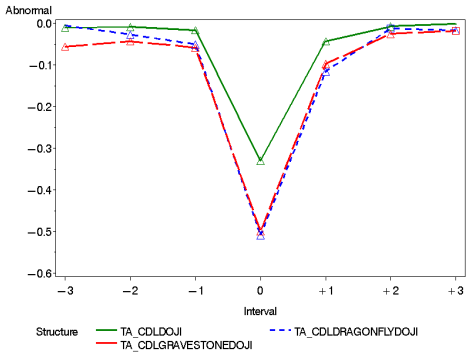
(b) Number of sell trades



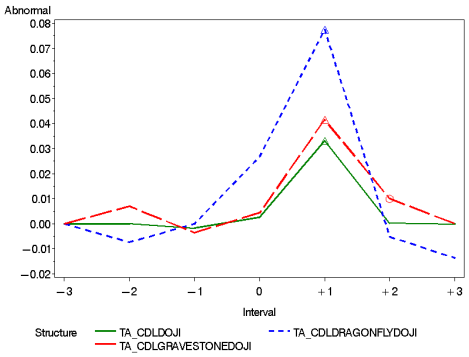
(c) Imbalance



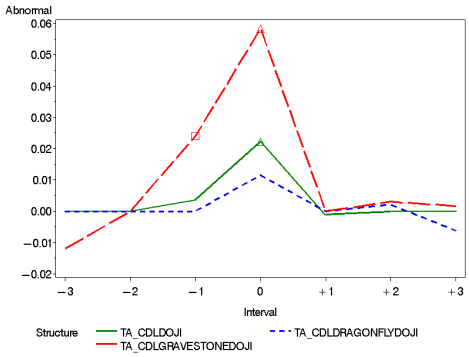
(d) High-Low



(e) Buyers's aggressiveness



(f) Sellers's aggressiveness



Full, dotted and dashed lines represent the intra-window median pattern for this trading activity measure respectively for the consensus (TACDLDOJI), the top consensus (TACDLDRAGONFLYDOJI) and, the bottom consensus (TACDLGRAVESTONEDOJI). Triangles (△), squares (□), and circles (○) indicate a rejection of the null hypothesis respectively at the 99%, 95% and 90% confidence levels.

2.6 Detecting the Presence of Informed Trading

As outlined in the previous section, when a consensus configuration occurs, liquidity is higher and trading activity is lower. As shown in Section 2.2, Bloomfield et al. (2005) discuss the change in the behavior of informed traders after the price discovery process, driving them to the use of limit orders to take profit from their informational advantage and earn the spread. As the intraday return is close to zero, consensus configurations seem to be related to moments without significant news. Informed traders should not actively trade in the market at these particular moments. However, they are most likely to be present in the order book, waiting for being picked off by noise and liquidity traders. If some private information is made available that enable them to update their beliefs about the asset's value, they place their orders on the same side, generating high trade imbalances, as well as a significant price movement.

In the following analysis, we use the PIN indicator to check whether there is less informed trading when a consensus structure occur. The PIN measure, which quantifies the probability of information-based trading, has already been extensively discussed in the literature [Easley and O'Hara (1987), Easley and O'Hara (1992), Easley et al. (1996), Easley et al. (1996), Easley et al. (1997), Easley et al. (1998), Easley et al. (2002), Easley et al. (2008) and, Easley et al. (2012)].[1] It is based on the trade imbalance that the presence of informed traders creates. There are many possible computations of the PIN measure. In this paper, we use the one provided in Easley et al. (1996), as it is usually done in the literature. We also use the factorization method of the log-likelihood function presented in Easley et al. (2008). We nevertheless discuss in Section 2.2 that informed traders may still be present even if they do not trade, behaving like market makers. In our case, we expect the PIN to be lower when a consensus configuration appears, since there is less informed trading, even if informed traders are still submitting limit orders.

[1] Details on the computation of the PIN are not presented here but are available in these papers.

We compute the PIN associated to each day for each stock of our sample, as it is done in the literature. The methodology only differs in the period length. While previous research uses longer time interval by aggregating daily information to form monthly or annual PIN observations, we use 15-minute interval to generate daily estimates for the PIN. We compare days with consensus structures to days without consensus structures and check whether days with consensus configurations exhibit a lower PIN. We count the number of consensus structures per days and exclude days where there are more than 10 occurrences. As the sample is not filtered in this analysis as we did it in the event study, we only keep days where these configurations are most likely to be isolated and not a consequence of non-trading. Then, we conduct a non-parametric comparison Kruskal-Wallis test whose null hypothesis postulates that the mean scores of the subsamples are equal while the alternative hypothesis postulates that they are different. The results of the test are presented in Table 2.4.

Table 2.4: Non-parametric tests on the relationship between PIN and the number of consensus configurations per day.

Number of occurrences	Count	Sum of Scores	Mean Score
0	28	79624	2843.714
1	117	278484.5	2380.209
2	283	671048.5	2371.196
3	487	1121739	2303.364
4	690	1614494	2339.846
5	732	1776181	2426.477
6	692	1672226	2416.512
7	546	1318298	2414.464
8	429	1075123	2506.113
9	309	807164.5	2612.183
10	198	519593	2624.207

This table presents the results of the non-parametric Kruskal-Wallis test which tests whether there are differences in the mean scores accross the different subsamples, stratified by the number of consensus structures per day. The first column denotes the 10 subsamples that correspond to a fix number of consensus configurations per day. The second column presents the number of days in the subsample. The third and fourth column respectively display the sum of scores and the mean scores of the Kruskal-Wallis test. The p-value associated to the Chi-square statistic of the test is equal to 0.0062 indicating that the null hypothesis is rejected, i.e. the scores of the 10 subsamples are significantly different from each other.

These results show that the mean score of the first row is higher than the nine other rows, indicating that the PIN is effectively higher when there is no consensus structure. This finding states that the lower trading activity occurring with the consensus configurations is due to less informed trading as we demonstrate it in Section 2.2. The p-value of the test is also highly significant (0.0062).

2.7 Granger causality

In this section, we test the Granger causality between each liquidity proxy and a price movement variable that represents the occurrence of a consensus structure. This variable measures the absolute value of the difference between opening and closing prices and therefore characterizes the Close-Open range, i.e. $OC_t = \frac{|Close_t - Open_t|}{RangeMid_t}$, where $RangeMid_t = \frac{High_t + Low_t}{2}$. These tests enable us to check whether the consensus structure is a cause or consequence of a change in liquidity or both of these relationships (Bidirectional causality).

Toda and Yamamoto (1995) present a procedure to produce efficient and unbiased estimators of VAR models when the processes are integrated. This methodology helps to produce unbiased Granger causality tests, as it is similar to testing zero-restrictions on a given set parameters of a VAR specification. Their first step suggests that we check for the order of integration of the time series. We use stationarity tests, ADF and Phillips-Perron statistics, and conclude that all our time series were stationary. In this case, the Wald test statistic may be directly used instead of going through the process presented in Toda and Yamamoto (1995).

Our unrestricted VAR models are specified as follows:

$$\begin{aligned} OC_t = \alpha_0 + \alpha_1 OC_{t-1} + \alpha_2 OC_{t-2} + ... + \alpha_p OC_{t-p} \\ +\beta_1 L_{t-1} + \beta_2 L_{t-2} + ... + \beta_p L_{t-p} + u_t, \end{aligned} \tag{2.7.1}$$

$$\begin{aligned} L_t = \alpha_0 + \alpha_1 L_{t-1} + \alpha_2 L_{t-2} + ... + \alpha_p L_{t-p} \\ +\beta_1 OC_{t-1} + \beta_2 OC_{t-2} + ... + \beta_p OC_{t-p} + u_t, \end{aligned} \tag{2.7.2}$$

where L_t denotes one of the liquidity proxies that is investigated, p denotes the number of lags and u_t is an error term. The optimal lag length, p, is determined through an optimization process based on Akaike's Information Criterion (AIC).

The null hypothesis of Granger Causality tests characterizes non-causality, i.e. "L_t does not Granger-cause OC_t" for the first VAR model and "OC_t does not Granger-cause L_t" for the second one. This test consists in testing that all the β_i of the models equal 0.

We compute the F-test statistic as follows:

$$S_1 = \frac{(RSS_r - RSS_u)/p}{RSS_u/(T-2p-1)} \sim F_{p,T-2p-1}, \tag{2.7.3}$$

where RSS_r and RSS_u are the residual sums of squares respectively for the restricted and unrestricted models. T and p respectively denote the number of observations and the lag length. If the result of this test statistic is greater than the specified critical value, we reject the null hypothesis that specifies that there is no Granger-Causality.

An asymptotically equivalent test is specified as follows:[1]

[1]When lagged dependent variables are included, the test is only valid asymptotically. Further information may be found in Lütkepohl (2006).

$$S_1 = \frac{T(RSS_r - RSS_u)}{RSS_u} \sim \chi^2(p). \tag{2.7.4}$$

Table 2.5 presents the results of the Granger causality tests, based on the asymptotically equivalent statistic. The first column indicates the p-value of the causality from the prices to the book while the second one displays the p-value from the book to the prices. The results show that the p-values are more significant from liquidity to HLOC prices meaning that HLOC prices dynamics are a response to a particular state of liquidity. This is consistent with the findings of Kavajecz and Odders-White (2004). The results are however not valid at the first limit and for displayed imbalances. We also observe a bidirectional causality for relative spread, dispersion and slope measures. In these cases, liquidity causes modifications of OC_t which in turns causes changes in liquidity.

Table 2.5: Granger causality tests

L_t	$OC \rightarrow Liquidity$	$Liquidity \rightarrow OC$
Quantities at the Best Bid	0.32754	0.51789
Quantities at the Best Ask	0.07330	0.56609
Hidden Quantities at the Best Bid	0.02897	0.80399
Hidden Quantities at the Best Ask	0.23496	0.04968
Depth First Limits (Bid+ask)	0.00507	0.40818
Hidden Depth First Limits (Bid+ask)	0.88486	0.11814
Displayed Depth 5 Best Bid	0.05838	0.00015
Displayed Depth 5 Best Ask	0.10212	0.00010
Hidden Depth 5 Best Bid	0.24773	0.38075
Hidden Depth 5 Best Ask	0.28121	0.00003
Depth 5 First Limits (Bid+ask)	0.13820	0.00000
Hidden Depth 5 First Limits (Bid+ask)	0.72686	0.00005
First Limits Imbalance	0.39725	0.37709
First limits Hidden Imbalance	0.29162	0.00000
5 Best Limits Imbalance	0.15460	0.27104
5 Best Limits Hidden Imbalance	0.57127	0.00000
Dispersion	0.00780	0.00024
Slope	0.00017	0.00000
Relative Spread	0.00000	0.00000

This table presents the results of the Granger causality tests conducted for both directions, i.e. HLOC prices to liquidity and liquidity to HLOC prices. The values are the p-values returned by the asymptotically equivalent test. "Quantities at the best bid (ask)" denotes the amount of shares displayed at the best bid (ask) limit. "Hidden" indicates the quantities that are not displayed. "Displayed Depth 5 Best Bid (Ask)" represents the total amount displayed at the five best bid (ask) limits. "Hidden Depth 5 Best Bid" denotes the total hidden amount at the five best bid (ask) limits. "Depth First Limits (Bid+ask)" is the sum of displayed best bid and offer quantities. "Hidden Depth First Limits (Bid+ask) only takes into account hidden quantities. "Depth 5 First Limits (Bid+ask)" and "Hidden Depth 5 First Limits (Bid+ask)" are computed across the five best price limits. "First limits Imbalance" is the best limits displayed imbalance ($Imbalance_{i,t} = \frac{DepthBid_{i,t} - DepthAsk_{i,t}}{DepthBid_{i,t} + DepthAsk_{i,t}}$, where i denotes a given security and t a given interval.). "First limits Hidden Imbalance" only considers hidden quantities. The same measures are computed for the five best limits ("5 Best Limits Imbalance" and "5 Best Limits Hidden Imbalance").

All in all, these results confirm our expectations and the explanations we provide in Section 2.2 on the change in informed traders' behavior when a price discovery process has made the price sufficiently efficient to bring it inside the best quotes, as it is the case when we observe a consensus. For spread, dispersion and slope measures, there is a bidirectional

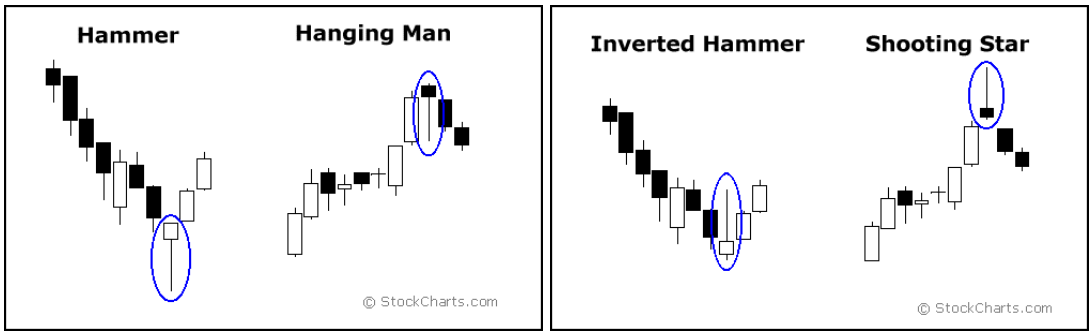
causality which suggests that the current price is the consequence of a particular state of previous liquidity which in turns affects the current state of liquidity. Previous liquidity drives the price discovery process and, in a second step, the price is efficient and drives the state of liquidity.

2.8 Robustness checks

In this section, we conduct two robustness checks. We first investigate other types of dynamics with a lower level of consensus, i.e. closing and opening prices may be significantly different. These structures are called Hammer-like configurations and are characterized by a small Close-Open range that is situated close to the highest or lowest prices of the interval and that may appear at the end of an uptrend or a downtrend.[1] Among Hammer-like structures, there are four structures that are characterized by a long shadow and a small real body. The Hammer appears at the end of a downtrend and is made of a very small real body with (almost) no upper shadow and a very long lower shadow. The same structure may appear at the end of an uptrend but, in that case, it is called a Hanging Man. Inverting the shadows, i.e. the upper shadow becomes the lower shadow and vice-versa, we obtain an Inverted Hammer at the end of a downtrend or a Shooting Star at the end of an uptrend. Figure 5.3 shows these configurations. As these figures are said to be strong reversal structures in the literature on Japanese candlesticks, they should have an influence on market timing cost, if EMH does not hold: for purchases (sales), Hammer and Inverted Hammer should lead to higher (lower) market timing cost, while Hanging Man and Shooting star should lead to lower (higher) market timing cost. As such, they indicate a high level of reversal potential. Table 2.6 presents these four configurations:

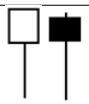
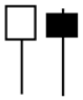
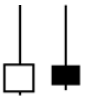
[1] A description of the presented structures is available in appendix.

Figure 2.7: Hammer-like structures



The Hammer and the Hanging Man appear when sellers dominate the first part of the session and buyers, the second part. By construction, they present a long lower shadow and almost no upper shadow. The Hammer occurs at the end of a downtrend while the Hanging Man puts an end to an uptrend. The Inverted Hammer and the Shooting Star are made with a small real body, a very long upper shadow and almost no lower shadow. The Inverted Hammer appears at the end of a downtrend and the Shooting Star occurs at the end of an uptrend. These structures are said to be strong reversal ones.

Table 2.6: Reversal configurations

Reversal	HLOC	Position in the trend	Representation
Top-downtrend - Hammer	Open ≈ Close and Close or Open ≈ High	Bottom	
Top-uptrend - Hanging Man	Open ≈ Close and Close or Open ≈ High	Top	
Bottom-downtrend - Inverted Hammer	Open ≈ Close and Close or Open ≈ Low	Bottom	
Bottom-uptrend - Shooting Star	Open ≈ Close and Close or Open ≈ Low	Top	

This table describes the different reversal configurations that may appear on HLOC charts. The second column indicates the rules that the structure must satisfy. The third column presents the position in the existing trend and, the third one, the technical term for the configuration. The graphical representation appears in the last column. More details are available in appendix.

This group of configurations is interesting for many purposes. Regarding liquidity, the results should be linked to previous findings on consensus configurations since the Close-Open range is also very small. Indeed, a top consensus may be a particular reversal structure (top-downtrend or top-uptrend) while a bottom consensus may be a bottom-downtrend or bottom-uptrend, depending on their position on the price chart. We expect the results to be less significant as the Close-Open range may be different from zero for these structures. We also expect a difference between bullish and bearish signals.

Then, we conduct the same analysis on HLOC patterns generated from 30-minutes and 60-minutes price series. By doing this, we create two new sets of events. This enables us to check whether our findings are the consequence of the choice of the time interval.[1]

[1] All the graphs of this analysis are not presented here but are available upon request.

2.8.1 Reversal configurations

To sum up, our results states that reversal structures really break the intra-window pattern. We observe a higher liquidity just before the formation of a top-downtrend but this liquidity is not provided by more depth: the book is more dense and the spread is lower. The top-uptrend also shows similar outcomes but only after it has fully appeared on the chart. The results we discuss here may come from the configuration type, i.e. top or bottom. This is confirmed by the order imbalance results which show that the order imbalance is more in favor of the ask when the Close-Open range is on top. Trading activity and aggressiveness measures are totally in line with the bearish or bullish reversal potential of the structure. We also observe lower volatility when these structures appear. The results are summarized in Table 2.7.

Table 2.7: Liquidity dynamics around reversal configurations

	Top-downtrend (625)			Bottom-downtrend (175)			Top-uptrend (469)			Bottom-uptrend (91)		
	-1	0	1	-1	0	1	-1	0	1	-1	0	1
Spread	–^{3}	- ^{2}		+++ ^{3}		- ^{2}		- - ^{1}	- ^{2}	+++ ^{2}		
Bid depth	- ^{3}	- - ^{3}	-^{1}	+++ ^{3}				++			- -	
Ask depth	+++ ^{3}						+++ ^{3}					
Imbalance	- - - ^{3}			+++ ^{2}			- - - ^{3}					
Dispersion - Bid				+++ ^{1}		- - ^{2}		- - - ^{3}	- ^{2}		+ ^{1}	
Dispersion - Ask	- - - ^{3}			+++ ^{3}						+++ ^{3}		
Slope - Bid	+++ ^{3}			- - - ^{3}			++ ^{3}	+++ ^{3}		- - ^{2}		
Slope - Ask	+++ ^{3}			- - - ^{3}		+++ ^{3}		+++ ^{3}	++ ^{2}	- - ^{2}		

This table presents the results obtained for the four reversal structures for each liquidity measure. Each panel represents a [-1;+1] time window around the occurrence of the event. "+" and "-" signs denote positive and negative values for the abnormal measure. "++" and "–" signs denote bigger positive and negative variations. "+++" and "—" signs denote peaks and trough over the time window. The exponents denote the significance: 1 for 10% significance, 2 for 5% significance and 3 for 1% significance.

The results clearly indicate that liquidity is higher before the apparition of the top-downtrend and when the top-uptrend occurs. Liquidity seems also to be lower for the bottom-downtrend and the bottom-uptrend. Depth results suggest that changes in liquidity around these structures are only one-sided. These outcomes are similar to those of the consensus structures for which bid and ask quantities evolve differently depending on the position of the price on the Close-Open range (near the highest or lowest of the interval). The imbalance significantly drops just before a top-downtrend or a top-uptrend. The book seems to be imbalanced in favor of the ask side when these configurations appear, confirming previous results. Regarding dispersion, we observe that the top-downtrend present a sharp drop in dispersion at the ask just before its occurrence. This may suggest that price excursions beyond and below opening and closing prices have also an impact on dispersion. A top-downtrend is likely to occur when the density of the ask side increases, i.e. price limits are closer from each other. This is consistent with an agreement on the minimum price and the end of the bearish rally. As expected, these structures exhibit changes in dispersion given their high reversal potential. Slope results are totaly in line with dispersion results. Number of buy and sell trades, trade imbalance, aggressiveness and volatility measures are consistent with the price pressures that drive these patterns. They are thus not reported here.

2.8.2 Changing the time interval

Table 2.8 presents the number of occurrences for each structure for both 30-minutes and 60-minutes price series.

Table 2.8: Events count for 30 and 60-minutes price series

Name	Bull/Bear	Count 30 minutes	Count 60 minutes
Consensus	1	1511	1103
Top consensus	1	254	198
Bottom consensus	1	283	197
Top-downtrend	1	268	225
Top-uptrend	-1	209	159
Bottom-downtrend	1	73	47
Bottom-uptrend	-1	42	27

If we consider 30-minutes intervals, consensus structures display very similar liquidity, trading activity and volatility patterns over the whole window. The conclusions of 15-minutes price series are also applicable. This is also true for 60-minutes intervals even if the patterns are more noisy. The results support all our findings. This indicates that the relationships between price dynamics and liquidity are still significant for longer periods.

While looking at reversal configurations, the conclusions of 30-minutes and 60-minutes intervals are very similar. Liquidity measures outcomes are much less significant for all measures, except for the dispersion and the slope whose conclusions remain unchanged. The spread still drops in case of a top-downtrend or a top-uptrend. We however observe fluctuations in the time window but without significance. Trading activity, aggressiveness and volatility measures display similar patterns as for 15-minutes price series but with much more noisy results. When intervals are longer, buying and selling pressures are always struggling, leading to more noise in the results.

To sum up, robustness checks results confirm the outcomes of our core study. The relationships between liquidity and price movements outlined in our study are thus applicable

to time intervals up to 60 minutes. The consensus structures do not present different results. This confirms that the consensus on the price that appears on the chart also appears in the order book whatever the time interval from 15 to 60 minutes. However, we also observe that Hammer-like configurations patterns are not as significant for longer intervals as for smaller ones.

2.9 Conclusion

In this paper, we investigate the relationship between price movements and liquidity in order to check whether it is possible to have a quick view on the state of liquidity in the limit order book by extracting information out of price dynamics. We focus on HLOC prices to characterize price dynamics. We use an event study methodology on 15-minute intervals in order to check how liquidity is affected by the occurrence of HLOC patterns. After filtering out the contagious effects and non-relevant events, we focus on three consensus configurations. These structures imply a consensus between buyers and sellers on the price of the security. We relate these consensus configurations to particular market microstructure events implied by a change in the behavior of informed traders. As outlined by Chakravarty and Holden (1995) and Bloomfield et al. (2005), informed traders becomes market makers when the fundamental value of the security is inside the quotes, supplying liquidity to other types of traders. As a result, traders seem to agree on the price of the security and only a few of them, liquidity or noise traders, is willing to hit the best opposite quote. Informed traders compete to provide the best limit and increase their probability of being picked off. This results in dramatically lower spread and dispersion, while the quantities are larger and the slope steeper at the moment of the event. A liquidity-taking order placed at that particular moment may thus incur a lower market impact, implying lower implicit costs.

We look at several liquidity and trading activity measures: spread, depth, order imbal-

ance, dispersion, slope, trade imbalance, aggressiveness and volatility. We disentangle bid and ask sides as we expect the book to be affected differently on each side. Liquidity seems to be higher over all dimensions when a consensus configuration appears. We also find that there is less trading activity at that moment. This is consistent with the limit order book models (Parlour, 1998; Foucault, 1999; Foucault et al., 2005) as well existing theoretical and empirical literature (Chakravarty and Holden, 1995; Bloomfield et al., 2005).

We also outline that the position of the Close-Open range on the High-Low range seems to have an impact on the behavior of liquidity. A top consensus is likely to be linked to a bigger depth variation on the ask side and a bottom consensus to a bigger depth variation on the bid side just before their occurrences. This also enables us to argue on causality, as in Kavajecz and Odders-White (2004). These structures seem to be the answer to particular liquidity and trading activity dynamics. The pattern may be related to the price discovery that is completed after the beginning of these types of consensus structures. These results have to be further discussed as these configurations may appear at the end of either an uptrend or a downtrend. It would be interesting to analyze how valid is this pattern if we disentangle bullish signals from bearish signals. Our results also show that consensus configurations that gap up or down are linked to less liquid states of the order book than other consensus configurations.

We further investigate what are the possible determinants of these outcomes by analyzing whether consensus configurations are negatively related to informed trading. No particular information events should arise as the return is close to zero. We therefore compute the PIN measure, which identifies the probability of informed trading during a given period, as proposed in Easley et al. (1996) and compare days with consensus structures to days without consensus structures. We find that the PIN is lower when consensus patterns occur and even more lower when the number of configurations during the day increases. These findings are consistent with the previous literature which demonstrates that informed traders behave as dealers when the value of their information is close to zero, as in Chakravarty and Holden

(1995) and Bloomfield et al. (2005). Informed traders are therefore still present but there presence does not impact trade imbalances that the PIN measures. This is coherent with our findings that relate these structures to liquidity through the behavior of informed traders.

We also conduct Granger causality tests which indicate that changes in liquidity proxies causes changes in the Close-Open range. This measure is similar to identify the occurrence of consensus configurations. We find that our price measure is a consequence of the state of liquidity and is therefore an indicator of liquidity in the short run. The results are consistent with the findings of Kavajecz and Odders-White (2004) which show that prices help characterize the state of the limit order book.

We perform two types of robustness checks. We first look at other types of HLOC reversal patterns. We observe interesting results on top-downtrend and top-uptrend configurations which indicate that liquidity is higher before their occurrence. However, this increase in liquidity is not provided by more depth. Placing a large liquidity-taking order at the time of the top-downtrend may thus not necessarily cost less as a higher density does not provide lower market impact, if quantities do not increase. The top-uptrend presents similar outcomes but it only displays higher liquidity at the time of occurrence. We do not investigate further trading activity and aggressiveness measures as they are totaly in line with the buying and selling price pressures that drive the signals. Finally, our results confirm previous findings on consensus structures which suggest that the position of the Close-Open range in comparison to the highest and lowest price of the interval has a one-sided influence on liquidity. When the Close-Open range is near the highest, there seem to be more liquidity at the ask and conversely, liquidity is higher at the bid when the Close-Open range is near the lowest price of the time interval. This is even more true when the Close-Open range is short, i.e. for consensus configurations.

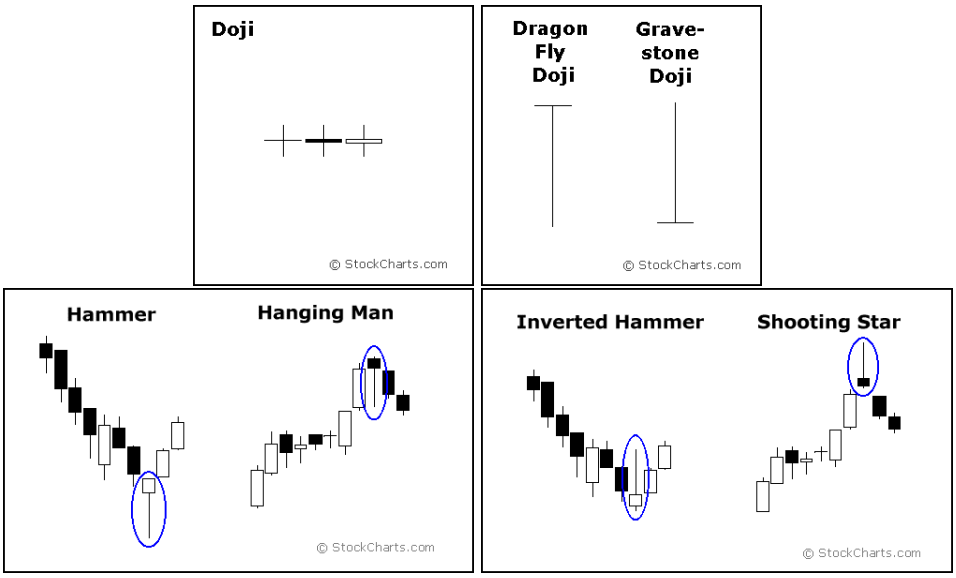
We then change our interval length in order to validate our results for longer time intervals. With 30-minutes and 60-minutes price series, the results are very similar for consensus

structures. Hammer-like configurations do not present very different patterns but display much less significance, except for dispersion and slope measures. The patterns are not as significant for this second category. As expected, all the patterns contain more noise, given the longer period taken into consideration, but still outline a relationship between price dynamics and our measures.

All our results suggest that market participants may benefit from HLOC analysis as a way to better time their order submissions and improve their transaction costs management. Consensus configurations are likely to summarize the information content of the four liquidity dimensions present in the limit order book. All things being equal, placing a marketable order when a consensus appears is likely to lead to a better order execution. The magnitude of the potential gains on transaction costs and the economic significance as well as the resulting optimal execution are beyond the scope of this analysis and are left for further research.

2.10 Appendix

Figure 2.8: Consensus and reversal structures



The Doji (consensus) presents a closing price (almost) equal to the opening price. It occurs when there is an agreement on the fair value of the asset and where markets are 'on a rest'. The Doji indicates the end of the previous trend. The most traditional Doji is a 'plus' sign but Dragonfly (top consensus) and Gravestone Doji (bottom consensus) are also frequent. A Dragonfly Doji appears when a strong buying pressure directly follows a strong selling pressure implying an upper shadow almost equal to zero. The Gravestone Doji occurs when the buyers have dominated the first part of the session and the sellers, the second one. The Hammer (top-downtrend) and the Hanging Man (top-uptrend) appear when sellers dominate the first part of the session and buyers, the second part. By construction, they present a long lower shadow and almost no upper shadow. The Hammer occurs at the end of a downtrend while the Hanging Man puts an end to an uptrend. The Inverted Hammer (bottom-downtrend) and the Shooting Star (bottom-uptrend) are made with a small real body, a very long upper shadow and almost no lower shadow. The Inverted Hammer appears at the end of a downtrend and the Shooting Star occurs at the end of an uptrend.

CHAPTER

THREE

INTRADAY LIQUIDITY, PRICE DYNAMICS AND UNCERTAINTY IN CAP-BASED PORTFOLIOS

3.1 Introduction

The quick and accurate estimation of liquidity has always been a particular challenge in finance. In early research, liquidity was reduced to immediacy, i.e. the immediate conversion of an asset into cash at the best available price (Demsetz 1968). As the literature on market microstructure expanded, a more comprehensive definition of liquidity was proposed. For example, Harris (2003) defines liquidity as 'the ability to trade large size quickly, at low cost, when you want to trade'. In this definition, four liquidity dimensions can be identified: immediacy, width, depth and resiliency. Given the multi-dimensional feature of liquidity, it has become particularly challenging to obtain a reliable snapshot of the dynamics of liquidity. In two seminal papers, Chordia et al. (2000) or Hasbrouck and Seppi (2001) investigate the commonalities in liquidity. Using different samples and methodologies, they both show

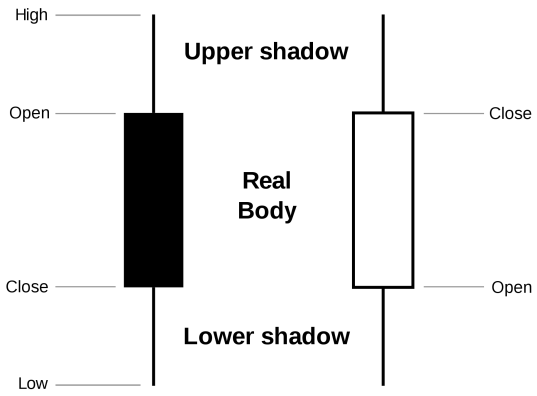
that liquidity co-moves across different stocks. An impressive body of research has then demonstrated that liquidity dynamics can be complex, all the more so in stressful market conditions when there is strong price uncertainty.

In this paper, we focus on Euronext and investigate how informative is price uncertainty, in terms of excess fundamental volatility, to estimate contemporaneous intraday liquidity for small, mid, and large caps separately, since liquidity has proven to vary across capitalization groups.[1] The key research question is therefore whether price uncertainty, as measured by easy-to-observe price movements, is instructive when it comes to evaluating liquidity for these three market cap portfolios. On the U.S. Treasury market, Engle et al. (2012) show that price uncertainty matters. They find that liquidity suppliers reduce their supply in response to price uncertainty. However, they focus on market depth only and the definition of price uncertainty is rough. There is price uncertainty when the dummy for no price change across all tiers on both sides of the order book, is equal to zero.

To circumvent these shortcomings, we use a comprehensive set of both book-based and trade-based liquidity proxies and we measure price uncertainty by studying the information content of High-Low-Open-Close prices (henceforth HLOC prices), which are particularly appropriate to characterize the magnitude of price fluctuations and the occurrence of price jumps. A graphical representation is given in Figure 3.1. The key advantage to using such a representation is that it goes beyond the use of the closing price only, as it is still the case in many research papers, e.g. in Chordia et al. (2000). As shown on Figure 1, price uncertainty is likely to be underestimated when only changes in closing prices matter. In this paper, we show that including the highest, the lowest and the opening prices gives additional information on the behaviors of buyers and sellers.

[1] We do not define uncertainty in the sense of Knight (1921), i.e. a risk that is not measurable, but by the amplitude of price variation between the opening and closing prices.

Figure 3.1: High, low, opening and closing (HLOC) price dynamics



High, low, opening and closing (HLOC) price dynamics are easily depicted in charts that include such a representation for each interval.

As suggested by Fiess and MacDonald (2002), we expect HLOC prices and trading activity measures to be related. We also expect HLOC prices to be related to liquidity proxies for the obvious reason that trading activity measures are by construction dependent on the state of the order book: a trade occurs when supply meets demand, the matching being realized in the limit order book. Furthermore, as liquidity is inversely related to volatility, HLOC price variations should also be significantly associated with liquidity.[1]

To capture the liquidity dynamics in the limit order book, we use book-based liquidity proxies related to depth (over the five best limits, at the ask and bid sides), relative (quoted) spread, slope and dispersion. We also investigate several trade-based proxies: the number of buyer and seller-initiated trades, the total number of trades, the average trade size as well as Amihud (2002)'s illiquidity ratio.

[1]The negative relationship between liquidity and volatility has been extensively discussed in the literature, e.g. Benston and Hagerman (1974), Stoll (1978a), Stoll (1978b), Amihud and Mendelson (1980), Ho and Stoll (1981), Copeland and Galai (1983), Admati and Pfleiderer (1988), Pagano (1989), Foster and Viswanathan (1990), Stoll (2003), Pastor and Stambaugh (2003), Harris (2003), and Beaupain et al. (2010).

To capture the HLOC price dynamics, we include the following variables in our models: the Open-Close (OC) range, the High-Low (HL) range, and an interaction variable (OCHL), i.e. the ratio of the OC to the HL ranges. All else equal, a large OC range indicates that the price discovery process is driven by strong buying or selling pressure, so that the price moves towards a new fundamental value. The HL range measures the total price movement. For a given OC range, a large HL range indicates that price uncertainty has been strong, without being necessarily captured by volatility which relies on closing prices only. The OCHL ratio measures the *relative* amplitude of the price movement beyond the Open-Close range. When the OCHL ratio is close to zero, the price discovery process is very much polluted by uncertainty that prevails around the true stock value. These three variables contain much more information than price returns or traditional volatility measures. They indeed provide more details on the nature of the price discovery process as well as on the behavior of both buyers and sellers. The return is computed on closing prices only. As a result, both OC and HL ranges convey additional information. The HL range is a measure of total price variation and it is not the traditional volatility estimator since researchers prefer using realized volatility proxies, e.g. De Vilder and Visser (2008). The OCHL variable is one of the main contributions of this paper since it compares the net price movements to the total price movements and gives further indication on the different dynamics that have occurred during the interval. We also use dummy variables related to the occurrence of zero returns, the bullish or bearish movement during the interval, and the occurrence of price gaps.[1]

We use Euronext intraday data and split the complete database of 701 stocks into three categories based on market capitalization. The literature has provided valuable evidence which suggests that market capitalization has an direct impact on liquidity: Less liquid stocks often belong to the small caps segment. We select the first hundred stocks in each category, i.e. small, mid and large caps, at the beginning of the sample and define three cap-based portfolios accordingly. For each of these three cap-based portfolio, we study the sensitivity of each of the liquidity proxies to all the price movement variables defined above. We estimate the regressions on 15-minute intervals by OLS with adjusted standard errors. We further

[1]Price gaps occur when the previous high (low) is below (above) the current low (high).

address this relationship by implementing the robust and median regression techniques that deal with the presence of outliers in the sample. We also investigate endogeneity issues by conducting VARX analyzes. As a robustness check, we also use 10-minute and 20-minute time intervals.

HLOC price dynamics are found to provide additional information on the behavior of buyers and sellers and on the way trading activity evolves. The results suggest that, whatever the liquidity dimension, price movements are informative to characterize liquidity dynamics. Positive changes in price ranges for both HL and OC ranges are related to negative variations in liquidity proxies. In other words, the wider the price range, the lower the liquidity. All else equal, liquidity is further reduced when the OCHL ratio decreases, which means that the price discovery process is accompanied by wider price variations beyond the OC range. Correspondingly, liquidity improves when the total variation observed during the interval does not differ much from the OC price range. When prices are less likely to reverse, liquidity providers are dominated by liquidity takers as imbalances between liquidity supply and demand tend to be higher. We also confirm that liquidity is lower when price gaps occur. Finally, HLOC price movements are found to be strong determinants of liquidity dynamics, even after adding realized volatility as a control variable.

All in all, the information content of price movements for intraday liquidity estimation is found to be significant. We conclude that the visual inspection of price movements may offer a good snapshot of liquidity dynamics, even for the smallest caps in the sample.

The remainder of the paper is organized as follows. In Section 3.2, we describe the sample and the model specification. We also define all the liquidity and price movement variables used in the empirical section. We report the empirical findings and the robustness checks in Section 3.3. The final section concludes.

3.2 Data and methodology

We use tick-by-tick Euronext data for 61 trading days from February 1, 2006 to April 30, 2006. The complete database of 701 stocks is divided into three cap-based portfolios. Large, mid, and small caps respectively represent those companies with a market capitalization larger than EUR 1 billion, between EUR 150 millions and EUR 1 billion, and below EUR 150 millions. We select the first hundred stocks in each category based on their market cap at the beginning of the sample.[1]

The use of this dataset presents two key advantages. First, we have information on the full order book, including undisclosed data on hidden orders and market members' ID. By using market members' ID, we are able to disentangle buyer-initiated and seller-initiated trades without any error margin. In many market microstructure studies, the Lee and Ready (1991) algorithm is used to categorize buyer and seller-initiated trades. Although this algorithm has proved to be relatively efficient, misclassification still occurs. In our dataset, there is none since we know the order that initiates the transaction. Second, we avoid the volume shift and market fragmentation that have been occurring since the implementation phase of MiFID. As today's trading environment is much more decentralized than before, more recent datasets are often less representative and less reliable. In a number of recent studies, there is often insufficient information on the level of trading activity that prevails on the competing Multilateral Trading Facilities (MTFs) and dark pools.

3.2.1 Liquidity proxies

At the end of each trading interval, we calculate the relative (quoted) spread (RS) and the quantities outstanding at the five best limits (i.e. depth) for the ask side (QA), the bid

[1] This selection is done to ensure that each portfolio contains hundred stocks.

side (QB), and the sum of the bid and ask (Q).

Finally, we use dispersion and slope measures. The dispersion measure indicates how remote limit orders are from one another. It is computed as follows:

$$Dispersion_{i,t} = \frac{1}{2}\left(\frac{\sum_{j=1}^{5} w_{i,j,t}^{Bid} Dst_{i,j,t}^{Bid}}{\sum_{j=1}^{5} w_{i,j,t}^{Bid}} + \frac{\sum_{j=1}^{5} w_{i,j,t}^{Ask} Dst_{i,j,t}^{Ask}}{\sum_{j=1}^{5} w_{i,j,t}^{Ask}}\right), \tag{3.2.1}$$

where $Dst_{i,j,t}^{Bid} = (Price_{i,j-1,t}^{Bid} - Price_{i,j,t}^{Bid})$ and $Dst_{i,j,t}^{Ask} = (Price_{i,j,t}^{Ask} - Price_{i,j-1,t}^{Ask})$ for security i, interval t, and the j^{th} price limit. For the distance at the first best limit, the midquote is used instead of the price at the next limit. $w_{i,j,t}$ are the quantities displayed at the offer or bid size, normalized by the total depth at the five best limits. As Kang and Yeo (2008) outline, dispersion is small under fierce competition since each trader wants to gain price priority. By gaining price priority, market members reduce the spread and the distance between each price limit, increasing the density of the order book.

As defined in Næs and Skjeltorp (2006), the slope measure is computed by averaging the price elasticities of quantities available at the five best quotes:

$$Slope_{i,t} = \frac{DE_{i,t} + SE_{i,t}}{2}, \tag{3.2.2}$$

where $DE_{i,t}$ and $SE_{i,t}$ are respectively the demand and supply elasticities. They are computed as follows:

$$DE_{i,t} = \frac{1}{5}\left(\frac{v_1^B}{|p_1^B/p_0 - 1|} + \sum_{\tau=1}^{4} \frac{v_{\tau+1}^B/v_\tau^B - 1}{|p_{\tau+1}^B/p_\tau^B - 1|}\right), \tag{3.2.3}$$

$$SE_{i,t} = \frac{1}{5}\left(\frac{v_1^A}{p_1^A/p_0 - 1} + \sum_{\tau=1}^{4} \frac{v_{\tau+1}^A/v_\tau^A - 1}{p_{\tau+1}^A/p_\tau^A - 1}\right). \tag{3.2.4}$$

p_τ^B and p_τ^A are the bid and ask prices displayed at quote τ. p_0 denotes the quoted midpoint. Finally, v_τ^B and v_τ^A are the natural logarithm of accumulated total share volume at limit τ for the bid and the ask respectively.[1]

A steep slope indicates that volumes in the order book are concentrated at a given limit (low elasticity) while a gentle slope suggests that volumes in the order book are not aggregated at a given limit (high elasticity). A steep slope indicates that traders would agree about the value of the security while a more gentle slope indicates that traders would rather have different views on the price of the security. As outlined by Glosten (1994), Goldstein and Kavajecz (2004) and Næs and Skjeltorp (2006), the slope is also negatively correlated with volatility, hence positively related to liquidity.

We also analyze trade-based proxies. We first look at the number of buyer and seller-initiated trades, NB and NS respectively. Using our database, we find the sign of the transaction by matching the orders that generate the trade and by comparing submission time of the orders. With this method, all trades are correctly categorized without any error margin. We also analyze the total number of trades, $Ntrades$, as well as the average trade size, ATS. We finally include Amihud (2002)'s illiquidity ratio that is specified as follows:

$$IL_t = \frac{|Return_t|}{Volume_t}. \tag{3.2.5}$$

This ratio is positively related to illiquidity. The higher the ratio, the more illiquid the security. Even if the return is present at the numerator of Amihud (2002)'s illiquidity ratio,

[1] By accumulated, we mean the sum of the quantities outstanding at that limit as well as the sum of all quantities outstanding at each better quote.

this measure is not closely related to the price movement variables that we investigate. First, the absolute return is adjusted for the volume. Second, the return is computed on a close-close basis, contrary to the price movement variables that are presented in the next section.

Descriptive statistics for both book-based and trade-based liquidity proxies defined above are given in the first panel of Table 3.1. As expected, the distributions of the liquidity proxies are heavily skewed and leptokurtic. Interestingly, we observe significant differences between the cap-based portfolios, suggesting that the characterization of liquidity may be different across these subsamples. In general, both skewness and kurtosis decrease when we move to large caps. The average percentage change for each liquidity proxy is however very low. This comes from the narrow time intervals that we use (15 minutes). For instance, dispersion changes on average by 1.68% for small caps, 8.21% for mid caps and 4.57% for large caps. Trading activity measures display negative signs for the first subsample, -0.1119 and -0.1143 for ΔNB or ΔNS, respectively. This is the consequence of some zero volume intervals that display a ΔNB or ΔNS equal to -1.

3.2.2 Price movement variables

For each of the 300 selected stocks, we rebuild HLOC prices over the whole sample period. As tick data are not appropriate for HLOC analysis, we use 15-minute intervals, leading to 34 intervals per day.[1] HLOC price dynamics are captured in different ways.

[1]This interval length is often used in the literature, e.g. Fiess and MacDonald (2002). In the robustness check section, we investigate whether our results are sensitive to the interval length and we apply the same methodology on 10-minute and 20-minute intervals.

Table 3.1: Descriptive statistics

	Small Caps				Mid Caps				Large Caps			
	Mean	STD	Skewness	Kurtosis	Mean	STD	Skewness	Kurtosis	Mean	STD	Skewness	Kurtosis
$\Delta Dispersion$	0.0168	0.25	33.777	3593.65	0.0821	25.63	557.47	310832	0.0457	0.42	34.12	4031.1
ΔRS	0.173	1.47	21.754	971.18	0.2923	2.16	98.54	27449.9	0.2671	1.4	51.295	10777.63
$\Delta Slope$	0.1067	0.84	13.23	336.67	0.1768	1.04	9.431	170.28	0.2009	0.92	6.281	139.17
ΔQB	0.0889	15.61	455.764	229267.8	1.2765	658.4	557.278	310684.6	0.1416	3.38	327.787	128235.5
ΔQA	0.0657	7.13	313.611	116066	1.245	610.29	553.733	307922.4	0.1383	4.69	503.639	268874.2
ΔQ	0.0543	9.17	373.24	163288.7	1.212	634.56	556.677	310226.4	0.0715	2.49	417.121	198502.2
$\Delta Ntrades$	0.0809	2.62	21.106	1089.19	0.3168	2.48	28.606	3535.27	0.2866	1.7	25.564	2085.72
ΔNB	-0.1119	2.72	36.546	2822.02	0.1666	2.37	12.552	405.49	0.4431	2.27	15.048	588.61
ΔNS	-0.1143	2.46	24.225	1477.76	0.1637	2.56	54.463	9487.19	0.3957	2.02	13.056	489.82
ΔATS	6.1549	442.76	185.764	38062.76	1.9356	34.23	86.631	10257.07	0.5125	10.52	105.565	15265.35
$\Delta Amihud$	15.7391	549.88	89.983	9635.14	27.7327	5221.07	335.447	114753.4	17.6412	7220.07	480.857	231313.4
ΔOC_t	-0.08832	2.09343	16.5353	530.26	0.17121	2.73511	14.6255	575.32	0.34819	2.19259	7.2744	142.79
$\Delta OCHL_t$	0.03881	1.4436	19.0019	872.42	0.17727	1.92326	15.3708	602.45	0.14639	1.38583	5.6033	78.22
ΔHL_t	0.05388	2.35805	16.129	438.41	0.28748	3.16102	30.1193	2588.7	0.29835	1.69443	26.0014	2738.66
ΔRV_t	0.28273	2.91399	16.3488	572.21	0.47407	3.06253	23.6848	1899.85	0.22444	1.42489	19.7264	1134.58

This table presents the descriptive statistics of both endogenous (first panel) and exogenous (second panel) variables of our models. The variables are presented in their first differences, as it is done in the models, for each subsample. QB and QA are the sums of the quantities outstanding at the five best limits, on bid and ask sides, respectively. Q is the sum of the two. $Slope$ denotes the slope of the book computed following Næs and Skjeltorp (2006) while $Dispersion$ stands for the dispersion measure computed following Kang and Yeo (2008). RS denotes the relative spread. NB and NS respectively denote the number of buy and sell trades. $Ntrades$ is the total number of trades and ATS is the average trade size of these trades. $Amihud$ denotes the Amihud ratio. $OC_t = \frac{|Close_t - Open_t|}{RangeMid_t}$ is the absolute value of the net movement. $OCHL_t = \frac{|Close_t - Open_t|}{High_t - Low_t}$ is the percentage of net variation on total fluctuation. HL_t denotes the High-Low range of the interval. The RV_t variable is the realized volatility.

Firstly, we use the absolute value of the price variation during interval t, i.e. $OC_t = \frac{|Close_t - Open_t|}{RangeMid_t}$, where $RangeMid_t = \frac{High_t + Low_t}{2}$. We take the absolute value in order to control for the bullish or bearish pressure that occurs during the interval. As outlined by Chakravarty and Holden (1995), Bloomfield et al. (2005), and Mazza (2013), a small difference between opening and closing prices may be related to a more liquid order book. When the fundamental value of the stock is located inside the spread, informed traders act less aggressively, take less liquidity in the book, and may even become dealers. As informed trading and liquidity demand are expected to be lower when the OC range is small, we expect OC_t to be negatively associated with liquidity. Therefore, all the liquidity proxies should exhibit a negative sign, except for the spread, dispersion, and Amihud variables.[1] The first hypothesis that we test is stated as follows:

H1: The wider the OC range implied by the price discovery process, the lower the liquidity.

Secondly, we include the High-Low range. It is computed as $HL_t = \frac{High_t - Low_t}{RangeMid_t}$, where $RangeMid_t = \frac{High_t + Low_t}{2}$. The HL range measures the total price movement. For a given OC range, a large HL range indicates that price uncertainty has been high, without being necessarily captured by volatility which relies on closing prices only.[2] We therefore expect HL_t and liquidity to be negatively related.[3]

The second hypothesis is formulated as follows:

H2: The wider the HL range (for a given OC range), the more uncertain the price discovery process, and the lower the liquidity.

[1]The relationship is expected to be positive in these three cases. The smaller the open-close range, the smaller the expected spread, the smaller the expected dispersion, and the lower the illiquidity ratio.

[2]The negative relationship between volatility and liquidity is documented in Benston and Hagerman (1974), Stoll (1978b), Stoll (1978a), Amihud and Mendelson (1980), Ho and Stoll (1981), Copeland and Galai (1983), Admati and Pfleiderer (1988), Foster and Viswanathan (1990), Stoll (2003), Pastor and Stambaugh (2003), Harris (2003) and, Beaupain et al. (2010), among others

[3]See Alizadeh et al. (2002) and De Vilder and Visser (2008).

Thirdly, we include an interaction term between the two previous variables by dividing the absolute value of the OC range by the HL range, i.e. $OCHL_t = \frac{|Close_t - Open_t|}{High_t - Low_t}$. This measure is equal to 1 when there is no price variation beyond the opening and closing prices. In this case, the two ranges coincide and the price discovery process is rather smooth, unidirectional, and not much affected by uncertainty. All else equal, liquidity is expected to improve in this environment. It is the opposite when the ratio is close to 0. When the price discovery process implies very large fluctuations relative to the open-close price change observed during the interval, liquidity is expected to deteriorate. All else equal, a significantly positive coefficient for the interaction variable would indicate that liquidity is further decreased when the OC range is smaller than the HL range. While controlling for the variation in both the OC and HL ranges, the OCHL variable is therefore positively related to liquidity. The third hypothesis is:

H3: The lower the OCHL ratio (for a given OC and HL ranges), the higher the negative impact of price uncertainty on liquidity.

Descriptive statistics for the price movement variables are given in the second panel of Table 3.1. As for liquidity proxies, these descriptive statistics clearly outline strong differences between the different quartiles.

Finally, we add four dummy variables. The "Zero Return" dummy (ZR_t) controls for the presence of zero returns. It is equal to 1 when there is a zero return and 0 otherwise. Lesmond et al. (1999) indicate that zero returns are associated with high transactions costs. They argue that informed traders do not trade as the value of the new information does not exceed the cost to trade. However, based on Chakravarty and Holden (1995) and Bloomfield et al. (2005), Mazza (2013) finds that (quasi) zero returns are moments when there is less informed trading and more liquidity. As a result, informed traders do not trade to move price towards their fundamental value but behave as dealers to earn the spread, supplying liquidity, which in turn produces a zero return. As a consequence, liquidity provision is actually higher.

We can state the fourth hypothesis as follows.

H4: When there is a zero return, liquidity is higher.

The "Upward Window" dummy (UW_t) controls for the presence of a price jump between the highest price observed during interval $t-1$ and the lowest price observed during interval t. The dummy is equal to 1 when there is an upward window, and 0 otherwise. The "Downward Window" dummy (DW_t) controls for the occurrence of a price jump between the lowest price observed during interval $t-1$ and the highest price observed during interval t. The dummy is equal to 1 in case of a downward window and 0 otherwise. The "Upward Window" and "Downward Window" dummies are expected to be negatively related to liquidity as the formation of a window is typically linked to an illiquid state of the order book, as outlined by Boudt and Petitjean (2013) among others.

H5: When there has been a price jump, liquidity is lower.

The last dummy is a "Direction" dummy (D_t) which is equal to 1 when there is a positive price movement and 0 otherwise. This dummy is added to control for the link between price trend and liquidity. In comparison with downward price movements, we expect upward price movements to lead to a decrease in both depth and order imbalances.[1] In the case of an upward price movement, QB becomes higher since many market participants want to participate to the upward discovery process led by informed traders. QA drops as sellers do not want to trade with better informed traders. Furthermore, since liquidity is abundant on the other side, sellers will submit aggressive marketable orders that generate more seller-initiated trades, i.e. NS is higher and NB is lower.

H6: When there is an upward price movement, both depth and order imbalances de-

[1]Order imbalance is equal to the number of buyer minus seller-initiated trades, i.e. NB - NS. Depth imbalance means the quantities outstanding at the five best limits for the ask side (QA) minus the quantities outstanding at the five best limits for the bid side (QB).

crease.

3.2.3 Model specification

The fixed-effect panel specification that we estimate is the following:

$$\begin{aligned}\Delta L_{i,t} = \alpha_0 + \alpha_1 \Delta OC_{i,t} + \alpha_2 \Delta OCHL_{i,t} + \alpha_3 \Delta HL_{i,t} \\ +\alpha_4 UW_{i,t} + \alpha_5 DW_{i,t} + \alpha_6 D_{i,t} + \sum_{ß=2}^{100} \beta_i S_i + \nu_{i,t}\end{aligned} \tag{3.2.6}$$

where $L_{i,t}$ is the liquidity proxy measured at interval t for stock i, and Δ indicates the percentage change of the variable, i.e. $\Delta L_{i,t} = \frac{L_{i,t} - L_{i,t-1}}{L_{i,t-1}}$. The $n-1$ S_i dummies denote the stocks' fixed effects that are present in each portfolio. As outlined in Chordia et al. (2000), we use first differences in order to deal with econometric issues that affects liquidity variables, such as high persistence and/or non-stationarity.

In order to control for volatility in the analysis, we run the same regressions with an additional variable, RV_t, which is the realized volatility. It has been shown to be an efficient estimate of volatility (De Vilder and Visser, 2008). We compute the realized volatility by aggregating 1-minute returns over 15 minutes. The objective of this specification is to check whether HLOC dynamics remain significant when the effect of volatility on liquidity is controlled for. This new model is specified as follows:

$$\begin{aligned}\Delta L_{i,t} = \alpha_0 + \alpha_1 \Delta OC_{i,t} + \alpha_2 \Delta OCHL_{i,t} + \alpha_3 \Delta HL_{i,t} \\ +\alpha_4 UW_{i,t} + \alpha_5 DW_{i,t} + \alpha_6 D_{i,t} + \alpha_7 RV_{i,t} + \sum_{ß=2}^{100} \beta_i S_i + \nu_{i,t}\end{aligned} \tag{3.2.7}$$

As multicollinearity may be an issue, Table 4.2 presents Pearson correlation coefficients

between the variables included in these specifications. By using the first differences, we observe that correlation typically drops and never goes above 0.60 (Table 3.3).

Table 3.2: Correlation Matrix - levels

	OC_t	$OCHL_t$	HL_t	RV_t
OC_t	1.00			
$OCHL_t$	0.36	1.00		
HL_t	0.77	0.03	1.00	
RV_t	0.52	0.02	0.70	1.00

This table presents the correlation between the numeric variables of the model. OC_t denotes the relative absolute value of the net movement, $OCHL_t$ denotes the ratio of net movement to total fluctuation and HL_t is the High-Low range. RV_t is the realized volatility.

Table 3.3: Correlation Matrix - first differences

	ΔOC_t	$\Delta OCHL_t$	ΔHL_t	ΔRV_t
ΔOC_t	1.00			
$\Delta OCHL_t$	0.50	1.00		
ΔHL_t	0.60	-0.04	1.00	
ΔRV_t	0.37	0.01	0.51	1.00

This table presents the correlation between the first differences of the numeric variables of the model. OC_t denotes the relative absolute value of the net movement, $OCHL_t$ denotes the ratio of net movement to total fluctuation and HL_t is the High-Low range. RV_t is the realized volatility.

3.3 Empirical analysis

Before running the panel regression models in the next section, we more closely examine the distribution of HLOC price movement variables by zooming on the first, fifth and tenth decile of their distribution. The first decile includes the narrowest price movements while

the tenth decile includes the largest price movements. For each of these three deciles, we compute all the liquidity proxies defined in the previous section. We then conduct non parametric median tests to assess whether the three subsamples display significantly different medians for each liquidity proxy. Under the null hypothesis, medians are equal. The goal is to compare the state of liquidity when HLOC price variations are significantly different. We carry out this analysis for the three portfolios of small, mid and large caps.

Table 3.4 reports the median scores. All the results of the median tests are highly significant at a 1% significance level. Clearly, ex-ante liquidity deteriorates as price ranges increase. In general, the mean scores of (il)liquidity proxies drop (rise) when the price range becomes larger. As an illustration, we show in Figure 3.2 the change in median scores for the relative spread across the three portfolios and the three deciles.

For large caps, the relative spread more than doubles when narrow and wide HLOC price variations are compared. A wide OC range indicates that the price discovery process has led to a new market value that is significantly different from the market value that prevailed before. When compared to a narrow OC range, there is a significant widening of the spread for large caps. The same result is found when we look at the effect of total price variation on the relative spread. In comparison to narrow HL ranges, wide HL ranges increase the relative spread significantly. The dispersion is also positively impacted by a rise in total price variation. It is nevertheless much more affected when the price discovery process is more intense: If narrow and wide OC ranges are compared, the dispersion goes from 0.1986 to 0.5925. Slope for large caps follow the oppositive pattern than the spread, confirming that liquidity worsens when HLOC price variations are wider. We observe that depth is much less affected by variations in the HLOC prices than the previous (il)liquidity proxies. It decreases with the intensity of the price discovery process but slightly rises with total price variation. Regarding the number of trades and the average trade size, there is a sharp and undeniable rise when total price variation increases, which explains the rise in (total) volume. However, the positive change in the Amihud ratio indicates that the rise in total volume is insufficient

to counterbalance the rise in the absolute price change that occurs when narrow and wide HLOC price variations are compared.

For mid caps, the overall picture is very similar: liquidity drops when HLOC price variations become wider. The only noticeable difference is the slight decrease in the Amihud ratio when narrow and wide total price variations compared.

Small caps exhibit more positive findings with respect to ex-post liquidity only. When HLOC price changes widen, the relative spread, dispersion, and slope point to a sharp deterioration in ex-ante liquidity. However, the resulting sharp rise in total volume (due to both the number of trades and the average trade size) leads to a lower Amihud ratio, which is especially noticeable when changes in total price variation are considered.

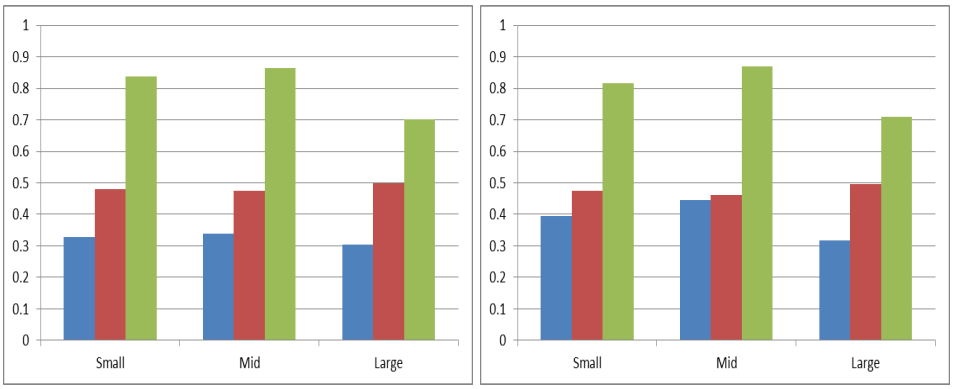
Table 3.4: A first look at liquidity and HLOC price movements

	Range size	OC_t			HL_t		
Proxy		Small	Mid	Large	Small	Mid	Large
RS	Narrow	0.3285	0.3397	0.3042	0.3947	0.4447	0.3174
	Medium	0.4792	0.4745	0.4994	0.4735	0.4606	0.4968
	Wide	0.8376	0.8643	0.7001	0.8167	0.8704	0.7081
Dispersion	Narrow	0.7552	0.3935	0.1986	0.8063	0.5573	0.5452
	Medium	0.4887	0.5003	0.5261	0.4826	0.4897	0.4821
	Wide	0.3354	0.6038	0.5925	0.3328	0.5248	0.5982
Slope	Narrow	0.6588	0.6627	0.7008	0.6033	0.5425	0.6699
	Medium	0.5216	0.5218	0.5012	0.5262	0.5359	0.5066
	Wide	0.1687	0.1631	0.2894	0.1875	0.1707	0.2773
Q	Narrow	0.1483	0.4730	0.5071	0.1336	0.3958	0.3931
	Medium	0.5134	0.5056	0.5066	0.5150	0.5073	0.5231
	Wide	0.7443	0.4819	0.4399	0.7463	0.5455	0.4218
Ntrades	Narrow	0.4589	0.4775	0.4507	0.3488	0.2311	0.1640
	Medium	0.5062	0.5012	0.4894	0.5051	0.5162	0.5147
	Wide	0.4911	0.5128	0.6340	0.6106	0.6392	0.7183
ATS	Narrow	0.1454	0.4318	0.5753	0.1383	0.3642	0.3372
	Medium	0.5062	0.4968	0.4874	0.4954	0.4921	0.5160
	Wide	0.7864	0.6046	0.5259	0.7842	0.6837	0.5320
Volume	Narrow	0.3602	0.4688	0.4901	0.3225	0.2759	0.1974
	Medium	0.5066	0.4990	0.4883	0.5056	0.5083	0.5198
	Wide	0.5867	0.5394	0.6036	0.6327	0.6581	0.6446
Amihud	Narrow	0.4979	0.3681	0.2584	0.5961	0.5262	0.4467
	Medium	0.5111	0.5062	0.5052	0.5109	0.4995	0.4896
	Wide	0.4130	0.5929	0.7042	0.3570	0.4812	0.6384

This table presents the outcomes of the non-parametric median test for the three market capitalization portfolios: small, mid, and large. For each variable, Narrow, Medium, and Wide respectively denote the first 10 percentiles, the mid of the sample, and the last 10 percentiles of the subsamples based on the range size. QB and QA are the sums of the quantities outstanding at the five best limits, on bid and ask sides, respectively. Q is the sum of the two. $Slope$ denotes the slope of the book computed following Næs and Skjeltorp (2006) while $Dispersion$ stands for the dispersion measure computed following Kang and Yeo (2008). RS denotes the relative spread. NB and NS respectively denote the number of buy and sell trades. $Ntrades$ is the total number of trades and ATS is the average trade size of these trades. $Amihud$ denotes the Amihud ratio.

Figure 3.2: HLOC price movements and relative spread

(a) Close-Open Range **(b)** High-Low Range



This figure presents the mean score of the median tests for both Close-Open and High-Low ranges for the relative spread. For each capitalization group (Small, Mid and Large), the three bars represent from the left to the right, the mean scores of the first decile, the mid of the sample and the last decile.

3.3.1 General panel regressions

For each liquidity proxy defined in Section 3.2.1, we study the relationship between liquidity and price movements by estimating the fixed-effect panel regression model defined in Section 3.2.3. We first estimate the model by using the clustered standard errors approach that is described in Petersen (2009) to correct for the bias in the OLS standard errors that arises for panel estimation. The standard errors produced by this method are robust to within cluster correlation. As outlined by Petersen (2009), this method produces unbiased standard errors when there is a firm-specific effect and is superior to White, Newey-West, and Fama-MacBeth correction methods in such a case. We split the sample by market capitalization and run the analysis separately on each of the three portfolios. We also employ robust and median regression techniques in order to deal with some of the limits of the traditional OLS technique, among which non-normality.

The clustered OLS results suggest that there is a statistically relevant relationship between HLOC price dynamics and liquidity (see Table 3.5). We validate our six hypotheses. The magnitude of this relationship varies among the liquidity dimensions but the results are similar and only differ in the significance level in most cases.

Table 3.5: Regressions with clustered standard errors

ΔL_t	ΔOC_t	ΔHL_t	$\Delta OCHL_t$	ZR_t	UW_t	DW_t	D_t
Panel A : Small Caps							
ΔRS	-0.001	0.179^{***}	0.011	0.246^{***}	-0.003	0.138^{*}	0.056^{***}
$\Delta Dispersion$	0.009^{**}	0.016^{***}	-0.002	0.031^{***}	0.079^{***}	0.085^{***}	0.005^{*}
$\Delta Slope$	0.001	-0.026^{***}	0.034^{***}	-0.011	0.163^{***}	0.030	0.004
ΔQB	0.006	0.009^{**}	0.000	-0.001	0.002	0.099^{***}	0.075^{***}
ΔQA	0.006^{**}	0.004^{**}	-0.000	0.087^{***}	0.116^{***}	-0.036^{*}	-0.080^{***}
ΔQ	0.004^{**}	0.004^{**}	-0.001	0.039^{***}	0.022	0.011	-0.010^{***}
ΔNB	0.143^{***}	0.297^{**}	-0.068^{***}	0.637^{***}	2.648^{***}	-0.625^{***}	-0.524^{***}
ΔNS	0.015	0.195^{***}	-0.014	0.109	-0.013	1.642^{***}	0.513^{***}
$\Delta Ntrades$	0.076^{***}	0.199^{***}	-0.043^{***}	0.344^{***}	1.161^{***}	0.612^{***}	-0.039
ΔATS	-0.004	-0.012	-0.025^{**}	0.155	-0.091	0.001	-0.430
$\Delta Volume$	0.184^{***}	0.271^{***}	-0.129^{***}	1.503^{**}	1.839^{***}	0.727^{**}	-0.285^{**}
$\Delta Amihud$	0.008	0.059	0.643^{***}	1.041^{*}	5.950^{**}	2.131^{**}	-1.203
Panel B : Mid Caps							
ΔRS	0.070^{***}	0.066^{**}	-0.061^{***}	-0.001	0.029	0.108^{*}	0.054^{**}
$\Delta Dispersion$	0.037	0.009^{*}	-0.043	0.003	-0.216	-0.033	-0.407
$\Delta Slope$	-0.021^{***}	-0.009	0.041^{***}	-0.020	0.233^{***}	0.079^{**}	-0.024^{**}
ΔQB	0.004	0.003	-0.001	-0.052^{***}	-0.005	0.093^{***}	0.082^{***}
ΔQA	0.005^{*}	0.001	0.006	0.048^{***}	0.163^{***}	-0.016	-0.120^{***}
ΔQ	0.002	0.001	0.003	0.008	0.041^{***}	0.015	-0.021^{***}
ΔNB	0.089^{***}	0.056	-0.088^{***}	0.306^{***}	0.998^{***}	-0.513^{***}	-0.584^{***}
ΔNS	0.027^{*}	0.134^{***}	-0.054^{***}	-0.266^{***}	-0.400^{***}	0.729^{***}	0.610^{***}
$\Delta Ntrades$	0.023^{**}	0.119^{***}	-0.045^{***}	0.040	0.376^{***}	0.199^{***}	-0.005
ΔATS	-0.000	0.010	-0.034^{***}	-0.149	-0.034	0.112	0.084
$\Delta Volume$	0.053^{*}	0.249^{***}	-0.145^{***}	-0.042	0.553^{***}	0.237	-0.027
$\Delta Amihud$	0.494^{**}	-0.259^{*}	0.665^{**}	1.305^{**}	6.509^{***}	7.222^{***}	0.482
Panel C : Large Caps							
ΔRS	0.037^{***}	0.142^{***}	-0.043^{***}	-0.111^{***}	0.109^{***}	-0.015	0.033^{***}
$\Delta Dispersion$	0.015^{***}	0.026^{***}	-0.015^{***}	-0.018^{***}	0.093^{***}	0.054^{***}	0.003
$\Delta Slope$	-0.023^{***}	-0.075^{***}	0.057^{***}	-0.023^{**}	0.124^{***}	0.197^{***}	-0.021^{***}
ΔQB	0.000	-0.005	0.008^{*}	-0.071^{***}	-0.076^{***}	0.012	0.034^{***}
ΔQA	0.002	-0.006^{**}	0.005	-0.025^{***}	-0.031	-0.097^{***}	-0.055^{***}
ΔQ	-0.001	-0.007^{***}	0.005^{**}	-0.020^{***}	-0.072^{***}	-0.057^{***}	-0.010^{***}
ΔNB	0.022	0.329^{***}	-0.075^{***}	0.248^{***}	0.068	-0.050	-0.538^{***}
ΔNS	-0.004	0.295^{***}	-0.054^{***}	-0.270^{***}	-0.080^{*}	-0.018	0.469^{***}
$\Delta Ntrades$	0.019^{*}	0.246^{***}	-0.052^{***}	-0.026^{**}	0.083^{*}	-0.002	-0.009
ΔATS	-0.009^{*}	0.092^{***}	0.002	0.090	-0.049	0.057	-0.022
$\Delta Volume$	-0.040	0.632^{***}	-0.028	0.183	0.304	-0.062	-0.064
$\Delta Amihud$	0.353^{***}	-0.167	0.752^{***}	0.367^{***}	19.938^{***}	8.470^{***}	-0.193

This table presents the different models that are estimated with clustered standard errors. Panels A, B and C display the estimates for small, mid and large caps, respectively. The first column, L_t, indicates the dependant variable and the following columns are the exogenous variables. QB and QA are the sums of the quantities outstanding at the five best limits, on bid and ask sides, respectively. $Q5$ is the sum of the two. $Slope$ denotes the slope of the book computed following Næs and Skjeltorp (2006) while $Dispersion$ stands for the dispersion measure computed following Kang and Yeo (2008). RS denotes the relative spread. NB and NS respectively denote the number of buy and sell trades. $Ntrades$ is the total number of trades and ATS is the average trade size of these trades. $Amihud$ denotes the Amihud ratio. $OC_t = \frac{|Close_t - Open_t|}{RangeMid_t}$ is the absolute value of the net movement. $OCHL_t = \frac{|Close_t - Open_t|}{High_t - Low_t}$ is the percentage of net variation on total fluctuation. HL_t denotes the High-Low range of the interval. ZR_t is a dummy variable that control for the presence of zero returns. UW_t is equal to 1 when an upward window occurs on the chart and 0 otherwise, while DW_t is equal to 1 when a downward window appears and 0 otherwise. The D_t dummy equals 1 when the net price movement is positive and 0 otherwise. The fixed effect dummies are not presented here.

In the discussion below, we check whether each of the six hypotheses stated in Section 3.2.2 is verified or not.

H1: The wider the OC range implied by the price discovery process, the lower the liquidity.

Liquidity drops when the OC range increases, validating the first hypothesis. When significant, the relative spread, dispersion, slope and Amihud estimates have all the expected signs. The large cap portfolio exhibits the most robust results: all four proxies are significant. The relative spread for mid caps is particularly affected by a rise in the OC range. Interestingly, liquidity in small caps seems to be less sensitive to the intensity of the price discovery process. However, depth responds to changes in the OC range to a larger extent for the small cap portfolio than for the other two. Given that small caps are typically less followed by analysts, we conjecture that a more intense price discovery process is viewed by traders as an indication that the market is more active, encouraging traders to submit a higher number of limit orders. Trading activity measures point in the same direction. Overall, a more intense price discovery for small caps and mid caps leads to a larger number of trades and a higher trading volume, while the average trade size remains unaffected. For large caps, such evidence is much weaker. Finally, ex-post liquidity, as measured by the Amihud ratio, is also negatively affected by the intensity of the price discovery, except for small caps where it has no effect. All else equal, the intensity of the price discovery process lowers ex-ante and ex-post liquidity, the relationship being the strongest for large and mid caps.

H2: The wider the HL range (for a given OC range), the more uncertain the price discovery process, and the lower the liquidity.

The second hypothesis is also verified since HL_t is negatively correlated with liquidity for all portfolios, as clearly indicated by the relative spread and the dispersion measures. As shown in Table 3.5, the most significant results are found for the large and small cap

portfolios. Overall, the level of liquidity for small caps is found to be much more sensitive to the total price variation than to the price discovery process itself. For small caps, hypothesis 2 is indeed verified across a higher number of liquidity proxies than hypothesis 1. There is no big difference between hypotheses 1 and 2 with respect to depth and trading activity, except maybe for large caps. The higher level of trading volume for large caps is nevertheless insufficient for the Amihud ratio to be negative *and* statistically significant. For mid caps, the Amihud ratio displayed is negative *and* statistically significant, but at 10% only. All else equal, a wider total price variation deteriorates ex-ante liquidity, while it does not seem to the case for ex-post liquidity.

H3: The lower the OCHL ratio (for a given OC and HL ranges), the higher the negative impact of price uncertainty on liquidity.

The third hypothesis is validated for the ex-ante liquidity proxies only. When OC and HL ranges variations are controlled for, the interaction variable (measuring 'relative price certainty') displays a positive relationship to ex-ante liquidity. Although it holds true for any of the three portfolios, large caps display the most convincing results. The relative spread, dispersion, and slope measures for the large cap portfolio are all three significant at 1%, with the expected sign. Although the OCHL interaction term has very weak influence on depth, it is clearly the opposite with respect to trading activity. The number of trades increases when 'relative' price uncertainty increases (i.e. when OCHL decreases). ATS and Volume follow the same pattern for the small and mid caps only. The 'volume' effect is sufficiently strong for the Amihud ratio to display a positive and significant sign. As a consequence, the negative effect on ex-post liquidity implied by a more intense price discovery could be counterbalanced by the high level of 'relative price uncertainty' that might prevail at the same time.

H4: When there is a zero return, liquidity is higher.

Evidence for the fourth hypothesis is mixed. Supporting results are found for large caps. The relative spread and the dispersion decreases when there is a zero return for large caps. However, the slope of the book decreases, pointing to a higher price elasticity of the quantities. A zero return leads to lower available quantities and a lower number of trades, indicating that there are fewer trades and available quantities when a zero return occurs. ATS and Volume increases but the effect is not significant. Ex-ante liquidity for mid caps does not seem to be sensitive to the occurrence of a zero return, in contrast to the small cap portfolio. Dispersion, spread, depth, and number of trades all increase when a zero return occurs for small caps, while it was the opposite for large caps. As indicated in Table 3.5, the sum of quantities over the two sides, ΔQ, shows a significant positive estimate (0.039^{**}) which suggests that depth increases on average for small caps when a zero return occurs. The improving liquidity provision, accompanied by a rise in the number of trades, is rather inconsistent with the findings of Lesmond et al. (1999) but in line with the propositions of Chakravarty and Holden (1995), Bloomfield et al. (2005) and Mazza (2013).

The fourth hypothesis is not validated with respect to the Amihud ratio for all three portfolios since the estimates are significant and positive. However, the relation between the Amihud ratio and the ZR_t dummy must be interpreted with caution. The absolute return in the Amihud ratio is based on closing prices only while the ZR_t dummy controls for moments when the OC range is equal to zero. As a consequence, the absolute return may be different from zero even if the ZR_t dummy is equal to 1. In other words, ex-post liquidity may be affected by the difference between the last closing price the current opening price, which is independent of the characterization of a zero return based on the OC range.

H5: When there is a price jump, liquidity is lower.

The fifth hypothesis is clearly validated when it comes to ex-post liquidity. For all three portfolios, there is a sharp increase in the Amihud ratio, which is affected by construction since the numerator of the ratio is equal to the absolute return. For both small and mid

caps, the rise in volume is not sufficient to counterbalance the increase in the absolute return. Ex-ante liquidity also seems to be negatively affected by the occurrence of price window dummies, especially for large caps where the relative spread and dispersion increase significantly. The slope estimate is positive and significant, suggesting that volumes in the order book are more concentrated at a given limit after the occurrence of a price jump. Although both the dispersion and the relative spread increase, the agreement among traders about the value of the security would be higher after the occurrence of a price jump (than without). Depth proxies indicate that the variation in depth is dependent on the direction of the window. When an upward window occurs, QB seems to drop while QA seems to rise. When a downward window occurs, QA seems to fall while QB seems to increase. As expected, trading activity estimates are significant for both NB_t and NS_t and the signs are opposed to the corresponding depth proxies. For instance, depth at the bid drops and the number of buyer-initiated trades is higher when an upward window occurs, pointing to higher trading aggressiveness. Correspondingly, depth at the ask drops and the number of seller-initiated trades is higher in the case of a downward window.

H6: When there is an upward price movement, both depth and order imbalances decrease.

The sixth hypothesis is validated. First, we observe that QB (QA) is higher (lower) when there is an upward price movement, implying that depth imbalance decreases. Second, NB also drops while NS increases, reducing the order imbalance. Such a result is explained by the increase in aggressive trading when liquidity is abundant on the other side of the book.

Overall, there is strong support for the six hypothesis stated in Section 3.2.2. We first observe that liquidity drops when OC_t or HL_t increases. Liquidity is also negatively affected by the occurrence of a price jump, whatever its direction. All else equal, ex-ante liquidity is positively related to the evolution of $OCHL_t$. All this is consistent with the literature on the relations between liquidity, volatility and price dynamics. Trading activity measures are

also related to the evolution of HLOC dynamics and the evolution of the Amihud illiquidity ratio broadly supports the findings related to the book-based proxies. Although the study has proven to be relevant for all the market capitalization portfolios, supporting evidence is stronger for large caps than for small or mid caps since results become more significant when we move from small or mid caps to the large cap portfolio.

3.3.2 Controlling for volatility

The results obtained when the realized volatility is included in the regressions are presented in Table 3.6.

Even when volatility is controlled for, the parameter estimates of the HLOC price movement variables remain significant. All these results suggest that the dynamics of HLOC price variations capture changes in liquidity that cannot be explained by the realized volatility. Overall, all dimensions of liquidity are significantly affected by HLOC price movements. HLOC price dynamics may therefore contain additional information when it comes to explaining liquidity on the stock market.

Let us study the six hypotheses again in more details to identify the most striking changes.

H1: The wider the OC range implied by the price discovery process, the lower the liquidity.

There is no significant change in the results between Tables 3.5 and 3.6, confirming that 'the intensity of the price discovery process lowers ex-ante and ex-post liquidity, the relationship being the strongest for large and mid caps'.

H2: The wider the HL range (for a given OC range), the more uncertain the price discovery process, and the lower the liquidity.

All else equal, a wider total price variation deteriorates ex-ante liquidity, as previously outlined. In Table 3.6, results are even slightly strengthened since the slope for mid caps is negative and statistically significant (while it was not in 3.5). Table 3.6 also reinforces the previous findings regarding the Amihud ratio: ex-post liquidity does not deteriorate. On the contrary, the ratio is negative and statistically significant at 5% for both mid and large caps.

Table 3.6: Controlling for volatility

ΔL_t	ΔOC_t	ΔHL_t	$\Delta OCHL_t$	ZR_t	UW_t	DW_t	D_t	ΔRV_t
Panel A : Small Caps								
ΔRS	0.032	0.173^{***}	-0.031^{**}	-0.152^{***}	-0.133^{*}	0.050	0.090^{**}	-0.034
$\Delta Dispersion$	0.015^{***}	0.008^{**}	-0.010^{***}	-0.019^{**}	0.067^{***}	0.079^{***}	0.007	0.001
$\Delta Slope$	-0.006	-0.052^{***}	0.029^{**}	-0.045	0.074	-0.070	-0.020	0.045^{**}
ΔQB	0.006	0.008	-0.004	-0.123^{***}	-0.012	0.026	0.166^{***}	-0.000
ΔQA	0.003	0.005	-0.004^{**}	0.066^{***}	0.000	-0.038	-0.201^{***}	-0.005^{**}
ΔQ	0.003^{*}	0.006^{**}	-0.002	0.008	-0.018	0.002	-0.029^{***}	-0.003^{*}
ΔNB	0.115^{***}	0.237^{**}	-0.098^{***}	0.343^{***}	1.963^{***}	-0.961^{***}	-1.008^{***}	0.129^{*}
ΔNS	0.006	0.136^{***}	-0.061^{***}	-0.739^{***}	-0.384^{**}	1.103^{***}	0.970^{***}	0.052^{**}
$\Delta Ntrades$	0.061^{***}	0.159^{***}	-0.076^{***}	-0.144^{**}	0.734^{***}	0.287^{***}	-0.082	0.050
ΔATS	-0.002	-0.008	-0.026^{**}	0.154	-0.065	0.020	-0.418	-0.011
$\Delta Volume$	0.155^{***}	0.197^{**}	-0.186^{***}	0.737	1.065^{**}	0.180	-0.457^{*}	0.108
$\Delta Amihud$	-0.117	-0.211	0.689^{***}	1.206^{**}	4.761^{*}	1.467	-1.235	0.822^{***}
Panel B : Mid Caps								
ΔRS	0.069^{***}	0.085^{***}	-0.067^{***}	-0.126^{***}	-0.082	-0.005	0.084^{***}	0.023^{*}
$\Delta Dispersion$	0.028^{*}	-0.010	-0.049	-0.024^{**}	-0.436	-0.161	-0.598	0.074
$\Delta Slope$	-0.022^{***}	-0.013^{*}	0.038^{***}	-0.055^{***}	0.193^{***}	0.077^{**}	-0.031^{**}	-0.012^{***}
ΔQB	0.004	0.002	-0.003	-0.099^{***}	-0.036^{**}	0.066^{***}	0.106^{***}	0.006
ΔQA	0.004	0.001	0.005	0.038^{***}	0.133^{***}	-0.022	-0.163^{***}	0.002
ΔQ	0.001	0.001	0.003	-0.003	0.025^{*}	0.012	-0.030^{***}	0.001
ΔNB	0.060^{***}	0.072^{***}	-0.089^{***}	0.149^{***}	0.660^{***}	-0.775^{***}	-0.755^{***}	0.087^{***}
ΔNS	0.014	0.105^{***}	-0.069^{***}	-0.608^{***}	-0.617^{***}	0.448^{***}	0.754^{***}	0.060^{***}
$\Delta Ntrades$	0.009	0.139^{***}	-0.053^{***}	-0.170^{***}	0.150^{***}	-0.010	-0.020	0.030^{***}
ΔATS	0.001	0.020^{*}	-0.034^{***}	-0.147	-0.024	0.118	0.086	-0.015
$\Delta Volume$	0.028	0.323^{***}	-0.159^{***}	-0.465^{*}	0.122	-0.173	-0.067	0.011
$\Delta Amihud$	0.419^{**}	-0.444^{***}	0.700^{***}	1.397^{**}	5.837^{***}	6.544^{***}	0.469	0.690^{***}
Panel C : Large Caps								
ΔRS	0.032^{***}	0.093^{***}	-0.037^{***}	-0.113^{***}	-0.077^{**}	-0.155^{***}	0.032^{***}	0.119^{***}
$\Delta Dispersion$	0.014^{***}	0.019^{***}	-0.014^{***}	-0.018^{***}	0.058^{***}	0.027	0.003	0.021^{***}
$\Delta Slope$	-0.021^{***}	-0.066^{***}	0.054^{***}	-0.026^{***}	0.170^{***}	0.239^{***}	-0.022^{***}	-0.029^{***}
ΔQB	0.000	-0.002	0.007^{*}	-0.073^{***}	-0.068^{***}	0.022	0.035^{***}	-0.007^{*}
ΔQA	0.003	-0.002	0.004	-0.027^{***}	-0.015	-0.081^{***}	-0.057^{***}	-0.012^{***}
ΔQ	-0.001	-0.002	0.005^{**}	-0.021^{***}	-0.053^{***}	-0.041^{***}	-0.010^{***}	-0.013^{***}
ΔNB	0.020	0.282^{***}	-0.072^{***}	0.240^{***}	-0.100	-0.163^{**}	-0.542^{***}	0.093^{***}
ΔNS	-0.005	0.277^{***}	-0.054^{***}	-0.281^{***}	-0.142^{***}	-0.075	0.473^{***}	0.034^{***}
$\Delta Ntrades$	0.019^{*}	0.236^{***}	-0.053^{***}	-0.035^{**}	0.043	-0.028	-0.009	0.016
ΔATS	-0.010^{*}	0.088^{***}	0.003	0.091	-0.063	0.046	-0.022	0.010
$\Delta Volume$	-0.038	0.633^{***}	-0.031	0.174	0.305	-0.048	-0.064	-0.016
$\Delta Amihud$	0.312^{***}	-0.803^{**}	0.828^{***}	0.375^{***}	17.819^{***}	6.970^{***}	-0.204	1.360^{**}

This table presents the different models that are estimated with clustered standard errors with a control variable for the realized volatility. Panels A, B and C display the estimates for small, mid and large caps, respectively. The first column, L_t, indicates the dependant variable and the following columns are the exogenous variables. QB and QA are the sums of the quantities outstanding at the five best limits, on bid and ask sides, respectively. $Q5$ is the sum of the two. $Slope$ denotes the slope of the book computed following Næs and Skjeltorp (2006) while $Dispersion$ stands for the dispersion measure computed following Kang and Yeo (2008). RS denotes the relative spread. NB and NS respectively denote the number of buy and sell trades. $Ntrades$ is the total number of trades and ATS is the average trade size of these trades. $Amihud$ denotes the Amihud ratio. $OC_t = \frac{|Close_t - Open_t|}{RangeMid_t}$ is the absolute value of the net movement. $OCHL_t = \frac{|Close_t - Open_t|}{High_t - Low_t}$ is the percentage of net variation on total fluctuation. HL_t denotes the High-Low range of the interval. ZR_t is a dummy variable that control for the presence of zero returns. UW_t is equal to 1 when an upward window occurs on the chart and 0 otherwise, while DW_t is equal to 1 when a downward window appears and 0 otherwise. The D_t dummy equals 1 when the net price movement is positive and 0 otherwise. The fixed effect dummies are not presented here.

H3: The lower the OCHL ratio (for a given OC and HL ranges), the higher the negative impact of price uncertainty on liquidity.

Results in Tables 3.5 and 3.6 are very similar. The validation of H3 is even stronger for the small cap portfolio since the relative spread and the dispersion are both negative and significant at 1% (while they were not before).

H4: When there is a zero return, liquidity is higher.

Evidence for the fourth hypothesis in Table 3.6 is much stronger than before. Controlling for realized volatility, both the relative spread and the dispersion are now significant and display the expected signs for all three portfolios. In Table 3.5, H4 was only validated for large caps.

H5: When there has been a price jump, liquidity is lower.

Evidence in support of the fifth hypothesis is weaker than before. The relative spread is now negative and significant for small and large caps. This finding is somewhat counterbalanced by the slope variable which is no longer positive and significant in the small cap portfolio. As before, the Amihud ratio increases overall, again validating H5 when it comes to ex-post liquidity. Results are nevertheless weaker for the small cap portfolio since the Amihud ratio is only significant at 10% after the occurrence of a downward window.

H6: When there is an upward price movement, both depth and order imbalances decrease.

There is no difference between Tables 3.5 and 3.6. The sixth hypothesis is again validated.

3.3.3 Robust and median regressions

We apply the robust and median regression techniques to our panel models in order to identify whether the results are affected by the presence of outliers. These types of regressions are useful when some hypotheses of the ordinary least squares estimation are potentially violated. This is the case in our model as shown by the descriptive statistics presented in Table 3.1. These two specifications also provide an interesting snapshot to circumvent the mean as a measure of central tendency for the estimates. The outcomes are presented in Table 3.7 and 3.8, respectively. We report the results with realized volatility as a control variable since it was significant in many cases in Table 3.6.

Table 3.7: Robust regressions with realized volatility

ΔL_t	ΔOC_t	ΔHL_t	$\Delta OCHL_t$	ZR_t	UW_t	DW_t	D_t	ΔRV_t
Panel A : Small Caps								
ΔRS	-0.000	0.064^{***}	-0.011^{***}	0.044^{***}	-0.034	0.034	0.004	-0.012^{***}
$\Delta Dispersion$	0.006^{***}	0.006^{***}	-0.005^{***}	0.003	0.024^{**}	0.027^{***}	0.007^{*}	-0.002^{**}
$\Delta Slope$	-0.005^{**}	-0.035^{***}	0.012^{***}	0.006	0.033^{*}	-0.015	-0.017^{*}	0.005^{**}
ΔQB	-0.003^{**}	0.004^{**}	0.001	0.000	-0.013	-0.012	0.063^{***}	-0.003^{*}
ΔQA	-0.003^{**}	0.004^{***}	0.001	0.056^{***}	-0.032^{**}	-0.054^{***}	-0.070^{***}	-0.004^{***}
ΔQ	-0.002^{*}	0.004^{***}	0.000	0.031^{***}	-0.018^{*}	-0.028^{**}	-0.016^{***}	-0.004^{***}
ΔNB	0.021^{***}	0.070^{***}	-0.025^{***}	0.145^{***}	0.405^{***}	-0.549^{***}	-0.442^{***}	0.015^{***}
ΔNS	0.023^{***}	0.039^{***}	-0.038^{***}	-0.321^{***}	-0.495^{***}	0.552^{***}	0.467^{***}	0.014^{***}
$\Delta Ntrades$	0.020^{***}	0.133^{***}	-0.034^{***}	-0.058^{***}	0.075^{***}	0.161^{***}	0.007	0.007^{**}
ΔATS	0.006^{**}	0.004	-0.002	0.002	0.027	-0.034	-0.023^{**}	0.007^{**}
$\Delta Volume$	0.032^{***}	0.091^{***}	-0.034^{***}	-0.045^{**}	0.129^{***}	0.122^{***}	-0.016	0.012^{***}
$\Delta Amihud$	0.051^{***}	-0.058^{***}	0.047^{***}	-0.211^{***}	0.403^{***}	0.462^{***}	0.024	0.097^{***}
Panel B : Mid Caps								
ΔRS	0.013^{***}	0.039^{***}	-0.019^{***}	0.001	-0.088^{***}	0.005	0.005	0.002
$\Delta Dispersion$	0.005^{***}	0.002^{***}	-0.005^{***}	0.000	0.016^{**}	0.030^{***}	0.006^{**}	0.001^{*}
$\Delta Slope$	-0.011^{***}	-0.027^{***}	0.019^{***}	0.014^{**}	0.048^{***}	0.014	-0.014^{***}	-0.003^{**}
ΔQB	-0.001	0.001	-0.001	-0.013^{***}	-0.035^{***}	-0.008	0.046^{***}	-0.004^{***}
ΔQA	-0.002^{**}	0.000	0.001	0.038^{***}	-0.025^{***}	-0.017^{**}	-0.049^{***}	-0.004^{***}
ΔQ	-0.000	0.001	-0.000	0.013^{***}	-0.022^{***}	-0.013^{**}	-0.007^{***}	-0.003^{***}
ΔNB	0.020^{***}	0.050^{***}	-0.034^{***}	0.092^{***}	0.238^{***}	-0.450^{***}	-0.352^{***}	0.017^{***}
ΔNS	0.010^{***}	0.077^{***}	-0.033^{***}	-0.240^{***}	-0.401^{***}	0.254^{***}	0.335^{***}	0.008^{***}
$\Delta Ntrades$	0.016^{***}	0.126^{***}	-0.038^{***}	-0.055^{***}	0.010	-0.002	-0.004	0.006^{***}
ΔATS	0.004^{***}	0.007^{***}	-0.008^{***}	-0.039^{***}	-0.023	-0.035^{**}	-0.013^{**}	0.005^{***}
$\Delta Volume$	0.027^{***}	0.085^{***}	-0.042^{***}	-0.075^{***}	-0.016	-0.030	-0.017^{*}	0.009^{***}
$\Delta Amihud$	0.071^{***}	-0.063^{***}	0.038^{***}	-0.196^{***}	0.383^{***}	0.497^{***}	0.022^{**}	0.076^{***}
Panel C : Large Caps								
ΔRS	0.007^{***}	0.052^{***}	-0.015^{***}	0.011^{***}	-0.082^{***}	-0.106^{***}	0.009^{***}	0.012^{***}
$\Delta Dispersion$	0.006^{***}	0.021^{***}	-0.009^{***}	-0.001	-0.014^{***}	-0.031^{***}	0.004^{***}	0.010^{***}
$\Delta Slope$	-0.009^{***}	-0.040^{***}	0.015^{***}	0.035^{***}	0.032^{***}	0.086^{***}	-0.011^{***}	-0.031^{***}
ΔQB	-0.005^{***}	-0.005^{***}	0.002^{*}	0.002	-0.091^{***}	-0.022^{**}	0.020^{***}	-0.017^{***}
ΔQA	-0.004^{***}	-0.004^{***}	0.002	0.023^{***}	-0.074^{***}	-0.074^{***}	-0.035^{***}	-0.019^{***}
ΔQ	-0.005^{***}	-0.003^{***}	0.003^{***}	0.011^{***}	-0.079^{***}	-0.046^{***}	-0.009^{***}	-0.018^{***}
ΔNB	0.011^{***}	0.195^{***}	-0.031^{***}	0.133^{***}	-0.115^{***}	-0.191^{***}	-0.336^{***}	0.006^{***}
ΔNS	0.005^{***}	0.189^{***}	-0.027^{***}	-0.187^{***}	-0.205^{***}	-0.058^{***}	0.308^{***}	-0.001
$\Delta Ntrades$	0.007^{***}	0.236^{***}	-0.026^{***}	-0.016^{***}	-0.129^{***}	-0.065^{***}	0.002	0.003^{**}
ΔATS	0.003^{***}	0.066^{***}	-0.010^{***}	-0.025^{***}	-0.085^{***}	-0.074^{***}	-0.016^{***}	-0.007^{***}
$\Delta Volume$	0.005^{***}	0.308^{***}	-0.027^{***}	-0.031^{***}	-0.197^{***}	-0.149^{***}	-0.013^{***}	-0.012^{***}
$\Delta Amihud$	0.307^{***}	-0.074^{***}	0.165^{***}	-0.057^{***}	0.494^{***}	0.380^{***}	0.011^{***}	0.038^{***}

This table presents the different models that are estimated with the robust regression technique. Panels A, B and C display the estimates for small, mid and large caps, respectively. The first column, L_t, indicates the dependant variable and the following columns are the exogenous variables. QB and QA are the sums of the quantities outstanding at the five best limits, on bid and ask sides, respectively. $Q5$ is the sum of the two. $Slope$ denotes the slope of the book computed following Næs and Skjeltorp (2006) while $Dispersion$ stands for the dispersion measure computed following Kang and Yeo (2008). RS denotes the relative spread. NB and NS respectively denote the number of buy and sell trades. $Ntrades$ is the total number of trades and ATS is the average trade size of these trades. $Amihud$ denotes the Amihud ratio. $OC_t = \frac{|Close_t - Open_t|}{RangeMid_t}$ is the absolute value of the net movement. $OCHL_t = \frac{|Close_t - Open_t|}{High_t - Low_t}$ is the percentage of net variation on total fluctuation. HL_t denotes the High-Low range of the interval. ZR_t is a dummy variable that control for the presence of zero returns. UW_t is equal to 1 when an upward window occurs on the chart and 0 otherwise, while DW_t is equal to 1 when a downward window appears and 0 otherwise. The D_t dummy equals 1 when the net price movement is positive and 0 otherwise. The fixed effect dummies are not presented here.

Table 3.8: Median regressions with realized volatility

ΔL_t	ΔOC_t	ΔHL_t	$\Delta OCHL_t$	ZR_t	UW_t	DW_t	D_t	ΔRV_t
Panel A : Small Caps								
ΔRS	0.002	0.102^{***}	-0.010^{**}	0.005	0.007	0.044^{**}	0.004	-0.016^{***}
$\Delta Dispersion$	0.007^{***}	0.009^{***}	-0.006^{***}	0.002	0.026^{**}	0.032^{***}	0.002	-0.002
$\Delta Slope$	-0.002	-0.036^{***}	0.013^{***}	0.005	0.007	-0.038^{**}	-0.006^{*}	0.009^{***}
ΔQB	-0.002	0.005^{**}	0.000	-0.006	0.003	0.019	0.032^{***}	-0.002
ΔQA	0.000	0.003^{*}	-0.001	0.023^{***}	-0.006	-0.009	-0.042^{***}	-0.003^{**}
ΔQ	-0.001	0.005^{***}	0.000	0.014^{***}	-0.002	-0.003	-0.012^{***}	-0.004^{***}
ΔNB	0.031^{***}	0.120^{***}	-0.028^{***}	0.173^{***}	0.615^{***}	-0.531^{***}	-0.506^{***}	0.020^{*}
ΔNS	0.020^{***}	0.101^{***}	-0.036^{***}	-0.398^{***}	-0.476^{***}	0.825^{***}	0.584^{***}	0.016^{**}
$\Delta Ntrades$	0.035^{***}	0.176^{***}	-0.040^{***}	-0.061^{***}	0.148^{***}	0.219^{***}	-0.014	0.008
ΔATS	0.008^{***}	0.003	-0.003	0.018	0.031	-0.045	-0.040^{***}	0.012^{***}
$\Delta Volume$	0.044^{***}	0.192^{***}	-0.037^{***}	-0.010	0.221^{***}	0.139^{**}	-0.032	0.029^{**}
$\Delta Amihud$	0.030^{*}	-0.051^{***}	0.287^{***}	-0.056	0.508^{***}	0.768^{***}	0.026	0.207^{***}
Panel B : Mid Caps								
ΔRS	0.020^{***}	0.059^{***}	-0.023^{***}	-0.003^{**}	-0.073^{***}	0.002	0.001	-0.000
$\Delta Dispersion$	0.008^{***}	0.003^{**}	-0.006^{***}	0.002	0.021^{***}	0.032^{***}	0.000	0.001
$\Delta Slope$	-0.013^{***}	-0.024^{***}	0.020^{***}	0.008^{***}	0.059^{***}	-0.004	-0.001	-0.002^{**}
ΔQB	0.001	0.001	-0.000	-0.010^{***}	-0.007	0.011	0.022^{***}	-0.001
ΔQA	0.001	-0.001	-0.000	0.016^{***}	0.003	0.001	-0.026^{***}	-0.002^{**}
ΔQ	0.001	0.000	-0.001	0.002	-0.003	-0.000	-0.003^{***}	-0.002^{***}
ΔNB	0.022^{***}	0.088^{***}	-0.034^{***}	0.129^{***}	0.327^{***}	-0.509^{***}	-0.427^{***}	0.029^{***}
ΔNS	0.011^{**}	0.116^{***}	-0.033^{***}	-0.296^{***}	-0.477^{***}	0.345^{***}	0.411^{***}	0.022^{***}
$\Delta Ntrades$	0.015^{***}	0.152^{***}	-0.038^{***}	-0.070^{***}	0.044^{**}	-0.005	-0.009	0.017^{***}
ΔATS	0.005^{**}	0.011^{***}	-0.007^{***}	-0.039^{***}	-0.020	-0.014	-0.013^{*}	0.008^{***}
$\Delta Volume$	0.018^{***}	0.212^{***}	-0.036^{***}	-0.079^{***}	0.009	-0.062^{**}	-0.029^{***}	0.017^{***}
$\Delta Amihud$	0.102^{***}	-0.092^{***}	0.231^{***}	-0.082^{***}	0.869^{***}	0.996^{***}	0.041^{***}	0.203^{***}
Panel C : Large Caps								
ΔRS	0.001^{***}	0.000^{***}	-0.001^{***}	-0.001^{***}	-0.003^{***}	0.001^{*}	0.002^{***}	0.000^{***}
$\Delta Dispersion$	0.008^{***}	0.018^{***}	-0.010^{***}	-0.003^{***}	0.007	-0.018^{**}	0.003^{***}	0.013^{***}
$\Delta Slope$	-0.011^{***}	-0.027^{***}	0.017^{***}	0.007^{***}	0.035^{***}	0.078^{***}	-0.006^{***}	-0.023^{***}
ΔQB	-0.004^{***}	-0.004^{**}	0.002	-0.009^{***}	-0.063^{***}	-0.007	0.020^{***}	-0.012^{***}
ΔQA	-0.004^{***}	-0.003^{*}	0.003^{**}	0.013^{***}	-0.053^{***}	-0.061^{***}	-0.030^{***}	-0.015^{***}
ΔQ	-0.004^{***}	-0.003^{**}	0.004^{***}	0.007^{***}	-0.061^{***}	-0.035^{***}	-0.008^{***}	-0.013^{***}
ΔNB	0.005^{*}	0.251^{***}	-0.028^{***}	0.152^{***}	-0.142^{***}	-0.212^{***}	-0.365^{***}	0.037^{***}
ΔNS	-0.001	0.248^{***}	-0.024^{***}	-0.193^{***}	-0.192^{***}	-0.085^{***}	0.330^{***}	0.011^{**}
$\Delta Ntrades$	0.007^{***}	0.259^{***}	-0.027^{***}	-0.015^{***}	-0.112^{***}	-0.051^{**}	0.000	0.019^{***}
ΔATS	0.002	0.076^{***}	-0.009^{***}	-0.022^{***}	-0.075^{***}	-0.086^{***}	-0.018^{***}	-0.002
$\Delta Volume$	-0.002	0.407^{***}	-0.025^{***}	-0.023^{***}	-0.193^{***}	-0.137^{***}	-0.020^{***}	0.008
$\Delta Amihud$	0.337^{***}	-0.057^{***}	0.373^{***}	-0.022^{***}	1.566^{***}	0.998^{***}	0.014^{***}	0.095^{***}

This table presents the different models that are estimated in the framework of a median regression. Panels A, B and C display the estimates for small, mid and large caps, respectively. The first column, L_t, indicates the dependant variable and the following columns are the exogenous variables. QB and QA are the sums of the quantities outstanding at the five best limits, on bid and ask sides, respectively. $Q5$ is the sum of the two. $Slope$ denotes the slope of the book computed following Næs and Skjeltorp (2006) while $Dispersion$ stands for the dispersion measure computed following Kang and Yeo (2008). RS denotes the relative spread. NB and NS respectively denote the number of buy and sell trades. $Ntrades$ is the total number of trades and ATS is the average trade size of these trades. $Amihud$ denotes the Amihud ratio. $OC_t = \frac{|Close_t - Open_t|}{RangeMid_t}$ is the absolute value of the net movement. $OCHL_t = \frac{|Close_t - Open_t|}{High_t - Low_t}$ is the percentage of net variation on total fluctuation. HL_t denotes the High-Low range of the interval. ZR_t is a dummy variable that control for the presence of zero returns. UW_t is equal to 1 when an upward window occurs on the chart and 0 otherwise, while DW_t is equal to 1 when a downward window appears and 0 otherwise. The D_t dummy equals 1 when the net price movement is positive and 0 otherwise. The fixed effect dummies are not presented here.

The use of the robust regression technique does not change the overall picture. There are nevertheless some points worth mentioning.

The validation of H1 in Table 3.7 is stronger for small caps when the robust regression technique is used instead of the method based on clustered standard errors and used in Table 3.6. Both the Amihud ratio and the slope variable become positive and significant. Depth for small caps also become negative and significant, as it was (and still is) the case for mid and large caps. Finally, the variables related to trading activity (i.e. volume, number of trades, and average trade size) become more (positively) significant than before, whatever the market capitalization of the stocks.

Regarding H2, mid caps display better results than before: The slope is significant at 1% (instead of 10%) and the dispersion measure is now positive and significant. The only unexpected result is the negative and now significant sign of the Amihud ratio for small caps.

Results for H3 are very much unchanged. They are slightly reinforced for mid caps as the dispersion measure turns out to be negatively significant while it was insignificant before.

H4 is still validated, although it seems to be more sensitive to the regression technique employed. H4 is now validated when it comes to the Amihud ratio, irrespective of the market capitalization of the stocks. Another good news is that the slope variable turns out to be positive and significant for mid and large caps. The bad news is that the relative spread behaves against our expectations for small and large caps.

We confirm that H5 is validated with respect to ex-post liquidity but results are rather disappointing with respect to ex-ante liquidity (although it was already the case in Table 3.6).

The validation of H6 is unchanged.

Empirical results obtained by running median regressions and reported in Table 3.8 are very much in line with Table 3.6.

We reassert the validation of H1, with two improvements: the coefficient of the Amihud ratio for small caps becomes positive and significant (while it was insignificant); and the coefficient of the dispersion measure for mid caps turns out to be significant at 1% (versus 10% before). Finally, variables related to trading activity for mid caps (i.e. volume, number of trades, and average trade size) as well as depth variables for large caps are now all significant (while they were not before).

The validation of H2 is reinforced for mid caps for two reasons. Firstly, the coefficient of the dispersion variable displays the correct sign and becomes significant at 5% (while it was not in Table 3.6). Secondly, the significance of the slope variable is improved, lowering from 10% to 1%. We confirm that ex-post liquidity does not deteriorate when HL rises: the coefficient of the Amihud ratio is now negative and statistically significant at 1%, whatever the market cap.

H3 is unaffected by the use of the median regression technique.

Results for H4 are overall improved. H4 is now validated when it comes to the Amihud ratio, irrespective of the market capitalization of the stocks (as it was also the case when the robust regression technique was used). Finally, the slope variable turns out to be positive and significant for mid and large caps. H4 is nevertheless not confirmed with respect to the ex-ante liquidity of small caps: the relative spread, the dispersion, and the slope are not significant.

H5 is clearly validated with respect to the Amihud ratio. The main improvement occurs for small caps where the ratio becomes significant at 1% when any of the two windows occurs. With respect to ex-ante liquidity, results are also improved for small caps since there

is no unexpected significant coefficient anymore. The relative spread and the slope variables even show the expected sign and significance when the median regression technique is used. Regarding large caps, trading activity (characterized by volume, average trade size, and number of trades) turns out to be clearly negatively affected by the occurrence of windows. The other positive change is the relative spread which is positive and significant at 10% (while it was not in table 3.6). Mid caps display mixed results. The dispersion variables becomes positive and significant at 1%, in line with our expectations, but the relative spread and the dispersion variables exhibit the wrong sign when there has been an upper window.

Results for H6 are again unchanged.

3.3.4 Addressing causality and endogeneity issues

Since a causal relationship exists between the dependent and the independent variables, as outlined by Mazza (2013), endogeneity issues arise in traditional regressions. This is quite obvious since the price is expected to move after changes in the state of the limit order book: Informed traders, through price discovery processes, move stocks' price towards their fundamental value. In this section, we address these issues by estimating VARX models that control for the endogeneity that results from a loop of causality. We estimate these VARX models for each 100-stock portfolio and check whether previous results still hold when endogeneity is controlled for.

For each liquidity proxy L_t, we run a multivariate VARX model which includes each numerical price movements variables, OC, HL and $OCHL$, as well as the analyzed liquidity proxy. As a result, we conduct several VARX(4) specifications, one for each liquidity proxy L_t. We use VARX models rather than traditional VAR models to control for the presence of fixed effects which are exogenous, i.e. they are determined outside of the VARX system. VARX models are specified as restricted VAR models including equations for each

exogenous variable but with right-hand-side coefficients restricted to zero.[1]

This multivariate VARX model is specified as follows:

$$\Delta L_t = \beta_{10} + \beta_{11}\Delta L_{t-1} + \beta_{12}\Delta L_{t-2} + \alpha_{11,1}\Delta OC_{t-1} + \alpha_{12,1}\Delta OC_{t-2} + \alpha_{11,2}\Delta HL_{t-1} + \alpha_{12,2}\Delta HL_{t-2} + \alpha_{11,3}\Delta OCHL_{t-1} + \alpha_{12,3}\Delta OCHL_{t-2} + \sum_{ß=2}^{100}\gamma_i S_{it} + u_{1t}, \tag{3.3.1}$$

$$\Delta OC_t = \beta_{20} + \beta_{21}\Delta L_{t-1} + \beta_{22}\Delta L_{t-2} + \alpha_{21,1}\Delta OC_{t-1} + \alpha_{22,1}\Delta OC_{t-2} + \alpha_{21,2}\Delta HL_{t-1} + \alpha_{22,2}\Delta HL_{t-2} + \alpha_{21,3}\Delta OCHL_{t-1} + \alpha_{22,3}\Delta OCHL_{t-2} + \sum_{ß=2}^{100}\gamma_i S_{it} + u_{2t}, \tag{3.3.2}$$

$$\Delta HL_t = \beta_{30} + \beta_{31}\Delta L_{t-1} + \beta_{32}\Delta L_{t-2} + \alpha_{31,1}\Delta OC_{t-1} + \alpha_{32,1}\Delta OC_{t-2} + \alpha_{31,2}\Delta HL_{t-1} + \alpha_{32,2}\Delta HL_{t-2} + \alpha_{31,3}\Delta OCHL_{t-1} + \alpha_{32,3}\Delta OCHL_{t-2} + \sum_{ß=2}^{100}\gamma_i S_{it} + u_{3t}, \tag{3.3.3}$$

$$\Delta OCHL_t = \beta_{40} + \beta_{41}\Delta L_{t-1} + \beta_{42}\Delta L_{t-2} + \alpha_{41,1}\Delta OC_{t-1} + \alpha_{42,1}\Delta OC_{t-2} + \alpha_{41,2}\Delta HL_{t-1} + \alpha_{42,2}\Delta HL_{t-2} + \alpha_{41,3}\Delta OCHL_{t-1} + \alpha_{42,3}\Delta OCHL_{t-2} + \sum_{ß=2}^{100}\gamma_i S_{it} + u_{4t}, \tag{3.3.4}$$

where L_t denotes one of the liquidity proxies that is investigated, OC_t, HL_t and $OCHL_t$ are the price movement variables, S_{it} are fixed-effect dummies and u_{it} is a white noise dis-

[1]For the ease of presentation and computation, we only include two lags in the VARX specification. Using intraday databases for a large number of stocks makes the use of the traditional optimal lag choice methods, e.g. minimization of the AIC information criterion, impossible.

turbance term with $E[u_{it}] = 0$, $E[u_{1t}u_{2t}] = 0$ and $i = (1, 2)$. There are 8 variables in each equation and four equations in each system. As a result, 32 parameters have to be estimated for each liquidity proxy L_t.

Table 3.9, 3.10 and 3.11 present the results obtained for Small, Mid and Large cap portfolios, respectively. The parameter estimates of the four equations of the system are showed.

Table 3.9: Multivariate VARX(4,2) model: Small cap portfolio

ΔL_t	$Equation$	ΔL_{t-1}	ΔOC_{t-1}	ΔHL_{t-1}	$\Delta OCHL_{t-1}$	ΔL_{t-2}	ΔOC_{t-2}	ΔHL_{t-2}	$\Delta OCHL_{t-2}$
ΔQB	ΔQB	-0.114^{***}	0.000	0.004^{**}	0.001	-0.022^{***}	-0.000	0.002	0.002
	ΔOC	-0.004	-0.095^{***}	0.028^{***}	0.022^{**}	-0.004	0.010	0.006	0.009
	ΔHL	-0.034^{*}	-0.076^{***}	-0.002	0.086^{***}	-0.027	-0.015^{*}	0.005	0.021^{**}
	$\Delta OCHL$	0.023^{*}	-0.083^{***}	0.083^{***}	-0.015^{**}	0.023^{*}	0.041^{***}	-0.009^{*}	-0.020^{***}
ΔQA	ΔQA	-0.132^{***}	-0.003^{*}	0.008^{***}	0.005^{**}	-0.040^{***}	0.000	0.004^{**}	0.003
	ΔOC	0.007	-0.095^{***}	0.028^{***}	0.022^{**}	0.031	0.010	0.006	0.009
	ΔHL	-0.035	-0.076^{***}	-0.002	0.086^{***}	0.000	-0.015^{*}	0.005	0.021^{**}
	$\Delta OCHL$	0.017	-0.083^{***}	0.083^{***}	-0.015^{**}	-0.007	0.041^{***}	-0.009^{*}	-0.020^{***}
ΔQ	ΔQ	-0.165^{***}	-0.002^{**}	0.005^{***}	0.002^{*}	-0.041^{***}	-0.001	0.003^{***}	0.001
	ΔOC	-0.015	-0.094^{***}	0.028^{***}	0.022^{**}	0.011	0.010	0.006	0.009
	ΔHL	-0.078^{**}	-0.076^{***}	-0.002	0.086^{***}	-0.039	-0.015^{*}	0.005	0.021^{**}
	$\Delta OCHL$	0.044^{*}	-0.083^{***}	0.083^{***}	-0.015^{**}	-0.023	0.041^{***}	-0.009^{*}	-0.020^{***}
$\Delta Slope$	$\Delta Slope$	-0.107^{***}	-0.000	0.009^{**}	0.020^{***}	-0.011^{**}	0.007^{**}	0.006^{*}	0.010^{**}
	ΔOC	0.037^{***}	-0.095^{***}	0.029^{***}	0.021^{**}	0.077^{***}	0.010	0.008	0.006
	ΔHL	0.046^{***}	-0.076^{***}	-0.001	0.085^{***}	0.048^{***}	-0.015^{*}	0.006	0.019^{**}
	$\Delta OCHL$	-0.000	-0.083^{***}	0.083^{***}	-0.015^{**}	-0.004	0.041^{***}	-0.009^{*}	-0.020^{***}
$\Delta Dispersion$	$\Delta Dispersion$	-0.142^{***}	-0.003^{***}	0.005^{***}	0.001	-0.045^{***}	-0.000	0.000	0.001
	ΔOC	0.022	-0.095^{***}	0.028^{***}	0.022^{**}	0.089^{**}	0.009	0.005	0.009
	ΔHL	0.001	-0.076^{***}	-0.002	0.086^{***}	-0.008	-0.015^{*}	0.005	0.021^{**}
	$\Delta OCHL$	0.025	-0.083^{***}	0.083^{***}	-0.015^{**}	0.094^{***}	0.040^{***}	-0.010^{**}	-0.020^{***}
ΔRS	ΔRS	-0.082^{***}	0.000	0.018^{***}	0.013^{*}	-0.012^{**}	-0.008	0.004	0.014^{**}
	ΔOC	0.031^{***}	-0.094^{***}	0.023^{***}	0.022^{**}	0.011	0.009	0.004	0.009
	ΔHL	0.010	-0.076^{***}	-0.004	0.086^{***}	0.011	-0.015^{*}	0.003	0.021^{**}
	$\Delta OCHL$	0.017^{***}	-0.083^{***}	0.080^{***}	-0.015^{**}	0.010^{**}	0.041^{***}	-0.010^{**}	-0.020^{***}
ΔNB	ΔNB	-0.024^{***}	-0.020	-0.002	0.056^{***}	0.003	0.006	-0.000	-0.011
	ΔOC	-0.000	-0.105^{***}	0.023^{**}	0.018	0.015^{***}	0.001	-0.009	0.013
	ΔHL	0.001	-0.071^{***}	-0.023^{**}	0.075^{***}	0.003	-0.020^{**}	-0.005	0.017^{*}
	$\Delta OCHL$	-0.003	-0.072^{***}	0.073^{***}	-0.030^{***}	0.026^{***}	0.051^{***}	-0.018^{**}	-0.020^{**}
ΔNS	ΔNS	-0.063^{***}	-0.003	-0.011	0.018^{*}	0.035^{***}	0.002	-0.009	-0.001
	ΔOC	0.011	-0.095^{***}	0.011	0.013	0.021^{**}	0.018	-0.026^{**}	0.041^{***}
	ΔHL	-0.004	-0.067^{***}	-0.015	0.075^{***}	-0.001	-0.010	-0.005	0.006
	$\Delta OCHL$	0.019^{***}	-0.086^{***}	0.082^{***}	-0.020^{**}	0.020^{***}	0.071^{***}	-0.044^{***}	-0.017^{*}
$\Delta Ntrades$	$\Delta Ntrades$	-0.023^{***}	-0.024^{***}	-0.005	0.045^{***}	0.017^{**}	-0.001	0.004	0.004
	ΔOC	0.004	-0.110^{***}	0.027^{***}	0.025^{**}	0.029^{***}	0.003	-0.014	0.013
	ΔHL	-0.003	-0.088^{***}	-0.006	0.097^{***}	0.001	-0.014	-0.002	0.013
	$\Delta OCHL$	0.006	-0.076^{***}	0.072^{***}	-0.029^{***}	0.055^{***}	0.042^{***}	-0.029^{***}	-0.017^{**}
$\Delta Volume$	$\Delta Volume$	-0.008	-0.047	0.000	0.055	-0.003	-0.016	0.012	-0.006
	ΔOC	-0.000	-0.109^{***}	0.028^{***}	0.025^{**}	0.001	0.004	-0.008	0.012
	ΔHL	-0.000	-0.088^{***}	-0.006	0.097^{***}	-0.001	-0.014	-0.002	0.012
	$\Delta OCHL$	-0.000	-0.075^{***}	0.073^{***}	-0.029^{***}	0.003^{***}	0.046^{***}	-0.018^{**}	-0.020^{**}
ΔATS	ΔATS	-0.008	-0.047	0.000	0.055	-0.003	-0.016	0.012	-0.006
	ΔOC	-0.000	-0.109^{***}	0.028^{***}	0.025^{**}	0.001	0.004	-0.008	0.012
	ΔHL	-0.000	-0.088^{***}	-0.006	0.097^{***}	-0.001	-0.014	-0.002	0.012
	$\Delta OCHL$	-0.000	-0.075^{***}	0.073^{***}	-0.029^{***}	0.003^{***}	0.046^{***}	-0.018^{**}	-0.020^{**}
$\Delta Amihud$	$\Delta Amihud$	-0.008	-0.047	0.000	0.055	-0.003	-0.016	0.012	-0.006
	ΔOC	-0.000	-0.109^{***}	0.028^{***}	0.025^{**}	0.001	0.004	-0.008	0.012
	ΔHL	-0.000	-0.088^{***}	-0.006	0.097^{***}	-0.001	-0.014	-0.002	0.012
	$\Delta OCHL$	-0.000	-0.075^{***}	0.073^{***}	-0.029^{***}	0.003^{***}	0.046^{***}	-0.018^{**}	-0.020^{**}

This table presents the parameter estimates of the multivariate VARX specification which includes, along with the analyzed liquidity proxy, OC, HL and $OCHL$ as endogenous variables. The exogenous variables that are determined outside of the VARX process are the stocks' fixed-effects dummies. Each system is presented by four-line blocks in the table and each line represents an equation of the VARX process. The first column, ΔL_t, indicates the liquidity proxy that is analyzed. The second column, $Equation$, denotes the dependent variable that is being investigated in this part of the VARX process. The different columns show the parameters estimates for each lag of the four variables. QB and QA are the sums of the quantities outstanding at the five best limits, on bid and ask sides, respectively. Q is the sum of the two. $Slope$ denotes the slope of the book computed following Hasbrouck and Skjeltorp (2006) while $Dispersion$ stands for the dispersion measure computed following Kang and Yeo (2008). RS denotes the relative spread. NB and NS respectively denote the number of buy

Table 3.10: Multivariate VARX(4,2) model: Mid cap portfolio

ΔL_t	$Equation$	ΔL_{t-1}	ΔOC_{t-1}	ΔHL_{t-1}	$\Delta OCHL_{t-1}$	ΔL_{t-2}	ΔOC_{t-2}	ΔHL_{t-2}	$\Delta OCHL_{t-2}$
ΔQB	ΔQB	-0.001	0.008	-0.003	0.016	-0.001	0.017	-0.008	0.016
	ΔOC	0.000	-0.117^{***}	0.033^{***}	0.012^{*}	0.000	-0.023^{***}	0.014^{***}	0.015^{**}
	ΔHL	-0.000	-0.099^{***}	0.008	0.104^{***}	-0.000	-0.013^{**}	-0.006	0.013^{*}
	$\Delta OCHL$	0.000	-0.069^{***}	0.053^{***}	-0.045^{***}	0.000	-0.012^{***}	0.020^{***}	0.015^{***}
ΔQA	ΔQA	-0.001	0.021	-0.009	0.048	-0.001	0.051	-0.022	0.066
	ΔOC	0.000	-0.117^{***}	0.033^{***}	0.012^{*}	0.000	-0.023^{***}	0.014^{***}	0.015^{**}
	ΔHL	-0.000	-0.099^{***}	0.008	0.104^{***}	0.000	-0.013^{**}	-0.006	0.013^{*}
	$\Delta OCHL$	0.000	-0.069^{***}	0.053^{***}	-0.045^{***}	0.000	-0.012^{***}	0.020^{***}	0.015^{***}
ΔQ	ΔQ	-0.001	0.010	-0.005	0.025	-0.001	0.026	-0.012	0.032
	ΔOC	0.000	-0.117^{***}	0.033^{***}	0.012^{*}	0.000	-0.023^{***}	0.014^{***}	0.015^{**}
	ΔHL	-0.000	-0.099^{***}	0.008	0.104^{***}	-0.000	-0.013^{**}	-0.006	0.013^{*}
	$\Delta OCHL$	0.000	-0.069^{***}	0.053^{***}	-0.045^{***}	0.000	-0.012^{***}	0.020^{***}	0.015^{***}
$\Delta Slope$	$\Delta Slope$	-0.148^{***}	0.009^{***}	0.001	0.000	-0.027^{***}	0.007^{***}	0.000	0.002
	ΔOC	0.043^{***}	-0.116^{***}	0.033^{***}	0.011	0.030^{***}	-0.023^{***}	0.015^{***}	0.014^{**}
	ΔHL	0.060^{***}	-0.098^{***}	0.008	0.102^{***}	0.030^{***}	-0.013^{**}	-0.005	0.012
	$\Delta OCHL$	-0.001	-0.069^{***}	0.053^{***}	-0.045^{***}	0.026^{***}	-0.011^{***}	0.020^{***}	0.015^{***}
$\Delta Dispersion$	$\Delta Dispersion$	-0.001	0.006	-0.006	-0.003	-0.001	0.062	-0.003	-0.047
	ΔOC	-0.000	-0.117^{***}	0.033^{***}	0.012^{*}	-0.000	-0.023^{***}	0.014^{***}	0.015^{**}
	ΔHL	-0.000	-0.099^{***}	0.008	0.104^{***}	-0.000	-0.013^{**}	-0.006	0.013^{*}
	$\Delta OCHL$	-0.000	-0.069^{***}	0.053^{***}	-0.045^{***}	-0.000	-0.012^{***}	0.020^{***}	0.015^{***}
ΔRS	ΔRS	-0.041^{***}	-0.014^{**}	0.009^{*}	0.014^{*}	-0.007^{*}	-0.004	-0.004	0.006
	ΔOC	0.009^{***}	-0.117^{***}	0.032^{***}	0.012^{*}	0.015^{***}	-0.024^{***}	0.013^{***}	0.016^{**}
	ΔHL	0.002	-0.099^{***}	0.008	0.104^{***}	0.005	-0.013^{**}	-0.006	0.013^{*}
	$\Delta OCHL$	0.017^{***}	-0.070^{***}	0.052^{***}	-0.045^{***}	0.009^{***}	-0.012^{***}	0.019^{***}	0.016^{***}
ΔNB	ΔNB	-0.099^{***}	-0.011^{**}	-0.001	0.023^{***}	-0.010^{**}	-0.000	-0.002	-0.001
	ΔOC	-0.003	-0.126^{***}	0.047^{***}	0.016^{**}	0.010^{*}	-0.006	-0.009	-0.011
	ΔHL	-0.011^{**}	-0.088^{***}	0.006	0.088^{***}	-0.002	-0.013^{**}	-0.010	-0.005
	$\Delta OCHL$	0.015^{***}	-0.090^{***}	0.075^{***}	-0.029^{***}	0.011^{***}	0.001	0.006	-0.005
ΔNS	ΔNS	-0.113^{***}	-0.009^{*}	-0.004	0.030^{***}	-0.013^{***}	0.004	-0.007^{*}	0.000
	ΔOC	-0.003	-0.109^{***}	0.019^{***}	0.004	-0.004	-0.012^{*}	0.014^{**}	0.003
	ΔHL	-0.013^{***}	-0.081^{***}	-0.016^{***}	0.082^{***}	-0.002	-0.011^{**}	-0.005	0.006
	$\Delta OCHL$	0.018^{***}	-0.070^{***}	0.061^{***}	-0.041^{***}	0.009^{**}	0.000	0.009^{**}	0.005
$\Delta Ntrades$	$\Delta Ntrades$	-0.120^{***}	-0.023^{***}	0.010^{***}	0.038^{***}	-0.013^{***}	-0.002	-0.002	0.004
	ΔOC	-0.009	-0.122^{***}	0.035^{***}	0.010	0.027^{***}	-0.012^{*}	0.003	0.005
	ΔHL	-0.046^{***}	-0.102^{***}	0.013^{**}	0.104^{***}	-0.003	-0.008	-0.008	-0.001
	$\Delta OCHL$	0.041^{***}	-0.073^{***}	0.052^{***}	-0.045^{***}	0.027^{***}	-0.002	0.006^{*}	0.013^{**}
$\Delta Volume$	$\Delta Volume$	-0.007^{*}	-0.081^{**}	0.000	0.088^{*}	-0.003	-0.002	-0.027	-0.064
	ΔOC	-0.000	-0.122^{***}	0.034^{***}	0.011	0.000	-0.011^{*}	0.006	0.004
	ΔHL	-0.001	-0.103^{***}	0.007	0.106^{***}	-0.000	-0.008	-0.008	-0.002
	$\Delta OCHL$	0.000	-0.072^{***}	0.056^{***}	-0.046^{***}	0.000	-0.002	0.009^{***}	0.012^{**}
ΔATS	ΔATS	-0.004	-0.041	0.009	0.072	-0.003	0.005	-0.027	-0.047
	ΔOC	-0.000	-0.122^{***}	0.030^{***}	0.008	-0.000	-0.007	0.013^{**}	0.001
	ΔHL	-0.000	-0.104^{***}	0.005	0.104^{***}	-0.000	0.001	0.001	-0.008
	$\Delta OCHL$	-0.001	-0.071^{***}	0.055^{***}	-0.050^{***}	-0.000	-0.005	0.007^{*}	0.014^{***}
$\Delta Amihud$	$\Delta Amihud$	-0.009^{**}	-0.092	0.162^{*}	-0.049	-0.004	-0.104	0.082	0.017
	ΔOC	-0.000	-0.118^{***}	0.021^{***}	0.014^{*}	0.000	-0.004	0.022^{***}	0.000
	ΔHL	0.000	-0.097^{***}	-0.007	0.098^{***}	-0.000	0.000	0.008	-0.003
	$\Delta OCHL$	-0.000	-0.071^{***}	0.057^{***}	-0.042^{***}	0.001^{***}	-0.003	0.009^{**}	0.009^{*}

This table presents the parameter estimates of the multivariate VARX specification which includes, along with the analyzed liquidity proxy, OC, HL and $OCHL$ as endogenous variables. The exogenous variables that are determined outside of the VARX process are the stocks' fixed-effects dummies. Each system is presented by four-line blocks in the table and each line represents an equation of the VARX process. The first column, ΔL_t, indicates the liquidity proxy that is analyzed. The second column, $Equation$, denotes the dependent variable that is being investigated in this part of the VARX process. The different columns show the parameters estimates for each lag of the four variables. QB and QA are the sums of the quantities outstanding at the five best limits, on bid and ask sides, respectively. Q is the sum of the two. $Slope$ denotes the slope of the book computed following Nes and Skjeltorp (2006) while $Dispersion$ stands for the dispersion measure computed following Kang and Yeo (2008). RS denotes the relative spread. NB and NS respectively denote the number of buy

Table 3.11: Multivariate VARX(4,2) model: Large cap portfolio

ΔL_t	$Equation$	ΔL_{t-1}	ΔOC_{t-1}	ΔHL_{t-1}	$\Delta OCHL_{t-1}$	ΔL_{t-2}	ΔOC_{t-2}	ΔHL_{t-2}	$\Delta OCHL_{t-2}$
ΔQB	ΔQB	-0.010^{***}	0.008	0.004	0.007	-0.000	-0.001	0.006	0.021^{**}
	ΔOC	-0.000	-0.115^{***}	-0.076^{***}	-0.106^{***}	-0.000	-0.028^{***}	-0.004	-0.011^{*}
	ΔHL	-0.002^{**}	-0.085^{***}	-0.103^{***}	0.102^{***}	-0.000	-0.024^{***}	-0.026^{***}	0.022^{***}
	$\Delta OCHL$	0.002^{*}	-0.083^{***}	0.092^{***}	-0.124^{***}	-0.000	0.001	0.020^{***}	-0.036^{***}
ΔQA	ΔQA	-0.151^{***}	0.007^{***}	0.001	0.002	-0.025^{***}	0.003^{*}	0.003	0.007^{**}
	ΔOC	-0.011^{**}	-0.115^{***}	-0.076^{***}	-0.106^{***}	0.017^{***}	-0.028^{***}	-0.004	-0.011^{*}
	ΔHL	-0.025^{***}	-0.085^{***}	-0.103^{***}	0.102^{***}	-0.003	-0.024^{***}	-0.026^{***}	0.022^{***}
	$\Delta OCHL$	0.014^{***}	-0.083^{***}	0.092^{***}	-0.124^{***}	0.011^{***}	0.001	0.020^{***}	-0.036^{***}
ΔQ	ΔQ	-0.056^{***}	0.004^{**}	0.002	0.002	-0.009^{***}	0.000	0.004	0.007^{**}
	ΔOC	-0.010^{*}	-0.115^{***}	-0.076^{***}	-0.106^{***}	0.000	-0.028^{***}	-0.004	-0.011^{*}
	ΔHL	-0.015^{***}	-0.085^{***}	-0.103^{***}	0.102^{***}	0.000	-0.024^{***}	-0.026^{***}	0.022^{***}
	$\Delta OCHL$	0.008^{**}	-0.083^{***}	0.092^{***}	-0.124^{***}	-0.001	0.001	0.020^{***}	-0.036^{***}
$\Delta Slope$	$\Delta Slope$	-0.290^{***}	0.011^{***}	0.015^{***}	-0.001	-0.076^{***}	0.015^{***}	0.013^{***}	-0.001
	ΔOC	0.007	-0.115^{***}	-0.075^{***}	-0.106^{***}	0.005	-0.028^{***}	-0.004	-0.011^{*}
	ΔHL	-0.012^{***}	-0.085^{***}	-0.104^{***}	0.102^{***}	0.006^{*}	-0.023^{***}	-0.026^{***}	0.022^{***}
	$\Delta OCHL$	0.001	-0.083^{***}	0.092^{***}	-0.124^{***}	-0.001	0.001	0.020^{***}	-0.036^{***}
$\Delta Dispersion$	$\Delta Dispersion$	-0.283^{***}	0.001	0.002^{*}	0.001	-0.092^{***}	-0.001	-0.001	0.003^{**}
	ΔOC	0.118^{***}	-0.117^{***}	-0.080^{***}	-0.105^{***}	0.069^{***}	-0.029^{***}	-0.006	-0.010^{*}
	ΔHL	0.050^{***}	-0.086^{***}	-0.104^{***}	0.102^{***}	-0.019^{**}	-0.023^{***}	-0.026^{***}	0.022^{***}
	$\Delta OCHL$	0.080^{***}	-0.084^{***}	0.090^{***}	-0.123^{***}	0.084^{***}	-0.001	0.018^{***}	-0.035^{***}
ΔRS	ΔRS	-0.258^{***}	-0.000	0.011^{***}	0.015^{***}	-0.064^{***}	0.001	-0.001	0.007^{**}
	ΔOC	0.022^{***}	-0.116^{***}	-0.079^{***}	-0.106^{***}	0.015^{***}	-0.029^{***}	-0.006	-0.010^{*}
	ΔHL	-0.013^{***}	-0.085^{***}	-0.101^{***}	0.102^{***}	-0.028^{***}	-0.023^{***}	-0.022^{***}	0.021^{***}
	$\Delta OCHL$	0.048^{***}	-0.084^{***}	0.085^{***}	-0.123^{***}	0.040^{***}	-0.001	0.015^{***}	-0.035^{***}
ΔNB	ΔNB	-0.114^{***}	-0.020^{***}	-0.045^{***}	0.051^{***}	-0.016^{***}	-0.022^{***}	-0.005	0.015^{***}
	ΔOC	-0.016^{***}	-0.106^{***}	-0.098^{***}	-0.117^{***}	-0.007^{**}	-0.026^{***}	-0.009	-0.014^{**}
	ΔHL	-0.007^{***}	-0.073^{***}	-0.127^{***}	0.085^{***}	-0.002	-0.023^{***}	-0.036^{***}	0.020^{***}
	$\Delta OCHL$	-0.004^{**}	-0.085^{***}	0.099^{***}	-0.120^{***}	-0.004^{*}	0.001	0.025^{***}	-0.036^{***}
ΔNS	ΔNS	-0.135^{***}	-0.017^{***}	-0.025^{***}	0.041^{***}	-0.017^{***}	-0.022^{***}	0.008^{*}	0.021^{***}
	ΔOC	-0.023^{***}	-0.113^{***}	-0.078^{***}	-0.109^{***}	-0.013^{***}	-0.028^{***}	-0.001	-0.011^{*}
	ΔHL	-0.013^{***}	-0.082^{***}	-0.107^{***}	0.096^{***}	-0.006^{***}	-0.025^{***}	-0.026^{***}	0.024^{***}
	$\Delta OCHL$	-0.005^{**}	-0.083^{***}	0.093^{***}	-0.122^{***}	-0.004^{*}	0.001	0.022^{***}	-0.035^{***}
$\Delta Ntrades$	$\Delta Ntrades$	-0.064^{***}	-0.025^{***}	-0.041^{***}	0.047^{***}	-0.005^{*}	-0.015^{***}	-0.002	0.011^{***}
	ΔOC	-0.018^{***}	-0.113^{***}	-0.078^{***}	-0.109^{***}	-0.007^{*}	-0.027^{***}	-0.004	-0.012^{**}
	ΔHL	-0.014^{***}	-0.084^{***}	-0.104^{***}	0.100^{***}	-0.005^{**}	-0.024^{***}	-0.025^{***}	0.021^{***}
	$\Delta OCHL$	-0.001	-0.082^{***}	0.090^{***}	-0.125^{***}	-0.002	0.001	0.021^{***}	-0.036^{***}
$\Delta Volume$	$\Delta Volume$	-0.001	-0.043	-0.219	0.032	-0.000	-0.041	-0.032	0.075
	ΔOC	-0.000	-0.114^{***}	-0.083^{***}	-0.108^{***}	-0.000	-0.027^{***}	-0.005	-0.012^{**}
	ΔHL	-0.000	-0.084^{***}	-0.107^{***}	0.101^{***}	0.000	-0.024^{***}	-0.026^{***}	0.021^{***}
	$\Delta OCHL$	-0.000	-0.082^{***}	0.090^{***}	-0.125^{***}	-0.000	0.001	0.020^{***}	-0.036^{***}
ΔATS	ΔATS	-0.011^{***}	-0.012	-0.020^{*}	0.010	-0.000	-0.009	-0.002	0.005
	ΔOC	-0.001	-0.113^{***}	-0.085^{***}	-0.109^{***}	-0.001	-0.028^{***}	-0.007	-0.013^{**}
	ΔHL	-0.001	-0.084^{***}	-0.109^{***}	0.100^{***}	-0.000	-0.024^{***}	-0.027^{***}	0.021^{***}
	$\Delta OCHL$	-0.000	-0.082^{***}	0.090^{***}	-0.124^{***}	-0.000	0.001	0.019^{***}	-0.036^{***}
$\Delta Amihud$	$\Delta Amihud$	-0.001	-0.147	0.206	-0.150	-0.001	-0.050	0.133	-0.067
	ΔOC	-0.000	-0.104^{***}	-0.082^{***}	-0.102^{***}	0.000	-0.021^{***}	-0.006	-0.020^{***}
	ΔHL	0.000	-0.073^{***}	-0.105^{***}	0.085^{***}	0.000	-0.018^{***}	-0.026^{***}	0.015^{***}
	$\Delta OCHL$	-0.000	-0.081^{***}	0.091^{***}	-0.111^{***}	0.000	0.001	0.021^{***}	-0.037^{***}

This table presents the parameter estimates of the multivariate VARX specification which includes, along with the analyzed liquidity proxy, OC, HL and $OCHL$ as endogenous variables. The exogenous variables that are determined outside of the VARX process are the stocks' fixed-effects dummies. Each system is presented by four-line blocks in the table and each line represents an equation of the VARX process. The first column, ΔL_t, indicates the liquidity proxy that is analyzed. The second column, $Equation$, denotes the dependent variable that is being investigated in this part of the VARX process. The different columns show the parameters estimates for each lag of the four variables. QB and QA are the sums of the quantities outstanding at the five best limits, on bid and ask sides, respectively. Q is the sum of the two. $Slope$ denotes the slope of the book computed following Hasbrouck and Skjeltorp (2006) while $Dispersion$ stands for the dispersion measure computed following Kang and Yeo (2008). RS denotes the relative spread. NB and NS respectively denote the number of buy

The results clearly indicate that liquidity is most likely to drive price changes, whatever the market capitalization segment. This is derived from the very significant parameter estimates observed in the column ΔL_{t-1}. This is however not the case of $\Delta Volume$, ΔATS and $\Delta Amihud$. These outcomes suggest that prices move in response to a change in liquidity. A price discovery process led by informed traders is an example of this relationship. As also outlined by Mazza (2013), we observe bidirectional causality for some liquidity proxies, e.g. $Dispersion$ and RS, but this is not generalized to all the results. This may be related to the fact that changes in these proxies cause price variations which in turn influence these proxies. Interestingly, ex post liquidity proxies exhibit opposite signs and less significant estimates. This may be explained by the fact that liquidity is no longer analyzed contemporaneously, since they measure liquidity that has been removed from the order book. Contemporaneous ex post liquidity proxies denote liquidity that is currently being removed from the order book while lagged ex post liquidity proxies refer to liquidity that has already been removed from the order book and, as a result, has already influenced the prices.

Lagged price movements seem to have a weaker impact on contemporaneous liquidity than liquidity on prices. As a consequence, they may not be used to predict changes in liquidity. The contemporaneous relationship outlined in the previous section is however still valid. The deeper analysis of this relationship, that may be addressed through simultaneous equations, is left for further research.

As a conclusion, HLOC price variables change after the state of the limit order book is modified, even if some outcomes suggest a bidirectional relationship. This result is consistent with the previous literature on price discovery processes where informed traders move prices by consuming outstanding liquidity in the limit order book, pushing them towards their fundamental value.

3.3.5 Time intervals

In order to check the robustness of our results even further, we focus on different interval lengths but keep the sample unchanged regarding the stocks and dates. The results are presented in Table 3.12 and Table 3.13 for 10-minute and 20-minute intervals, respectively. Following the recommendations of Petersen (2009), we estimate the same regressions by OLS with clustered standard errors and control for within correlation across stocks.

Table 3.12: Robustness check : 10-minute interval regressions with clustered standard errors

ΔL_t	ΔOC_t	ΔHL_t	$\Delta OCHL_t$	ZR_t	UW_t	DW_t	D_t	ΔRV_t
Panel A : Small Caps								
ΔRS	0.033	0.169^{***}	-0.048^{**}	-0.088^{*}	-0.015	0.065	0.060	-0.034
$\Delta Dispersion$	0.017^{***}	0.006^{*}	-0.013^{***}	-0.014^{*}	0.062^{***}	0.045^{***}	0.013^{*}	-0.000
$\Delta Slope$	-0.016^{**}	-0.029^{***}	0.045^{***}	-0.082^{***}	-0.015	-0.037	-0.073^{***}	0.005
ΔQB	0.004	0.011	0.002	-0.143^{***}	-0.007	0.039	0.164^{***}	-0.003
ΔQA	0.007	0.005	-0.006^{*}	0.064^{***}	0.060	-0.023	-0.183^{***}	-0.007^{*}
ΔQ	0.004	0.005^{*}	-0.004^{*}	-0.010	0.017	0.010	-0.021^{**}	-0.005^{**}
ΔNB	0.053	0.100^{***}	-0.075^{***}	0.176^{***}	1.066^{***}	-0.648^{***}	-0.913^{***}	0.031
ΔNS	0.040^{*}	0.109^{***}	-0.089^{***}	-0.920^{***}	-0.532^{***}	1.080^{***}	1.034^{***}	0.039
$\Delta Ntrades$	0.056^{***}	0.113^{***}	-0.084^{***}	-0.240^{***}	0.500^{***}	0.324^{***}	-0.012	0.005
ΔATS	0.005	0.001	-0.019^{***}	0.163	-0.022	0.165	-0.034	0.001
$\Delta Volume$	0.073^{*}	0.199^{***}	-0.151^{***}	-0.215	0.890^{*}	1.150^{*}	-0.273	-0.007
$\Delta Amihud$	0.446	-0.374^{**}	0.424^{*}	1.339^{**}	0.230	1.148	-3.482	0.860^{***}
Panel B : Mid Caps								
ΔRS	0.011	0.174^{***}	-0.038^{***}	-0.130^{***}	-0.124^{*}	-0.126^{*}	-0.012	0.004
$\Delta Dispersion$	0.017^{***}	0.005^{*}	-0.011^{***}	-0.020^{***}	0.047^{***}	0.056^{***}	0.010^{**}	0.004
$\Delta Slope$	-0.017^{***}	-0.030^{***}	0.041^{***}	-0.075^{***}	0.153^{***}	0.094^{***}	-0.018	-0.003
ΔQB	0.001	0.008^{***}	0.007	-0.087^{***}	-0.015	0.037^{*}	0.120^{***}	0.001
ΔQA	0.007	0.009^{**}	-0.003	0.024^{**}	0.117^{***}	-0.026	-0.157^{***}	-0.006
ΔQ	0.002	0.003^{**}	0.003	-0.009	0.020	-0.005	-0.017^{***}	-0.002
ΔNB	0.078^{***}	0.069^{***}	-0.102^{***}	0.214^{***}	0.584^{***}	-0.767^{***}	-0.787^{***}	0.071^{***}
ΔNS	0.031^{**}	0.104^{***}	-0.087^{***}	-0.612^{***}	-0.623^{***}	0.440^{***}	0.806^{***}	0.055^{***}
$\Delta Ntrades$	0.033^{**}	0.120^{***}	-0.076^{***}	-0.174^{***}	0.142^{***}	0.027	-0.009	0.033^{***}
ΔATS	-0.002	0.059^{**}	0.010	0.012	-0.121	0.592	-0.352^{***}	-0.020
$\Delta Volume$	0.018	0.471^{***}	-0.142^{***}	-0.077	0.334	0.078	-0.692^{***}	-0.015
$\Delta Amihud$	0.288^{*}	-0.449^{***}	0.563^{***}	0.114	5.303^{***}	2.613^{***}	0.360	0.880^{***}
Panel C : Large Caps								
ΔRS	0.031^{***}	0.141^{***}	-0.032^{***}	-0.096^{***}	-0.069	-0.138^{***}	0.038^{***}	0.063^{***}
$\Delta Dispersion$	0.013^{***}	0.025^{***}	-0.010^{***}	-0.015^{***}	0.059^{***}	0.028^{***}	0.005^{*}	0.014^{***}
$\Delta Slope$	-0.018^{***}	-0.080^{***}	0.052^{***}	-0.031^{***}	0.187^{***}	0.246^{***}	-0.025^{***}	-0.013^{***}
ΔQB	0.003	0.000	0.004	-0.073^{***}	-0.110^{***}	-0.014	0.045^{***}	-0.003
ΔQA	0.008^{***}	-0.000	0.000	-0.013^{*}	-0.010	-0.082^{***}	-0.057^{***}	-0.012^{***}
ΔQ	0.002	-0.002	0.002	-0.016^{***}	-0.068^{***}	-0.058^{***}	-0.007^{**}	-0.008^{***}
ΔNB	-0.001	0.393^{***}	-0.074^{***}	0.239^{***}	-0.059	-0.098	-0.645^{***}	0.048^{***}
ΔNS	-0.004	0.314^{***}	-0.071^{***}	-0.355^{***}	-0.030	-0.060	0.572^{***}	0.032^{***}
$\Delta Ntrades$	0.013^{*}	0.289^{***}	-0.054^{***}	-0.044^{***}	0.108^{***}	0.013	0.001	0.012
ΔATS	-0.013^{***}	0.070^{***}	-0.014	-0.014	-0.136	0.131	-0.165	0.025
$\Delta Volume$	-0.066^{**}	0.680^{***}	-0.043	0.018	0.097	0.399	-0.228	0.033
$\Delta Amihud$	0.371^{***}	-0.320^{***}	0.758^{***}	0.477^{***}	5.420^{***}	5.374^{***}	0.058	0.573^{***}

This table presents the different models that are estimated by OLS with the clustering approach for a 10-minute intervals sample. Panels A, B and C display the estimates for small, mid and large caps, respectively. The first column, L_t, indicates the dependant variable and the following columns are the exogenous variables. QB and QA are the sums of the quantities outstanding at the five best limits, on bid and ask sides, respectively. Q is the sum of the two. $Slope$ denotes the slope of the book computed following Næs and Skjeltorp (2006) while $Dispersion$ stands for the dispersion measure computed following Kang and Yeo (2008). RS denotes the relative spread. NB and NS respectively denote the number of buy and sell trades. $Ntrades$ is the total number of trades and ATS is the average trade size of these trades. $Amihud$ denotes the Amihud ratio. $OC_t = \frac{|Close_t - Open_t|}{RangeMid_t}$ is the absolute value of the net movement. $OCHL_t = \frac{|Close_t - Open_t|}{High_t - Low_t}$ is the percentage of net variation on total fluctuation. HL_t denotes the High-Low range of the interval. ZR_t is a dummy variable that control for the presence of zero returns. UW_t is equal to 1 when an upward window occurs on the chart and 0 otherwise, while DW_t is equal to 1 when a downward window appears and 0 otherwise. The D_t dummy equals 1 when the net price movement is positive and 0 otherwise. The fixed effect dummies are not presented here.

Table 3.13: Robustness check : 20-minute interval regressions with clustered standard errors

ΔL_t	ΔOC_t	ΔHL_t	$\Delta OCHL_t$	ZR_t	UW_t	DW_t	D_t	ΔRV_t
Panel A : Small Caps								
ΔRS	0.050^{**}	0.111^{***}	-0.064^{***}	-0.168^{***}	-0.283^{**}	0.032	0.149^{***}	0.034
$\Delta Dispersion$	0.016^{***}	0.007^{***}	-0.013^{***}	-0.012	0.062^{***}	0.039^{*}	0.015^{**}	0.000
$\Delta Slope$	-0.019^{***}	-0.029^{***}	0.052^{***}	-0.157^{***}	0.095^{*}	0.033	-0.007	0.005
ΔQB	0.006^{*}	0.004	0.004	-0.128^{***}	-0.015	0.041	0.170^{***}	-0.001
ΔQA	0.011^{***}	-0.002	-0.000	0.084^{***}	0.019	-0.075^{***}	-0.174^{***}	0.000
ΔQ	0.006^{**}	0.000	0.001	0.010	-0.004	-0.017	-0.013	-0.001
ΔNB	0.085^{***}	0.137^{***}	-0.113^{***}	0.227^{***}	1.327^{***}	-0.593^{***}	-0.882^{***}	0.072^{**}
ΔNS	0.027	0.074^{***}	-0.100^{***}	-0.837^{***}	-0.390^{*}	0.949^{***}	0.969^{***}	0.104^{***}
$\Delta Ntrades$	0.050^{***}	0.117^{***}	-0.098^{***}	-0.285^{***}	0.580^{***}	0.275^{***}	0.028	0.067
ΔATS	0.009	-0.001	-0.036^{***}	-0.134	-0.054	-0.109	0.003	0.001
$\Delta Volume$	0.076^{*}	0.156^{***}	-0.219^{***}	-0.637^{***}	0.518	-0.023	-0.077	0.163
$\Delta Amihud$	0.028	-0.068	1.341^{**}	1.615^{**}	5.721^{**}	2.596^{***}	-0.685	0.508^{***}
Panel B : Mid Caps								
ΔRS	0.051^{***}	0.087^{***}	-0.050^{***}	-0.172^{***}	-0.081	-0.000	0.074^{***}	0.013
$\Delta Dispersion$	0.012^{***}	0.005^{***}	-0.007^{***}	-0.036^{***}	0.060^{***}	0.046^{***}	0.012^{**}	0.002
$\Delta Slope$	-0.020^{***}	-0.022^{***}	0.035^{***}	-0.062^{***}	0.132^{***}	0.047^{*}	-0.041^{**}	-0.001
ΔQB	0.005^{**}	0.002	-0.004^{*}	-0.101^{***}	-0.040^{*}	0.007	0.107^{***}	0.006
ΔQA	0.006^{***}	0.000	-0.004^{**}	0.030^{**}	0.151^{***}	-0.025	-0.153^{***}	-0.000
ΔQ	0.003^{***}	-0.000	-0.002	-0.007	0.027^{*}	-0.003	-0.027^{***}	0.001
ΔNB	0.082^{***}	0.090^{***}	-0.103^{***}	0.179^{***}	1.114^{***}	-0.802^{***}	-0.784^{***}	0.094^{***}
ΔNS	0.055^{***}	0.085^{***}	-0.093^{***}	-0.647^{***}	-0.381^{***}	0.653^{***}	0.794^{***}	0.072^{***}
$\Delta Ntrades$	0.056^{***}	0.098^{***}	-0.084^{***}	-0.182^{***}	0.516^{***}	0.048	-0.020	0.040^{***}
ΔATS	0.001	0.016	-0.017	-0.325	-0.144	0.810	0.134	0.005
$\Delta Volume$	0.071	0.331^{***}	-0.145^{**}	-0.459	1.266^{**}	0.963	-0.031	0.127^{**}
$\Delta Amihud$	0.564^{**}	-0.379^{**}	0.898^{***}	1.778^{**}	4.284^{***}	6.741^{**}	0.104	0.489^{***}
Panel C : Large Caps								
ΔRS	0.015^{***}	0.134^{***}	-0.016^{***}	-0.107^{***}	-0.139^{***}	-0.152^{***}	0.034^{***}	0.024
$\Delta Dispersion$	0.007^{***}	0.026^{***}	-0.007^{***}	-0.024^{***}	0.017	0.022	0.002	0.006^{**}
$\Delta Slope$	-0.011^{***}	-0.071^{***}	0.032^{***}	-0.037^{***}	0.232^{***}	0.202^{***}	-0.026^{***}	-0.000
ΔQB	0.001	-0.007^{**}	0.005^{*}	-0.073^{***}	-0.075^{***}	-0.018	0.033^{***}	-0.003
ΔQA	0.002	-0.006	0.004	-0.033^{***}	-0.029	-0.095^{***}	-0.059^{***}	-0.004
ΔQ	0.000	-0.008^{**}	0.003^{*}	-0.022^{***}	-0.066^{***}	-0.067^{***}	-0.012^{***}	-0.006^{*}
ΔNB	0.038^{**}	0.294^{***}	-0.077^{***}	0.192^{***}	0.965^{***}	-0.034	-0.490^{***}	0.111^{***}
ΔNS	0.016^{*}	0.240^{***}	-0.057^{***}	-0.277^{***}	0.462^{***}	0.500^{***}	0.434^{***}	0.104^{***}
$\Delta Ntrades$	0.034^{***}	0.214^{***}	-0.057^{***}	-0.056^{***}	0.638^{***}	0.293^{***}	-0.008	0.078^{***}
ΔATS	0.001	0.057^{***}	-0.011^{***}	-0.007	-0.073	0.023	-0.029^{***}	-0.008^{*}
$\Delta Volume$	0.015	0.421^{***}	-0.065^{***}	0.003	0.774^{***}	0.393	-0.065^{***}	0.125^{***}
$\Delta Amihud$	0.269^{***}	-0.073	0.726^{***}	0.366^{***}	3.611^{***}	3.678^{***}	0.038	0.212^{**}

This table presents the different models that are estimated by OLS with the clustering approach for a 20-minute intervals sample. Panels A, B and C display the estimates for small, mid and large caps, respectively. The first column, L_t, indicates the dependant variable and the following columns are the exogenous variables. QB and QA are the sums of the quantities outstanding at the five best limits, on bid and ask sides, respectively. Q is the sum of the two. $Slope$ denotes the slope of the book computed following Næs and Skjeltorp (2006) while $Dispersion$ stands for the dispersion measure computed following Kang and Yeo (2008). RS denotes the relative spread. NB and NS respectively denote the number of buy and sell trades. $Ntrades$ is the total number of trades and ATS is the average trade size of these trades. $Amihud$ denotes the Amihud ratio. $OC_t = \frac{|Close_t - Open_t|}{RangeMid_t}$ is the absolute value of the net movement. $OCHL_t = \frac{|Close_t - Open_t|}{High_t - Low_t}$ is the percentage of net variation on total fluctuation. HL_t denotes the High-Low range of the interval. ZR_t is a dummy variable that control for the presence of zero returns. UW_t is equal to 1 when an upward window occurs on the chart and 0 otherwise, while DW_t is equal to 1 when a downward window appears and 0 otherwise. The D_t dummy equals 1 when the net price movement is positive and 0 otherwise. The fixed effect dummies are not presented here.

These tables show that the results are very similar to the core analysis (Table 3.6). The highlighted dynamics seem to be valid when we consider longer or shorter time frames. The estimates are even more significant for 20-minute intervals than for 10-minute intervals but this may be attributed to the increase in noise that affects smaller time frames, especially for small caps.

There are very few differences in the significance levels but the signs remain consistent in general. For instance, the slope displays more significant estimates, for each portfolio, in both samples in comparison to Table 3.6. Depth proxies show significant estimates that are inconsistent with H1 for small and mid caps in the 20-minute sample but remain consistent for large caps. To sum up, the conclusions are unchanged when the interval length is modified.

3.4 Conclusion

Liquidity measurement is not an easy task. On the one hand, liquidity is empirically specified in several dimensions. On the other hand, very few theoretical models (if any) can be applied in real-time by practitioners. In this context, efficient intraday liquidity estimation remains a thorny issue in finance. Not surprisingly, the art of making accurate and quick estimations of intraday liquidity is still in its infancy and represents a serious challenge for both practitioners and academics. In this paper, we ask the following question: Can we improve the state of the art in real-time liquidity measurement by looking at the HLOC price dynamics on an intraday basis? Although price movements are recognized as useful for evaluating trading activity and liquidity in the existing literature, no study has investigated the information content of HLOC intraday price dynamics for liquidity evaluation. A relation between liquidity and volatility has been well documented but HLOC dynamics contain much more information than the usual volatility proxies.

Using Euronext order book data, we define three cap-based portfolios of 100 stocks each and compute several proxies for liquidity, price movements, and trading activity. The dependent variables characterize the different dimensions of liquidity. We include several proxies for order-based measures: depth, spread, slope and dispersion. We also test trade-based proxies: number of trades, mean size of trades and Amihud (2002)'s illiquidity ratio. We estimate the regression using clustered standard errors that control for within correlation across stocks. In a second step, we include a proxy for the realized volatility in order to control for volatility dynamics that may affect price movements. We also conduct robust and median regressions to control for the presence of outliers in the sample. As a further robustness check, we test our models at the 10-minute and 20-minute intervals.

Our findings suggest that price movements help better characterize liquidity dynamics whatever the liquidity dimension considered. Firstly, we conclude that the intensity of the price discovery process (that we characterize by the OC range) lowers ex-ante and ex-post liquidity. Secondly, a wider total price variation that results from higher price uncertainty (i.e. a higher HL range) deteriorates ex-ante liquidity even further. Thirdly, controlling for the intensity of the price discovery process and the level of price uncertainty, we find that higher 'relative price certainty' (that we measure by the interaction variable of the OC range divided by the HL range) displays a positive relationship to ex-ante liquidity. Fourthly, liquidity seems to be higher when there is a zero return, noticing that the inclusion of realized volatility as a control variable has a positive effect on this relation. The results are nevertheless relatively sensitive to the estimation method. The most convincing results with respect to ex-ante and ex-post liquidity are found when the robust and clustered OLS techniques are used, respectively. Fifthly, ex-post liquidity is lower when there has been a price jump. Weak evidence is found for ex-ante liquidity. Finally, both depth and order imbalances decrease when there is an upward price movement.

All in all, positive changes in price ranges for both Close-Open and High-Low ranges are related to negative variations in liquidity proxies. All else equal, liquidity is higher when

the two ranges coincide (i.e. when the net movement measured by the OC range does not differ from the total variation during the interval), pointing to a high level of 'relative price certainty'. Liquidity seems to be higher when there is a zero return. It is the opposite when price gaps occur, at least regarding ex-post liquidity. All these results are very much confirmed across all three cap-based portfolios. The large cap portfolio nevertheless display the most convincing results, which may be explained by the higher number of analysts following these stocks.

We conclude that the investigation of the dynamics of HLOC prices at the intraday level helps practitioners characterize the state of the limit order book more efficiently than by trying to measure each liquidity dimension in turn. In this respect, HLOC prices constitute a useful source of information, notwithstanding the fact that they can be displayed in any trading platform around the world.

CHAPTER

FOUR

RETHINKING ZERO RETURNS IN THE LIQUIDITY PUZZLE OF A LIMIT ORDER MARKET

4.1 Introduction

The multidimensionality of liquidity has always been a particular concern in market microstructure research. Harris (2003) defines liquidity as "the ability to trade large size quickly, at low cost, when you want to trade".[1] In this definition, three liquidity dimensions can be identified: immediacy, width, and depth. An additional dimension, resiliency, is often referred to and is related to the recovery following a liquidity shock. All these dimensions are associated with multiple proxies. Many research studies deal with the quest for the best liquidity proxy or transaction costs estimator (Goyenko et al., 2009; Aitken and Comerton-Forde, 2003). Liquidity proxies broadly fall into two categories: order book-based and trade-based measures. Order book-based measures are computed from the order book on

[1] Harris (2003), pp. 394.

high frequency data and characterize ex-ante liquidity, i.e. before that trading occurs. Trade-based proxies are computed ex-post with information on trading activity: volumes, number of trades, average trade size, etc. Another strand of the literature associates liquidity to asset price movements. This paper relates to this literature by examining whether the illiquidity measure based on the occurrence of zero returns, as outlined in Lesmond et al. (1999), is an adequate proxy for market liquidity. Several papers examine the information content of price series to create easy-to-compute liquidity or transaction cost measures. Roll (1984)'s effective spread estimator or Amihud (2002)'s illiquidity ratio are among the best examples. The main advantage of such estimators is easy to understand since price series are more widely available, even for emerging markets, than order book and trade data.

Both theoretical and empirical evidence have been provided in the literature to justify such a relationship. For instance, Chordia et al. (2001) empirically find that liquidity and trading activity are influenced by market returns and volatility. Chordia et al. (2002) also find that order imbalances seriously affect market returns. Furthermore, as shown in Kyle (1985)'s model, a high order imbalance means the presence of some private information on which informed traders base their order submissions, creating a temporarily reduced liquidity which in turn makes prices more informative. Some research studies also relate returns to liquidity in an attempt to include it in asset pricing models. Amihud and Mendelson (1986) theoretically find that the return is an increasing function of the bid-ask spread. Pastor and Stambaugh (2003) also propose to include liquidity in asset pricing models as they identify a relationship between returns and a liquidity proxy based on order flow estimation on a 34-year sample. Other papers, such as Brennan and Subrahmanyam (1996), Brennan et al. (1998) and Datar et al. (1998), show that returns and liquidity are negatively related. More recently, Kavajecz and Odders-White (2004) identify a strong relationship between liquidity and technical analysis through the use of support and resistance, as well as moving average indicators.

Lesmond et al. (1999) present a new method for the estimation of transaction costs using

the incidence of zero returns in daily data, as opposed to high frequency and order book data that are usually required to address transaction costs and liquidity issues. They argue that the more frequent the zero returns, the lower the liquidity. Their reasoning is based on the hypothesis that transaction costs constitute a threshold that must be exceeded by the value of the upcoming information. A zero return will therefore occur when traders do not consider the information available to them sufficiently valuable to cover transaction costs. The value associated to the information set may be related to its time horizon. Bekaert et al. (2007) use this proxy on emerging markets data as their dataset only contains price series. Goyenko et al. (2009) also test whether this measure, along with others, is an accurate liquidity proxy. They find that zero returns are not good at capturing the effective spread and that the performance of zero return measures has not deteriorated after the change to decimals.

In this paper, we investigate whether the frequency of zero returns, as presented in Lesmond et al. (1999), is an appropriate illiquidity proxy for limit order markets by using intraday data, comparing it to order book-based liquidity proxies. In this case, it would provide additional information to intraday datasets that do not include details on pending orders or the state of the limit order book beyond the best quotes. These details are much less frequent in the datasets and even more difficult to gather given the recent emergence of Alternative Trading Systems (ATS), Multilateral Trading Facilities (MTF) and cross-listing. Using an event study on Euronext market data that cover 701 stocks, we find that 15-minute zero returns are not related to illiquidity but rather to liquidity, measured during the same time interval. The results are valid across small, mid and large cap subsamples. We further run conditional logit regressions with a dummy variable that captures zero returns as the dependent variable. The set of regressors includes spread, depth, dispersion and slope proxies as well as trading activity variables. The outcomes corroborate our previous findings. We therefore argue against the use of zero returns as a measure of illiquidity and instead provide evidence of a positive relationship to liquidity in order-driven markets. We nevertheless confirm that the frequency of zero returns may still be considered as a proxy for the absence of informed trading, as proposed by Lesmond et al. (1999) and discussed by Bekaert et al. (2007).

In order to further check our intraday results, we carry on the same analysis on different other timeframes: 20, 30 and 60 minutes. The results are consistent with the findings of 15-minute intervals. The outcomes clearly confirm that zero returns are related to liquidity measured in the order book. In a second step, we consider daily zero returns and how they are related to the same liquidity variables. Since Lesmond et al. (1999) use daily occurrences of zero returns, we examine whether daily zero returns are associated with liquidity or illiquidity. Even if the results are less significant, zero returns are more likely to characterize liquidity rather than illiquidity in daily datasets.

The remainder of the paper is organized as follows. Section 4.2 discusses zero-returns as an indicator of the state of liquidity. Section 4.3 describes the dataset and the different liquidity measures that are used. Section 4.4 presents the event study methodology and the results. Section 4.5 reports the methodology and the outcomes of the fixed-effects logit regressions. The final section concludes.

4.2 Zero returns and (il)liquidity

4.2.1 The zero return measure

Lesmond et al. (1999) relate zero returns to illiquidity by arguing that zero returns (or zeros) occur when informed traders are not (or less) willing to trade. They also argue that less liquid stocks are more likely to exhibit zero volume days, hence zero returns. Following their reasoning, zero returns occur when the value of the new information set does not exceed the cost of trading. Therefore, informed traders do not react to the information signal. Lesmond et al. (1999) find that firm size is negatively related to the frequency of zero returns and positively related to both Roll (1984)'s measure and quoted spread. They intuitively argue that informed traders do not trade after the evaluation of the new information set's value,

which would lead to a zero return, as the effect of other types of traders is assimilated to noise.

Their tobit model assumes that the unobserved true return on day t for the stock i (R^*_{it})is given by:[1]

$$R^*_{it} = \beta_j R_{mt} + \epsilon_{it}, \tag{4.2.1}$$

with

$$\begin{aligned} R_{it} &= R^*_{it} - \alpha_{1i} \quad \text{if} \quad R^*_{it} < \alpha_{1i}, \\ R_{it} &= 0 \qquad\quad\;\; \text{if} \quad \alpha_{1i} < R^*_{it} < \alpha_{2i}, \\ R_{it} &= R^*_{it} - \alpha_{2i} \quad \text{if} \quad R^*_{it} > \alpha_{2i}, \end{aligned} \tag{4.2.2}$$

where R_{mt} is the market return for day t and ϵ_{it} is a public information shock for stock i at day t.

The threshold that the value of the new information set should exceed is denoted as $\alpha_{1i} \leq 0$ for negative information and $\alpha_{2i} \geq 0$ for positive information, for stock i. This model is estimated by maximum likelihood, using the function described in Lesmond et al. (1999).

This model follows the framework of Copeland and Galai (1983), Glosten and Milgrom (1985) and Kyle (1985) and implies that the arrival of new information is correctly valued by each informed trader as well as that their reaction timings are similar.

Based on this limited dependent variable model, Lesmond et al. (1999) build two types of

[1]For more information on the model, please refer to Lesmond et al. (1999).

measures. The first one is presented as the difference between the thresholds, $\alpha_{2i} - \alpha_{1i}$, and the second one is computed as the proportion of zero returns on a given time interval. These measures have been put in question in several papers. Goyenko et al. (2009) test the usefulness of these measures by comparing them to a large set of liquidity proxies, emphasizing that the frequency of zero returns does not constitute an accurate proxy for liquidity. Bekaert et al. (2007) use the frequency of zero returns to address the relationship between liquidity and asset pricing on emerging market data, since they only dispose on price series. They also discuss the limitations of this measure. Levine and Schmukler (2006) study the relationship between liquidity and cross-listing using the frequency of zero returns. Liu (2006) discuss and modify the measure but Chang et al. (2010) assess that Liu (2006)'s modification is not significantly related to stock returns in the Japanese market, as opposed to zeros. Lang et al. (2010) apply the proxy to evaluate whether liquidity, transparency and valuation are related to each other.

4.2.2 Zero returns, informed trading and liquidity

In a dealership market, Lesmond et al. (1999)'s model may still hold since the market maker will probably widen the spread if a new information comes and if informed traders have failed to stay hidden. In this type of market, the dealer acts as a counterparty for all order submissions. The dealer is a monopolist with regard to the determination of the bid and ask quotes. The detection of informed traders, or of their willingness to trade, forces the market maker to enlarge the spread, which in turn impacts the return that informed traders would have made without the dealer's intervention. They are therefore not willing to trade anymore. The market becomes more illiquid and the probability of a zero return is higher. However, Lesmond et al. (1999) use data from NYSE/AMEX individual stocks which are attributed designated market makers, also called specialists. The specialists may trade if all limit orders outstanding at the best quotes have been fulfilled. Traders are therefore more likely to trade with each other, as it is the case in a traditional limit order market.

When we more closely examine the applicability of this model in a limit order book market, some hypotheses do not hold anymore. One of the main assumptions behind this model is that only informed traders move prices as zero returns are said to proxy zero "informed" volume. However, uninformed traders may also significantly influence prices and recent research evidence has shown that informed traders do not always trade aggressively, e.g. Bloomfield et al. (2005). Bekaert et al. (2007) also argue in favor of that point even if zero returns may still be considered as a measure of the lack of informed trading, in its most usual form. Mazza (2013) also confirms that there are less informed traders when opening and closing prices are very close, as measured by the PIN indicator of Easley et al. (1996). Lesmond et al. (1999) did not test this hypothesis directly as it is a condition for their model to hold, i.e. the zero return is the consequence of informed traders' non-willingness to trade after the upcoming of a new information whose value does not exceed the threshold. Bekaert et al. (2007) nevertheless argue that news are associated to shocks which are related to excess volatility. As a consequence, if there are news that do not enable trading or no news, there is no excess volatility, which implies a higher liquidity. This has been suggested by Pagano (1989), among others, who identifies a positive relationship between illiquidity and volatility. This link is opposed to the proposition of Lesmond et al. (1999) who argue that zero returns imply higher transaction costs, hence lower liquidity. Bekaert et al. (2007) also present cross-listing of firms as a serious limitation of the model, since local liquidity may be dramatically different from foreign liquidity.

Another reason that puts the usefulness of the zero return measure in limit order markets in question resides in the connection between zeros and illiquidity. Lesmond et al. (1999) establish the link by making the hypothesis that transaction costs are higher when informed traders do not want to trade after the arrival of a new information set because the value of this new information does not exceed the cost of trading. If we question this hypothesis, as we should for limit order markets with a public and visible order book, the model does not hold anymore. The only proposition that would still apply is that zeros are linked to less informed trading. Lesmond et al. (1999), Bekaert et al. (2007) and Mazza (2013) agree on that point. Their reasoning only differ in the way informed trading is associated with

illiquidity. Lesmond et al. (1999) argue that a high level of transaction costs, hence an illiquid state of the order book, is what forces informed traders not to trade. Mazza (2013) however suggests that informed traders trade less because there is a consensus on the fair value of the security and prices are efficient. As informed traders do not participate to the current session, i.e. they do not trade to move the price towards the fundamental value, the order book presents a higher liquidity. There is an extensive literature on order submissions made by informed traders which suggests that the absence of informed trading should result in a higher liquidity. This argument is opposed to the reasoning of Lesmond et al. (1999). For instance, Harris (1998) suggests that informed traders' use of market orders is higher when they believe their information is short-lived. Harris (2003) further suggests that "their most important decision is whether to trade aggressively or not".[1] This decision implies the minimization of trading costs: if the stock is (not) liquid, they may (not) trade aggressively. If they know that their informational advantage will last long enough, they will prefer trading slowly to diminish their market impact. Anand et al. (2005) also empirically investigate the changes in the trading strategies of informed traders and find that they place liquidity-taking orders earlier in the day while they provide liquidity later in the same day.

On that account, Mazza (2013) proposes some theoretical justifications to characterize the higher liquidity that occurs when there is a temporary consensus between buyers and sellers on the fundamental value of the security. He bases his reasoning on Bloomfield et al. (2005) who show that informed traders provide liquidity when they estimate their information has a low value, rather than consuming it from the book, as they do when they estimate this value as high. In their model, informed traders do not stop trading when the value of their information goes down, but submit aggressive limit orders to earn the spread. In the hypothesis of asymmetric information, they incur a lower risk when they submit limit orders as they do not bear the cost of trading against a better informed trader. Behaving as dealers when they can not profitably trade on their information set is their best choice to make money. Bloomfield et al. (2005) also suggest that their profits are even higher in this last case. In addition, Harris (1998) and Bloomfield et al. (2005) demonstrate that liquidity

[1] Harris (2003), pp. 225, Section 10.4

traders are also most likely to change their way of trading throughout the day: They initially try to meet their target by submitting limit orders, which are less expensive, but they become more aggressive since their non-execution risk increases near the end of the day. As a result, they hit the liquidity offered by informed traders. These propositions are consistent with Anand et al. (2005). Furthermore, symmetrically informed traders will not trade with each others, as outlined by Harris (2003). This implies that trading activity is driven by liquidity or noise traders who pick informed traders' pending limit orders off. Informed traders will try to gain price priority by reducing the spread and come closer to the fundamental value to increase the probability of being picked off by other traders. There is a competition effect that makes the order book more informative, more dense, and the spread narrower. Trading is reactivated when informed traders estimate that the spread is sufficiently low to hit the best opposite quote or when the fundamental value has changed, as explained in Mazza (2013). As a result, when a price discovery process has been sufficiently efficient to bring an agreement on the fundamental value of the security that is located inside the spread, informed traders provide liquidity to earn the spread.

Furthermore, Lesmond et al. (1999) use zero returns as low frequency liquidity proxies, aiming at measuring liquidity when volumes, trades and order data are not available. Based on high frequency data, we nevertheless argue that zero returns are not always related to less trading volume, since liquidity may abound on both sides of the book, implying that prices do not move before the outstanding liquidity has completely dried out at the best quotes. There is clearly a need to disentangle zero volume from zero returns for large as well as small caps.

In addition, if we consider the widely used Amihud illiquidity ratio, presented in Amihud (2002), which is defined as $IL_t = \frac{|Return_t|}{Volume_t}$, i.e. the lower the return for a given volume, the higher the liquidity. When returns are closer to zero, for a given volume, the ratio is lower and the liquidity is higher. As a consequence, the interpretation of the ratio is opposed to the rationale that Lesmond et al. (1999) present.

Given these limitations, we question the use of zero returns as a proxy for illiquidity or for higher transaction costs in a limit order market. In this paper, we analyze the connection between intraday zero returns and different order book-based and trade-based liquidity proxies. Our study focuses on intraday zero returns rather than daily to examine the justification of the proxy on lower time frames. Addressing the occurrence of zero returns and their implications towards market microstructure is different from the initial measure proposed by Lesmond et al. (1999). Nevertheless, their justifications are grounded on the interpretation of each individual zero return that they aggregate to form a low-frequency proxy. As a result, their reasoning should also be verified on higher frequency data. The possible pitfall of moving to intraday intervals resides in the fact that intraday liquidity is U-shaped (Biais et al., 1995). In order to make our results robust, we take this U-shaped pattern into account in our event study methodology. To the best of our knowledge, this paper is the first to thoroughly analyze the information content of zero returns on an intraday time frame. As robustness checks, we consider different interval lengths and investigate whether the relationship does still hold for daily data.

4.3 Data and liquidity proxies

4.3.1 Data

We test the relationship between zeros and liquidity using Euronext market data on 701 stocks. This unique and rich dataset contains all orders and trades for 61 trading days from February 1 to April 30, 2006. Since the implementation phase of MiFID starting in November 2007, volumes have been shifting from national exchanges to Multilateral Trading Facilities (MTF) due to cross-listing. The key advantage of this dataset is to avoid that phenomenon. In order to still be representative of market activity, more recent datasets must include sufficient information from MTF and market data, which has become extremely dif-

ficult in today's decentralized trading environment. In addition, our dataset includes market members' ID that we use to disentangle buyer-initiated and seller-initiated trades, without any error margin. Traditionally, in market microstructure studies, the Lee and Ready (1991) algorithm is used to categorize buyer and seller-initiated trades. This algorithm has proved to be sufficiently efficient but misclassification may still occur. In our case, we are able to precisely match the two orders that generate each transaction. Finally, we are also provided with undisclosed data on hidden orders.[1] We build 15-minute-intervals from 9:00 AM to 5:30 PM (CET), which leads to 34 intervals per day.[2] In the robustness checks section, we investigate whether our findings still hold for different time frames.

In our event study, we need to control for contagious events, i.e. more than one event in a window. An event is the occurrence of a zero return. In order to avoid contagion effects between events, we do not consider occurrences in the previous and next three periods around the occurrence of the zero return. We only consider [-3,+3] windows where there is only one event occurring at time $t = 0$. Before filters, the sample contains 671843 zero returns. After filters, only 6302 events remain. This is the consequence of non-trading that often affects small caps stocks. The event study methodology requires considering one single event in each window. Controlling for contagious events in our case implies that most of the zero returns related to illiquidity are by construction removed from the sample since they are expected to be clustered on a given trading day. The remaining zero returns are therefore more likely to be related to liquidity. In any case, illiquid zero returns may easily be identified on a given security and may be easily excluded. Furthermore, in the conditional Logit specification, the whole sample is used to circumvent this sample selection issue. We split the complete database into three capitalization-based portfolios, motivated by Lesmond et al. (1999) who outline that firm size has an impact on the relationship between zero returns and liquidity. Large, mid, and small cap companies respectively exhibit a market capitalization larger than EUR 1 billion, between EUR 150 millions and EUR 1 billion, and

[1] Hidden orders are orders that gradually display part of their total amount. For instance, a hidden order of 500 could appear on the book with a quantity of 100 and will automatically be refilled when 100 shares have been consumed.

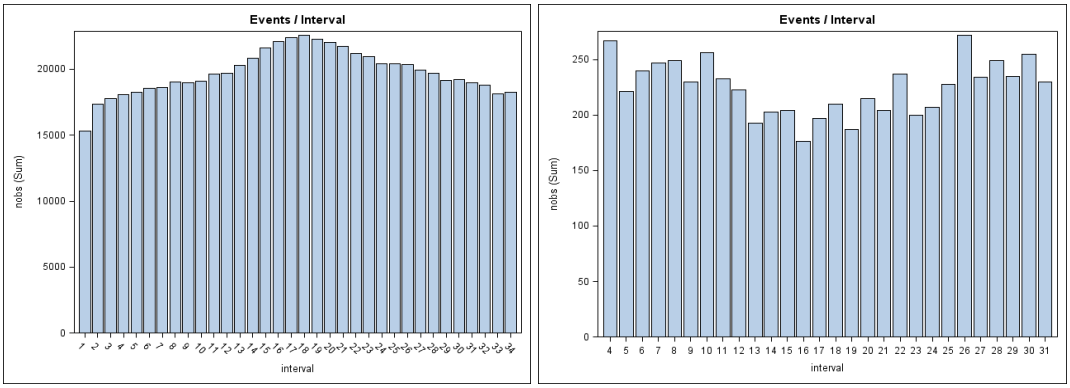
[2] This interval length is commonly used in the literature, e.g. in Mazza (2013).

below EUR 150 millions.

Figure 4.1 shows the distribution of our events among the 34 intervals. We observe that the two distributions are roughly uniform. The unfiltered sample however displays a peak around midday which comes from non-trading that occurs at lunch time. These zero returns are ruled out when we control for contagion in events.

Figure 4.1: Zero returns by interval

(a) No filter **(b)** Filter



This figure displays the number of events in each time interval. Panel (a) and (b) respectively present the distribution of the events before and after filters have been applied. The 34 intervals correspond to 15-minute intervals starting at 9:00 AM until 5:30 PM.

4.3.2 Liquidity proxies

We analyze the relationship between zero returns and liquidity proxies by measuring liquidity at the end of each 15-minute interval. This allows us to directly link liquidity to zero returns. We first analyze order book-based liquidity measures such as the relative spread and depth. We consider two levels of depth: depth at the best quotes and depth beyond the best quotes up to the fifth limit. These variables are computed in numbers of shares. We then include dispersion and slope measures that are respectively presented in Næs and Skjeltorp (2006) and Kang and Yeo (2008). The dispersion measures how far from each other are the

price limits: the more distant the quotes, the higher the dispersion. This proxy is computed as follows:

$$Dispersion_{i,t} = \frac{1}{2}\left(\frac{\sum_{j=1}^{5} w_{i,j,t}^{Bid} Dst_{i,j,t}^{Bid}}{\sum_{j=1}^{5} w_{i,j,t}^{Bid}} + \frac{\sum_{j=1}^{5} w_{i,j,t}^{Ask} Dst_{i,j,t}^{Ask}}{\sum_{j=1}^{5} w_{i,j,t}^{Ask}}\right), \tag{4.3.1}$$

where, for security i and interval t, $w_{i,j,t}^{Bid/Ask}$ are the weights which are equal to ask and bid sizes, at the j^{th} price limit normalized by the total depth of the five best limits, $Dst_{i,j,t}^{Bid} = (Price_{i,j-1,t}^{Bid} - Price_{i,j,t}^{Bid})$ and, $Dst_{i,j,t}^{Ask} = (Price_{i,j,t}^{Ask} - Price_{i,j-1,t}^{Ask})$. The midquote is used for the distance of the first best limits.

The dispersion is a liquidity proxy as for a given large market order, the resulting transaction costs are lower if the book is more dense, i.e. less disperse. It is also a measure of traders' willingness to provide liquidity. The book becomes more dense when traders compete to supply liquidity at the best quotes, and obtain price priority. This is an interesting measure as it combines both prices and quantities.

The slope is a measure of the elasticities $\partial q/\partial p$. As outlined by Næs and Skjeltorp (2006), the slope is negatively correlated to both volatility and trading activity. Glosten (1994) and Goldstein and Kavajecz (2004) also argue in favor of a negative relationship between order book slope and volatility. The main explanation resides in the adjustment of prices towards new equilibria when a new information arrives. Næs and Skjeltorp (2006) also find a negative relationship between the slope and the coefficient of variation in analysts' earnings forecasts. As a result, the slope also proxies for the disagreement among analysts on the value of the security: the more gentle the slope, the higher the level of disagreement. Volumes are concentrated when the slope is steep, as analysts agree on the value of the security. The slope is then positively related to liquidity. The slope should also increase when a zero return occurs, since it materializes a strong agreement among informed traders

on the fundamental value of the security. Demand ($DE_{i,t}$) and supply ($SE_{i,t}$) elasticities are respectively calculated, for stock i and interval t, as follows:

$$DE_{i,t} = \frac{1}{5}\left(\frac{v_1^B}{|p_1^B/p_0 - 1|} + \sum_{\tau=1}^{4} \frac{v_{\tau+1}^B/v_\tau^B - 1}{|p_{\tau+1}^B/p_\tau^B - 1|}\right), \tag{4.3.2}$$

$$SE_{i,t} = \frac{1}{5}\left(\frac{v_1^A}{p_1^A/p_0 - 1} + \sum_{\tau=1}^{4} \frac{v_{\tau+1}^A/v_\tau^A - 1}{p_{\tau+1}^A/p_\tau^A - 1}\right). \tag{4.3.3}$$

p_τ^B and p_τ^A are the prices, respectively at the bid and at the ask, appearing at the quote τ. p_0 denotes the quoted midpoint. Finally, v_τ^B and v_τ^A are the natural logarithm of accumulated total share volume at the limit τ respectively for the bid and the ask sides.[1] In each equation, the first term represents the slope of the first line of the book to the midquote while the second one is the sum of the four remaining local slopes. The slope is then obtained by averaging both supply and demand elasticities.[2]

Table 4.1 presents the descriptive statistics of these liquidity proxies, as well as information on market capitalization, volume and number of trades. We may observe that the distributions of these liquidity proxies are heavily skewed and leptokurtic.

Using these proxies, we establish different hypotheses that we test in our empirical analyzes. We also formulate assumptions on trading activity.

Hypothesis 1. *Depth is higher when a zero return occurs.*

As depth is positively related to liquidity, we expect this relationship to hold for each

[1] "By accumulated", we mean the sum of the quantities outstanding at that limit and the sum of all quantities outstanding at each better quote.

[2] For more details on the slope, see also Næs and Skjeltorp (2006).

Table 4.1: Descriptive statistics

	Group	*Mean*	*STD*	*Median*	*Skewness*	*Kurtosis*	*CV*
Market Cap (kEUR)	*Small*	66049.591	42219.508	60015.690	0.416	−1.054	63.921
	Mid	461084.347	244309.592	400465.334	0.596	−0.843	52.986
	Large	11280287.064	18571741.538	4593171.461	3.764	17.913	164.639
Volume	*Small*	5583.599	77167.448	0.000	55.419	4275.049	1382.038
	Mid	6739.115	175117.735	100.000	204.604	67361.596	2598.527
	Large	49681.859	215485.532	6080.000	30.373	2165.499	433.731
Number of trades	*Small*	2.751	10.824	0.000	17.321	521.367	393.499
	Mid	4.973	14.277	1.000	26.460	1929.963	287.062
	Large	49.198	79.122	22.000	6.741	143.908	160.825
Depth BBO	*Small*	230983.000	1567896.644	931.000	10.211	126.613	678.793
	Mid	462602.351	6815596.565	499.000	16.820	290.320	1473.316
	Large	49579.580	311513.514	1986.000	9.492	107.840	628.310
Depth 5 limits	*Small*	1133794.079	7262446.241	7800.000	9.077	90.080	640.544
	Mid	2400065.297	37698131.568	3780.000	20.556	443.553	1570.713
	Large	250027.301	1377816.872	12053.500	7.996	73.537	551.067
Dispersion	*Small*	0.050	0.083	0.024	25.513	2090.019	165.336
	Mid	0.108	3.257	0.054	630.524	399147.148	3020.423
	Large	0.114	0.412	0.042	13.814	287.887	361.015
Slope	*Small*	491.457	571.736	333.126	5.549	48.892	116.335
	Mid	1003.658	1120.889	636.069	3.584	17.940	111.680
	Large	3186.220	2701.555	2485.429	1.436	2.036	84.789
Relative Spread	*Small*	1.279	2.078	0.867	8.009	76.363	162.405
	Mid	0.590	0.979	0.410	87.427	17153.975	165.875
	Large	0.211	0.412	0.111	279.359	134775.222	195.705

This table presents the descriptive statistics for market capitalization, volume, number of trades, as well as liquidity variables. These statistics are computed for each market capitalization segment: Small, mid and large caps. Market capitalization is computed in kEUR, volume in number of shares and the number of trades is the sum of the numbers of buyer and seller-initiated trades. Depth BBO denotes the depth at the first limit while Depth 5 limits is the sum of bid and ask quantities up to the fifth limit. The dispersion and slope measures are computed as in Kang and Yeo (2008) and Næs and Skjeltorp (2006), respectively.

capitalization group. However, if the book is becoming more dense and the spread lower, the quantities outstanding at the best quotes may actually be smaller. This effect should be minimized for the largest stocks that are more followed by analysts.

Hypothesis 2. *The relative spread is lower when a zero return occurs.*

The spread is negatively related to liquidity and as a consequence, we expect it to drop when a zero return takes place. Again, it is appropriate to distinguish large caps from small caps, given the tick size effect, i.e. large caps obviously display a much narrower spread than small caps do. As a matter of fact, if this relation is verified for smaller caps, the trough must be sharper than for larger caps, since there is much more room for improvement when the spread is usually large.

Hypothesis 3. *The dispersion is lower when a zero return occurs.*

The competition driven by informed traders who are willing to earn the spread forces the spread to be lower when a zero return occurs. The density of the book becomes higher and quotes are very close to each other. This hypothesis is much related to Proposition 2.

Hypothesis 4. *The slope is steeper when a zero return occurs.*

This is a corollary of Proposition 3, as the shape of the book is linked with its dispersion. We expect a steeper slope at the moment of the zero return, since a steep slope denotes a high level of agreement among analysts and traders about the fair value of the security, as previously mentioned.

Hypothesis 5. *There is less trading activity when a zero return occurs.*

If liquidity is effectively higher and the number of analysts and traders following the stock constant, a zero return should result in lower trading activity, since informed traders submit passive limit orders inside the existing quotes and do not trade aggressively anymore. We measure trading activity with the number of buyer and seller-initiated transactions that we classify after having matched orders and trades.

Hypothesis 6. *Volatility is lower when a zero return occurs.*

This hypothesis may seem trivial but is inferred from two phenomena. First, liquidity is inversely related to volatility. So, we expect volatility to be lower. Second, a zero return is also characterized by a lower volatility. Price excursions outside the close-open range should also be reduced as passive limit orders are submitted by informed traders around the quotes to absorb larger trades. We measure volatility by using the high-low range.

4.4 Event study

4.4.1 Methodology

To verify whether the assumptions of Lesmond et al. (1999) are correct, we first run an event study of liquidity proxies around zero returns. Our event is a dummy variable that equals 1 when the return is zero and 0 otherwise.

The event study is built using abnormal measures of the different liquidity measures presented in the previous section. We consider a [-3,+3] window containing 7 15-minute intervals: The event as well as three period before and after the event.

We compute the abnormality for liquidity proxy l and stock i at interval t as follows:

$$Abnormal_{i,t,l} = \frac{Liquidity_{i,t,l} - Median^{NE}_{i,t,l}}{Median^{NE}_{i,t,l}}, \tag{4.4.1}$$

where $Liquidity_{i,t,l}$ is the liquidity proxy for stock i at interval t and $Median^{NE}_{i,t,l}$ is the median of the liquidity proxy l for stock i across all non-events occurring during the time interval t.

This approach has also been tested in Boudt and Petitjean (2013) and Mazza (2013). The median is more robust to represent the central tendency of the distribution, since the distributions of liquidity proxies are heavily skewed and leptokurtic as shown in Table 4.1. The mean is much more affected by extreme values.

After the computation of the abnormality for each proxy, we quantify the distance between the current observation and the abnormal measure. We then aggregate the events to

form graphs for each liquidity proxy. We finally test whether a particular type of pattern occurs during the [-3,+3] window and check if abnormality is significantly different from zero. As in Mazza (2013), we use a standard non parametric sign test since it does not need any assumption about the shape of the distribution. The null hypothesis specifies that the abnormal measure has a median equal to zero. The alternative hypothesis postulates the opposite. The M test statistic is computed as: $M = \frac{N_+ - N_-}{2}$, where M follows a binomial distribution, N_+ is the number of positive values and N_- is the number of negative values. Values equal to zero are discarded.

We then analyze the p-values of each time interval of the window and check whether there are significant differences. If the p-value at the event interval is significant, the identified abnormal value is significantly different from zero, meaning that zero returns are associated to the corresponding configuration of the limit order book. If p-values are significant before the event, the zero return may be the consequence of a particular state of liquidity. If p-values are significant after the zero return, it may be the cause of the current state of liquidity, as measured by a given proxy.

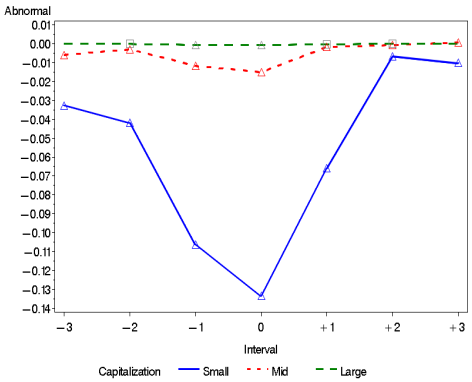
4.4.2 Results

The event study analysis reveals that liquidity is effectively higher when an intraday zero return occurs, whatever the market capitalization group. Relative spread, slope and dispersion confirm this outcome. For the depth proxy, the results are less clear but are more significant for the largest caps. We therefore do not reject Propositions 1 to 4. Regarding trading activity and volatility, we also validate Propositions 5 and 6. Trading activity is significantly lower for the two proxies when a zero return takes place. Volatility is also significantly dropping in $t = 0$.

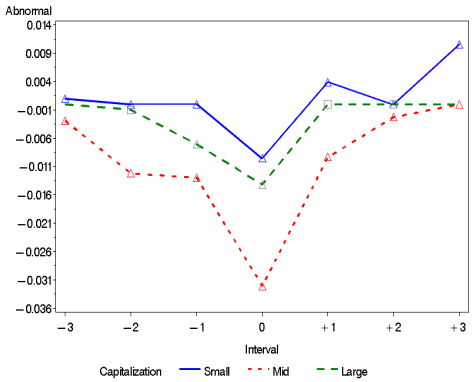
The results of the event study of liquidity proxies are presented in Figure 4.2.

Figure 4.2: Abnormal liquidity around zero returns

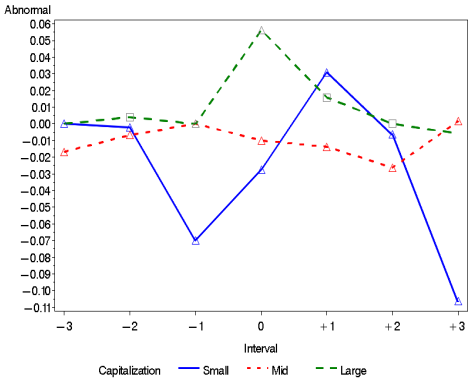
(a) Relative Spread



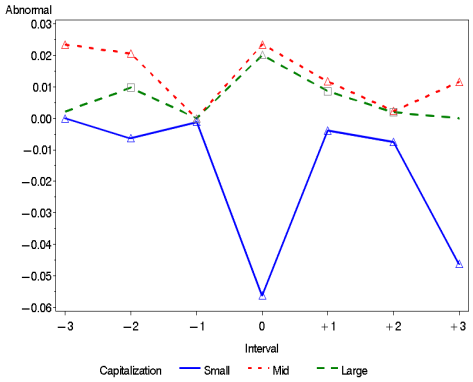
(b) Dispersion



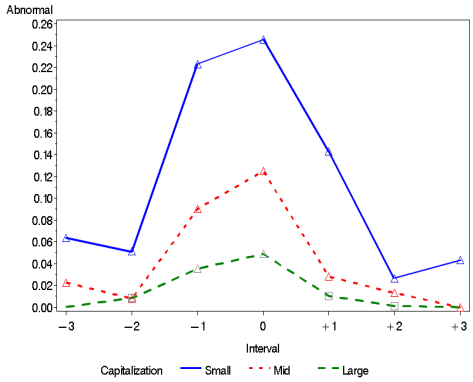
(c) Depth BBO



(d) Depth 5 limits



(e) Slope



These lines represent the intra-window median pattern for abnormal liquidity for the three market capitalization groups: Small, mid and large caps. Triangles (△), squares (□), and circles (○) indicate a rejection of the null hypothesis of the sign test respectively at the 99%, 95% and 90% confidence levels.

The relative spread drops more significantly for small caps, since they are less likely to display a narrow spread unlike large caps. As a consequence, the sharp drop in the abnormal

relative spread for small caps is much more pronounced. The relationship between zero returns and liquidity is also very strong as the recovery to normal values takes place quickly, implying that zero returns are associated to liquidity shocks that are short-term and highly resilient.

The dispersion drops sharply when the event occurs and is somehow anticipating the trough in $t-1$. The shock also reverts very quickly to normal values in $t+1$. This confirms our intuition that the competition among traders effectively takes place when a zero return is observed. A smaller spread, a higher dispersion and more depth confirm the hypothesis of an increased presence of informed traders in the supply of liquidity.

The relationship for depth measures is not verified for small caps. It seems however that there is a significant increase in these proxies when a zero return occur as far as large cap stocks are concerned.

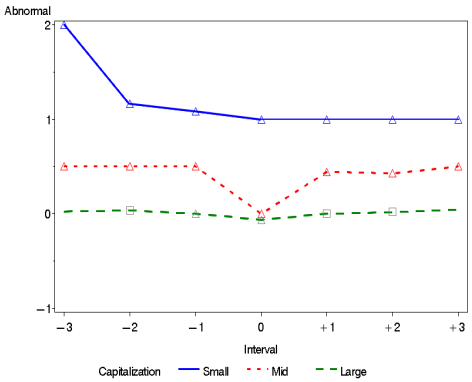
Zero returns also affect depth calculated with the outstanding quantities of the five best limits. The slope also increases significantly when a zero return occurs, confirming the outcomes of the dispersion.

These results do not confirm the explanation that stands behind the proxy presented in Lesmond et al. (1999). They rather point to the contrary: informed traders act as market makers to profit from their informational advantage and earn the spread that liquidity and noise traders are willing to pay to meet their needs. These results are in line with the expectations we built in Propositions 1 to 4.

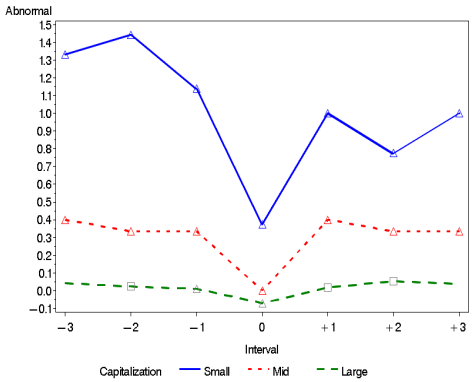
Figure 4.3 presents the outcomes of the event studies on trading activity measures.

Figure 4.3: Abnormal trading activity around zero returns

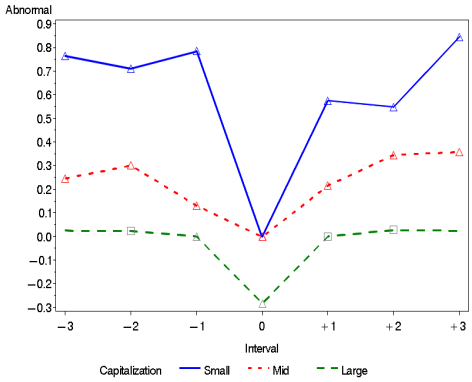
(a) Number of buys



(b) Number of sells



(c) High-Low



These lines represent the intra-window median pattern for abnormal trading activity for the three market capitalization groups: Small, mid and large caps. Triangles (△), squares (□), and circles (○) indicate a rejection of the null hypothesis of the sign test respectively at the 99%, 95% and 90% confidence levels.

These graphs show that a drop appears at time $t = 0$, when the zero return occurs, for both trading activity and volatility proxies, even if the patterns are clearer for sell trades than buy trades. Volatility drops and therefore confirms that liquidity should be higher.[1] This may also be a clue of the lack of informed trading during this period of time. The graphs also emphasize that trading activity tend to be lower for all capitalization groups, which may be an additional clue of the absence of aggressive trading from the informed traders. The higher

[1]The literature has extensively discussed the negative relationship between liquidity and volatility, e.g. Benston and Hagerman (1974), Stoll (1978a), Stoll (1978b), Amihud and Mendelson (1980), Ho and Stoll (1981), Copeland and Galai (1983), Admati and Pfleiderer (1988), Pagano (1989), Foster and Viswanathan (1990), Stoll (2003), Pastor and Stambaugh (2003), Harris (2003) and, Beaupain et al. (2010).

the market capitalization, the lower the magnitude of the variation. As the number of trades is lower, the liquidity that informed traders provide seems not to be totally hit by liquidity traders and noise traders.

All in all, the event study clearly indicates that zero returns are much more related to liquidity than to illiquidity, as previously suggested by Lesmond et al. (1999). In the next section, we confirm the results of this non-parametric analysis by estimating fixed-effects logit regressions, with the occurrence of zero returns as the binary response variable.

4.5 Logit regressions

4.5.1 Methodology

In a second step, we run logit regressions on the unfiltered sample with $ZR_{i,t}$, the occurrence of zero returns for stock i at interval t, as the dependent variable:

$$Prob(ZR_{i,t} = 1|\boldsymbol{x'_{i,t}}, c_i) = \frac{exp(\boldsymbol{x'_{i,t}\beta} + c_i)}{1 + exp(\boldsymbol{x'_{i,t}\beta} + c_i)}, \tag{4.5.1}$$

where $ZR_{i,t}$ is the response variable, $\boldsymbol{x'_{i,t}}$ is a $1 \times (k+1)$ vector of the k explanatory variables (including intercept), $\boldsymbol{\beta}$ is a $(k+1) \times 1$ vector of coefficients (including intercept), c_i is the unobserved time invariant effect of stock i and

$$\begin{aligned}\boldsymbol{x'_{i,t}\beta} = \beta_0 + \beta_1 RS_{i,t} + \beta_2 Q5_{i,t} + \beta_3 Dispersion_{i,t} + \beta_4 Slope_{i,t} \\ +\beta_5 IT_{i,t} + \beta_6 Volume_{i,t}.\end{aligned} \tag{4.5.2}$$

In this equation, $RS_{i,t}$ stands for the relative spread of stock i for interval t, $Q5_{i,t}$ for depth at the five best limits, $Dispersion_{i,t}$ for the dispersion, $Slope_{i,t}$ for the slope and $Volume_{i,t}$ for the volume. $IT_{i,t}$ is a dummy variable that captures the presence of informed trading through trade imbalance. It is equal to one when the absolute trade imbalance is higher than 50% and 0 otherwise. The absolute value of the trade imbalance is computed as: $abstrd_{i,t} = \frac{|NBuys_{i,t} - Nsells_{i,t}|}{(NBuys_{i,t} + Nsells_{i,t})/2}$. This measure controls for informed trading, assuming that trade imbalance becomes larger when informed traders trade aggressively.[1] $Volume$ denotes the volume that is transacted during time interval t for stock i, that we need to control to avoid scale effects.

Table 4.2 presents Pearson's correlation coefficients between each non-dummy variable of the model. The values clearly indicate that the measures are not correlated.

Table 4.2: Correlation matrix

	$Volume$	RS	$Q5$	$Dispersion$	$Slope$
$Volume$	1.00				
RS	-0.01	1.00			
$Q5$	0.07	0.07	1.00		
$Dispersion$	-0.01	0.01	0.00	1.00	
$Slope$	0.12	-0.08	-0.02	-0.02	1.00

This table presents Pearson's correlation coefficients between the non-categorical variables of the model. $Volume$ denotes the volume. $Q1$ and $Q5$ respectively denote the quantities outstanding at the best limit and at the the five best limits. $Slope$ and $Dispersion$ respectively denote the slope and dispersion proxies.

We may not estimate this panel model by likelihood maximization as it is usually done for logit regressions. In nonlinear models, the fixed effects, c_i, are not removed by differenciation, like in linear models. In addition, maximum likelihood estimators are valid only asymptotically. This assumes that the number of parameters does not increase as the sample

[1] The 50% is an arbitrary level. However, we test different levels in the conditional logit regressions and the results are not affected by the arbitrary choice of the threshold. These results are available upon request.

gets larger. This is not the case with panel data, since the number of individuals increases when we consider more records in the dataset. Estimating the c_i results in biased $\boldsymbol{\beta}$ estimators too. The biases are even greater when the number of time points per individual is small. This is called the incidental parameters problem (Neyman and Scott, 1948; McFadden, 1973; Chamberlain, 1980). One method to deal with this concern is to apply conditional maximum likelihood estimation which basically consists in conditioning the traditional likelihood function on the change of the state of the dependent variable $ZR_{i,t}$ between all time periods of the sample. The "sufficient statistic" of the conditional logit model is $\sum_{t=1}^{Np} ZR_{i,t}$ for c_i, where Np is the number of time intervals. In this approach, stocks are discarded when $\sum_{t=1}^{Np} ZR_{i,t} = 0$ or $\sum_{t=1}^{Np} ZR_{i,t} = Np$, i.e. stocks that never change state, since they do not contribute to the likelihood function. Conditioning the estimation on changes of the dependent variable involves the removal of the c_i from the likelihood function.[1] The likelihood is then maximized using Newton-Raphson optimization.

Conditional logit controls for firms' fixed effects without estimation biases but presents some drawbacks. First, it does not evaluate the impact of variables that are constant over time, e.g. market capitalization. This is a major concern since Lesmond et al. (1999) outlines that market capitalization is inversely related to the occurrence of zero returns. We bypass this disadvantage by running conditional logit regressions on three subsamples based on the market capitalization: Small, mid and large caps. Another drawback of this method that does not affect our study is that individuals are discarded if they display the same response variable level across all time periods of the subsample. This is not influential in our study since at least one zero return occurs for each stock.

4.5.2 Results

The results of the conditional logit regressions are displayed in Table 4.3.

[1] For more details on this calculation step, please refer to Maddala (1987) (pp.316).

Table 4.3: Conditional logit regressions

$Variable$	$Small$		Mid		$Large$	
	$Estimate$	$Odds$	$Estimate$	$Odds$	$Estimate$	$Odds$
RS	0.316^{***}	1.371	-0.029^{***}	0.971	-0.208^{***}	0.812
$Q5$	0.000^{***}	1.000	0.000^{***}	1.000	0.000^{***}	1.000
$Dispersion$	-0.501^{***}	0.606	-0.090^{***}	0.914	-0.078^{***}	0.925
$Slope$	0.050^{***}	1.052	0.090^{***}	1.094	0.094^{***}	1.098
IT	-2.521^{***}	0.080	-1.984^{***}	0.137	-0.681^{***}	0.506
$Volume$	-0.002^{***}	0.998	-0.001^{***}	0.999	-0.010^{***}	0.990

This table presents the results of the estimation of the conditional logit model $Prob(ZR_{i,t} = 1|\boldsymbol{x'_{i,t}}, c_i) = \frac{exp(\boldsymbol{x'_{i,t}}\boldsymbol{\beta}+c_i)}{1+exp(\boldsymbol{x'_{i,t}}\boldsymbol{\beta}+c_i)}$, where $\boldsymbol{x'_{i,t}}\boldsymbol{\beta} = \beta_0 + \beta_1 RS_{i,t} + \beta_2 Q5_{i,t} + \beta_3 Dispersion_{i,t} + \beta_4 Slope_{i,t} + \beta_5 IT_{i,t} + \beta_6 Volume_{i,t}$. The model is estimated using three subsamples based on the market capitalization. The maximum likelihood estimation is made under the assumption that the state of the response variable $ZR_{i,t}$ changes for each stock across all time periods of the sample. We use Newton-Raphson's optimization algorithm for likelihood maximization. Both parameter estimates and odds-ratio are shown. $*$, $**$ and $***$ respectively denote 90%, 95% and 99% confidence levels for the parameter estimates.

These outcomes clearly confirm the findings of the event study. There are very few differences due to the fact that the whole sample of zero returns is used. For example, the significantly positive coefficient for the spread for small caps is most probably the consequence of non-trading.

All variables behave consistently with the expectations presented in Propositions 1 to 4. First, as in Lesmond et al. (1999), we observe some variation in the parameter estimates across the different subsamples, even if the sign are always consistent, except the relative spread for small caps stocks. The relative spread and the dispersion display negative and highly significant estimates while depth and slope exhibit positive and highly significant parameter estimates, even if they are very low. As a result, positive modifications in spread and dispersion proxies negatively affect the log-odds of a return equal to zero, while positive variations in depth and slope positively influence the log-odds of a zero return.

All these findings indicate that, in a limit order market such as Euronext, intraday zero returns are more likely to occur in liquid states of the book, rather than in illiquid states as proposed by Lesmond et al. (1999). Since liquidity is inversely correlated with transaction costs, we expect them to be lower when zero returns appear.

Nevertheless, we confirm that the occurrence of zero returns is an indicator of the absence of aggressive informed trading, given that IT presents significantly negative estimates. We agree with Lesmond et al. (1999), Bekaert et al. (2007) and Mazza (2013) on this point. For small caps, the difference in odds-ratio is expected to be 0.08 when the trade imbalance is higher than 50%, all else equal. For large caps, the odds-ratio moves to 0.506.[1]

In a nutshell, there is less informed trading when a zero return occurs. However, informed traders may still be present as our measure only captures the impact on trade imbalances that informed traders create, such as the PIN indicator.[2] However, as outlined by Bloomfield et al. (2005), informed traders may still be present in the order book, trying to earn the spread that liquidity traders would agree to pay, even if they do not generate excess trade imbalance.

4.5.3 Robustness checks

In this section, we test whether the results of the conditional logit regressions obtained on 15-minute intervals do still holds for other interval lengths. We apply the same methodology to 20, 30 and 60-minute intervals. The outcomes are presented in Table 4.4.

[1] We conduct similar regressions by changing the threshold in order to check whether the results are robust to a change in this value. We set the threshold at 20, 30, 40, 50, 60, 70 and 80 % in different logit regressions and the coefficients of the $IT_{i,t}$ variable are always negative and highly significant. The results of these regressions are available upon request.

[2] This measure has already been extensively discussed in the literature (Easley and O'Hara, 1987, 1992; Easley et al., 1996, 1997, 1998, 2002, 2008, 2012).

Table 4.4: Conditional logit regressions: Robustness checks

Variable	*Small*		*Mid*		*Large*	
	Estimate	*Odds*	*Estimate*	*Odds*	*Estimate*	*Odds*
Panel A: 20-minute intervals						
RS	-0.009	0.991	-0.082^{***}	0.921	-0.204^{***}	0.815
$Q5$	0.001^{***}	1.001	0.000^{***}	1.000	0.000^{***}	1.000
$Dispersion$	-0.429^{***}	0.651	-0.114^{***}	0.892	-0.070^{***}	0.933
$Slope$	0.016	1.016	0.073^{***}	1.076	0.094^{***}	1.098
IT	-2.303^{***}	0.100	-1.791^{***}	0.167	-0.624^{***}	0.536
$Volume$	0.006^{***}	1.006	-0.002^{***}	0.998	-0.008^{***}	0.992
Panel B: 30-minute intervals						
RS	0.087^{***}	1.091	-0.237^{***}	0.789	-0.111^{***}	0.895
$Q5$	-0.000^{***}	1.000	0.000^{***}	1.000	0.000^{***}	1.000
$Dispersion$	-0.404^{***}	0.667	-0.222^{***}	0.801	-0.091^{***}	0.913
$Slope$	0.066^{***}	1.068	0.052^{***}	1.053	0.106^{***}	1.112
IT	-1.905^{***}	0.149	-1.487^{***}	0.226	-0.532^{***}	0.588
$Volume$	-0.008^{***}	0.992	-0.001^{***}	0.999	-0.007^{***}	0.993
Panel C: 60-minute intervals						
RS	-0.015^{*}	0.985	-0.057^{***}	0.944	-0.003	0.997
$Q5$	0.000^{***}	1.000	0.000^{***}	1.000	0.000	1.000
$Dispersion$	-0.263^{***}	0.768	-0.039	0.962	-0.039	0.962
$Slope$	0.004	1.004	0.027^{***}	1.028	0.014^{***}	1.014
IT	-0.424^{***}	0.654	-0.213^{***}	0.808	-0.050^{***}	0.951
$Volume$	-0.002^{***}	0.998	-0.000^{**}	1.000	-0.001^{***}	0.999

This table presents the results of the estimation of the conditional logit model $Prob(ZR_{i,t} = 1|\boldsymbol{x'_{i,t}}, c_i) = \frac{exp(\boldsymbol{x'_{i,t}}\boldsymbol{\beta}+c_i)}{1+exp(\boldsymbol{x'_{i,t}}\boldsymbol{\beta}+c_i)}$, where $\boldsymbol{x'_{i,t}}\boldsymbol{\beta} = \beta_0 + \beta_1 RS_{i,t} + \beta_2 Q5_{i,t} + \beta_3 Dispersion_{i,t} + \beta_4 Slope_{i,t} + \beta_5 IT_{i,t} + \beta_6 Volume_{i,t}$. Panel A, B and, C respectively present the results obtained with 20, 30 and 60-minute samples. The model is estimated using three subsamples based on the market capitalization. The maximum likelihood estimation is made under the assumption that the state of the response variable $ZR_{i,t}$ changes for each stock across all time periods of the sample. We use Newton-Raphson's optimization algorithm for likelihood maximization. Both parameter estimates and odds-ratio are shown. $*$, $**$ and $***$ respectively denote 90%, 95% and 99% confidence levels for the parameter estimates.

These results clearly indicate that the relationships outlined using 15-minute intervals are validated when changing interval lengths. The vast majority of the estimates exhibits the same signs and significance levels as the core analysis. Even if the relationship seems to deteriorate for large caps for 60-minute intervals, the measure may still be considered as a

proxy for the absence of aggressive informed trading, as outlined by Lesmond et al. (1999) and Bekaert et al. (2007).

4.5.4 Daily zero returns

Lesmond et al. (1999) use zero returns to characterize the state of liquidity on a daily basis. Even if intraday zero returns seem to point to a higher liquidity, this relationship has to be tested on daily frequencies in order to provide a complete analysis of the proxy. Table 4.5 presents the results of the conditional logit regressions estimated on daily data that cover the same sample of stocks.

Table 4.5: Conditional logit regressions: Daily data

$Variable$	$Small$		Mid		$Large$	
	$Estimate$	$Odds$	$Estimate$	$Odds$	$Estimate$	$Odds$
RS	-0.135^{***}	0.874	-0.095	0.910	-0.180	0.835
$Q5$	0.001^{***}	1.001	0.000^{**}	1.000	0.001^{*}	1.001
$Dispersion$	-3.028^{**}	0.048	0.096	1.101	-0.012	0.988
$Slope$	-0.070	0.932	0.111^{***}	1.118	0.093^{***}	1.097
IT	-0.020	0.980	-0.083	0.920	0.145	1.156
$Volume$	-0.003^{***}	0.997	-0.000	1.000	-0.001^{***}	0.999

This table presents the results of the estimation of the conditional logit model $Prob(ZR_{i,t} = 1|\boldsymbol{x'_{i,t}}, c_i) = \frac{exp(\boldsymbol{x'_{i,t}\beta}+c_i)}{1+exp(\boldsymbol{x'_{i,t}\beta}+c_i)}$, where $\boldsymbol{x'_{i,t}\beta} = \beta_0 + \beta_1 RS_{i,t} + \beta_2 Q5_{i,t} + \beta_3 Dispersion_{i,t} + \beta_4 Slope_{i,t} + \beta_5 IT_{i,t} + \beta_6 Volume_{i,t}$. The model is estimated using three subsamples based on the market capitalization. The maximum likelihood estimation is made under the assumption that the state of the response variable $ZR_{i,t}$ changes for each stock across all time periods of the sample. We use Newton-Raphson's optimization algorithm for likelihood maximization. Both parameter estimates and odds-ratio are shown. $*$, $**$ and $***$ respectively denote 90%, 95% and 99% confidence levels for the parameter estimates.

The results suggest that the relationship to liquidity is much weaker than for intraday datasets. However, these outcomes do not point to illiquidity either. The relative spread, the depth, the dispersion and the slope display some significance, in particular for small caps.

Interestingly, we observe that the relationship to the IT variable seems to no longer hold. This may also be explained by the scarcity of high trade imbalance on daily data.

In a nutshell, these findings indicate that daily zero returns are also not associated with higher illiquidity in the order book, as opposed to Lesmond et al. (1999).

4.6 Conclusion

Liquidity estimation has been at the core of several papers in the recent literature. Various liquidity proxies have been put forward and some of them relate liquidity to price returns. Given the multidimensionality of liquidity, a consensus on the best proxies has been nevertheless impossible to reach.

In this paper, we discuss one of these liquidity proxies, which is based on the occurrence of zero returns. Lesmond et al. (1999) justify the use of this proxy on the intuition that informed traders will not trade if the value of a new information set does not exceed the cost of transacting. The resulting price dynamics is a zero return since the impact of noise and liquidity traders is assumed to be negligible on average. So, if transactions costs are higher and informed traders not reacting, then liquidity should be lower, which explains why their illiquidity measure is grounded on the occurrence of zero returns.

This model, however, raises some important issues which depend on the market under scrutiny. The model is valid in dealership markets as the dealer will probably increase the spread, assuming that there is an excessive price pressure of informed traders, leading to a zero return. This relation does not hold in limit order book markets for several reasons. First, prices may be moved, even if informed traders are not transacting. Second, zero returns are not always related to zero volume as liquidity may be abundant at the first limits, implying

that the volume is positive but the return stays unchanged. In other words, noise and liquidity traders may still trade when informed traders are out of the market. Third, the fact that informed traders do not actually trade does not necessarily lead to lower liquidity in a pure limit order book. When there is less (or no) informed trading, transaction costs need not necessarily be higher. Bloomfield et al. (2005) also demonstrate that informed traders change their way of trading when the current trading price comes near the fundamental value of the security. As they precisely know the true value, they submit aggressive limit orders near the existing quotes and wait for being picked off by liquidity traders to meet their immediacy needs. Mazza (2013) also finds that (quasi) zero returns are positively correlated with liquidity. The Amihud illiquidity measure also contrasts with Lesmond et al. (1999)'s intuition since the Amihud ratio is lower (pointing to a higher liquidity) when returns are closer to zero.

Using Euronext market data, we apply an intraday event study methodology that addresses the behavior of liquidity around zero returns in a window ranging from 45 minutes before and after the zero return. We use liquidity measures from the limit order book: relative spread, depth, dispersion and slope. We find that liquidity is effectively higher when a zero return occurs. The effect is also found to be rather short-lived. These results are verified for the three market capitalization-based subsamples, i.e. small, mid and large caps, of our 701-stock sample. Trading activity is also much lower at the time of the event, as well as volatility.

We further check this relationship by running conditional logit regressions, with the probability of encountering a zero return as the dependent variable. The outcomes are consistent with the event study results and present highly significant parameters for liquidity variables. We also confirm that there is less informed trading when a zero return occurs, since the proxy of informed trading exhibits negative and highly significant estimates, for all subsamples. Considering 20, 30, or 60-minute intervals does not significantly influence the results. We also examine whether daily zero returns are related to liquidity in the same manner. The

results indicate that the relationship is still validated even if it shows less significance levels and dot not corroborate the findings of Lesmond et al. (1999) that associate daily zero returns to illiquidity.

All in all, we conclude that intraday zero returns are most likely related to liquidity in a pure limit order book market, rather than to illiquidity. This puts in question the proxy proposed by Lesmond et al. (1999) which is based on the proportion of zero returns on a given time interval. Even if less liquid stocks may indeed present a high number of zero return days, Lesmond et al. (1999)'s rationale seems not to be validated in limit order markets.

CHAPTER

FIVE

DO JAPANESE CANDLESTICKS HELP SOLVE THE TRADER'S DILEMMA?

5.1 Introduction

Transaction cost management has always been a major concern of the implementation of trading decisions. There are different components in transaction costs. Explicit costs, which can be determined before the execution of the trade, refer to brokerage commissions, market fees, clearing costs, settlement costs and taxes. Implicit costs, which are not visible and cannot be evaluated ex-ante, consist of bid-ask spread, market impact and opportunity costs.[1] Bid-ask spread is a compensation for the supply of liquidity. Market impact is the cost incurred for consuming more than the liquidity available at the best opposite quote (BOQ hereafter). Opportunity costs are due to the price movement that materializes between

[1] Opportunity costs include three different components: operational opportunity costs, market timing opportunity costs and missed trade opportunity costs. Operational opportunity costs arise when the delay required to trade is operational, the second component is due to the market timing under the control of the broker and the missed trade opportunity costs occur when the total size has not been fully executed.

the trade decision and the trade itself. Market timing costs belong to the latter category and denote the additional cost incurred when traders split their orders to lower market impact.

The main challenge when implementing trade decisions lies in the impossibility to reduce all costs components simultaneously: traders have to strike a balance to identify the best trade-off for their execution needs. The most tricky issue is linked to the so-called trader's dilemma. When traders place market orders, they always have to decide whether they should split their orders over time to reduce market impact, or submit them directly which probably incurs the cost of drying out quantities outstanding at the BOQ for large order sizes. When they split an order, traders are however exposed to a potential adverse price evolution that may hinder their performance, i.e. market timing opportunity cost. For instance, if a trader wants to buy a large quantity, and therefore decide to split the order over 3 days, a positive price appreciation during this period will significantly hinder the order execution. This cost may be viewed as a risk since the price evolution may positively or negatively affect the execution.

The question that this paper addresses is the following: Is it possible to solve the trader's dilemma by using Japanese candlesticks? Japanese candlesticks are an Eastern charting technique that represents High-Low-Open-Close price movements. Candlestick charts give market participants a quick snapshot of buying and selling pressures, as well as turning points. Candlesticks are related to transaction costs for several reasons. First, as outlined by Kavajecz and Odders-White (2004), price dynamics, which are easily characterized by candlesticks, are expected to be related to modifications in the state of the limit order book and to the supply of liquidity. Transactions costs and liquidity are negatively correlated: market impact rises (drops) rapidly when liquidity is low (high). Wang et al. (2012) also outline that order submission strategies are related to technical analysis in the Taiwan Stock Exchange. They also argue on causality indicating that technical analysis drives changes in order submission strategies. Second, Mazza (2013) finds that liquidity is higher when some particular candlestick structures occur, suggesting that a relationship does exist between limit

order book variables and price movements. He also associates this relation to a change in informed traders's behavior, as proposed by Bloomfield et al. (2005). Third, according to the literature on Japanese candlesticks, some structures may help forecast future prices, which determines market timing costs. This argument stands directly against the efficient market hypothesis (henceforth EMH) of Fama (1970). In this paper, we restrict our analysis to Doji and Hammer-like configurations which are described in the following section.

Using market data on a sample of European stocks belonging to three national indexes (BEL20, AEX, CAC40), we study sequences of orders and estimate fixed-effects panel regression models including market impact or market timing costs as the dependant variable. Dummy variables that identify the occurrence of candlestick structures as well as a set of control variables are explanatory. We establish different types of relationship with contemporaneous and lagged signals in order to check whether it is possible to lower transaction costs after the apparition of a potential signal. In a second step, in order to further assess whether candlesticks are useful or not in this regard, we compare the market impact cost of an average quantity submitted after the apparition of a signal to the market impact cost of the same quantity submitted randomly along the day.

Our results suggest that market impact is lower at the time of a Doji and after its occurrence. There is no impact for Hammer-like configurations. Market timing cost is not lower when these structures occur. As market timing costs are mainly determined by price movements, we question the usefulness of candlesticks in predicting future stock prices and contribute to the previous literature on the efficient market hypothesis. Order processing simulations also show that transaction costs are lower when large orders are not split over time after the occurrence of a signal. All our results suggest that candlesticks may help traders solve their dilemma by identifying the right moment for submitting aggressive orders.

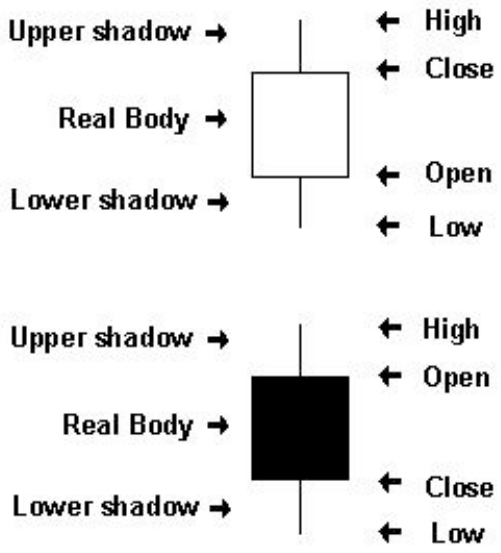
The remainder of the paper is organized as follows. Section 5.2 provides a literature review as well as a description of Japanese candlesticks. Section 5.3 describes the dataset.

Section 5.4 presents the methodology and the results of the panel regressions. Section 5.5 reports the methodology and the results of the simulations. The final section concludes.

5.2 Japanese candlesticks and liquidity

Japanese candlesticks are a technical analysis charting technique based on High-Low-Open-Close prices.[1] The formation process of candlesticks appears in Figure 5.1. Depending on the length of the shadows and the size and color of the bodies, many structures may be identified, from one to five candles. These candlesticks emphasize what happened in the market at that particular moment. Each configuration can be translated into traders' behaviors through price dynamics implied by buying and selling pressures.

Figure 5.1: Candlestick formation process



Japanese candlesticks are interesting because they summarize a lot of information in one single chart: the closing price, the opening price as well as the lowest and highest prices. With the raising interest in high frequency trading and the narrowing of trading intervals,

[1] Steve Nison was the first to bring this method to the west in the nineties, even if Japanese candlesticks have been used for centuries in eastern countries. See Nison (1991), Nison (1994), Morris (1995) and Bigalow (2001) for more information on candlestick charting.

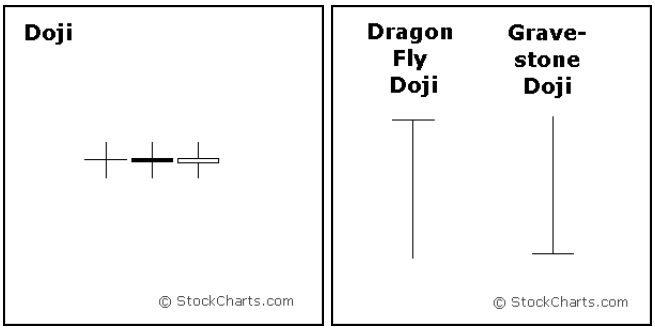
they have been increasingly used by practitioners to capture short term price dynamics. Papers addressing candlesticks enter in the "stock return predictability" category. For example, Marshall et al. (2006) and Marshall et al. (2008) find no evidence that candlesticks have predictive value for the Dow Jones Industrial Average stocks and for the Japanese equity market, respectively. They replicate daily data with a bootstrap methodology similar to the one used in Brock et al. (1992). However, intraday data is more relevant as traders do not typically wait for the closing of the day to place an order. Nevertheless, using intraday candlesticks charts on two future contracts (the DAX stock index contract and the Bund interest rate future), Fock et al. (2005) still find no evidence which suggests that candlesticks, alone or in combination with other methods, have a predictive ability. Duvinage et al. (2013) also address the profitability of candlestick trading rules on DJIA's stocks and conclude that it is not profitable to trade using these signals. However, none of these papers looks at the relationships between candlestick configurations and the transaction costs of trade sequences. To our knowledge, this paper is the first research study that investigates the information content of HLOC price movements for execution purposes.

In this paper, we investigate two categories of candlesticks structures. The first one is the Doji category. The Doji is one of the core structures of the literature on Japanese candlesticks.[1] A Doji appears when the closing price is (almost) equal to the opening price. We observe different types of Doji. The most frequent Doji is a "plus", i.e. no real body and almost equal shadows. If both closing and opening prices are also equal to the highest price of the interval, the Doji becomes a Dragonfly Doji. By contrast, it becomes a Gravestone Doji when both closing and opening prices are equal to the lowest price of the interval. In essence, the Doji is not an indicator of price reversal: it only helps detect the end of the current trend. Our signals are based on these three Doji structures, i.e. traditional, Dragonfly and Gravestone, and are disentangled in bullish and bearish signals: the Doji is bullish (bearish) when the previous candle is black (white) and the next candle is white (black). These configurations are presented in Figure 5.2. If these structures are able to forecast future short-term returns, bullish (bearish) signals should result in higher (lower) market timing cost when the

[1] Nison (1991), Nison (1994) and Morris (1995).

trader buys. The opposite should also be verified for sales.

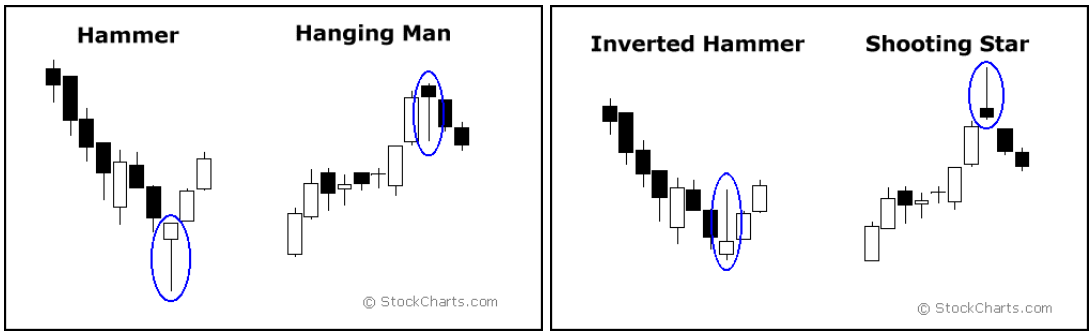
Figure 5.2: Doji



The Doji presents a closing price (almost) equal to the opening price. It occurs when there is an agreement on the fair value of the instrument and when markets are 'on a rest'. The Doji indicates the end of the previous trend. The most traditional Doji is a 'plus' sign but Dragonfly and Gravestone Doji are also frequent. A Dragonfly Doji appears when a strong buying pressure directly follows a strong selling pressure implying an upper shadow almost equal to zero. The Gravestone Doji occurs when the buyers have dominated the first part of the session and the sellers, the second one.

The second category contains Hammer-like configurations. Among Hammer-like structures, there are four structures that are characterized by a long shadow and a small real body. The Hammer appears at the end of a downtrend and is made of a very small real body with (almost) no upper shadow and a very long lower shadow. The same structure may appear at the end of an uptrend but, in that case, it is called a Hanging Man. Inverting the shadows, i.e. the upper shadow becomes the lower shadow and vice-versa, we obtain an Inverted Hammer at the end of a downtrend or a Shooting Star at the end of an uptrend. Figure 5.3 shows these configurations. As these figures are said to be strong reversal structures in the literature on Japanese candlesticks, they should have an influence on market timing cost, if EMH does not hold: for purchases (sales), Hammer and Inverted Hammer should lead to higher (lower) market timing cost, while Hanging Man and Shooting star should lead to lower (higher) market timing cost.

Figure 5.3: Hammer-like structures



The Hammer and the Hanging Man appear when sellers dominate the first part of the session and buyers, the second part. By construction, they present a long lower shadow and almost no upper shadow. The Hammer occurs at the end of a downtrend while the Hanging Man puts an end to an uptrend. The Inverted Hammer and the Shooting Star are made with a small real body, a very long upper shadow and almost no lower shadow. The Inverted Hammer appears at the end of a downtrend and the Shooting Star occurs at the end of an uptrend. These structures are said to be strong reversal ones.

As outlined by Duvinage et al. (2013) and Marshall et al. (2006), candlestick-based strategies fail to beat a Buy-and-Hold strategy and therefore are not able to predict future short-term returns, confirming EMH. As a result, we do not expect market timing to be improved around the occurrence of these structures.

Market liquidity has been extensively discussed in the literature. Harris (2003) defines liquidity as "the ability to trade large size quickly, at low cost, when you want to trade".[1] Three dimensions of liquidity can be identified in this definition: immediacy, width, and depth. Given this multidimensionality, it should come as no surprise that several studies have attempted to find the best liquidity proxy. For instance, Goyenko et al. (2009) propose a comprehensive study on low and high frequency measures. They conclude that the Amihud (2002) illiquidity ratio as well as their own effective and realized spread measures better

[1] Harris (2003), pp. 394.

characterize price impact and spread. They nevertheless do not investigate order-book based proxies, which are the main transaction cost drivers. As outlined by Aitken and Comerton-Forde (2003), liquidity measures usually fall into two categories: trade-based and order-based measures. Trade-based liquidity proxies rely on volumes, returns, number of trades, and trade imbalances. Order-based liquidity proxies include relative spread, depth, order imbalances, and dispersion or slopes. While trade-based measures estimate past (or ex-post) liquidity, order-book measures estimate current and future liquidity. In this paper, we control for the state of liquidity in our transaction cost regressions by using relative spread and depth, as they are directly related to trading costs. Our depth proxy is a measure of total depth that is outstanding at the opposite side over the five best limits.

There is also an extensive literature on the possible relationships between liquidity and price fluctuations. For example, Chordia et al. (2001), Chordia et al. (2002), and Chordia and Subrahmanyam (2004) find that there is a relationship between liquidity, trading activity and market returns. More recently, Cao et al. (2009) find that liquidity facilitates price discovery in the ASX and helps in predicting future returns. Kavajecz and Odders-White (2004) also propose an interesting study that relates liquidity to technical analysis using limit order book data. They show that moving averages as well as support and resistance are associated with liquidity. Extending the study by Kavajecz and Odders-White (2004), Mazza (2013) investigates whether HLOC prices, which are fully captured by Japanese candlesticks, are related to liquidity. The key finding is that liquidity is effectively higher when a Doji occurs. The theoretical justification is related to the behavior of informed traders, as shown by Chakravarty and Holden (1995) and Bloomfield et al. (2005). Following the latter study, informed traders are likely to provide liquidity when they cannot profitably trade on their information set. This is the case when the fundamental value lies inside the quotes. Informed traders behave like dealers and earn the spread that liquidity and noise traders are willing to pay to meet their liquidity needs. Bloomfield et al. (2005) further argue that liquidity and noise traders are most likely to simultaneously change their way of trading as they have to satisfy their clients' requirements. The Doji characterizes the trading situation outlined by Bloomfield et al. (2005) above since it occurs when the stock is 'on a rest' and there is a

consensus on the value of the stock. The Doji marks the end of a price discovery process that helps locate the fundamental price inside the quotes. Informed traders then supply liquidity to benefit from their information. As they are informed, they do not incur adverse selection costs, which motivates them to behave as dealers and provide liquidity.

As a matter of fact, we expect that market timing costs are not improved when a candlestick structure appears. This is the logical consequence of EMH, as future returns are not predictable. Recent researches on Japanese candlesticks also argue in favor of that point (Duvinage et al., 2013; Marshall et al., 2006). Furthermore, as liquidity is higher when a Doji appears, following the reasoning of Mazza (2013) and Bloomfield et al. (2005), market impact costs should be lower. This should particularly be the case for the Doji since the consensus on the price is clearer.

5.3 Data

5.3.1 Sample

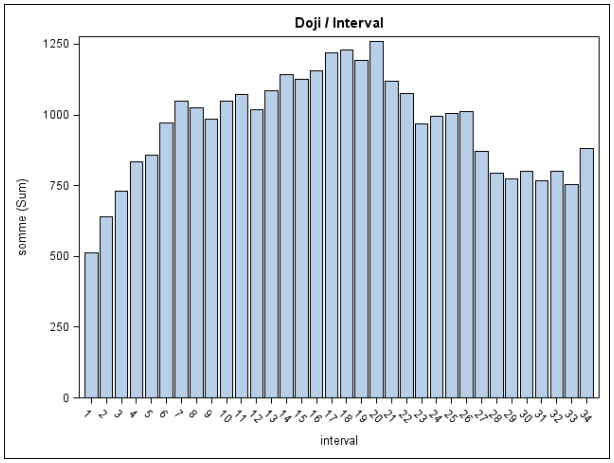
We use Euronext market data on 81 stocks belonging to three national indexes: BEL20, AEX or CAC40. We have tick-by-tick data for 61 trading days from February 1, 2006 to April 30, 2006, including information on hidden orders and market members' ID. This dataset contains all orders, trades and order book data over the whole period.

We have rebuilt High-Low-Open-Close prices from this database for the 81 stocks over the whole sample period. As tick data are not adapted for candlestick analysis, we build 15-minute-intervals which leads to 34 intervals a day. This interval length is the best trade-off which allows to include intraday trends and to avoid noisy candlestick patterns resulting from non-trading intervals. We use the HLOC prices calculated above in order to identify

candlestick configurations based on TA-Lib.[1] We obtain a total of 167068 records (81 firms, 61 days, 34 intervals/day). From this dataset, we remove non-trading patterns as they result in 'Four Prices Doji'.[2]

We look at the occurrences of the identified structures and check whether Doji appear at a particular moment during the day. Figure 5.4 shows that the distribution of Doji is roughly uniform with the most significant peaks occurring during lunch time. Doji also seem to not occur frequently during the first two intervals of the day. This may be explained by the strong unidirectional movement that appears at that moment, as trends are at their very beginning. This should not influence our results. Table 5.1 presents the number of occurrences of each structure which is identified in our dataset through the TA-lib.

Figure 5.4: Doji by interval



This figure displays the number of Doji in each time interval.

[1]The TA-lib library is compatible with the MATLAB Software. For each type of configuration and for each record, it returns "1" if the bullish part of the structure is identified, "-1" for the bearish part and "0" otherwise. As the structures are bullish, bearish or both, for each event type, the values that may appear are [0 ; 1], [-1 ; 0] or [-1 ; 0 ; 1]. The TA-lib allows some flexibility in the recognition of the configurations. As it is an open source library, we have been able to check the parametrization of the structures. The structures are recognized according to the standard flexibility rules presented in Nison (1991) and Morris (1995). The TA-lib contains 61 pre-programmed structures.

[2]A Four Prices Doji occurs when all the prices are equal. When they occur in daily, weekly or monthly charts, they are a strong clue of a potential reversal. However, in intraday price charts, they represent non-trading intervals.

Table 5.1: Number of signals

Structure	Count
Hammer	4487
Inverted Hammer	2264
Shooting Star	972
Hanging Man	5145
Doji	29828
Bearish Doji	18031
Bullish Doji	11797
Dragonfly Doji	7071
Gravestone Doji	7557
Bullish Dragonfly Doji	2575
Bearish Dragonfly Doji	4496
Bullish Gravestone Doji	3013
Bearish Gravestone Doji	4544

This table presents the count of each structure over the whole sample.

5.3.2 Sequences of orders

In order to study the magnitude of transaction costs around candlestick signals, we divide the trading days in 15-minute intervals and compute the transaction costs of sequences of orders submitted by market members.

We make the following assumptions when building our sequences: Firstly, we only consider principal orders so that, in a given sequence, every order is submitted by the same market member for his own account. Secondly, we do not consider orders that provide liquidity because they do not generate transaction costs.

Then, we use the market member's identity code[1] to construct the sequences of orders for each stock. For a given market member, a sequence is initiated with a first marketable order and cumulates the following marketable orders in the same direction. The sequence stops when the market member submits a passive order,[2] changes order direction, or when it reaches the end of the interval.

The descriptive statistics of the sequences are presented in Table 5.2. $Duration$ and $volume$ respectively denote the duration and the volume of the sequence, while N denotes the number of orders in the sequence. We observe that the distributions of the number of orders are similar across the three different exchanges. The duration is however lower for the CAC40 and the AEX compared to the BEL20.

Table 5.2: Sequences - Descriptive statistics

CAC40	Mean	Median	Maximum	Minimum	Standard Deviation
$Duration$	00:04:23	00:02:49	00:15:00	00:00:01	00:04:20
$Volume$	203191	96391	14234329	23	362943
N	3.45	3.00	428.00	2.00	3.25
AEX					
$Duration$	00:04:26	00:02:51	00:15:00	00:00:01	00:04:23
$Volume$	192262	95249	30830446	31	337990
N	3.30	2.00	121.00	2.00	2.47
BEL20					
$Duration$	00:06:19	00:05:19	00:15:00	00:00:01	00:04:59
$Volume$	88801	52762	6460514	41	140701
N	3.44	2.00	100.00	2.00	4.64

Cross-sectional statistics on the sequences are reported for the three indexes of the whole sample. N refers to the sequence's number of orders. $Volume$ is the sequence's volume expressed in currency units. $Duration$ refers to the execution time of the sequences.

[1] Actually, these ID codes are numerical in order to ensure market members' anonymity but allow us to isolate the whole set of orders or trades associated with a given member from the other orders and trades in the sample.

[2] By passive order, we mean an order that is neither a market order nor a marketable limit order.

5.3.3 Transaction costs measures

The market impact of an order i is computed as the signed difference between the average execution price (AEP_i) and the BOQ prevailing at the order i submission time (BOQ_i), expressed in percentage of the BOQ:

$$MI_i^{buy/sell} = \frac{|AEP_i - BOQ_i|}{BOQ_i} * 100 \tag{5.3.1}$$

The market impact of a sequence j of n orders is expressed in percentage of the total amount that the investor would pay without any transaction costs, i.e. the amount if the entire volume of the sequence executes at the BOQ prevailing at the beginning of the sequence (BOQ_1). Practically, for a sequence j of n orders, we compute the sum of the market impact of the n orders in EUR that we divide by the total quantity executed in the sequence j multiplied by the BOQ prevailing at the submission of the first order (BOQ_1).

$$MI_j^{buy/sell} = \frac{\sum_{i=1}^{n} Q_i * BOQ_i * MI_i}{\sum_{i=1}^{n} Q_i * BOQ_1} * 100 \tag{5.3.2}$$

Let us assume a sequence that is made of two buy orders of 100 units respectively. The BOQ at the submission time of the first order is equal to 84.5 and its AEP is equal to 84.75. The BOQ at the submission time of the second order is equal to 85 and its AEP is equal to 85.25. The market impact of the first order and the second order are equal to 0.295% and 0.294% respectively. The market impact of the entire sequence is equal to:

$$MI = \frac{(100 * 84.5 * 0.295\%) + (100 * 85 * 0.294\%)}{(200 * 84.5)} = 0.2954\% \tag{5.3.3}$$

The market timing of an order i is computed as the difference between the BOQ_i pre-

vailing just before the submission of the order and the BOQ_1 prevailing at the submission of the first order of the sequence and is expressed as a percentage of the BOQ_1. The market timing of the first order of the sequence is thus equal to zero.

$$MT_i^{buy} = \frac{(BOQ_i - BOQ_1)}{BOQ_1} * 100 \tag{5.3.4}$$

$$MT_i^{sell} = \frac{(BOQ_1 - BOQ_i)}{BOQ_1} * 100 \tag{5.3.5}$$

The market timing of a sequence j of n orders is then expressed in percentage of the total amount the investor pays if the entire volume of the sequence executes at the BOQ_1 prevailing at the beginning of the sequence. Practically, for a sequence j of n orders, we compute the sum of the market timing cost of the n orders in EUR and we divide it by the total quantity executed in the sequence j multiplied by the BOQ_1 prevailing at the submission of the first order.

$$MT_j^{buy/sell} = \frac{\sum_{i=2}^{n} MT_i * Q_i * BOQ_1}{\sum_{i=1}^{n} Q_i * BOQ_1} * 100 = \frac{\sum_{i=2}^{n} MT_i * Q_i}{\sum_{i=1}^{n} Q_i} * 100 \tag{5.3.6}$$

In the example mentioned above, the market timing cost of the second order is equal to $\frac{85-84.5}{84.5} = 0.5917\%$. And the market timing cost of the entire sequence is equal to:

$$MT = \frac{0.5917\% * 100}{200} = 0.2958\% \tag{5.3.7}$$

5.4 Panel regressions

5.4.1 Methodology

We test the impact of candlestick structures on both market timing and market impact costs components through different fixed-effects panel regression models in order to control for stock-specific effects. The robustness of standard errors is a major concern in panel regressions. Based on Petersen (2009), we apply the clustering approach that makes standard errors heteroscedasticity-consistent. As outlined by Petersen (2009), this method produces unbiased standard errors when a firm effect does exist, as opposed to White, Newey-West, and Fama-MacBeth correction methods. Clusters are used to control for common factors in the fixed effects. For instance, macroeconomic news may evenly affect all the stocks that are present in an index. Omitting to control for common factors may lead to potential biases.

In our fixed-effect panel regression model, transaction costs are the dependent variable. We establish different regressions for the two components that we investigate, i.e. market timing cost and market impact respectively. We first include dummy variables for each of the four candlestick structures, i.e. Hammer (H), Inverted Hammer (IH), Hanging Man (HM) and Shooting Star (SS). These dummies are equal to 1 when the structure has been detected during the interval and 0 otherwise. We also include some control variables that are key determinants of the sequence's costs: the number of orders ($Orders$) of the sequence, its duration in minutes ($Duration$) and its volume in thousands (V). The market timing cost is more likely to increase when there are more orders over a longer duration while market impact cost decreases in the same conditions. A higher volume to trade results in higher costs whatever the component. We then control for the state of liquidity at the beginning of the sequence by including the depth in thousands ($Depth$), and the relative spread in percentage, (RS). The $Depth$ proxy sums the quantities outstanding at the five best opposite quotes, i.e. $Depth = \sum_{i=1}^{5} QB_i$, in case of sell orders and $Depth = \sum_{i=1}^{5} QA_i$, in case of buy orders,

where QB_i and QA_i are respectively the bid and ask quantities outstanding at the limit i.

The model that we estimate is specified as follows:

$$\begin{aligned} M_{i,s,t}^{buy} &= \alpha_0 + \alpha_1 Orders_s + \alpha_2 Duration_s + \alpha_3 V_s + \alpha_4 Depth_s \\ &+\alpha_5 RS_s + \alpha_6 H_{i,t} + \alpha_7 IH_{i,t} + \alpha_8 HM_{i,t} + \alpha_9 SS_{i,t} + \nu_s, \end{aligned} \tag{5.4.1}$$

where $M_{i,s,t}^{buy}$ is the transaction cost component, measured for the buying sequence s that begins during interval t for stock i, that can be either market impact or market timing cost.

The effect estimated in this regression is contemporaneous. We also conduct a similar regression with lagged signals, i.e. the dummy identification variables H, IH, HM and SS are lagged once:

$$\begin{aligned} M_{i,s,t}^{buy} &= \alpha_0 + \alpha_1 Orders_s + \alpha_2 Duration_s + \alpha_3 V_s + \alpha_4 Depth_s \\ &+\alpha_5 RS_s + \alpha_6 H_{i,t-1} + \alpha_7 IH_{i,t-1} + \alpha_8 HM_{i,t-1} + \alpha_9 SS_{i,t-1} + \nu_s, \end{aligned} \tag{5.4.2}$$

We conduct this regression in order to assess whether we can effectively base a strategy on the signal once it has fully appeared.

We apply the same methodology to Doji configurations, separately for all types Doji (D) and Dragonfly (DF) and Gravestone (GR) Doji. The models are specified as follows for contemporaneous effects on market impact:

$$MI^{buy}_{i,s,t} = \alpha_0 + \alpha_1 Orders_s + \alpha_2 Duration_s + \alpha_3 V_s + \alpha_4 Depth_s + \alpha_5 RS_s + \alpha_7 D_{i,t} + \nu_s, \tag{5.4.3}$$

where $MI^{buy}_{i,s,t}$ is the market impact cost of the sequence s for stock i at interval t and $D_{i,t}$ is a dummy variable indicating the presence of a Doji and:

$$MI^{buy}_{i,s,t} = \alpha_0 + \alpha_1 Orders_s + \alpha_2 Duration_s + \alpha_3 V_s + \alpha_4 Depth_s + \alpha_5 RS_s + \alpha_6 DF_{i,t} + \alpha_7 GR_{i,t} + \nu_s, \tag{5.4.4}$$

where $DF_{i,t}$ is a dummy variable indicating the presence of a Dragonfly Doji and $GR_{i,t}$ is a dummy variable indicating the presence of a Gravestone Doji. The same regression specifications are also implemented for lagged signals.

However, for market timing cost, we need to know which evolution of future prices the signal should lead to. We disentangle bullish and bearish Doji by investigating the previous trend, i.e. if the previous trend is negative (positive), the Doji is a bullish (bearish) signal. This process is only applicable to market timing costs as market impact is not affected by future price movements.

$$MT^{buy}_{i,s,t} = \alpha_0 + \alpha_1 Orders_s + \alpha_2 Duration_s + \alpha_3 V_s + \alpha_4 Depth_s + \alpha_5 RS_s + \alpha_6 Dbull_{i,t} + \alpha_7 Dbear_{i,t} + \nu_s, \tag{5.4.5}$$

where $MT^{buy}_{i,s,t}$ is the market timing cost of the sequence s for stock i at interval t and $Dbull_{i,t}$ is a dummy variable indicating the presence of a bullish Doji and $Dbear_{i,t}$, the presence of a bearish Doji and:

$$MT_{i,s,t}^{buy} = \alpha_0 + \alpha_1 Orders_s + \alpha_2 Duration_s + \alpha_3 V_s + \alpha_4 Depth_s \\ +\alpha_5 RS_s + \alpha_6 DFbull_{i,t} + \alpha_7 DFbear_{i,t} + \alpha_8 GRbull_{i,t} + \alpha_9 GRbear_{i,t} + \nu_s, \tag{5.4.6}$$

where $DFbull_{i,t}$ and $DFbear_{i,t}$ respectively refer to dummy identification variables of bullish and bearish Dragonfly Doji while $GRbull_{i,t}$ and $GRbear_{i,t}$ respectively denote dummy identification variables of bullish and bearish Gravestone Doji. The same regression specifications are also implemented for lagged signals.

We expect market impact to be lower when one of these structures occur, implying a negative sign for the dummy variables. This explanation is consistent with a higher liquidity supply in the order book around technical signals, as outlined by Mazza (2013) or Kavajecz and Odders-White (2004). If the signal is also an indicator of future price movements, which goes against the EMH, dummy variables associated with future price should lead to a decrease (increase) in market timing costs for buy (sell) sequences of orders. An opposite process should apply for signals of positive future price evolutions, if EMH does not stand. As the performance of candlesticks in predicting returns has been seriously questioned in the literature, e.g. Duvinage et al. (2013) and Marshall et al. (2006), we do not expect any significant effect in market timing regressions.

$Orders$ and $Duration$ should be negatively correlated with market impact and potentially positively correlated with market timing cost, depending on the future price evolution. Splitting orders over a long time logically reduces market impact but increases market timing cost. The trading volume of the sequence, V, should be positively correlated with both market timing and market impact costs. Liquidity should be negatively related to transaction costs. Therefore, we expect $Depth$ to be negatively related to market impact and RS positively related to this cost. The effect should be less significant for market timing cost as the order size is smaller when orders are split, implying a lower liquidity effect.

5.4.2 Results

Table 5.3 to 5.6 present the results of the fixed-effects panel regression models by cluster of firms. First of all, the behavior of the control variables is consistent with our expectations for all models, for both lagged and contemporaneous effects. The market impact cost is lower when $Orders$ and $Duration$ increase, with a less significant effect for the latter as the main evolution is captured through the $Orders$ variable. Market timing cost presents opposite results, as expected. V exhibits strongly significant positive parameters indicating an evolution in the same direction as transactions costs. $Depth$'s negative effect is strongly significant, even for market timing costs. A possible explanation is that, when depth is large, traders execute more volume against the depth available at the start of the sequence and are therefore less exposed to the market timing cost. The RS variable is strongly significant for market impact and show mixed results for market timing cost for both purchases and sales models. Concerning the market impact, the results are in line with our expectations on the liquidity impact of the spread. With regards to the market timing cost, a wide spread, as outlined by Glosten and Milgrom (1985), through a large adverse selection component, could reveal the presence of informed trading and impact the market timing cost. It seems that the results are significant only for sell sequences of traders whose market timing decrease with the spread, revealing the presence of informed buyers.

As expected, Table 5.3 shows that EMH holds and that market timing costs (Panels A and B) may not be better managed by looking at Hammer-like structures, whatever contemporaneous or lagged signals. Some parameters are significant but exhibit the opposite sign. To refute the hypothesis that candlesticks may help predict returns, Hammer and Inverted Hammer should exhibit positive (negative) signs for purchases (sales) while Hanging Man and Shooting Star should exhibit positive (negative) signs for sales (purchases). The strong differences between the parameters and their significance indicate that we may not base any market timing strategy on these signals. Panels C and D also show that market impact results are consistent with Mazza (2013) who outlines a relationship between liquidity and the

occurrence of Hammer and Hanging Man configurations. We however observe that this relationship is only valid for purchases as the parameters for sales models do not present any significance. The effect seems to be very short-lived as lagged models display less significant results.

Table 5.3: Market impact and market timing - Hammer-like configurations

$Model$	$Orders$	$Duration$	V	$Depth$	RS	H	IH	HM	SH
Panel A : Market Timing - Purchases									
t	0.057^{***}	0.072^{***}	0.037^{***}	-0.004^{***}	-1.902	0.034	-1.843^{***}	-0.111	0.173
$t-1$	0.057^{***}	0.072^{***}	0.038^{***}	-0.004^{***}	-1.910	-0.077	-0.578^{***}	-0.148	0.990^{**}
Panel B : Market Timing - Sales									
t	0.043^{**}	0.072^{***}	0.040^{***}	-0.005^{***}	-4.641^{***}	-0.445^{***}	0.429^{**}	-1.125^{***}	-0.159
$t-1$	0.043^{**}	0.072^{***}	0.040^{***}	-0.005^{***}	-4.595^{***}	0.010	0.017	-0.308^{*}	-0.446
Panel C : Market Impact - Purchases									
t	-0.006^{***}	-0.001^{*}	0.007^{***}	-0.001^{***}	1.176^{***}	-0.038^{***}	-0.019	-0.049^{***}	0.037
$t-1$	-0.006^{***}	-0.001^{**}	0.007^{***}	-0.001^{***}	1.175^{***}	-0.040^{**}	0.005	-0.009	0.017
Panel D : Market Impact - Sales									
t	-0.004^{***}	-0.001^{**}	0.005^{***}	-0.001^{***}	0.796^{***}	-0.006	-0.032^{**}	-0.006	-0.005
$t-1$	-0.004^{***}	-0.001^{**}	0.005^{***}	-0.001^{***}	0.797^{***}	-0.008	0.006	-0.025	0.126^{*}

This table presents the results of different panel regression models. Panels A and B present the results for the market timing cost for purchases and sales respectively. Panels C and D display parameter estimates for market impact models respectively for purchases and sales. t and $t-1$ stand for contemporaneous and lagged signals respectively. $Orders$ indicate the number of order of a sequence s, $Duration$ its duration and V its volume. $Depth$ and RS are liquidity proxies respectively for depth and relative spread. H, IH, HM and SH are candlesticks identification dummies, respectively for Hammer, Inverted Hammer, Hanging Man and Shooting Star, that equal 1 when the structure occurs for contemporaneous models (t) and when the structure has occurred during previous interval for lagged models ($t-1$). These dummies equal 0 otherwise. *, ** and, *** respectively denote the 10%, 5% and 1% significance levels.

Tables 5.4 and 5.5 show that Doji configurations do not help reduce market timing costs as the parameters show inconsistent signs and for contemporaneous signals only. For lagged signals, the parameters are not significant anymore. This is consistent with EMH, suggesting that candlesticks cannot predict future short term price evolution, in line with Duvinage et al. (2013) and Marshall et al. (2006).

Table 5.4: Market timing - Doji

$Model$	$Orders$	$Duration$	V	$Depth$	RS	$DBear$	$DBull$
Panel A : Purchases							
t	0.057^{***}	0.075^{***}	0.037^{***}	-0.004^{***}	-1.962	-0.714^{***}	-0.671^{***}
$t-1$	0.057^{***}	0.072^{***}	0.038^{***}	-0.004^{***}	-1.926	-0.070	-0.240^{***}
Panel B : Sales							
t	0.043^{**}	0.074^{***}	0.039^{***}	-0.005^{***}	-4.594^{***}	-0.652^{***}	-0.539^{***}
$t-1$	0.043^{**}	0.072^{***}	0.040^{***}	-0.005^{***}	-4.579^{***}	-0.186^{**}	0.141

This table presents the results of different panel regression models. Panels A and B present the results for the market timing cost for purchases and sales respectively. t and $t-1$ stand for contemporaneous and lagged signals respectively. $Orders$ indicate the number of order of a sequence s, $Duration$ its duration and V its volume. $Depth$ and RS are liquidity proxies respectively for depth and relative spread. $Dbull$ and $Dbear$ are candlesticks identification dummies, respectively for Bullish Doji and Bearish Doji, that equal 1 when the structure occurs for contemporaneous models (t) and when the structure has occurred during previous interval for lagged models ($t-1$). These dummies equal 0 otherwise. *, ** and, *** respectively denote the 10%, 5% and 1% significance levels.

Table 5.5: Market timing - Dragonfly and Gravestone Doji

$Model$	$Orders$	$Duration$	V	$Depth$	RS	$DFbull$	$DFbear$	$GRbull$	$GRbear$
Panel A : Purchases									
t	0.057^{***}	0.074^{***}	0.037^{***}	-0.004^{***}	-1.937	-0.329^{*}	-0.669^{***}	-1.191^{***}	-0.830^{***}
$t-1$	0.057^{***}	0.072^{***}	0.038^{***}	-0.004^{***}	-1.897	-0.451^{**}	-0.374^{**}	-0.171	0.003
Panel B : Sales									
t	0.043^{**}	0.074^{***}	0.040^{***}	-0.005^{***}	-4.582^{***}	-1.254^{***}	-0.755^{***}	-0.373^{**}	-0.362^{***}
$t-1$	0.043^{**}	0.072^{***}	0.040^{***}	-0.005^{***}	-4.577^{***}	-0.040	-0.176	-0.151	-0.207

This table presents the results of different panel regression models. Panels A and B present the results for the market timing cost for purchases and sales respectively. t and $t-1$ stand for contemporaneous and lagged signals respectively. $Orders$ indicate the number of order of a sequence s, $Duration$ its duration and V its volume. $Depth$ and RS are liquidity proxies respectively for depth and relative spread. $DFbull$, $DFbear$, $GRbull$ and $GRbear$ are candlesticks identification dummies, respectively for Bullish Dragonfly Doji, Bearish Dragonfly Doji, Bullish Gravestone Doji and Bearish Gravestone Doji, that equal 1 when the structure occurs for contemporaneous models (t) and when the structure has occurred during previous interval for lagged models ($t-1$). These dummies equal 0 otherwise. *, ** and, *** respectively denote the 10%, 5% and 1% significance levels.

Tables 5.6 and 5.7 however show very interesting results which are consistent with pre-

vious findings, as in Mazza (2013). Doji structures are likely to help reduce market impact costs. Market impact is much lower for sequences beginning during the interval that contains a Doji and for sequences beginning during the next interval. These findings indicate that we may benefit from an execution strategy based on these candlesticks. The results also tell us that traders may place aggressive orders during the formation of the Doji or just after and therefore profit from a window ranging from 15 to 30 minutes.

Table 5.6: Market impact - Doji

$Model$	$Orders$	$Duration$	V	$Depth$	RS	D
Panel A : Purchases						
t	-0.006^{***}	-0.001^{*}	0.007^{***}	-0.001^{***}	1.173^{***}	-0.049^{***}
$t-1$	-0.006^{***}	-0.001^{*}	0.007^{***}	-0.001^{***}	1.177^{***}	-0.022^{**}
Panel B : Sales						
t	-0.004^{***}	-0.001^{*}	0.005^{***}	-0.001^{***}	0.796^{***}	-0.023^{***}
$t-1$	-0.004^{***}	-0.001^{**}	0.005^{***}	-0.001^{***}	0.799^{***}	-0.022^{***}

This table presents the results of different panel regression models. Panels A and B present the results for the market timing cost for purchases and sales respectively. t and $t-1$ stand for contemporaneous and lagged signals respectively. $Orders$ indicate the number of order of a sequence s, $Duration$ its duration and V its volume. $Depth$ and RS are liquidity proxies respectively for depth and relative spread. D is a Doji identification dummy that equals 1 when a Doji occurs for contemporaneous models (t) and when a Doji has occurred during previous interval for lagged models ($t-1$). This dummy equals 0 otherwise. *, ** and, *** respectively denote the 10%, 5% and 1% significance levels.

Table 5.7: Market impact - Dragonfly and Gravestone Doji

$Model$	$Orders$	$Duration$	V	$Depth$	RS	DF	GR
Panel A : Purchases							
t	-0.006^{***}	-0.001^{*}	0.007^{***}	-0.001^{***}	1.175^{***}	-0.062^{***}	-0.042^{***}
$t-1$	-0.006^{***}	-0.001^{*}	0.007^{***}	-0.001^{***}	1.179^{***}	-0.030^{**}	-0.035^{**}
Panel B : Sales							
t	-0.004^{***}	-0.001^{*}	0.005^{***}	-0.001^{***}	0.797^{***}	-0.033^{***}	-0.041^{***}
$t-1$	-0.004^{***}	-0.001^{**}	0.005^{***}	-0.001^{***}	0.799^{***}	-0.024^{*}	-0.016

This table presents the results of different panel regression models. Panels A and B present the results for the market impact cost for purchases and sales respectively. t and $t-1$ stand for contemporaneous and lagged signals respectively. $Orders$ indicate the number of order of a sequence s, $Duration$ its duration and V its volume. $Depth$ and RS are liquidity proxies respectively for depth and relative spread. DF and GR are candlesticks identification dummies, respectively for Dragonfly Doji and Gravestone Doji, that equal 1 when the structure occurs for contemporaneous models (t) and when the structure has occurred during previous interval for lagged models ($t-1$). These dummies equal 0 otherwise. *, ** and, *** respectively denote the 10%, 5% and 1% significance levels.

In a nutshell, the results display that market timing cost is not affected by the reversal potential that candlesticks contain. The results show some significant estimates but the signs of the estimates are opposite to what should be expected according to the literature on Japanese candlesticks which suggests that Hammer-like configurations are associated with a high reversal probability. This confirms our intuition that EMH holds and that Japanese candlesticks are not able to predict future price returns, as outlined in Duvinage et al. (2013) and Marshall et al. (2008).

The other main result of these panel regressions is the relationship between market impact cost and the occurrence of these structures. We find that market impact is much lower when a Doji occurs, whatever its type. This is consistent with Mazza (2013), which outlines that liquidity is higher when a Doji appears on a price chart. The effect is also long enough as sequences starting after the occurrence of these Doji still exhibit lower market impact costs. The results are also valid for Hammer and Hanging Man but for purchases only and with much less significance as well as short-lived effects.

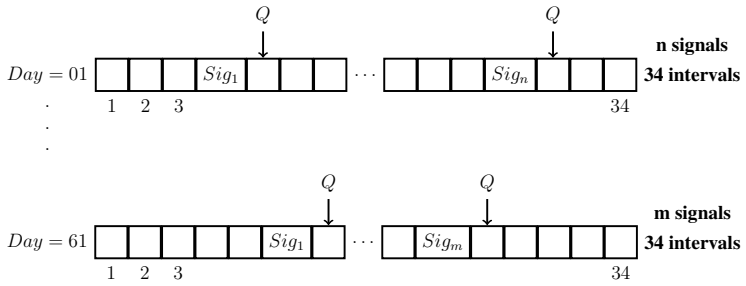
5.5 Order processing simulation

5.5.1 Methodology

In order to verify whether a trader benefits from candlesticks to reduce market impact, we simulate order processing and confront the random strategy to the original strategy based on candlesticks. The original strategy relies on the observed time series of prices and is therefore not simulated.

The candlestick-based strategy is presented in Figure 5.5. Each time a Doji occurs, it consists in placing a quantity Q, equal to the mean of the sizes of the sequences for the security i over the whole sample. A successful strategy would imply that candlesticks have some informational content for liquidity estimation and help reduce market impact. Market impact is computed by averaging all days for each stock separately. If there are $n_{d,i}$ signals occurring on the same day d for security i, $n_{d,i}$ market orders are submitted just after complete realization of the structures. It enables us to check the profitability of a trader who waits for the signal to be fully realized.

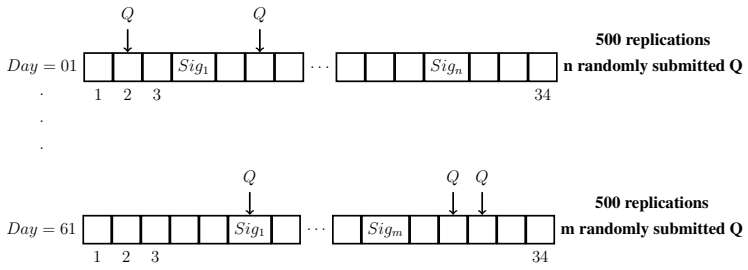
Figure 5.5: Candlestick-based strategy



This figure shows the process followed by a trader who adopts the candlestick-based strategy. The trader submits a quantity Q after the occurrence of a signal. Hence, he submits n orders on $Day = 01$, where n corresponds to the number of signals, i.e. Doji, that occur on $Day = 01$. On $Day = 61$, he submits m orders, each of a quantity Q, where m denotes the number of Doji occurring on $Day = 61$.

The random strategy is explained in Figure 5.6. It consists in submitting $n_{d,i}$ orders of size Q randomly, separately for each day and stock, where $n_{d,i}$ is the number of Doji that occur in the candlestick-based strategy, independently of the occurrence of these Doji. For each day d and each stock i, we generate $n_{d,i}$ orders randomly. The only element that differs between the two strategies is the moment at which $n_{d,i}$ orders are submitted. In the random strategy, the $n_{d,i}$ orders execute against the existing depth, which is therefore not simulated. The market impact is computed for each trade and averaged in the same way as for the original candlestick-based strategy.[1]

Figure 5.6: Random strategy



This figure shows the process followed by the random strategy. Hence, he submits n orders of a quantity Q on $Day = 01$, where n corresponds to the number of Doji that occur on $Day = 01$ in the candlestick-based strategy, independently of the occurrence of these Doji. On $Day = 61$, he randomly submits m orders, each of a quantity Q, where m denotes the number of Doji occurring on $Day = 61$. The process is replicated 500 times for each day and each stock.

To check the robustness of our results, we replicate the sample selection in the random strategy 500 times. In other words, for each day d and each stock i, we create 500 random samples of size $n_{d,i}$. It allows us to control for a possible lucky draw in the moment chosen to randomly submit the $n_{d,i}$ orders of size Q. We calculate the market impact for each replication and compare this cost to the candlestick-based strategy. We then count the number of times the replications beat the strategy and compute a p-value. For the original strategy to be

[1] Several other naive order placement strategies exist, e.g. placing aggregated orders at the beginning or at the end of the trading day. However, these strategies may not be used as benchmarks in our study since orders have to be split to allow the comparison with the candlestick-based strategy. Furthermore, the random strategy permits the simulation process by generating 500 random sequences of orders across the trading day.

significantly better than the random strategy at the 5% level, it must display a lower market impact than at least 95% of the 500 simulations.

5.5.2 Results

Table 5.8 compares the results of the two execution strategies: The Doji-based and the random strategies. The average market impact of the Doji's strategy is the average market impact of the $(n * d * i)$ orders submitted after the occurrence of a signal and the average market impact of the random strategy is the average market impact of the $500 * (n * d * i)$ orders submitted after the randomly generated signals. The p-value refers to the number of times the random strategy beats the Doji-based strategy divided by 500.

Table 5.8: Simulation

$MeanRandom$	$MeanDoji$	$p-value$
Panel A : Purchases		
3.2154	2.806	0.00
Panel B : Sales		
3.1626	2.883	0.00

This table presents the results for the simulations of the two strategies: random and Doji-based. Panels A and B present the results for purchases and sales respectively. $MeanRandom$ refers to the average of the market impact cost for the 500 replications of the random strategy over the 81 securities. $MeanDoji$ is the average of the market impact cost of the Doji-based strategy over the 81 securities. The $p-value$ is averaged across stocks and is computed as the number of times the market impact cost of the random replications for the 81 securities is lower than the market impact cost of the Doji-based strategy divided by the total number of replications (500).

The results show that, for both purchases and sales, the Doji-based strategy presents a significantly lower market impact cost than the random strategy. The p-value associated to the test is lower than 1%. The average market impact of the Doji-based strategy is equal to 2.806 basis points for buy orders and to 2.883 basis points for sell orders. These results

are significantly lower than the market impact paid by the random submission for both buy and sell orders, 3.2154 and 3.1626 respectively. These findings suggest that investors may benefit from candlesticks to reduce their market impact.

5.6 Conclusion

Transaction costs management is still a tricky issue, since it is not possible to improve all the costs components simultaneously. Hence, traders are confronted to the so-called trader's dilemma which is based on the choice of an execution strategy, namely splitting orders over time or not. This dilemma may be summarized by two transaction costs components: market impact, that arises when a large order is submitted, and market timing that arises when a large order is split into smaller ones that execute at different prices through time.

Based on a sample of 81 European stocks from three Euronext indexes, we investigate whether the two components of transaction costs of orders' sequences are impacted by the occurrence of particular candlestick structures. We focus on two categories of structures, Hammer-like and Doji configurations, as they are the best known single lines of the Japanese candlesticks literature. Furthermore, Mazza (2013) shows that these candlesticks are related to the state of liquidity in the limit order book, hence to trading costs. We estimate fixed-effects panel regressions of market impact or market timing on a set of explanatory variables that contain candlestick identification dummies. We also include as exogenous control variables the number of orders, duration, volume of the sequences, and liquidity proxies. In order to further assess whether a trader may benefit from the occurrence of a given signal, we also conduct order processing simulations in which we compare the market impact cost of a candlestick-based execution strategy to the market impact cost of a random execution strategy that we replicate 500 times.

Our results support the existing literature and present interesting findings which indicate that traders may benefit from candlesticks to improve their execution strategies. We find that market impact costs are significantly lower when and after that Doji structures occur. The effect is long enough to allow next sequences of orders to exhibit lower costs. This is also true for Hammer and Hanging Man but only for purchase orders. Those results indicate that market impact can be reduced by trading after the occurrence of these structures. Order processing simulations show similar outcomes, i.e. the Doji-based strategy exhibits significantly lower market impact cost than the random one. We also contribute to the existing literature on EMH by outlining that candlesticks fail to predict future returns, since market timing costs are not lower when or after that one of these configurations occurs. This is consistent with previous findings (Duvinage et al., 2013; Marshall et al., 2006).

As a conclusion, all these findings suggest that candlesticks are valuable tools that could help traders to implement trading decisions through an improvement of their order execution. This paper outlines an interesting feature of candlesticks: they help detect time windows where market impact costs are lower. As such, they may be particularly suitable for immediate liquidity demanders, such as institutional traders or hedgers.

CHAPTER

SIX

CONCLUSION

"The important thing is not to stop questioning"
Albert Einstein

The complexity of today's financial markets has raised many issues that are currently being addressed in the literature. Among these topics, market microstructure, which deals with the organization of financial markets and the understanding of trading mechanisms, has arisen as an unavoidable topic since October 1987's financial crash. Several research studies have tried to lift the veil on the origin of this crash by examining order submission and trade dynamics. For instance, Blume et al. (1989) highlight a strong correlation between stock price movements and order imbalances. More recently, since the advent of high frequency trading and the Flash Crash of May 2010, this branch of finance has become a relatively hot topic. Before that, several theoretical and empirical papers have addressed dealership markets, limit order markets, informed and uninformed trading, transparency, transactions costs, fragmentation or market liquidity, trying to understand how trading occurs and how asymmetrically informed traders trade with each other.

A specific strand of the market microstructure literature investigates the relation between liquidity and price dynamics. Liquidity has been previously defined as "*the ability to trade large size quickly, at low cost, when you want to trade*" (Harris, 2003).[1] Within this definition, liquidity is presented as a multidimensional concept. Four dimensions are traditionally assigned to liquidity: immediacy, width, depth and resiliency. Given these four dimensions, an accurate estimation of liquidity is a serious challenge. As a consequence, many research studies have tried to identify the best proxy for liquidity, among the plethora of measures that have been presented in the literature for almost three decades. For instance, Aitken and Comerton-Forde (2003) and Goyenko et al. (2009) compare different liquidity measures and try to assess which proxies perform best in determining the current state of liquidity. The other main issue with liquidity evaluation concerns the unavailability of accurate order and trade data. That is the reason why different measures are based on available data, among which price series. As an illustration, Roll (1984) creates a spread proxy based on daily high and low prices while Lesmond et al. (1999) identify the frequency of zero returns as an illiquidity proxy.

This four-essay doctoral study contributes to this very rich literature by investigating how liquidity and price dynamics are connected. Its rationale lies in the multidimensionality of liquidity. Even if the correlation between price dynamics and liquidity has been well documented in the literature, the way we characterize price dynamics is different from what is typically done in empirical research. Several papers, for instance Blume et al. (1989) and Chordia and Subrahmanyam (2004), address this relationship by examining closing prices or returns. In our study, we investigate how High-Low-Open-Close price dynamics (HLOC) are related to liquidity. HLOC price series are widely available and may be easily represented in charts that convey a high level of information regarding price pressures as well as trading activity determinants.

The reasons for such a relationship are manifold. First, the seminal paper of Kavajecz and Odders-White (2004) demonstrates that liquidity provision and technical analysis are

[1] Harris (2003), pp. 394.

strongly linked. They address moving averages as well as support and resistance levels and find that these technical indicators are associated with the relative position of depth in the order book and with peaks in depth in the order book, respectively. Furthermore, Wang et al. (2012) show that order submissions strategies are strongly related to the occurrence of technical signals in the Taiwan Stock Exchange. These two papers confirm the usefulness of price analysis for order submission and liquidity purposes. Yet, these works do not include the informational content of HLOC price dynamics that may be interesting since they provide indication on the behavior of buyers and sellers and give additional information on what happened during the chosen time interval. Our research study fills that gap. We also ground this relation on previous theoretical evidence, such as Chakravarty and Holden (1995) and Bloomfield et al. (2005). These papers show that a consensus on the fundamental value of the security, that is characterized by particular HLOC configurations, is associated with moments where informed traders supply liquidity instead of demanding it, leading to a higher level of liquidity in the order book. They outline that informed traders become dealers since they cannot profit from their information anymore to move the price towards its fundamental value which is close to the current price. Their knowledge of the fundamental value of the stock motivates them to behave as market makers by preventing them from trading with a better informed trader. Therefore, they earn the spread by submitting aggressive limit orders near the quotes.

This doctoral thesis is organized in four chapters, each of them corresponding to an essay addressing the relationship between HLOC price movements and liquidity.

The first chapter presents the results of the first essay that sheds light on the connection between liquidity and some particular 15-minute HLOC dynamics through the use of an event study. Our event denotes moments where there is a consensus on the price during the observed interval, i.e. the opening and closing prices are very close. There are different types of consensus configurations but they all suggest that opening and closing prices are almost equal. Bloomfield et al. (2005) propose some justification for the existence of such a

relationship. In their experimental design, they show that informed traders do not consume liquidity when they situate the fundamental value of the security inside the spread, which is the case when a consensus occurs. Indeed, as they do not profit from moving prices towards their fundamental value, they behave as dealers to earn the spread and to benefit from their knowledge of the fundamental value that allows them to submit aggressive limit orders. In this paper, we analyze different order book-based liquidity proxies (i.e. depth, spread, slope and dispersion) as well as trading activity variables. The results indicate that this relationship actually holds and liquidity is really higher when a consensus configuration occurs. We further confirm these findings with Granger causality and PIN analyzes, as well as different robustness checks, including different time frames and HLOC dynamics.

The second chapter corresponds to the second essay and investigate more deeply this relationship by examining several HLOC price variables that characterize all HLOC dynamics. These variables include the Open-Close range, the High-Low range, the interaction term between these two variables, as well as dummies that control for the occurrence of zero returns, price gaps and price evolution. This chapter is the logical step after having found that a relationship does exist. In this paper, we run fixed-effect panel regressions for three portfolios and for each liquidity proxy on a set of HLOC price dynamics explanatory variables. The three portfolios are constructed with the hundred largest market capitalizations in each market cap segment: Small, mid and large caps. The liquidity proxies that are investigated include both order book-based as well as trade-based measures: depth, spread, dispersion, slope, the number of trades, the average trade size and Amihud (2002)'s illiquidity ratio. Our results confirm the outcomes of the first essay, i.e. liquidity is higher when a consensus structure occurs. We further test the relationships with robust and median regression techniques as well as other time frames, i.e. 10 and 20 minute-intervals, showing that the identified relationships remain significant when we consider smaller or longer time frames, as well as when we control for the presence of outliers in the sample.

The third chapter addresses the particular case of intraday zero returns and how they are

related to liquidity. This essay has been motivated by the seminal work of Lesmond et al. (1999) which first suggests that the frequency of daily zero returns can be considered as a proxy for illiquidity. This measure has been used in several empirical papers, mostly related to asset pricing, e.g. Bekaert et al. (2007) who use it as a proxy for liquidity to address the relationship between asset pricing and liquidity in emerging markets. Lesmond et al. (1999) argue that zero returns occur when informed traders do not trade upon the arrival of a new information set as they value its informational advantage to not exceed the cost of trading, which results in zero volume and hence, zero returns. Using an event study methodology as well as conditional logit regressions, we provide evidence that suggests that intraday zero returns may be related to liquidity rather than to illiquidity. Our results also show that daily zero returns are more related to liquidity than illiquidity. We also discuss the influence that the market structure may constitute in this issue. Indeed, we argue that the proxy may still hold as a valid illiquidity measure in dealership markets since a dealer will enlarge the spread to cover the risk of being picked up by a better informed trader. Informed are not willing to trade anymore and the price does not move, implying a zero return. This would not be the case in a limit order market where informed traders may change their role.

The fourth chapter is devoted to the analysis of the informational content of Japanese candlesticks and HLOC price dynamics towards transaction costs management. In this essay, we question whether HLOC dynamics may help to solve the so-called trader's dilemma which is based on the impossibility to reduce all transaction costs' components simultaneously. When they submit an order, traders have to choose between submitting large quantities and incur a higher market impact by drying out best quotes quantities or submitting small orders to reduce market impact, bearing the risk of a possible future adverse price evolution. In this paper, we focus on particular HLOC configurations that have been proved to be related to liquidity in the previous essays of this thesis, i.e. consensus configurations. Our panel regressions' results suggest that consensus configurations are related to moments where the market impact costs are lower, due to a higher level of liquidity. We further analyze whether these dynamics may help to solve the trader's dilemma by comparing two order submission strategies: a random strategy and a candlestick-based strategy. The outcomes clearly show

that the candlestick-based strategy significantly yields in a lower market impact. As a result, the trader's dilemma may be solved by submitting liquidity demanding orders after the occurrence of consensus configurations.

All in all, this thesis has provided some valuable evidence to both academic and professional spheres. We first indicate that HLOC price dynamics may be used in the characterization of the state of liquidity. Based on recent academic research, this finding is highly relevant given the multidimensionality of liquidity that prevents the determination of a unique indicator. The issue of the availability of order book data may also be circumvented. We then dissociate this relation from the liquidity-volatility relationship that has been well documented in the past decades by including a proxy for realized volatility in a regression framework of liquidity on price dynamics. We also criticize a model based on the occurrence of zero returns and argue that the proposed theoretical justification is not as straightforward as usually thought, arguing that it does not suit all market structures. We finally find that these easy-to-observe configurations may help traders or portfolio managers to reduce transaction costs that are a deadweight loss which dramatically alters funds' performances. As a consequence, all these outcomes are useful for both academics and professionals and improve the conception of market liquidity issues for academic research and execution purposes.

This four-essay thesis opens different avenues for future research. First, it would be most interesting to conduct the same analysis on a different sample, e.g. U.S. stocks that may deliver different results. Investigating how other types of securities react to HLOC dynamics may also be useful. By using more recent samples, it would be possible to check whether high frequency as well as algorithmic trading affect the findings. This would also allow to reduce the interval length without increasing the noise in the patterns. A deeper analysis on how HLOC price dynamics and transaction costs are related, beyond the trader's dilemma, is also a future work that can be carried out. Finally, examining volume-based charts rather than time-based charts would bring another point of view and represent an interesting avenue for future research.

BIBLIOGRAPHY

Admati, A. and P. Pfleiderer (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial Studies 1*(1), 3–40.

Aitken, M. and C. Comerton-Forde (2003). How should liquidity be measured? *Pacific-Basin Finance Journal 11*, 45–59.

Alizadeh, S., M. W. Brandt, and F. X. Diebold (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance 57*(3), 1047–1091.

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets 5*, 31–56.

Amihud, Y. and H. Mendelson (1980). Dealership market: Market-making with inventory. *Journal of Financial Economics 8*(1), 31–53.

Amihud, Y. and H. Mendelson (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics 17*(2), 223–249.

Anand, A., S. Chakravarty, and T. Martell (2005). Empirical evidence on the evolution of liquidity: Choice of market versus limit orders by informed and uninformed traders. *Journal of Financial Markets 8*(3), 288 – 308.

Ball, C. A., W. N. Torous, and A. E. Tschoegl (1985). The degree of price resolution: The case of the gold market. *Journal of Futures Markets 5*(1), 29–43.

Beaupain, R., P. Giot, and M. Petitjean (2010). Volatility regimes and liquidity co-movements in cap-based portfolios. *Finance 31*(1), 55–79.

Bekaert, G., C. R. Harvey, and C. Lundblad (2007). Liquidity and expected returns: Lessons from emerging markets. *Review of Financial Studies 20*(6), 1783–1831.

Benston, G. J. and R. L. Hagerman (1974). Determinants of bid-asked spreads in the over-the-counter market. *Journal of Financial Economics 1*(4), 353–364.

Biais, B., P. Hillion, and C. Spatt (1995). An empirical analysis of the limit order book and the order flow in the paris bourse. *The Journal of Finance 50*(5), 1655–1689.

Bigalow, S. (2001). *Profitable Candlestick Trading: Pinpointing Market Opportunities to Maximize Profits*. Wiley Trading.

Bloomfield, R., M. O'Hara, and G. Saar (2005). The 'make or take' decision in an electronic market: Evidence on the evolution of liquidity. *Journal of Financial Economics 75*(1), 165–199.

Blume, M. E., A. C. Mackinlay, and B. Terker (1989). Order imbalances and stock price movements on october 19 and 20, 1987. *Journal of Finance 44*(4), 827–848.

Boudt, K. and M. Petitjean (2013). Intraday liquidity dynamics and news releases around price jumps: Evidence from the DJIA stocks. *Journal of Financial Markets* (forthcoming).

Brennan, M. J. and H. H. Cao (1997). International portfolio investment flows. *The Journal of Finance 52*(5), 1851–1880.

Brennan, M. J., T. Chordia, and A. Subrahmanyam (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics 49*(3), 345–373.

Brennan, M. J. and A. Subrahmanyam (1996). Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics 41*(3), 441–464.

Brock, W., J. Lakonishok, and B. LeBaron (1992). Simple technical trading rules and the stochastic properties of stock returns. *Journal of Finance 47*(5), 1731–64.

Cao, C., O. Hansch, and X. Wang (2009). The Information Content of an open Limit-Order Book. *Journal of Futures Markets 29*(1), 16–41.

Chakravarty, S. and C. W. Holden (1995). An integrated model of market and limit orders. *Journal of Financial Intermediation 4*(3), 213–241.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies 47*(0), 225–238.

Chan, K., A. J. Menkveld, and Z. Yang (2008). Information asymmetry and asset prices: Evidence from the china foreign share discount. *The Journal of Finance 63*(1), 159–196.

Chan, L. K. C. and J. Lakonishok (1995). The behavior of stock prices around institutional trades. *The Journal of Finance 50*(4), 1147–1174.

Chan, Y.-C. (2005). Price Movements Effects on the State of the Electronic Limit-Order Book. *The Financial Review 40*, 195–221.

Chang, Y. Y., R. Faff, and C.-Y. Hwang (2010). Liquidity and stock returns in japan: New evidence. *Pacific-Basin Finance Journal 18*(1), 90–115.

Chordia, T., R. Roll, and A. Subrahmanyam (2000). Commonality in Liquidity. *Journal of Financial Economics 56*, 3–28.

Chordia, T., R. Roll, and A. Subrahmanyam (2001). Market Liquidity and Trading Activity. *Journal of Finance 56*(2), 501–530.

Chordia, T., R. Roll, and A. Subrahmanyam (2002). Order Imbalance, Liquidity and Market Returns. *Journal of Financial Economics 65*, 111–130.

Chordia, T., R. Roll, and A. Subrahmanyam (2008). Liquidity and Market Efficiency. *Journal of Financial Economics 87*, 249–268.

Chordia, T. and A. Subrahmanyam (2004). Order Imbalance and Individual Stock Returns: Theory and Evidence. *Journal of Financial Economics 72*, 485–518.

Copeland, T. E. and D. Galai (1983). Information effects on the bid-ask spread. *The Journal of Finance 38*(5), 1457–1469.

Datar, V. T., N. Y. Naik, and R. Radcliffe (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets 1*(2), 203–219.

De Vilder, R. and M. Visser (2008). Ranking and combining volatility proxies for garch and stochastic volatility models. *MPRA paper*.

Demsetz, H. (1968). The Cost of Transacting. *Quarterly Journal of Economics 82*, 33–53.

Duvinage, M., P. Mazza, and M. Petitjean (2013). The intraday performance of market timing strategies and trading systems based on Japanese candlesticks. *Quantitative Finance*, forthcoming.

Easley, D., M. M. L. de Prado, and M. O'Hara (2012). Flow toxicity and liquidity in a high-frequency world. *Review of Financial Studies 25*(5), 1457–1493.

Easley, D., R. F. Engle, M. O'Hara, and L. Wu (2008, Spring). Time-varying arrival rates of informed and uninformed trades. *Journal of Financial Econometrics 6*(2), 171–207.

Easley, D., S. Hvidkjaer, and M. O'Hara (2002, October). Is information risk a determinant of asset returns? *Journal of Finance 57*(5), 2185–2221.

Easley, D., N. M. Kiefer, and M. O'Hara (1996, July). Cream-skimming or profit-sharing? the curious role of purchased order flow. *Journal of Finance 51*(3), 811–33.

Easley, D., N. M. Kiefer, and M. O'Hara (1997). One day in the life of a very common stock. *Review of Financial Studies 10*(3), 805–35.

Easley, D., N. M. Kiefer, M. O'Hara, and J. B. Paperman (1996, September). Liquidity, information, and infrequently traded stocks. *Journal of Finance 51*(4), 1405–36.

Easley, D., M. Lopez de Prado, and M. O'Hara (2010, November). The Microstructure of the "Flash Crash": Flow Toxicity, Liquidity Crashes and the Probability of Informed Trading. *The Journal of Portfolio Management 37*(2), 118–128.

Easley, D. and M. O'Hara (1987, September). Price, trade size, and information in securities markets. *Journal of Financial Economics 19*(1), 69–90.

Easley, D. and M. O'Hara (1992, June). Time and the process of security price adjustment. *Journal of Finance 47*(2), 576–605.

Easley, D., M. O'Hara, and J. Paperman (1998, August). Financial analysts and information-based trade. *Journal of Financial Markets 1*(2), 175–201.

Engle, R., M. Flemingz, E. Ghysels, and G. Nguyen (2012). Liquidity and Volatility in the U.S. Treasury Market: Evidence From A New Class of Dynamic Order Book Models. *Working Paper*.

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance 25*, 383–417.

Ferguson, N. (2008). *The Ascent of Money: A Financial History of the World*. A Penguin book : Business/History. Penguin Group (USA) Incorporated.

Fiess, N. and R. MacDonald (2002). Towards the fundamentals of technical analysis: analysing the information content of High, Low and Close prices. *Economic Modelling 19*(3), 353–374.

Fock, J., C. Klein, and B. Zwergel (2005). Performance of Candlestick Analysis on Intraday Futures Data. *The Journal of Derivatives 13*(1), 28–40.

Foster, F. and S. Viswanathan (1990). A theory of the interday variations in volume, variance, and trading costs in securities markets. *Review of Financial Studies 3*(4), 593–624.

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

Foucault, T., O. Kadan, and E. Kandel (2005). Limit order book as a market for liquidity. *Review of Financial Studies 18*(4), 1171–1217.

Glosten, L. R. (1994). Is the electronic open limit order book inevitable? *The Journal of Finance 49*(4), 1127–1161.

Glosten, L. R. and P. R. Milgrom (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14*(1), 71–100.

Goettler, R. L., C. A. Parlour, and U. Rajan (2005). Equilibrium in a dynamic limit order market. *The Journal of Finance 60*(5), 2149–2192.

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

Goldstein, M. A. and K. A. Kavajecz (2004). Trading strategies during circuit breakers and extreme market movements. *Journal of Financial Markets 7*(3), 301 – 333.

Goyenko, R., C. Holden, and C. Trzcinka (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics 92*(2), 153–181.

Grinblatt, M. and M. Keloharju (2001). How distance, language, and culture influence stockholdings and trades. *The Journal of Finance 56*(3), 1053–1073.

Harris, L. (1991). Stock price clustering and discreteness. *Review of Financial Studies 4*(3), 389–415.

Harris, L. (1998). Optimal dynamic order submission strategies in some stylized trading problems. *Financial Markets, Institutions and Instruments 7*(2), 1–76.

Harris, L. (2003). *Trading and exchanges: market microstructure for practitioners*. Financial Management Association survey and synthesis series. Oxford University Press.

Harris, L. and V. Panchapagesan (2005). The information content of the limit order book : evidence from NYSE specialist trading decisions. *Journal of Financial Markets 8*, 25–67.

Hasbrouck, J. and D. Seppi (2001). Common Factors in Prices, Order Flows, and Liquidity. *Journal of Financial Economics 59*, 383–411.

Hendershott, T., C. M. Jones, and A. J. Menkveld (2011). Does algorithmic trading improve liquidity? *The Journal of Finance 66*(1), 1–33.

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

Kandel, S., O. Sarig, and A. Wohl (2001). Do investors prefer round stock prices? evidence from israeli ipo auctions. *Journal of Banking and Finance 25*(8), 1543–1551.

Kang, J.-K. and R. Stulz (1997). Why is there a home bias? an analysis of foreign portfolio equity ownership in japan. *Journal of Financial Economics 46*(1), 3–28.

Kang, W. and W. Y. Yeo (2008). Liquidity beyond the best quote: A study of the nyse limit order book. *Working Paper*.

Kaniel, R. and H. Liu (2006). So what orders do informed traders use? *The Journal of Business 79*(4), 1867–1913.

Kavajecz, K. A. and E. R. Odders-White (2004). Technical analysis and liquidity provision. *Review of Financial Studies 17*(4), 1043–1071.

Keim, D. B. and A. Madhavan (1997). Transactions costs and investment style: an inter-exchange analysis of institutional equity trades. *Journal of Financial Economics 46*(3), 265–292.

Kissell, R., M. Glantz, and R. Malamut (2004). A practical framework for estimating transaction costs and developing optimal trading strategies to achieve best execution. *Finance Research Letters 1*(1), 35–46.

Knight, F. H. (1921). *Risk, uncertainty and profit*. Boston, New York, Houghton Mifflin Company.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica 53*(6), 1315–1335.

Lang, M., K. V. Lins, and M. Maffett (2010). Transparency, Liquidity, and Valuation: International Evidence. *Working Paper*.

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

Lesmond, D., J. Ogden, and C. Trzcinka (1999). A new estimate of transaction costs. *Review of Financial Studies 12*(5), 1113–1141.

Levine, R. and S. L. Schmukler (2006). Internationalization and stock market liquidity. *Review of Finance 10*(1), 153–187.

Liu, W. (2006). A liquidity-augmented capital asset pricing model. *Journal of Financial Economics 82*(3), 631 – 671.

Lütkepohl, H. (2006). *New Introduction to Multiple Time Series Analysis*. Berlin: Springer.

Maddala, G. S. (1987). Limited dependent variable models using panel data. *The Journal of Human Resources 22*(3), 307–338.

Marshall, B., M. Young, and R. Cahan (2008). Are candlestick technical trading strategies profitable in the japanese equity market? *Review of Quantitative Finance and Accounting 31*, 191–207.

Marshall, B., M. Young, and L. Rose (2006). Candlestick technical trading strategies: Can they create value for investors? *Journal of Banking & Finance 30*(8), 2303–2323.

Mazza, P. (2013). Price Dynamics and Market Liquidity: An Intraday Event Study on Euronext. *Working Paper*.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Z. (ed.) (Ed.), *Frontiers in Economics*, pp. 105–142. New York: Academic Press.

Morris, G. (1995). *Candlestick Charting Explained: Timeless Techniques for Trading Stocks and Futures*. New–York: McGraw-Hill Trade.

Neyman, J. and E. L. Scott (1948). Consistent estimates based on partially consistent observations. *Econometrica 16*(1), 1–32.

Nison, S. (1991). *Japanese Candlestick Charting Techniques: A Contemporary Guide to the Ancient Investment Technique of the Far East*. New–York: New York Institute of Finance.

Nison, S. (1994). *Beyond Candlesticks: New Japanese Charting Techniques Revealed*. New-York: John Wiley and Sons.

Næs, R. and J. A. Skjeltorp (2006). Order book characteristics and the volume-volatility relation: Empirical evidence from a limit order market. *Journal of Financial Markets 9*(4), 408–432.

Pagano, M. (1989). Endogenous market thinness and stock price volatility. *The Review of Economic Studies 56*(2), 269–287.

Parlour, C. (1998). Price dynamics in limit order markets. *Review of Financial Studies 11*(4), 789–816.

Pastor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111*(3), 642–685.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies 22*(1), 435–480.

Roşu, I. (2009). A dynamic model of the limit order book. *Review of Financial Studies 22*(11), 4601–4641.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance 39*(4), 1127–1139.

Stoll, H. R. (1978a). The pricing of security dealer services: An empirical study of nasdaq stocks. *The Journal of Finance 33*(4), 1153–1172.

Stoll, H. R. (1978b). The supply of dealer services in securities markets. *The Journal of Finance 33*(4), 1133–1151.

Stoll, H. R. (2003). Chapter 9 market microstructure. In M. H. G.M. Constantinides and R. Stulz (Eds.), *Corporate Finance*, Volume 1, Part A of *Handbook of the Economics of Finance*, pp. 553–604. Elsevier.

Toda, H. Y. and T. Yamamoto (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics 66*, 225–250.

Wang, Z.-M., C. Chiao, and Y.-T. Chang (2012). Technical analyses and order submission behaviors: Evidence from an emerging market. *International Review of Economics and Finance 24*(0), 109–128.

Weatherford, J. (2009). *The History of Money*. Crown Publishing Group.