



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

DETERMINANTS OF ARBITRAGE IN MULTI-MARKETS

TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN ECONOMÍA APLICADA

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL INDUSTRIAL

MARIO OSVALDO MORALES ALFARO

PROFESOR GUÍA:
MARCELA VALENZUELA BRAVO

PROFESOR CO-GUÍA:
ALEJANDRO BERNALES SILVA

COMISIÓN:
PATRICIO VALENZUELA AROS

SANTIAGO DE CHILE
2020

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL
Y AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA
POR: MARIO OSVALDO MORALES ALFARO
FECHA: 2020
PROF. GUÍA: MARCELA VALENZUELA BRAVO

DETERMINANTS OF ARBITRAGE IN MULTI-MARKETS

Debido a los avances tecnológicos que han experimentado los mercados financieros en las últimas décadas, hoy las transacciones de acciones equivalentes se distribuyen a través de múltiples bolsas con libros de órdenes electrónicos. Este escenario ha permitido que traders de alta frecuencia aprovechen momentos en que las diferencias de precios entre las bolsas admiten oportunidades de arbitraje rentables.

Esta tesis estudia las condiciones de mercado predominantes cuando ocurren estas oportunidades de arbitraje. Para ello, se usan datos de alta frecuencia de los libros de órdenes de 39 acciones transadas en dos bolsas europeas, Euronext y BATS Chi-X, durante todo el año 2014. Los datos se analizaron con frecuencias diarias e intradía.

En primer lugar, se detectaron las oportunidades de arbitraje existentes durante este periodo, con las cuales fuera posible realizar una ganancia al comprar una acción en un mercado y venderla en el otro. Luego, se usó una regresión logística, donde la variable dependiente identifica si es que en el periodo existía al menos una oportunidad y, las variables independientes son diferentes medidas de dimensiones del mercado: liquidez, incertidumbre, actividad transaccional y fragmentación de mercado.

Entre los principales resultados, se encuentra que la liquidez tiene una relación negativa con la aparición de las oportunidades, siendo más probables que estas aparezcan cuando los mercados son más ilíquidos. Además, es más relevante la distribución de la liquidez más allá de los mejores precios. Tanto la incertidumbre como la actividad transaccional muestran una relación positiva. Por último, el grado de fragmentación del mercado muestra una relación positiva con datos diarios, pero negativa con datos intradía, contrario a lo supuesto inicialmente.

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL
Y AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA
POR: MARIO OSVALDO MORALES ALFARO
FECHA: 2020
PROF. GUÍA: MARCELA VALENZUELA BRAVO

DETERMINANTS OF ARBITRAGE IN MULTI-MARKETS

Due to the technological advances that financial markets have experienced in the last decades, today the trading of equivalent stocks is distributed across multiple exchanges with electronic order books. This scenario has allowed high-frequency traders to take advantage of moments in which the price differences between exchanges admit profitable arbitrage opportunities.

This thesis studies the prevalent market conditions when these arbitrage opportunities occur. For that purpose, high-frequency data of the order books of 39 stocks traded at two European exchanges, Euronext and BATS Chi-X, during the entire year 2014 is used. The data was analyzed on a daily and intraday frequency.

First, the existing arbitrage opportunities during this period were detected, which allowed to make a profit by buying one stock at one market and selling it on another market. Then, a logistic regression was used, where the dependent variable identifies if there existed at least one arbitrage opportunity in that period, and the independent variables are different measures of market dimensions: liquidity, uncertainty, trading activity and market fragmentation.

Among the main results, liquidity has a negative relationship with the appearance of opportunities, being more likely for these to appear when markets are more illiquid. Additionally, it is more relevant the distribution of liquidity beyond the best prices. Both uncertainty and trading activity show a positive relationship. Finally, the degree of market fragmentation exhibits a positive relation with daily data, but a negative one with intraday data, contrary to what was initially expected.

Agradecimientos

En primer lugar, quiero agradecer a mi familia. A mi mamá María, mi papá Mario y mi hermana Rossana, por el amor incondicional y siempre creer en mí. A mis tíos, abuelos y primos, a quienes siempre tengo presentes a pesar de la distancia.

A mis amigos de la universidad con los que viví esta etapa. A Cris, Iván y Javi por toda la alegría y risas que me han brindado (además de soportarme cuando me amurro). A Álvaro, Iván, Luis y Osvaldo por todos los buenos momentos en el grupo, salidas a comer, viajes, trotes y trekking. A Daniela, Gustavo, Nicolás, Francisca y Mariana, que me han acompañado desde el principio, e incluso desde antes de entrar a la U. A Camilo y Felipe, por el apoyo hasta en los momentos más nefastos. A Giovanni y Felipe, por todas las risas, locuras y deconstrucción. A Cata y Caro, con quienes formé una amistad desde el comienzo de industrias. Y a todos con los que compartí en algún momento.

A mis amigos del colegio. A Feña, Lisset, Mohit, Rodrigo, Enzo, Romina, Coni, Caro, y todos con los que, a pesar de no tener tanto contacto, los sigo recordando con mucho cariño.

A mi profesora guía, Marcela Valenzuela, quien me brindó todo su apoyo, me animó y me aconsejó en el proceso de esta tesis.

Finalmente, agradecer el apoyo financiero para la realización de este trabajo del proyecto Fondecyt # 1190477.

Table of contents

- 1 Introduction** **1**

- 2 Arbitrage and market fragmentation** **3**
 - 2.1 Fragmentation and high-frequency trading 4
 - 2.2 Arbitrage and market quality 5
 - 2.3 Determinants of arbitrage 6
 - 2.3.1 Hypotheses 9

- 3 Data** **10**
 - 3.1 Arbitrage opportunities 12

- 4 Methodology** **15**
 - 4.1 Construction of variables 16
 - 4.1.1 Liquidity 16
 - 4.1.2 Uncertainty 18
 - 4.1.3 Trading activity 19
 - 4.1.4 Market fragmentation 19

- 5 Results** **20**
 - 5.1 Contemporaneous models 20
 - 5.1.1 Univariate models 20
 - 5.1.2 Multivariate models 23
 - 5.2 Predictive models 24
 - 5.2.1 Univariate models 24
 - 5.2.2 Multivariate models 26

- 6 Conclusion** **28**

- Bibliography** **30**

- 7 Appendix** **33**

List of Tables

3.1	Daily average of trading statistics for both venues	12
3.2	Daily average of trading statistics by index	12
3.3	Number and percentage of opportunities according to their duration	13
3.4	Number of opportunities and statistics about their duration for each 2-hour interval of the day	14
3.5	Profit rate of arbitrage opportunities by index	14
7.1	Daily average of trading statistics for the AEX stocks traded at Euronext.	33
7.2	Daily average of trading statistics for the AEX stocks traded at BATS Chi-X.	34
7.3	Daily average of trading statistics for the BEL20 stocks traded at Euronext.	35
7.4	Daily average of trading statistics for the BEL20 stocks traded at BATS Chi-X.	36
7.5	Depth as only regressor, univariate contemporaneous models with daily data.	37
7.6	Depth as only regressor, univariate contemporaneous models with 30-minute data.	38
7.7	Depth crossed as only regressor, univariate contemporaneous models with daily data.	39
7.8	Depth crossed as only regressor, univariate contemporaneous models with 30-minute data.	40
7.9	Bid-ask spread as only regressor, univariate contemporaneous models with daily data.	41
7.10	Bid-ask spread as only regressor, univariate contemporaneous models with 30-minute data.	42
7.11	Relative liquidity as only regressor, univariate contemporaneous models with daily data.	43
7.12	Relative liquidity as only regressor, univariate contemporaneous models with 30-minute data.	44
7.13	Slope as only regressor, univariate contemporaneous models with daily data.	45
7.14	Slope as only regressor, univariate contemporaneous models with 30-minute data.	46
7.15	Volatility variables as only regressor, univariate contemporaneous models with daily data.	47
7.16	Volatility variables as only regressor, univariate contemporaneous models with 30-minute data.	48
7.17	Crossed volatility variables as only regressor, univariate contemporaneous models with daily data.	49

7.18	Crossed volatility variables as only regressor, univariate contemporaneous models with 30-minute data.	50
7.19	Trading activity and market fragmentation as only regressor, univariate contemporaneous models with daily data.	51
7.20	Trading activity and market fragmentation as only regressor, univariate contemporaneous models with 30-minute data.	52
7.21	Multivariate contemporaneous models with daily data.	53
7.22	Multivariate contemporaneous models with 30-minute data.	54
7.23	Depth as only regressor, univariate predictive models with daily data.	55
7.24	Depth as only regressor, univariate predictive models with 30-minute data.	56
7.25	Depth crossed as only regressor, univariate predictive models with daily data.	57
7.26	Depth crossed as only regressor, univariate predictive models with 30-minute data.	58
7.27	Bid-ask spread as only regressor, univariate predictive models with daily data.	59
7.28	Bid-ask spread as only regressor, univariate predictive models with 30-minute data.	60
7.29	Relative liquidity as only regressor, univariate predictive models with daily data.	61
7.30	Relative liquidity as only regressor, univariate predictive models with 30-minute data.	62
7.31	Slope as only regressor, univariate predictive models with daily data.	63
7.32	Slope as only regressor, univariate predictive models with 30-minute data.	64
7.33	Volatility variables as only regressor, univariate predictive models with daily data.	65
7.34	Volatility variables as only regressor, univariate predictive models with 30-minute data.	66
7.35	Crossed volatility variables as only regressor, univariate predictive models with daily data.	67
7.36	Crossed volatility variables as only regressor, univariate predictive models with 30-minute data.	68
7.37	Trading activity variables as only regressor, univariate predictive models with daily data.	69
7.38	Trading activity variables as only regressor, univariate predictive models with 30-minute data.	70
7.39	Multivariate predictive models with daily data.	71
7.40	Multivariate predictive models with 30-minute data.	72
7.41	Fragmentation and volatility interaction models without control variables, daily data.	73
7.42	Fragmentation and volatility interaction models without control variables, 30-minute data.	74
7.43	Fragmentation and crossed volatility interaction models without control variables, daily data.	75
7.44	Fragmentation and crossed volatility interaction models without control variables, 30-minute data.	76
7.45	Fragmentation and volatility interaction models with all control variables, daily data.	77
7.46	Fragmentation and volatility interaction models with all control variables, 30-minute data.	78

7.47	Fragmentation and crossed volatility interaction models with all control variables, daily data.	79
7.48	Fragmentation and crossed volatility interaction models with all control variables, 30-minute data.	80
7.49	Fragmentation and volatility interaction models with one control variable, daily data.	81
7.50	Fragmentation and volatility interaction models with one control variable, 30-minute data.	82
7.51	Fragmentation and volatility interaction models with one control variable, daily data.	83
7.52	Fragmentation and volatility interaction models with one control variable, 30-minute data.	84
7.53	Fragmentation and volatility interaction models with one control variable, daily data.	85
7.54	Fragmentation and volatility interaction models with one control variable, 30-minute data.	86
7.55	Fragmentation and volatility interaction models with one control variable, daily data.	87
7.56	Fragmentation and volatility interaction models with one control variable, 30-minute data.	88
7.57	Fragmentation and volatility interaction models with one control variable, daily data.	89
7.58	Fragmentation and volatility interaction models with one control variable, 30-minute data.	90
7.59	Fragmentation and volatility interaction models with one control variable, daily data.	91
7.60	Fragmentation and volatility interaction models with one control variable, 30-minute data.	92
7.61	Fragmentation and crossed volatility interaction models with one control variable, daily data.	93
7.62	Fragmentation and crossed volatility interaction models with one control variable, 30-minute data.	94
7.63	Fragmentation and crossed volatility interaction models with one control variable, daily data.	95
7.64	Fragmentation and crossed volatility interaction models with one control variable, 30-minute data.	96
7.65	Fragmentation and crossed volatility interaction models with one control variable, daily data.	97
7.66	Fragmentation and crossed volatility interaction models with one control variable, 30-minute data.	98
7.67	Fragmentation and crossed volatility interaction models with one control variable, daily data.	99
7.68	Fragmentation and crossed volatility interaction models with one control variable, 30-minute data.	100
7.69	Fragmentation and crossed volatility interaction models with one control variable, daily data.	101

7.70	Fragmentation and crossed volatility interaction models with one control variable, 30-minute data.	102
7.71	Fragmentation and crossed volatility interaction models with one control variable, daily data.	103
7.72	Fragmentation and crossed volatility interaction models with one control variable, 30-minute data.	104

List of Figures

3.1 Arbitrage opportunities duration	13
--	----

Chapter 1

Introduction

Financial markets have seen major transformations over the last decades. Due to technological improvements, traditional trading floors have been replaced by electronic exchanges and markets have become deeply fragmented because of the multiple new venues that have emerged with these improvements.

Furthermore, new regulations promoting and encouraging the appearance of new actors have been promulgated both in the U.S. (RegNMS) and the European Union (MiFID), this under the assumption that stronger competition between exchanges would drive down transaction costs.

This new outlook of dispersed trading across a large number of venues has allowed the arise of instances in which mispricing occurs, thus traders with access to multiple markets may take advantage by arbitraging the mispriced stocks.

A particular kind of trader has been of particular attention to researchers and regulators: high-frequency traders. This term is commonly associated with traders running fast algorithms in order to profit. Thus, these traders are able to profit from arbitrage opportunities that last just a few seconds. Whether this activity is beneficial or harmful is still debated, with strong evidence supporting both sides of the discussion.

This work focus on the study of market conditions when arbitrage opportunities occur as a result of price differences between two exchanges. This is relevant because, by knowing when it would be more likely that these opportunities arise, regulators would be able to promote market designs that are the most beneficial for market quality and investor would be better informed when taking investment decisions.

It is first hypothesized that opportunities will appear when: markets are less liquid, because mispricing is more likely to occur when it is more difficult to trade assets; there is more uncertainty, because this would lead to more unexpected changes in prices, allowing mispricing moments; there is more trading activity, that is when there is more liquidity being consumed, and; markets are more fragmented, i.e., trading is more distributed across different venues.

The data used for this investigations comprises 1 year of tick-by-tick records of the Limit Order Book and Trades and Orders Book for all the constituents of two indexes: the Amsterdam Exchange Index (AEX) and Brussels Stock Exchange Index (BEL20). These companies stocks are traded at both Euronext and BATS Chi-X, two of the most important exchanges in Europe. Once the data is cleaned, it includes the information for 39 companies, which were part of the indexes during the whole year 2014.

Then, after identifying the relevant arbitrage opportunities for the study, logistic regressions were used for modeling the prevalent market conditions when these opportunities arise. This include contemporaneous models, in which the explanatory variables are measured at the same time than the dependant variable, and predictive models, in which the explanatory variables are lagged in 1 period.

As it was hypothesized, most of the liquidity variables show a negative correlation with the appearance of arbitrage opportunities, meaning that when markets are less liquid, it is more likely to observe this phenomenon. The exception was the crossed bid-ask spread, which exhibits a positive correlation. This makes sense since it is mechanically necessary for the arbitrage opportunities to occur when the crossed spread is negative.

In the same way, the uncertainty and trading activity variables show a positive correlation, thus corroborating the hypothesis that the presence of opportunities is greater when we are in presence of greater uncertainty about the assets' future prices, which is reflected in the volatility of the midquote, and when there is more liquidity being consumed as a consequence of more incoming orders.

On the other side, market fragmentation shows mixed results. When aggregating the data to a daily basis, the results are as expected with a positive relation, however this result reverses when analyzing intraday data.

This study is organized as follows: chapter 2 presents a review of relevant literature about arbitrage and market fragmentation, and the hypotheses emerging from previous research. Chapter 3 describes the data used in this investigation, as well as the processes for data cleansing and arbitrage opportunities detection. Chapter 4 describes the econometric methodology and the construction of variables. Chapter 5 presents the main results about the prevalent conditions when arbitrage opportunities occur. Chapter 6 concludes.

Chapter 2

Arbitrage and market fragmentation

Arbitrage is a concept that plays a central role in financial markets. One of the forms to define it is the “simultaneous purchase and sale of equivalent securities in two different markets in order to profit from discrepancies in their relationship” (Bodie, Kane and Marcus, 2009) [5].

It is directly related to the Law of One Price (LOP), which states that in well-functioning, efficient financial markets identical cash flows must trade at the same price. If not, the demand for the relatively cheap securities would rise, as well as the supply for the relatively expensive ones, driving the prices closer until they become the same. This way, arbitrageurs enforce the LOP once it breaks down and an arbitrage opportunity exists.

Theoretically, it requires no capital, is riskless and should disappear instantaneously. However, in reality there exist various frictions such as short-selling costs, funding constraints, idiosyncratic risk, etc., which make some arbitrage opportunities persistent (Gromb and Vayanos, 2010) [15].

A particular case of arbitrage arise in the presence of fragmented financial markets. When companies cross-list their shares in different exchanges, differences in the price of these assets on each market may occur over time, due to demand shocks or a delayed arrival of information, allowing arbitrageurs with access to both venues to exploit these opportunities and make a profit.

This case of arbitrage opportunities has become relevant because of the proliferation of new markets and platforms that has occurred over the last decades, and the consequent cross-listing of shares by companies. In the US, most stocks were listed on the New York Stock Exchange or the NASDAQ at the beginning of the 2000s. After a decade, there existed more than 50 different exchanges, dark pools and alternative trading venues (Arnuk and Saluzzi, 2012) [2].

Furthermore, new technologies and innovation have allowed exchanges to become completely electronic venues. Thus, services formerly provided by brokers, dealers and other actors have been replaced by computer algorithms. Although it has not been uniform, all asset classes have migrated up to a certain degree to electronic markets (Cardella et al.,

2014 [7]). Also, smart order routers are often used when determining where to send an order (Foucault and Menkveld, 2008) [11].

In this context of electronic venues and algorithmic trading, high-frequency traders have been of particular interest, which are typically associated with algorithms running on extremely fast computers. By 2012, high-frequency trading firms represented 2% of the trading firms in the US, but accounted for almost 70% of traded volume in US equity markets (Easley, López de Prado and O’Hara, 2012) [9].

However, there is no consensus about whether fragmentation and high-frequency trading are beneficial or harmful. The appearance of these new trading venues and the increasing competition that it carries may benefit investors by reducing costs and enhancing execution speed and prices efficiency (O’Hara and Ye, 2009 [25]). In response to this, regulators have established new laws and instructions that encourage the entrance of new actors. On 2005, the United States Securities and Exchange Commission promulgated the Regulation National Market System (RegNMS), meanwhile the European Union promulgated the Markets in Financial Instruments Directive (MiFID) in 2007.

On the other side, some argue that high-frequency trading is detrimental for investors’ welfare. Nobel laureate Joseph Stiglitz (2014) [29] has expressed his concern about high-frequency trading and doubts about its social value. Foucault, Kozhan, and Tham (2016) [10] find that it increases the adverse selection costs and harms liquidity, meanwhile Rösch (2017) [27] state that, although there is a detrimental effect, the benefits surpass the harm and arbitrage enhances market integration.

In the next sections, the main research about fragmentation, high frequency trading and arbitrage will be discussed.

2.1 Fragmentation and high-frequency trading

The discussion about the effects of market fragmentation and the arise of high-frequency trading remains a open. There is evidence evidencing both benefits and harms as a result of this phenomenon, but there is no agreement on whether the overall effect of this results improves or not social welfare.

First of all, the appearance of new venues and high-frequency traders has helped the adoption electronic trading. As a consequence, this has resulted into lower transaction costs, with a decrease on average retail commissions, and a higher traded volume, although it has decreased since the 2007-2008 financial crisis (Angel, Harris and Spatt, 2015 [1]). But, this does not mean that the social benefit is positive.

Pagnotta and Philippon (2015) [26] develop a theoretical model in which exchanges differentiate between themselves by investing in speed. They find an over-investment by the fastest type of exchange, since it privately benefits itself but diminishes fee competition.

Budish, Cramton and Shim (2015) [6] find that at a millisecond level correlations among

different exchanges break down allowing the appearance of mechanical arbitrage opportunities. This produces an inefficient arms race for speed which harms market liquidity, but as a consequence of market design and not directly from fragmentation.

More recently, Haslag and Ringgenberg (2016) [18] study how market fragmentation affects liquidity. They find that, for larger stocks, exchange competition reduces trading costs, meanwhile for smaller stocks, network externalities reduce market quality.

Other studies regarding the effects of high frequency traders on market quality establish different results depending of the characteristics of these kind of traders in relation with other, slower traders. Goettler, Parlour and Rajan (2009) [13] develop a model in which traders arrive randomly to the market and decided whether to become informed about the fundamental value or not. They find liquidity improvement when traders become more informed.

Bernales (2017) [4] extends the above model by adding speed dispersion among traders, which are randomly assigned to be fast or slow. He finds that waiting costs are reduced, but slow traders' profits get damaged. When faster traders have no informational advantage, they harm market liquidity, but when they are also better informed, market liquidity improves. Similarly, Jovanovic and Menkveld (2016) [21] find that faster and better-informed high-frequency traders improve welfare by reducing information asymmetry between other traders.

By studying the automation of the manually quoting NYSE limit order book, Hendershott, Jones and Menkveld (2011) [19] find that liquidity improves with narrower spreads for large cap stocks, since quotes become more informative after the increase of price discovery by algorithmic traders. Hasbrouck (2018) [17] interprets his findings as a narrower spread and quicker price recovery (after the spread widens following a market order arrival) as a results of increased high-frequency traders competition.

Benos et al. (2015) [3] study the commonality among high-frequency traders, and find that their correlated trading is associated with permanent price impacts. This would be the result of informed trades by high-frequency traders, who collectively buy and sell at the “right” time, and this activity “does not generally contribute to undue price pressure and price dislocations”.

Overall, theory and evidence is mixed. When high-frequency traders are faster that other traders, there is a reduction in market quality, but there is an improvement when they are better informed. When they are faster and better informed, the result is unclear with evidence for both beneficial and harmful effects.

2.2 Arbitrage and market quality

With the proliferation of new trading venues and cross-listing of stocks, high-frequency traders with access to multiple markets could benefit from differences in the prices of equivalent stocks. Although, in theory, this is beneficial because of the LOP enforcement, however, this could carry detrimental effects to market quality. Understanding these consequences is

important, since investors concentrate their trading activity in the overlapping hours of these venues (Menkveld, 2008) [23].

Theory predicts that when arbitrage opportunities appear as a result of transient demand (supply) shocks, arbitrageurs trade against market demand (supply), thus, decreasing inventory holding costs for liquidity providers. Gromb and Vayanos (2002) [14] develop a model in which arbitrageurs exploit discrepancies in prices. They find that this activity benefits all investors, because “through their trading, arbitrageurs bring prices closer to fundamentals and supply liquidity to the market”.

In a more recent paper, Gromb and Vayanos (2010) [15] show how arbitrage improves liquidity by studying large price deviations that last for months. This is explained because of the presence of various frictions, such as short-selling costs, funding constraints, idiosyncratic risk, etc.

However, short-lived opportunities may arise due to asynchronous adjustments in asset prices following information arrival. Thus arbitrage becomes “toxic” since dealers may be unaware that they are providing liquidity at a loss (Easley, Lopez de Prado, and O’Hara, 2012) [9]. This harms liquidity because, in order to cover the risk of being adversely selected, dealers charge larger bid-ask spreads (Copeland and Galai, 1983) [8].

Foucault, Kozhan, and Tham (2016) [10] study toxic arbitrage using data from three pairs of currencies. Following the methodology of Schultz and Shive (2010) [28], they use price patterns following the occurrence of arbitrage to sort them into two groups: toxic (staggered price movements in the same direction following the occurrence) and nontoxic (reversals in the rate of the currency pair that triggers the arbitrage opportunity).

They find a positive and significant relation between the daily fraction of toxic arbitrage opportunities and illiquidity and high-speed arbitrage can be a source of adverse selection.

Later, Rösch (2017) [27] studies arbitrage among 72 American Depositary Receipts (ADRs) and the (currency-adjusted) price of their home-market counterparts, using tick-by-tick data. He classifies arbitrage opportunities as toxic and non-toxic, and finds that most of these opportunities are a consequence of demand shocks, and as arbitrageurs trade against this demand, they improve liquidity.

2.3 Determinants of arbitrage

In order to fully understand this kind of arbitrage, it is necessary to study what are the prevalent characteristics of the markets when this events occur, and whether it is possible to predict the occurrence of these opportunities by observing the current market conditions.

One of the most important features of financial markets is liquidity. Although a precise definition will depend on each particular model, it refers to “the ease with which an asset can be traded without affecting the asset’s price” (Valenzuela et al., 2015 [30]).

According to Hasbrouck (2007) [16], liquidity is sometimes defined as “depth, breadth, and resiliency”. A deep market is such that when looking a little above the current market price, there is a large incremental quantity available for sale, while when looking a little below the current price, there is a large incremental quantity sought by one or more buyers. A broad market is the one with many participants, none of whom is presumed to exert significant market power. In a resilient market, the price effects from the trading process are small and quickly disappear.

In the case of an exchange, it is told to be liquid if traders are able to execute a large amount of orders quickly and at a low cost, meanwhile it will be illiquid if, by executing these orders, traders will face an increasing price per stock.

Gagnon and Karolyi (2010) [12] study the arbitrage between ADRs and their respective home-market shares from a sample of 506 companies from 35 countries. One of their hypothesis is that, for more illiquid shares there exist more price discrepancies between markets.

Although they don’t find conclusive results in all of their illiquidity measures, they find that for better-developed countries, illiquidity is reliably positive in the home-market and, for less-developed countries, US-market illiquidity is also positive. However, they use daily closing data, limiting the results by not considering the occurrence of arbitrage in different moments of the day.

Marshall, Nguyen and Visaltanachoti (2013) [22] focus on the study of two S&P 500 ETFs and the prevailing conditions when arbitrage opportunities between these assets arise. For their study they use high-frequency data.

They measure liquidity by the width of the bid-ask spread. The more wide the spread is, the more illiquid the market is. They find that it is more likely for arbitrage opportunities to occur when the market is relatively less liquid.

Foucault, Kozhan and Tham (2016) [10] study triangular arbitrage among three currency pairs (dollar-euro, euro-pound, dollar-pound) using high-frequency data. They measure illiquidity with effective spreads, slope of limit order books or a measure of adverse selection costs for dealers.

The arbitrage opportunities they consider meet four conditions: they are frequent, they are short lived, they are more efficiently exploited by machines and, they yield a small profit per opportunity. They find that “illiquidity is higher when arbitrage opportunities are more frequently due to asynchronous price adjustment than to price pressures”.

Uncertainty is another important concept when studying market quality. Many times uncertainty is used indistinctly from risk, however, there certain differences between these two ideas. In general, it is told to be in front of a risky situation when a variable may take different outcomes, but with a known probability of occurrence. When a situation is uncertain, probabilities can only be approximated but not exactly. Then, uncertainty is related with the inability to forecast future outcomes. In this investigation, uncertainty is given from the random processes that prices follow.

Because of the uncertainty of future prices, which translate into uncertainty in returns and other measures derived from prices, there is a constant worry from market agents about how high is the current uncertainty and the possible variations that these variables may have in the future.

In conclusion, higher uncertainty is related to higher future variations with respect to the current observations. Then, if prices get far away from current prices, investor won't be able to update their quotes and then there will be moments in which assets will be mispriced, originating arbitrage opportunities.

In their investigation, Marshall, Nguyen and Visaltanachoti (2013) [22] include the study of liquidity risk, measured as the second moment of liquidity changes, i.e., spread volatility. They find that arbitrage opportunities occur when liquidity risk is higher, and also when the VIX, and index that reflects the S&P 500 volatility, is higher as well.

From another point of view, Budish, Cramton and Shim (2015) [6] give a major importance to volatility. They see arbitrage as a result from the market structure, in which, at a millisecond level, correlations between markets break and volatility is relatively higher, thus, explaining the appearance of opportunities.

Another feature is trading activity. It refers to the amount of trades observed at a particular moment, namely, the difficulty with which buy and sell orders are executed. Trading activity may vary both by asset and along time, and is closely related to liquidity.

Trading activity is the result from the synergy between supply and demand. Since liquidity may be interpreted as the set of conditions that supply and demand show at a particular moment, then trading activity will be conditioned by liquidity. As Johnson (2008) [20] states, "large changes in liquidity cannot occur without a lot of population flux, and a small amount of flux must imply a small change in liquidity".

Then, when trading activity increases, the available liquidity in the market will suffer variations. Then, higher trading activity is related to higher volatility of the available liquidity. Then, when this volatility is higher, there may be instants with distorted prices, giving space to arbitrage opportunities.

Again, Marshall, Nguyen and Visaltanachoti (2013) [22] include the traded value in their research, and find it to be positively correlated with the appearance of arbitrage opportunities.

The last condition is market fragmentation. As it has been mentioned earlier, trading activity and liquidity has distributed across several venues. Now, investor not only have to decide when to trade but where. Menkveld (2008) [23] predicts that in presence of a mixed setting, in which due to timezone differences markets overlap and become fragmented at certain time of the day, traders concentrate their activity in the overlap.

Besides, as Foucault, Kozhan and Tham (2016) [10] explain, information may arrive asynchronously to both venues, leading to delayed adjustments in prices which may be exploited by arbitrageurs. This way, it is expected that as market fragmentation increases, more arbitrage opportunities would be observed.

2.3.1 Hypotheses

This investigation will focus on studying what conditions prevail when arbitrage opportunities arise. This will be done both on a contemporaneous basis, in which market conditions will be observed in the same period in which an arbitrage opportunity exists, and on a predictive basis, by analyzing the conditions just before opportunities appear.

Thus, the hypotheses to be tested are as follows:

- Arbitrage opportunities are more likely to appear when market liquidity is lower
- Arbitrage opportunities are more likely to appear when market uncertainty is higher
- Arbitrage opportunities are more likely to appear when trading activity is higher
- Arbitrage opportunities are more likely to appear when market fragmentation is higher

Chapter 3

Data

The data used in this investigation contains both the Limit Order Book (LOB) and the Trades and Orders Book (TAS) from two of the most important exchanges in Europe: Euronext and BATS Chi-X.

Euronext was founded in 2000, after the merger of the exchanges in Amsterdam, Paris and Brussels. It has since grown by acquiring additional exchanges, and today it is the largest stock exchange in continental Europe, with over 1,300 issuers.

BATS Chi-X is a low cost alternative for trading equities whose primary listing are in European exchanges, such as Euronext. Chi-X started its operations in 2007 and was the first multilateral trading venue that launched in anticipation to the MiFID. Then, in 2011 was bought by BATS Global Markets and now it offers access to more than 6,000 securities from 18 major European markets.

The design of both exchanges is similar and is known as Continuous Limit Order Book (CLOB), which is a record of limit orders entered by investors. A limit order specifies the price at which an investor wishes to make a deal over the also specified quantity of shares. The arrival of these orders is recorded in the TAS. The LOB is built over the orders with the best price: the sell order with the lowest price will define the Ask Price, and the buy order with the highest price will define the Bid Price.

Investors can also submit market orders, which are buy or sell orders at the current best price, thus executing immediately. Their arrival is also recorded in the TAS.

In this study, high-frequency data will be used. This means that consecutive records of arriving orders or updates in the books may be in less than one second. Specifically, the data is recorded in the order of milliseconds and it comprises all the information from 2014.

The LOB database contains up to three best prices for each side of the book (buy and sell) and their respective quantities, and a new record is inserted each time one of these prices changes after the entry of a new limit order or the execution of a market order. The TAS database contains the price and quantity of each market or limit order. Each record represents a new incoming order, and from this book it is possible to build up the LOB.

The assets that will be studied are the main components of the Amsterdam Exchange Index (AEX) and the Brussels Exchange Index (BEL20). The AEX includes the 25 most traded stocks primarily listed in Euronext Amsterdam, meanwhile the BEL20 consists of up to 20 representative Belgian companies.

These stocks are mainly traded in Euronext and BATS Chi-X. In order to be consistent in the analysis, only 39 stocks will be considered, 22 dutch and 17 belgian stocks, which are the ones that were part of the index during the whole year and were traded simultaneously on both venues.

Before performing any analysis, a data cleansing process was carried out. In the first place, only the data between 08:00 and 16:00 was considered. That is because that's the time interval in which both markets are simultaneously operating. Then, the records were filtered so that there was at least one strictly positive best price at each side of the book.

Days with different opening or closing hours were dropped, so that all days considered had the same length of trading hours. In this case, only three days were dropped: December 24th and December 31st, since their closing hours were earlier because of the respective holidays, and June 12th, day in which the trading started several minutes after 08:00.

Between March 30th and October 26th hours were added 1 hour in the database, because of the daylight saving time. Thus, the periods within these trading days are comparable with those outside this interval.

Observations in which an arbitrage opportunity existed within the same book, that is, the bid price was higher than the ask price, were also deleted. This cleansing is suggested by Schultz and Shive (2010) [28].

The next step was to select the last record for each second in which there were more than one observation. Then, missing seconds are filled with information from previous observation, so that every second in the day has a record. This way, the book will contain the same information an investor faces when she visits the limit order book.

Finally, for each stock, from both books, the one from Euronext and the one from BATS Chi X, are merged, leaving a consolidated book with the best prices from both exchanges. Additionally, a crossed limit order book is constructed. This one considers the best three prices, for both the buy and sell side, from all the prices available on both venues. Thus, this book represents what an investor would observe if she was able to trade on both markets.

As it can be seen in 3.1, on average, each day 2,464,589 shares of the involved stocks are traded in both markets, which accounts for an average value of 40,365,896 euros. The trading in Euronext represents almost three times what is traded in BATS Chi-X, with 1,809,772 shares traded each day and 29,371,382 euros in Euronext, meanwhile BATS Chi-X has a daily average volume of 654,817 shares that account for 10,994,514 euros.

The average number of incoming limit buy and sell orders is also larger for Euronext than BATS Chi-X. Both, the average number of submitted limit buy and limit sell orders is almost the same.

	Average volume	Average value	Average buy orders	Average sell orders
Euronext	1809772	29371382	36653	36518
BATS Chi-X	654817	10994514	50798	50778
Total	2464589	40365896	50798	50778

Table 3.1: Daily average of trading statistics for both venues

A larger share of the traded volume and value takes place on stocks from the AEX index, with 4,108,486 shares and 57,578,219 euros traded on average each day. In contrast, the companies from the BEL20 index trade only 337,194 shares and 18,091,127 euros each day.

The number of limit buy orders and limit sell orders submitted each day is similar for the companies belonging to both the AEX and BEL20 indexes.

	Average volume	Average value	Average buy orders	Average sell orders
AEX	4108486	57578219	67944	67756
BEL20	337194	18091127	28610	28807
Total	2464590	40365897	50798	50778

Table 3.2: Daily average of trading statistics by index

A more detailed description of the trading activity for each stock can be found in the Appendix. In Table 7.1, the same trading statistics are computed for each of the stocks traded at Euronext from the AEX index, meanwhile Table 7.2 contains the information for their counterparts traded at BATS Chi-X. Similarly, Table 7.3 shows the same information for the stocks from the BEL20 index traded at Euronext, and Table 7.4 contains the descriptive statistics for the stocks traded at BATS Chi-X.

3.1 Arbitrage opportunities

Arbitrage opportunities are defined by the next condition:

$$\frac{bid_i}{ask_j} > (1 + gap), i \neq j \quad (3.1)$$

This means that we are in presence of an arbitrage opportunity each time the bid price in book i is higher than the ask price in the book j , plus a certain gap. When this condition holds, an investor could buy the stock in j and sell it in i , earning a profit.

The *gap* component of the condition represents a minimum required earning so that the opportunity is profitable, and the costs associated with the trade, such as transaction costs, are all covered. When considering a $gap = 0$, a total of 10461 arbitrage opportunities are observed. Of these opportunities, most of them last less than one second, as it is shown in Figure 3.1.

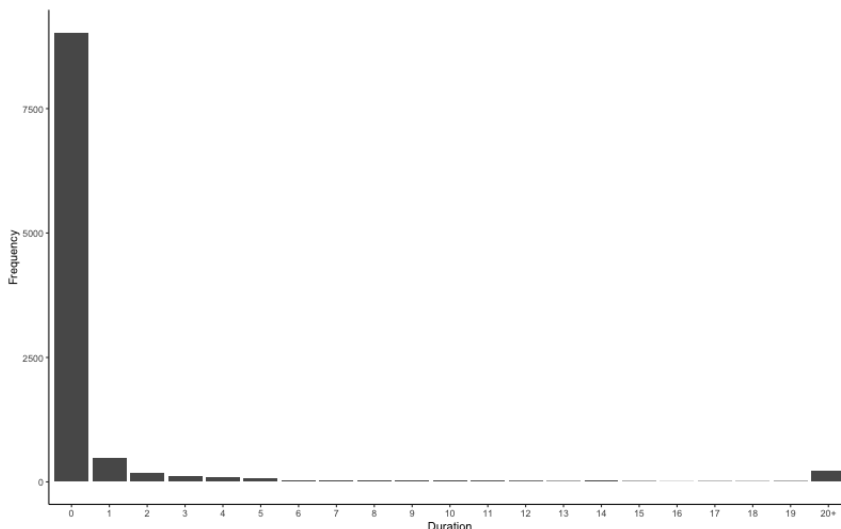


Figure 3.1: Arbitrage opportunities duration

The duration is approximated to the greatest integer less than or equal to the real duration, since we cannot know the exact duration of the opportunities after the data cleansing process. This way, an opportunity is said to last 0 seconds if it is only observed in one consecutive record in the database.

On average, the opportunities last for 2.48 seconds, with a median of 0 seconds. The opportunity with the maximum duration lasted 3,283 seconds, which is equal to 54 minutes and 43 seconds. On table 3.3, statistics about the distribution of the duration are shown, distinguishing between in which market the bid price was higher. As it can be seen, most opportunities last for less than 1 second.

	Opportunities	% of total
Less than 1 second	9033	86.35 %
Between 1 and 2 seconds	483	4.62 %
Between 2 and 3 seconds	184	1.76 %
Between 3 and 5 seconds	214	2.05 %
Between 5 and 10 seconds	199	1.9 %
More than 10 seconds	348	3.33 %
Total	10461	100 %

Table 3.3: Number and percentage of opportunities according to their duration

Most of the arbitrage opportunities occur in the first two hours of the trading day, from 8 : 00 to 10 : 00, with 43.1% of all the opportunities detected. The amount of opportunities decreases in the next two two-hour intervals to a similar level, 15.4% and 15.6% for the 10 : 00 to 12 : 00 interval and the 12 : 00 to 14 : 00 interval, respectively. Finally, the proportion increases again in the last interval with 25.9% of all the opportunities. Then, the arbitrage opportunities are more likely to happen at the open or the close of market.

For the purpose of this study, only those arbitrage opportunities which last less than one second will be considered. These subset meet the conditions Foucault, Kozhan and Tham

	N	%	Min	Median	Mean	Max
[08:00 ; 10:00]	4507	43.1	0	0	2.7	595
(10:00 ; 12:00]	1613	15.4	0	0	2.4	238
(12:00 ; 14:00]	1635	15.6	0	0	4.5	3283
(14:00 ; 16:00]	2706	25.9	0	0	1.0	116
Total	10461	100	0	0	2.5	3283

Table 3.4: Number of opportunities and statistics about their duration for each 2-hour interval of the day

(2016) [10] describe for high-frequency arbitrage opportunities:

- (i) they are frequent: on average, each company exhibits 0.92 opportunities a day. This frequency decreases dramatically for longer-than-one-second opportunities
- (ii) they are short lived: their duration is in the order of milliseconds
- (iii) they are more efficiently exploited by machines: due to their short duration, a trader would not be able to execute market orders fast enough to exploit these opportunities
- (iv) they yield a small profit per opportunity, as it will be shown next.

The potential profit rate of each opportunity is computed following the next formula, which is tantamount to the methodology used by Mitchel, Pulvino and Stafford (2002) [24] and also used by Marshall, Nguyen and Visaltanachoti (2013) [22]:

$$\text{Profit rate} = 100 \frac{\text{Bid-ask spread}}{\text{Midquote}} \quad (3.2)$$

On average, each opportunity yields a profit of 11 basis points, with a median of 2 basis points. When inspecting the profits from the companies within each index, the mean is higher for the companies from the BEL20 index, although the maximum profit is achieved by a stock in the AEX index, as it can be seen on table 3.5.

	N	Min	Median	Mean	Max
Total	9033	0.01	0.02	0.11	6.60
AEX	5239	0.01	0.02	0.09	6.60
BEL20	3794	0.01	0.02	0.13	6.21

Table 3.5: Profit rate of arbitrage opportunities by index

More than half of the opportunities on both indexes yield less than 2 basis points, representing a small profit per opportunity, just like the description in Foucalt, Kozhan and Tham (2016) [10].

The total profit in euros is at least 73,174.56, which may seem attractive enough for some (algorithmic) traders to exploit. This is a lower boundary since only the minimum quantity of the limit orders associated with each opportunity is considered. When this minimum amount of shares is traded, the bid and/or the ask prices change, enabling the arbitrage opportunity to disappear. Besides, a larger amount may be achieved when analyzing tick-by-tick data.

Chapter 4

Methodology

A logit model is used to statistically test the hypotheses described above. In this model, the dependent variable is 1 if there exists an arbitrage opportunity for stock pair i in period t . Only the arbitrage opportunities that last less than one second are considered.

$$y_{i,t} = \begin{cases} 1, & \text{if there exists an opportunity for stock pair } i \text{ in period } t \\ 0, & \text{if not} \end{cases} \quad (4.1)$$

Each stock pair i is composed by the stocks from the same company traded at both venues: Euronext and BATS Chi-X. That is, i represents the securities for which the cash flows they claim are the same.

Each period t represents, in the base model, a different trading day from year 2014. Then, $y_{i,t}$ will be 1 if there exists at least one arbitrage opportunity that lasted less than one second in day t for stocks i . However, in order to test the robustness of the results, the model is run under different periodicity configurations, then t may also represent periods of 2 hours, 1 hour, 30 minutes and 15 minutes long.

The independent variables are proxies of the different market dimensions mentioned in the precedent sections, meanwhile the methodology of their construction will be described in the next subsection.

Unless the definition of their construction says otherwise, each variable is measured in each second of the trading day. Then, when aggregating the data for each period configuration, each variable is averaged through each time interval. This means that, for example, the variables in the daily model represent the mean of the second-basis measures during each day.

For making the variable coefficients comparable between them, they are all standardized, subtracting the mean of each variable and dividing it by the standard deviation.

In addition to the independent variables, fixed-effect variables are also included for each

stock i and each period t . This way, the effects that each variable have over the arbitrage opportunities prevalence will be clean of effects from each company that are observed across all periods, and each period phenomena that could be observed across all companies at a specific time.

The first set of models have a contemporaneous configuration. This is, their focus is to study what are the prevalent characteristics of the market when the arbitrage opportunities occur. Thus, both the dependent and independent variables are measured during the same period in each model.

The second set of models have a predictive configuration. In this case, independent variables are lagged one period with respect to the dependent variable. This way, these models study whether there are certain characteristics that could anticipate the occurrence of arbitrage opportunities. When using intraday period configurations, the first observation of each day is dropped, since the lagged independent variable would represent information from the previous trading day, being inconsistent with the rest of the observations which consider only the information within the same day.

Within each set, the models were constructed first considering only one independent variable, this way studying how each measure is related individually to the appearance of the opportunities, and then they were constructed combining different measures of different aspects of the book.

4.1 Construction of variables

The dependent variable is defined by equation 4.1, and the existence of each arbitrage opportunity is defined by condition 3.1.

As it was said before, the independent variables are proxies of the different market dimensions studied, and their construction is described in the next subsections. It is worth mentioning that continuous variables were standardized around their global mean and global standard deviation, this way, coefficients are comparable across different variables.

4.1.1 Liquidity

1. Depth: this variable is one of the main measures of liquidity, and it refers to the amount of shares offered or demanded at a determined price. Depending on which prices are considered, the next variables are constructed:
 - (a) Depth ask: average of the amount of shares offered at the ask price of each book.
 - (b) Depth bid: average of the amount of shares demanded at the bid price of each book.
 - (c) Depth best prices: average of the shares offered and demanded at the ask and bid prices of each book.

- (d) Depth beyond ask: average of the amount of shares offered at the second and third best prices on the sell side of each book.
- (e) Depth beyond bid: average of the amount of shares demanded at the second and third best prices on the buy side of each book.
- (f) Depth beyond: average of the amount of shares offered and demanded at the second and third best prices of both the sell and buy side of the each book.

Since these variables are constructed from the data of each book alone, crossed-book versions are also constructed, considering the consolidated book with all prices:

- (a) Depth crossed ask: amount of shares offered at the crossed ask price.
 - (b) Depth crossed bid: amount of shares offered at the crossed bid price.
 - (c) Depth best crossed prices: amount of shares offered and demanded at the crossed ask and crossed bid prices.
 - (d) Depth beyond crossed ask: amount of shares offered at the second and third best crossed prices on the sell side.
 - (e) Depth beyond crossed bid: amount of shares demanded at the second and third best crossed prices on the buy side.
 - (f) Depth beyond crossed: amount of shares offered and demanded at the second and third crossed prices of both the sell and buy side.
2. Bid-ask spread: this is another of the main measures of liquidity, and refers to the distance between the best prices.
- (a) Bid-ask spread: average of the difference between the ask and bid prices of each book.
 - (b) Bid-ask spread standardized: average of the difference between the ask and bid prices divided by the midquote (average of the best prices) of each book.

$$\text{Bid-ask spread standardized} = \frac{Ask - Bid}{\frac{Ask + Bid}{2}} \quad (4.2)$$

Like the depth variables, crossed version of the bid-ask spread are also computed:

- (a) Crossed bid-ask spread: difference between the crossed ask and the crossed bid prices.
 - (b) Crossed bid-ask spread standardized: difference between the crossed ask and the crossed bid prices divided by the crossed midquote, as in equation 4.2.
3. Relative liquidity: this variable is proposed by Valenzuela et al. (2015) [30], and measures how distributed is the depth across the best prices, then, when the relative liquidity is greater the liquidity is more concentrated around the best prices, this is because the amount of shares available at the best price is relatively higher than that available at the next best prices, or because the next best prices a relatively closer to the best price. To construct this variable, the next steps are followed:
- (a) The difference in ticks between the best price and the following prices is computed:

$$\Delta_{i,t}^{j_n} = |p_{i,t}^{j_{best}} - p_{i,t}^{j_n}| \quad (4.3)$$

Where j represents each side of the book (buy or sell), j_{best} refers to the best price of side j and j_n the n -th best price.

- (b) The accumulated distribution from the best to the third best price $F_{i,t}^{j_n}$ is computed.
- (c) A weight for each of the price is calculated, in which:

$$g(\lambda, \Delta_{i,t}^{j_n}) = \frac{\exp(-\lambda \Delta_{i,t}^{j_n})}{\sum_n \exp(-\lambda \Delta_{i,t}^{j_n})} \quad (4.4)$$

λ is assumed to be 0.366, as it was first suggested by Valenzuela et al. (2015) [30].

- (d) Finally, relative liquidity is calculated as:

$$RLIQ_{i,t}^j = \sum_n F_{i,t}^{j_n} g(\lambda, \Delta_{i,t}^{j_n}) \quad (4.5)$$

- 4. Slope: this is proposed by Foucault, Kozhan and Tham (2016) [10] and is also associated to the distribution of liquidity. It measures how deep or thin is the book at the best prices and with a bigger slope, the book is thinner. It is constructed according to the next formula:

$$\text{Slope}_{i,t}^j = \frac{|p_{i,t}^{j_{best}} - p_{i,t}^{j_2}|}{\text{depth}_{i,t}^{j_{best}}} \quad (4.6)$$

4.1.2 Uncertainty

- 1. Midquote return volatility: the main measure of uncertainty is volatility. For this purpose, it is first computed the logarithmic return of the average midquote (between both venues) for each asset on a second-basis, that is, the difference between the logarithm of the midquote at time t and the logarithm of the midquote at $t - 1$.

$$r_{i,t} = \log(m_{i,t}) - \log(m_{i,t-1}) \quad (4.7)$$

Then, depending on the periodicity of the model, the volatility of the midquote return is computed as the standard deviation of the returns within that period.

- 2. Volatility quintile: this variable indicates in which quintile a company's midquote return volatility is located at a given moment.
- 3. High volatility quantile: from the midquote return volatility, a binary variable is created, where it takes value 1 if the stock i was in the top quantile of volatility at period t .
 - (a) High volatility 10%: whether the volatility of stock i at t is in the top 10%.
 - (b) High volatility 25%: whether the volatility of stock i at t is in the top 25%.
 - (c) High volatility 50%: whether the volatility of stock i at t is in the top 50%.
- 4. Crossed midquote return volatility: this variable is tantamount to the midquote return volatility, but computing the return of the crossed midquote, that is, the midquote between the crossed ask and crossed bid prices.
- 5. Crossed volatility quintile: this variable indicates in which quintile a company's crossed midquote return volatility is located at a given moment.

6. Crossed high volatility quantile: from the crossed midquote return volatility, a binary variable is created, where it takes value 1 if the stock i was in the top quantile of volatility at period t .
 - (a) Crossed high volatility 10%: whether the volatility of stock i at time t is in the top 10%.
 - (b) Crossed high volatility 25%: whether the volatility of stock i at time t is in the top 25%.
 - (c) Crossed high volatility 50%: whether the volatility of stock i at time t is in the top 50%.

4.1.3 Trading activity

1. Traded volume: the sum of the traded shares within a specific period. This is calculated by adding the quantities of shares of each market order executed in the period, considering all the orders within each second.
2. Traded value: first, the traded value is computed at each second as the product between the price and quantity of each executed market order. Then, this value is added through each second, and then through each specific period.
3. Submitted limit orders: this variable refers to the amount of limit orders submitted within each period. It is computed for both the buy orders and sell orders, and can be added by using the amount of individual orders or the amount of shares involved in each order. Then, four new variables are created.

4.1.4 Market fragmentation

1. Fragmentation: based on Haslag and Ringgenberg (2016) [18], a proxy for the level of market fragmentation is constructed using a Herfindahl-Hirschman Index (HHI) of trade volume for every asset during each period across both venues. The HHI measures the concentration of trade, ranging from 0 to 1. Then, it is subtracted from 1, thus indicating a higher fragmentation when this measures is higher.

$$\text{Fragmentation}_{i,t} = 1 - \sum_k s_{i,k,t}^2 \quad (4.8)$$

Where $s_{i,k,t}$ is the share of the total trades of stock i in venue k at period t . This variables is also constructed considering the traded value.

Chapter 5

Results

5.1 Contemporaneous models

The first group of models are the contemporaneous models. Their focus is to study the prevalent characteristics of the market when there exists an arbitrage opportunity. Then, the independent variables are measured at the same time that an arbitrage opportunity appears.

The first subsection presents the results from running the models with only one independent variable as a regressor, meanwhile the second subsection presents the results from different specifications considering several variables simultaneously.

All these models include a fixed effect on the stock and period, which could be the day in which arbitrage opportunities occurred, for the daily model, or the moment within the day, for the intraday models.

5.1.1 Univariate models

The first models were those containing liquidity variables. The result from regressing the dependant variable against the depth variables appear in Table 7.5. All the the coefficients are significant at a 95% confidence interval, and with a negative sign. This confirms the hypothesis that liquidity has a negative effect in the appearance of arbitrage opportunities.

The coefficients for the depth beyond variables are larger (in absolute value) than those of the depth at the best prices, meaning that opportunities are more likely to appear when the liquidity is decreasing beyond the bid and ask prices.

It is also worth noting that the magnitude of the coefficients for the buy side of the book are slightly larger than the ones of the sell side, corroborating this side is more informative, as it is stated by Valenzuela et al. (2015) [30].

Similar results are obtained when using the crossed-book depth variables. All of the coefficients obtained with these variables are negative and significant. Also, the coefficients for the buy side are larger in magnitude than those of the sell side, and the coefficients of the depth beyond the best prices is also larger in magnitude than those of the best prices. The results for the daily models are in Table 7.7.

The very same models were run again, but changing the periodicity of the data. All the results before mentioned are maintained, that is, a negative and significant coefficient for all variables, and the relations between their magnitude, with the buy side larger than the sell side, and the depth beyond larger than the depth at best prices, for both the within-book and crossed-book configurations.

All the results for the within-book depth models with 30-minute periodicity are shown in Table 7.6. The respective results for the crossed-book depth models are in Table 7.8.

These findings go in line with the proposed hypotheses. The depth variable aims to measure the liquidity available in the market, thus, a market is said to be liquid if there is a larger amount of shares available at each price. Since in all models the coefficients obtained were negative, when the market is less liquid it is more likely to observe arbitrage opportunities.

Another approach to measure liquidity is through the bid-ask spread. If the spread is wider the market would be less liquid. The results for this variables are less conclusive than those from the depth variables.

As Table 7.9 shows, in the daily-data configuration, the coefficients are negative and significant for both the within-book and crossed-book spreads, suggesting that arbitrage opportunities are more likely to occur when the market is more liquid.

However, when changing the periodicity of the data, results differ. In the case of the within-book spread, the coefficient becomes positive with intraday data, it is not significant with 120-minute intervals but it is with shorter intervals. On the other side, the crossed-book spread maintains its negative sign, but becomes not significant in the 15-minute interval for the standardized version.

The next variables are relative liquidity and slope, whose results can be found in Table 7.11 and 7.13 for the daily data. In the first case, robust results are obtained for the buy and the sell side. The coefficients for the relative liquidity are positive and significant with all specifications. When considering the relative liquidity of both sides as a whole, the coefficient obtained is again positive and significant for all the data configurations.

The relative liquidity variable indicates how concentrated is liquidity around the best prices. A larger value means the number of shares available around the best prices is larger than that of the prices beyond, or the best prices are nearer one from another.

This means that, for both the sell and the buy side, when there is a larger concentration of depth around the best prices, it is more likely to find arbitrage opportunities in that instant.

The slope measures how deep or thin is the book at the best prices. When the slope is larger, the book is thinner, reflecting a larger difference between the best and second-best

prices relative to the depth offered at the best price. In other words, when the slope is bigger a relatively large difference in price would be observed once the liquidity at the best price is consumed.

This time, the results are robust only for the crossed-book variables, with a positive and significant coefficient. Meanwhile, for the within-book variables, the results are not significant for the sell side in all configurations, and they are not for any within-book slope variable with daily data. Thus, when observing intraday intervals, arbitrage opportunities occur the most when the book is steeper in a crossed-book perspective.

The next group of variables are those related with uncertainty. The within-book results are presented in Table 7.15 for the daily configuration, and the subsequent tables show the results with intraday configurations.

The first variable is the midquote logarithmic return volatility, calculated as the standard deviation of the logarithmic return of the midquote during a determined time interval. This is directly related with uncertainty, with a larger value of the volatility indicating a larger uncertainty in that period. The coefficient is positive and significant for all periodicity, corroborating the hypotheses that arbitrage opportunities thrive when there is more uncertainty.

The following set of models aims to determine if those stocks with the largest volatility exhibit a larger probability of opportunities. Volatility quintile is a categorical variable which identifies in which quintile the stock belongs when they are ordered by the estimated volatility, with the first level representing the bottom 20%, and the fifth level the top 20%. The third level is omitted, so the results are relative to this particular level.

As it could be expected, being in the two bottom levels diminishes the likelihood of arbitrage, with the magnitude of the first quintile being even larger in magnitude than that of the second one. On the other side, the top quintile exhibits a positive coefficient, as well as the fourth quintile but with a lower magnitude. All these results are significant, and remain when reducing the time intervals of the data.

Then, another set of three models are specified, each with a dummy variable which indicates if the stock's volatility is among a top quantile during a specific interval. These quantiles are 10%, 25% and 50%. The resultant coefficients are positive and significant for all the configurations, going in line with the hypotheses.

The results for the crossed-book variables, in Table 7.17, is similar, maintaining all the signs mentioned above for the within-book version. One observation worth mentioning is that for the crossed high volatility quantile variables, the magnitude increases when the variable is constructed over a higher quantile, meaning that, when observing the crossed-book data, the highest the volatility is for certain stock, the higher the probability of witnessing arbitrage.

The last group of tables, from 7.19 onward, presents the results of the variables measuring trading activity and market fragmentation.

All the trading activity variables show a positive and significant coefficient along all differ-

ent time-frame configurations. This result validates the hypothesis that when trading activity is larger, it is more likely for opportunities to occur because of the illiquidity that may occur as a consequence of the executing orders. Between these variables, the submitted amount of limit buy and sell orders show a coefficient larger in magnitude than the ones of traded volume and traded value, being this proxy of trading activity more informative.

Finally, fragmentation measured both by traded volume and traded value show identical significant results in all models. However, mixed results are observed: although the obtained coefficients are positive with daily data, they are negative with intraday data, which is counter intuitive since it was expected that, with larger fragmentation, i.e. the traded volume and value were distributed among more exchanges, there would be more opportunities observed.

5.1.2 Multivariate models

The multivariate models consider a mix of variables from all the studied dimensions. The purpose of these specifications is to study if the relations found in the univariate models are preserved when all market dimensions are observed at once (see from Table 7.21 onward).

All the models include the crossed bid-ask spread, and in all cases the resulting coefficient is negative and significant. This makes sense for the same explanation given above, since the crossed bid-ask spread has to be negative in order to allow the appearance of an opportunity.

The crossed midquote return volatility is also included in all models, and the results are consistent with all data configurations, obtaining a significant positive result, validating the uncertainty hypothesis.

The liquidity variables present different results. All the models consider the depth at the best crossed prices, but there are two variations with the remaining variables: two models include the crossed relative liquidity and the other two include the depth beyond the best crossed prices.

Relative liquidity is always positive and significant, meaning that the more depth concentration there is around the best prices the more likely it is for arbitrage to arise.

The depth at the best prices is not significant with daily data. With intraday data, only the depth at the crossed bid price is always significant, and negative as it was hypothesized. However, these variables become always insignificant when in presence of the depth beyond the best prices, which, again, only the bid size remains always significant and negative, once again validating that the buy side is more informative.

Market fragmentation presents the same striking result from the univariate models, with positive and significant coefficients in the daily models, and negative and significant in the intraday models.

The trading activity variables show a positive and significant result in all the models and all the data configurations, validating again the hypothesis.

5.2 Predictive models

The next group of models are the predictive models. In these cases, the dependent variable is modeled against one-period-lagged independent variables describing the characteristics of the market. This way, it will be possible to conclude if there is a specific feature that prevails just before an arbitrage opportunity appears.

The first subsection presents the results from univariate models, that is, models with only one regressor, meanwhile the second subsection presents the results from multivariate model specifications. Once again, all the models include a fixed effect on the stock and period.

5.2.1 Univariate models

As it was seen in the previous section, the depth variables also exhibit a negative and significant relationship with the appearance of arbitrage opportunities. In Table 7.23 the results from modelling with daily data can be observed. This relationship is maintained when using shorter periods (see Table 7.24). This indicates that, when there is a lower depth in prices, then it is more likely that in the consecutive period there will be at least one arbitrage opportunity.

The same phenomenon occurs with the crossed-book depth variables, where they all get a negative and 99%-significant coefficient with all data configurations (see Tables 7.25 and 7.26). Both in the within-book and crossed-book models, the coefficients for the depth beyond the best prices variables always get a larger magnitude than those of the depth at the best prices. This indicates that a lower depth in the second and third best prices has a greater effect in the likelihood of opportunities appearance in the next period.

Next, the spread variables show a counter intuitive result: they have a negative (and significant) relationship with the dependent variable. Thinner spreads are thought to represent a more liquid market, since this would reflect a greater agreement on the price of the asset. Then, this result would indicate that, when the market is more liquid, it is more likely that an arbitrage opportunities will occur in the next period.

However, for an arbitrage opportunity to occur, it is necessary that the crossed bid-ask spread is negative, so there is a profit from buying and selling simultaneously in both markets. Then, the thinner the spread is, the more likely it is for prices to cross and admit a profitable opportunity. From this point of view, it makes sense the coefficients are negative and significant (see Tables 7.27 and 7.28).

For the relative liquidity, which measures how concentrated is the depth around the best prices, the results are positive and significant coefficients, and maintain these characteristics when using shorter time periods. Then, this means that arbitrage opportunities are more likely to happen right after the depth is more concentrated around the best prices. One explanation is that, when there is more depth concentration, it is more easy for mispricing to occur after a liquidity demand shock arrives.

On the other side, slope measures how distant are the best and second-best prices relative to the best-price depth. That is, with a higher slope, the difference in the best prices is higher. Thus, a higher slope is directly related with a higher relative liquidity. However, almost none of the results for this variable are significant, and they present varying coefficient signs when changing the data periodicity, leaving no conclusive result.

The next variables are the ones thought as uncertainty measures. The main uncertainty variable is the midquote logarithmic return volatility, measured as the standard deviation of the second-to-second logarithmic return of the midquote during the corresponding period. The next ones are dummy variables indicating if a particular stock in a particular period is among a specific quantile.

As it appears on Table 7.33, there exists a positive correlation between the uncertainty and arbitrage occurrence. The midquote return volatility shows a positive and significant coefficient for all data specifications.

The base level of the volatility quintile variable is the middle quintile. Then, corroborating the hypothesis, it is observed that the two bottom quintiles present a negative effect, meanwhile the opposite happens for the top two quintiles. It is interesting that, in both cases, the extreme quintiles show a greater and always significant coefficient, unlike the second and fourth quintiles which, although they maintain their sign, they are not significant for all periodicities.

In line with the previous observations, the variables indicating whether the stock is among the top 10%, 25% or 50% in terms of volatility, are also positive and significant to certain degree.

Then, looking at the crossed-book versions of the uncertainty measures, the same results are obtained, however, with different levels of significance. In particular, with daily data (Table 7.35), the fifth quintile of volatility becomes not significant.

It is worth noting that, with except of the 120-minute data models, the magnitude of the coefficient of the top quantile variables are greater the highest the quantile is. Then, added to the previous results, there is evidence that the highest (lowest) the volatility is in a certain period, relative to the other assets, the more (less) impact there is the likelihood of next-period arbitrage.

The trading activity variables, presented in Tables 7.37 and 7.38, show supportive results for the hypothesis. In all data configurations, the variables for traded volume, traded value and submitted limit buy/sell orders show a positive and significant impact in the likelihood of opportunities, probably because of the more liquidity being consumed with this activity.

An interesting result is that the coefficients of submitted limit buy/sell orders are larger than those of the aggregate traded volume and value. This result could be related to the competition among traders, since it is more likely that more competitiveness/aggressiveness would be reflected in more orders submitted rather than a larger number of shares being effectively traded.

Finally, the results from fragmentation variables are shown next to the trading activity variables. For the daily model, the results are positive and significant. Thus, a higher fragmentation would imply a greater probability of arbitrage in the next day, making sense with the proposed hypothesis.

However, with intraday data, the coefficient becomes negative and significant at a 99% confidence level. An alternative interpretation of this result is that, the less fragmented markets are, the more concentration of trading activity there is in one of them. Thus, because one market is less active than the other, it is more likely for mispricing to occur, admitting the appearance of arbitrage opportunities.

5.2.2 Multivariate models

In the first place, the same specifications used in the contemporaneous multivariate models are replicated with a predictive approach. That is, variables for all market dimensions are mixed (see from Table 7.39 onward).

The results are consistent with the contemporaneous models. The crossed bid-ask spread is negative and significant for all the models.

The crossed midquote return volatility is always positive and significant, with exception of the models in which trading activity variables are measured according to the traded value and with 120-minute data configuration. Besides these specific cases, the uncertainty variable shows consistent results.

For the liquidity variables, when using daily data the only significant coefficients are those for the crossed relative liquidity, in both the buy and sell sides, being consistent with previous results. The crossed depth variables are insignificant in all models, except for the depth beyond the crossed ask price in one of the models.

With intraday data, the depth at the crossed bid price is always significant and negative when using the relative liquidity variable. But, it becomes insignificant in all the models with the depth beyond the best prices. Instead, the depth beyond the crossed bid price is always negative and significant in the intraday models. The crossed depth ask in the first case is significant with 60, 30 and 15-minute data, as well as the depth beyond crossed ask in the second case.

The next group of models include an interaction between the fragmentation variable and one of the uncertainty variable, plus control variables representing liquidity and volume. These models will help to understand whether fragmentation has a differentiated predictive power for arbitrage opportunities on different uncertainty conditions. The coefficients of the control variables are not presented.

In the first place, the models were run with only the fragmentation and uncertainty variables, including the interaction term. With within-book data, the coefficients are similar to the ones in the univariate cases. Fragmentation shows a positive coefficient when using

daily data (see Table 7.41), but with intraday data, the coefficient becomes negative. The uncertainty variables show a positive and significant coefficient on all period configurations.

The interaction term in these cases shows no significant coefficient except on the 15-minute data models. When using the shortest period length, all the interaction terms (except one) become significant. The interaction between fragmentation and the volatility quantile variables is negative, thus indicating that stocks with lower fragmentation and higher midquote return volatility are more likely to exhibit arbitrage opportunities in the next 15-minute period.

The crossed-book data models present similar results. Fragmentation shows a positive coefficient with daily data (see Table 7.43) and negative with intraday data. Uncertainty variables show positive coefficients in all period configurations.

Although the interaction terms present some significant coefficients with daily a 120-minute data, it is in the 15-minute data that at least one term presents a significant coefficient in each model.

The interpretation of this results is that, when there is less fragmentation, the trading activity is more concentrated in one of the two analyzed markets, thus, when this happens and there is a higher volatility, i.e. there is higher activity, it is more likely that mispricing occurs in the other market, the one with less activity, allowing the appearance of arbitrage opportunities.

Next, control variables were added to test the robustness of these results. These variables include the remaining book dimensions, which are trading activity and liquidity. For trading activity, the selected variable is traded volume, meanwhile for liquidity, the selected variables are relative liquidity and depth beyond. Also, the midquote return volatility is included. The corresponding variables were used in both their within-book and crossed-book versions.

The first model includes all the control variables mentioned above. In the within-book version, most of the interaction terms become not significant. With daily data (see Table 7.45) only one of the interaction terms become significant and most of the volatility variables become not significant. With the intraday configurations, the volatility variables recover their significance, however the interaction terms don't, being only significant for the volatility quintiles interactions with 15-minute data. A similar result is obtained with crossed-book data (see Tables 7.47 and 7.48). But in this case, the interaction terms don't become significant with the shortest period length.

The following models were run considering only one control variable at a time. The results show the same trend as before, with fragmentation and uncertainty variables maintaining their sign as in the univariate cases, and mostly maintaining their significance, meanwhile the interaction terms are mostly not significant except for the 15-minute data models. Thus, the results presented without control variables are robust when using only the shortest periods.

Chapter 6

Conclusion

Technological advances have transformed financial markets in the last decades, which have seen major improvements in the speed at which assets are traded. Moreover, several new exchanges have emerged, leading to a fragmented scenario in which trading is distributed across numerous venues. Given this context, the focus of this investigation was to study the market conditions in which arbitrage opportunities are more likely to occur. For this purpose, the Limit Order Book and Trades and Orders Book of two exchanges for 39 European stocks were analyzed.

Liquidity, one of the most important features of financial markets, has shown a robust correlation with the appearance of arbitrage opportunities, when it is measure as the available depth. Opportunities are more likely to arise when markets are less liquid, that is, when there is less available depth, particularly at those prices beyond the best ones, with more informativeness from the buy side of the book.

Bid-ask spreads show mixed results. Contemporaneously, arbitrage opportunities appear the most when within-market spreads are wider, that is, when markets are locally less liquid. However, the crossed-book spread shows a negative relation with the studied phenomenon, although with different levels of significance. In the predictive models, both the within-book and crossed-book spreads show a negative coefficient. At first, this result seems to contradict the hypothesis, but it makes sense since spreads must be thinner and become negative to allow the appearance of opportunities.

The relative liquidity, which measures how liquidity is distributed through the book, shows a positive relation, both predictive and contemporaneously. That is, when liquidity is more concentrated around the best price or when the best prices are more distanced between themselves, arbitrage is more likely to appear. In this instances, liquidity is decreasing once the depth at the best price is consumed and prices suffer major variations, allowing for moments of mispricing.

Uncertainty has also shown to be a robust market condition that relates to and predicts the appearance of opportunities, since these are more likely to happen when markets are in presence of more uncertainty. The main proxy for measuring this feature is the volatility of

the logarithmic return of the midquote. This variable presents a positive coefficient in the models, corroborating the hypothesis.

Moreover, dummy variables distinguishing the position of a certain stock's volatility at a certain moment show supporting results. When a stock's volatility is among the top quintile it is more likely for the stock to exhibit arbitrage, meanwhile it is less likely when it belongs to the bottom quintile. Similar results are obtained when distinguishing just for the top 10th, 25th and 50th percentiles.

Trading activity shows a significant positive relation, in line with the proposed hypothesis, measured as traded volume, traded value and the number of arriving limit orders. When there is more trading activity, more liquidity is being consumed, which allows for moments in which stocks become mispriced at one or both venues, permitting arbitrage activity.

Market fragmentation shows mixed results. When using daily-aggregated data, a positive relation is obtained, which is consistent with the hypothesis that in presence of more fragmentation (i.e., trading activity is more distributed across both exchanges) opportunities will appear with a greater probability. However, with intraday data, the results reverses and a significant negative relation is evidenced.

One possible explanation is that, when using aggregate data it is possible to observe the effect of fragmentation under greater demand shocks, which, according to the reviewed literature, provokes beneficial arbitrage opportunities, meanwhile with intraday data, one observes the harmful effect of fragmentation that increases costs and lead traders to charge wider spreads, which leads to less arbitrage opportunities.

Bibliography

- [1] ANGEL, J. J., HARRIS, L. E., AND SPATT, C. S. Equity trading in the 21st century: An update. *The Quarterly Journal of Finance* 5, 01 (2015), 1550002.
- [2] ARNUK, S., AND SALUZZI, J. *Broken markets: how high frequency trading and predatory practices on Wall Street are destroying investor confidence and your portfolio*. FT Press, 2012.
- [3] BENOS, E., BRUGLER, J., HJALMARSSON, E., AND ZIKES, F. Interactions among high-frequency traders.
- [4] BERNALES, A. Algorithmic and high frequency trading in dynamic limit order markets. *Available at SSRN 2352409* (2017).
- [5] BODIE, Z., KANE, A., AND MARCUS, A. J. *Investments, eight edition*. 2009.
- [6] BUDISH, E., CRAMTON, P., AND SHIM, J. The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics* 130, 4 (2015), 1547–1621.
- [7] CARDELLA, L., HAO, J., KALCHEVA, I., AND MA, Y.-Y. Computerization of the equity, foreign exchange, derivatives, and fixed-income markets. *Financial Review* 49, 2 (2014), 231–243.
- [8] COPELAND, T. E., AND GALAI, D. Information effects on the bid-ask spread. *the Journal of Finance* 38, 5 (1983), 1457–1469.
- [9] EASLEY, D., LÓPEZ DE PRADO, M. M., AND O’HARA, M. Flow toxicity and liquidity in a high-frequency world. *The Review of Financial Studies* 25, 5 (2012), 1457–1493.
- [10] FOUCAULT, T., KOZHAN, R., AND THAM, W. W. Toxic arbitrage. *The Review of Financial Studies* 30, 4 (2016), 1053–1094.
- [11] FOUCAULT, T., AND MENKVELD, A. J. Competition for order flow and smart order routing systems. *The Journal of Finance* 63, 1 (2008), 119–158.
- [12] GAGNON, L., AND KAROLYI, G. A. Multi-market trading and arbitrage. *Journal of Financial Economics* 97, 1 (2010), 53–80.

- [13] GOETTLER, R. L., PARLOUR, C. A., AND RAJAN, U. Informed traders and limit order markets. *Journal of Financial Economics* 93, 1 (2009), 67–87.
- [14] GROMB, D., AND VAYANOS, D. Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of financial Economics* 66, 2-3 (2002), 361–407.
- [15] GROMB, D., AND VAYANOS, D. Limits of arbitrage. *Annu. Rev. Financ. Econ.* 2, 1 (2010), 251–275.
- [16] HASBROUCK, J. *Empirical market microstructure: The institutions, economics, and econometrics of securities trading*. Oxford University Press, 2007.
- [17] HASBROUCK, J. High-frequency quoting: Short-term volatility in bids and offers. *Journal of Financial and Quantitative Analysis* 53, 2 (2018), 613–641.
- [18] HASLAG, P. H., AND RINGGENBERG, M. The causal impact of market fragmentation on liquidity. In *Fourth Annual Conference on Financial Market Regulation* (2016).
- [19] HENDERSHOTT, T., JONES, C. M., AND MENKVELD, A. J. Does algorithmic trading improve liquidity? *The Journal of Finance* 66, 1 (2011), 1–33.
- [20] JOHNSON, T. C. Volume, liquidity, and liquidity risk. *Journal of Financial Economics* 87, 2 (2008), 388–417.
- [21] JOVANOVIĆ, B., AND MENKVELD, A. J. Middlemen in limit order markets. *Available at SSRN 1624329* (2016).
- [22] MARSHALL, B. R., NGUYEN, N. H., AND VISALTANACHOTI, N. Etf arbitrage: Intraday evidence. *Journal of Banking & Finance* 37, 9 (2013), 3486–3498.
- [23] MENKVELD, A. J. Splitting orders in overlapping markets: A study of cross-listed stocks. *Journal of Financial Intermediation* 17, 2 (2008), 145–174.
- [24] MITCHELL, M., PULVINO, T., AND STAFFORD, E. Limited arbitrage in equity markets. *The Journal of Finance* 57, 2 (2002), 551–584.
- [25] O’HARA, M., AND YE, M. Is market fragmentation harming market quality? *Journal of Financial Economics* 100, 3 (2011), 459–474.
- [26] PAGNOTTA, E. S., AND PHILIPPON, T. Competing on speed. *Econometrica* 86, 3 (2018), 1067–1115.
- [27] RÖSCH, D. The impact of arbitrage on market liquidity. In *Paris December 2014 Finance Meeting EUROFIDAI-AFFI Paper* (2017).
- [28] SCHULTZ, P., AND SHIVE, S. Mispricing of dual-class shares: Profit opportunities, arbitrage, and trading. *Journal of Financial Economics* 98, 3 (2010), 524–549.
- [29] STIGLITZ, J. E. Tapping the brakes: Are less active markets safer and better for the economy?

- [30] VALENZUELA, M., ZER, I., FRYZLEWICZ, P., AND RHEINLÄNDER, T. Relative liquidity and future volatility. *Journal of Financial Markets* 24 (2015), 25–48.

Chapter 7

Appendix

1. Daily average of trading statistics for the AEX stocks traded at Euronext

	Average volume	Average value	Average buy orders	Average sell orders
AIRF.PA	3023890.9	25930071	36541.05	36325.03
AEGN.AS	5363825.3	34128449	59841.46	59225.98
AHLN.AS	2229597.4	29828356	33344.18	32971.37
AKZO.AS	460077.3	25183218	45899.05	45565.96
ASML.AS	1043142.0	72374713	74859.90	74369.17
BOSN.AS	257308.8	10387503	31911.09	32069.19
COR.AS	253734.9	9244825	24180.70	23867.22
DSMN.AS	593720.4	29698349	45673.01	45280.54
FUGRc.AS	925135.2	21375560	34421.53	34338.86
HEIN.AS	603729.5	32106480	44536.51	44427.73
ING.AS	14265233.0	149829905	96865.17	96476.30
KPN.AS	11149168.2	28522810	34907.77	34759.68
ISPA.AS	7535794.5	82417257	80863.82	80565.37
PHG.AS	2471036.5	59332809	72337.23	72189.02
PTNL.AS	3154655.9	11121798	24939.38	24885.94
RAND.AS	507550.7	20252576	43877.29	43211.09
RDSa.AS	4753902.3	132502728	101435.76	100955.98
ELSN.AS	1378674.9	22949870	39469.53	39246.17
SBMO.AS	1649816.7	18921460	27432.28	27298.43
TNTE.AS	1227134.9	7261592	27878.21	27518.24
UNc.AS	3297732.1	100331152	95935.58	95980.23
WLSNc.AS	460691.2	9768417	27872.83	27811.06

Table 7.1: Daily average of trading statistics for the AEX stocks traded at Euronext.

2. Daily average of trading statistics for the AEX stocks traded at BATS Chi-X

	Average volume	Average value	Average buy orders	Average sell orders
AFp.CHI	1014274.4	8755462	15926.54	15525.40
AGNa.CHI	2169910.8	13771550	21774.82	21573.71
AHa.CHI	836350.8	11181301	9358.92	9547.39
AKZAa.CHI	243741.7	13275043	15103.95	14933.27
ASMLa.CHI	394800.7	27573976	25551.15	25051.42
BOKAa.CHI	105269.1	4296381	9107.94	9594.38
CORAA.CHI	63083.6	2271154	5326.84	5633.43
DSMa.CHI	223688.2	11204062	15095.49	15234.65
FURa.CHI	230510.5	5767285	10146.56	10350.85
HEIAa.CHI	271675.6	14636122	14924.91	14899.98
INGAa.CHI	4125666.0	43361538	31163.17	31119.81
KPNa.CHI	5184837.5	13291838	13414.51	13557.48
MTa.CHI	2331697.4	25512283	30448.72	30446.45
PHIAa.CHI	1094396.5	26135576	29857.28	30156.79
PNLa.CHI	878431.2	3083634	7175.10	7208.04
RANDa.CHI	204849.2	8200068	15163.26	15103.46
RDSAa.CHI	1416433.2	39324632	39637.44	39892.06
RENa.CHI	537430.2	8965254	11946.75	12149.81
SBMOa.CHI	468037.7	5309393	8192.89	8227.64
TNTEa.CHI	439493.2	2525930	6766.53	6766.21
UNAA.CHI	1305511.2	39673891	45723.00	46024.64
WKLla.CHI	241050.1	5134543	7933.90	8295.78

Table 7.2: Daily average of trading statistics for the AEX stocks traded at BATS Chi-X.

3. Daily average of trading statistics for the BEL20 stocks traded at Euronext

	Average volume	Average value	Average buy orders	Average sell orders
ABI.BR	1025046.79	83856278.0	76948.00	76878.83
ACKB.BR	25901.17	2368913.2	6304.45	6093.96
AGES.BR	340858.78	9818319.2	10145.61	10096.01
BEFB.BR	17289.33	973089.1	3015.71	2947.95
BEKB.BR	60540.35	1635104.4	8584.79	9165.95
BCOM.BR	450058.05	11462595.4	22822.76	23545.07
COFB.BR	21692.01	1938200.0	4652.88	4429.71
COLR.BR	97490.35	3701702.9	7330.80	7475.35
DELB.BR	293836.75	15100673.3	33785.62	33854.03
IETB.BR	30508.81	1004472.7	4633.81	4604.58
ELI.BR	107834.65	4038041.2	4047.60	4243.79
GBLB.BR	54842.77	3930239.5	6475.54	6464.71
KBC.BR	765368.78	32872554.7	66151.98	66366.01
SOLB.BR	133433.90	15237913.9	23696.07	23054.17
TNET.BR	60712.94	2644815.5	6765.20	6843.94
UCB.BR	170907.44	10557785.5	13056.57	13074.58
UMI.BR	319256.79	10873300.9	26010.20	25732.81

Table 7.3: Daily average of trading statistics for the BEL20 stocks traded at Euronext.

4. Daily average of trading statistics for the BEL20 stocks traded at BATS Chi-X

	Average volume	Average value	Average buy orders	Average sell orders
ABlb.CHI	482997.71	39535998.1	36367.76	36450.17
ACKBb.CHI	5778.76	531575.3	6462.33	6868.63
AGSb.CHI	150976.39	4325957.2	8946.12	9018.72
BEFBb.CHI	3942.35	224690.0	2918.77	2710.20
BEKBb.CHI	15325.94	410128.5	6990.17	7629.16
BELGb.CHI	191784.88	4975276.6	6926.00	7263.12
COFBb.CHI	7458.67	668597.8	4645.10	4199.35
COLRb.CHI	57870.48	2196185.5	6189.74	6605.48
DELBb.CHI	145931.48	7519848.8	11546.46	11755.50
DIEb.CHI	8215.62	262207.6	3738.35	3780.45
ELIb.CHI	5970.29	224700.6	3673.73	3854.95
GBLBb.CHI	25034.34	1797118.5	5999.81	6421.57
KBCb.CHI	340596.00	14618963.2	26933.25	27466.33
SOLBb.CHI	61818.97	7050426.3	10692.50	10432.17
TNETb.CHI	29458.25	1291041.7	6054.45	6343.78
UCBb.CHI	80545.42	5042177.2	6043.53	6238.68
UMIb.CHI	143020.02	4860258.3	7815.43	7802.15

Table 7.4: Daily average of trading statistics for the BEL20 stocks traded at BATS Chi-X.

5. Depth as only regressor, univariate contemporaneous models with daily data

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	1.9545*** (0.3931)	1.9139*** (0.3928)	1.8773*** (0.3936)	1.829*** (0.3921)	1.7328*** (0.3918)	1.746*** (0.3918)
Depth ask	-0.4887*** (0.0925)					
Depth bid		-0.5025*** (0.0969)				
Depth best prices			-0.6045*** (0.1035)			
Depth beyond ask				-0.8096*** (0.0883)		
Depth beyond bid					-0.8569*** (0.0857)	
Depth beyond						-0.8993*** (0.0901)
Marginal effects						
Depth ask	-0.0889***					
Depth bid		-0.0914***				
Depth best prices			-0.1098***			
Depth beyond ask				-0.1461***		
Depth beyond bid					-0.1543***	
Depth beyond						-0.1619***

Table 7.5: Depth as only regressor, univariate contemporaneous models with daily data.

6. Depth as only regressor, univariate contemporaneous models with 30-minute data

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	-0.1733*** (0.0572)	-0.2104*** (0.0575)	-0.2604*** (0.0583)	-0.3425*** (0.0567)	-0.3684*** (0.0566)	-0.3836*** (0.0567)
Depth ask	-0.4396*** (0.0391)					
Depth bid		-0.489*** (0.0394)				
Depth best prices			-0.5574*** (0.0416)			
Depth beyond ask				-0.7058*** (0.0373)		
Depth beyond bid					-0.7358*** (0.0369)	
Depth beyond						-0.7536*** (0.037)
Marginal effects						
Depth ask	-0.0195***					
Depth bid		-0.0217***				
Depth best prices			-0.0247***			
Depth beyond ask				-0.0312***		
Depth beyond bid					-0.0325***	
Depth beyond						-0.0333***

Table 7.6: Depth as only regressor, univariate contemporaneous models with 30-minute data.

7. Depth crossed as only regressor, univariate contemporaneous models with daily data

Variable	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Constant	1.9389*** (0.3937)	1.9079*** (0.3918)	1.8528*** (0.3936)	1.8479*** (0.3915)	1.7605*** (0.3918)	1.762*** (0.3917)
Depth crossed ask	-0.4962*** (0.0831)					
Depth crossed bid		-0.5253*** (0.0855)				
Depth best crossed prices			-0.6428*** (0.0936)			
Depth beyond crossed ask				-0.7218*** (0.0898)		
Depth beyond crossed bid					-0.7553*** (0.0873)	
Depth beyond crossed						-0.8138*** (0.0925)
Marginal effects						
Depth crossed ask	-0.0901***					
Depth crossed bid		-0.0954***				
Depth best crossed prices			-0.1166***			
Depth beyond crossed ask				-0.1306***		
Depth beyond crossed bid					-0.1365***	
Depth beyond crossed						-0.1471***

Table 7.7: Depth crossed as only regressor, univariate contemporaneous models with daily data.

8. Depth crossed as only regressor, univariate contemporaneous models with 30-minute data

Variable	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Constant	-0.1754*** (0.0564)	-0.2161*** (0.0566)	-0.2787*** (0.0574)	-0.2689*** (0.0567)	-0.2984*** (0.0567)	-0.3218*** (0.0569)
Depth crossed ask	-0.4654*** (0.0377)					
Depth crossed bid		-0.5111*** (0.0372)				
Depth best crossed prices			-0.5986*** (0.0397)			
Depth beyond crossed ask				-0.5982*** (0.0375)		
Depth beyond crossed bid					-0.6317*** (0.037)	
Depth beyond crossed						-0.6626*** (0.0375)
Marginal effects						
Depth crossed ask	-0.0206***					
Depth crossed bid		-0.0227***				
Depth best crossed prices			-0.0265***			
Depth beyond crossed ask				-0.0265***		
Depth beyond crossed bid					-0.028***	
Depth beyond crossed						-0.0293***

Table 7.8: Depth crossed as only regressor, univariate contemporaneous models with 30-minute data.

9. Bid-ask spread as only regressor, univariate contemporaneous models with daily data

Variable	Model 13	Model 14	Model 15	Model 16
Constant	2.1672*** (0.3899)	1.8361*** (0.3951)	2.1605*** (0.3912)	1.5111*** (0.4002)
Bid-ask spread	-0.3322*** (0.0923)			
Bid-ask spread standardized		-0.4203*** (0.0655)		
Crossed bid-ask spread			-0.6547*** (0.1119)	
Crossed bid-ask spread standardized				-0.6088*** (0.0654)
Marginal effects				
Bid-ask spread	-0.0605***			
Bid-ask spread standardized		-0.0763***		
Crossed bid-ask spread			-0.119***	
Crossed bid-ask spread standardized				-0.1099***

Table 7.9: Bid-ask spread as only regressor, univariate contemporaneous models with daily data.

10. Bid-ask spread as only regressor, univariate contemporaneous models with 30-minute data

Variable	Model 13	Model 14	Model 15	Model 16
Constant	0.0485 (0.0531)	0.0888* (0.0524)	0.1042* (0.054)	0.0961* (0.0524)
Bid-ask spread	0.0789*** (0.0138)			
Bid-ask spread standardized		0.0414*** (0.012)		
Crossed bid-ask spread			-0.0153 (0.0232)	
Crossed bid-ask spread standardized				-0.0477** (0.0195)
Marginal effects				
Bid-ask spread	0.0035***			
Bid-ask spread standardized		0.0018***		
Crossed bid-ask spread			-7e-04	
Crossed bid-ask spread standardized				-0.0021**

Table 7.10: Bid-ask spread as only regressor, univariate contemporaneous models with 30-minute data.

11. Relative liquidity as only regressor, univariate contemporaneous models with daily data

Variable	Model 17	Model 18	Model 19	Model 20	Model 21	Model 22
Constant	3.0758*** (0.4008)	3.148*** (0.3972)	3.2629*** (0.4004)	3.044*** (0.4035)	3.098*** (0.3996)	3.3542*** (0.4056)
Relative liquidity ask	0.7367*** (0.0482)					
Relative liquidity bid		0.7846*** (0.0479)				
Relative liquidity			0.8687*** (0.0515)			
Crossed relative liquidity ask				0.5036*** (0.0423)		
Crossed relative liquidity bid					0.6021*** (0.0425)	
Crossed relative liquidity						0.697*** (0.0468)
Marginal effects						
Relative liquidity ask	0.1307***					
Relative liquidity bid		0.1386***				
Relative liquidity			0.1531***			
Crossed relative liquidity ask				0.0904***		
Crossed relative liquidity bid					0.1073***	
Crossed relative liquidity						0.1239***

Table 7.11: Relative liquidity as only regressor, univariate contemporaneous models with daily data.

12. Relative liquidity as only regressor, univariate contemporaneous models with 30-minute data

Variable	Model 17	Model 18	Model 19	Model 20	Model 21	Model 22
Constant	0.2657*** (0.054)	0.2636*** (0.0539)	0.312*** (0.0544)	0.2591*** (0.0549)	0.2602*** (0.0549)	0.3356*** (0.056)
Relative liquidity ask	0.4078*** (0.0185)					
Relative liquidity bid		0.4109*** (0.0185)				
Relative liquidity			0.4946*** (0.0199)			
Crossed relative liquidity ask				0.1825*** (0.0166)		
Crossed relative liquidity bid					0.1833*** (0.0167)	
Crossed relative liquidity						0.2378*** (0.0178)
Marginal effects						
Relative liquidity ask	0.018***					
Relative liquidity bid		0.0181***				
Relative liquidity			0.0217***			
Crossed relative liquidity ask				0.0081***		
Crossed relative liquidity bid					0.0081***	
Crossed relative liquidity						0.0105***

Table 7.12: Relative liquidity as only regressor, univariate contemporaneous models with 30-minute data.

13. Slope as only regressor, univariate contemporaneous models with daily data

Variable	Model 23	Model 24	Model 25	Model 26	Model 27	Model 28
Constant	2.1836*** (0.389)	2.1886*** (0.3891)	2.189*** (0.389)	2.1923*** (0.3888)	2.1815*** (0.3894)	2.185*** (0.389)
Slope ask	-0.007 (0.035)					
Slope bid		0.0253 (0.0247)				
Slope			0.024 (0.0277)			
Crossed slope ask				0.1223*** (0.0393)		
Crossed slope bid					-0.0992** (0.0403)	
Crossed slope						0.0098 (0.0473)
Marginal effects						
Slope ask	-0.0013					
Slope bid		0.0046				
Slope			0.0044			
Crossed slope ask				0.0223***		
Crossed slope bid					-0.0181**	
Crossed slope						0.0018

Table 7.13: Slope as only regressor, univariate contemporaneous models with daily data.

14. Slope as only regressor, univariate contemporaneous models with 30-minute data

Variable	Model 23	Model 24	Model 25	Model 26	Model 27	Model 28
Constant	0.0939* (0.0524)	0.0908* (0.0524)	0.0905* (0.0524)	0.0932* (0.0524)	0.093* (0.0524)	0.0907* (0.0524)
Slope ask	0.0084 (0.0059)					
Slope bid		0.0383*** (0.0142)				
Slope			0.0247*** (0.0094)			
Crossed slope ask				0.0248** (0.0107)		
Crossed slope bid					0.0299*** (0.0101)	
Crossed slope						0.0416*** (0.0113)
Marginal effects						
Slope ask	4e-04					
Slope bid		0.0017***				
Slope			0.0011***			
Crossed slope ask				0.0011**		
Crossed slope bid					0.0013***	
Crossed slope						0.0018***

Table 7.14: Slope as only regressor, univariate contemporaneous models with 30-minute data.

15. Volatility variables as only regressor, univariate contemporaneous models with daily data

Variable	Model 29	Model 30	Model 31	Model 32	Model 33
Constant	2.1769*** (0.3898)	2.0841*** (0.3954)	2.1363*** (0.3897)	2.0919*** (0.3914)	1.8318*** (0.3929)
Midquote return volatility	0.2394*** (0.0325)				
Volatility quintile 1		-0.5656*** (0.0888)			
Volatility quintile 2		-0.2144*** (0.0792)			
Volatility quintile 4		0.2386*** (0.0792)			
Volatility quintile 5		0.4171*** (0.0923)			
High volatility 10%			0.4095*** (0.0909)		
High volatility 25%				0.4642*** (0.072)	
High volatility 50%					0.5076*** (0.0608)
Marginal effects					
Midquote return volatility	0.0434***				
Volatility quintile 1		-0.0999***			
Volatility quintile 2		-0.0387***			
Volatility quintile 4		0.0436***			
Volatility quintile 5		0.0763***			
High volatility 10%			0.0752***		
High volatility 25%				0.0849***	
High volatility 50%					0.0922***

Table 7.15: Volatility variables as only regressor, univariate contemporaneous models with daily data.

16. Volatility variables as only regressor, univariate contemporaneous models with 30-minute data

Variable	Model 29	Model 30	Model 31	Model 32	Model 33
Constant	-0.1705*** (0.0551)	-0.0501 (0.0584)	0.0515 (0.0531)	-0.1198** (0.0539)	-0.4983*** (0.0569)
Midquote return volatility	0.1687*** (0.01)				
Volatility quintile 1		-0.8478*** (0.0505)			
Volatility quintile 2		-0.3734*** (0.042)			
Volatility quintile 4		0.3148*** (0.0378)			
Volatility quintile 5		0.8855*** (0.0398)			
High volatility 10%			0.8864*** (0.0376)		
High volatility 25%				0.919*** (0.0299)	
High volatility 50%					0.8704*** (0.0291)
Marginal effects					
Midquote return volatility	0.0074***				
Volatility quintile 1		-0.0251***			
Volatility quintile 2		-0.0133***			
Volatility quintile 4		0.0147***			
Volatility quintile 5		0.0522***			
High volatility 10%			0.0514***		
High volatility 25%				0.0475***	
High volatility 50%					0.0372***

Table 7.16: Volatility variables as only regressor, univariate contemporaneous models with 30-minute data.

17. Crossed volatility variables as only regressor, univariate contemporaneous models with daily data

Variable	Model 34	Model 35	Model 36	Model 37	Model 38
Constant	2.2116*** (0.3902)	2.0276*** (0.3955)	2.1789*** (0.3904)	2.106*** (0.3915)	1.7888*** (0.3915)
Crossed midquote return volatility	0.6076*** (0.0545)				
Crossed volatility quintile 1		-0.6668*** (0.0931)			
Crossed volatility quintile 2		-0.1551* (0.0809)			
Crossed volatility quintile 4		0.3279*** (0.0801)			
Crossed volatility quintile 5		0.7391*** (0.1004)			
Crossed high volatility 10%			0.6432*** (0.0953)		
Crossed high volatility 25%				0.6125*** (0.0784)	
Crossed high volatility 50%					0.5506*** (0.0652)
Marginal effects					
Crossed midquote return volatility	0.1086***				
Crossed volatility quintile 1		-0.115***			
Crossed volatility quintile 2		-0.0277*			
Crossed volatility quintile 4		0.0593***			
Crossed volatility quintile 5		0.1337***			
Crossed high volatility 10%			0.118***		
Crossed high volatility 25%				0.1119***	
Crossed high volatility 50%					0.1***

Table 7.17: Crossed volatility variables as only regressor, univariate contemporaneous models with daily data.

18. Crossed volatility variables as only regressor, univariate contemporaneous models with 30-minute data

Variable	Model 34	Model 35	Model 36	Model 37	Model 38
Constant	-0.4513*** (0.0558)	-0.1086* (0.0586)	0.0691 (0.0528)	-0.1389*** (0.0538)	-0.5332*** (0.0572)
Crossed midquote return volatility	0.3908*** (0.0129)				
Crossed volatility quintile 1		-0.8557*** (0.0525)			
Crossed volatility quintile 2		-0.3448*** (0.0427)			
Crossed volatility quintile 4		0.3762*** (0.0383)			
Crossed volatility quintile 5		1.005*** (0.0406)			
Crossed high volatility 10%			1.0195*** (0.0376)		
Crossed high volatility 25%				0.9627*** (0.0303)	
Crossed high volatility 50%					0.9012*** (0.0297)
Marginal effects					
Crossed midquote return volatility	0.017***				
Crossed volatility quintile 1		-0.0241***			
Crossed volatility quintile 2		-0.0119***			
Crossed volatility quintile 4		0.0173***			
Crossed volatility quintile 5		0.0598***			
Crossed high volatility 10%			0.0612***		
Crossed high volatility 25%				0.05***	
Crossed high volatility 50%					0.0383***

Table 7.18: Crossed volatility variables as only regressor, univariate contemporaneous models with 30-minute data.

19. Trading activity and market fragmentation as only regressor, univariate contemporaneous models with daily data

Variable	Model 39	Model 40	Model 41	Model 42	Model 43	Model 44
Constant	2.3462*** (0.3899)	1.4413*** (0.3969)	-0.1652 (0.4127)	-0.2204 (0.4123)	2.2104*** (0.3889)	2.2105*** (0.3888)
Traded volume	0.6035*** (0.0608)					
Traded value		0.5504*** (0.0557)				
Submitted limit buy orders			1.5919*** (0.0826)			
Submitted limit sell orders				1.6113*** (0.0828)		
Fragmentation (volume)					0.1234*** (0.0329)	
Fragmentation (value)						0.1236*** (0.0329)
Marginal effects						
Traded volume	0.1086***					
Traded value		0.0991***				
Submitted limit buy orders			0.2753***			
Submitted limit sell orders				0.2782***		
Fragmentation (volume)					0.0225***	
Fragmentation (value)						0.0225***

Table 7.19: Trading activity and market fragmentation as only regressor, univariate contemporaneous models with daily data.

20. Trading activity and market fragmentation as only regressor, univariate contemporaneous models with 30-minute data

Variable	Model 39	Model 40	Model 41	Model 42	Model 43	Model 44
Constant	0.0684 (0.0524)	-0.4908*** (0.0576)	-0.7788*** (0.0579)	-0.7732*** (0.0579)	0.0411 (0.0528)	0.0411 (0.0528)
Traded volume	0.3236*** (0.0131)					
Traded value		0.3519*** (0.0124)				
Submitted limit buy orders			0.6716*** (0.0129)			
Submitted limit sell orders				0.6658*** (0.0128)		
Fragmentation (volume)					-0.1393*** (0.0153)	
Fragmentation (value)						-0.1393*** (0.0153)
Marginal effects						
Traded volume	0.0142***					
Traded value		0.0153***				
Submitted limit buy orders			0.0284***			
Submitted limit sell orders				0.0282***		
Fragmentation (volume)					-0.0062***	
Fragmentation (value)						-0.0062***

Table 7.20: Trading activity and market fragmentation as only regressor, univariate contemporaneous models with 30-minute data.

21. Multivariate contemporaneous models with daily data

Variable	Model 1	Model 2	Model 3	Model 4
Constant	2.296*** (0.4264)	1.1076*** (0.4059)	2.0626*** (0.429)	0.8739** (0.4079)
Crossed bid-ask spread standardized	-0.6662*** (0.0755)	-0.8399*** (0.0743)	-0.6556*** (0.0759)	-0.8306*** (0.0746)
Crossed midquote return volatility	0.7132*** (0.0652)	0.7908*** (0.0656)	0.705*** (0.0652)	0.7821*** (0.0656)
Crossed relative liquidity ask	0.1652*** (0.0523)		0.1675*** (0.0522)	
Crossed relative liquidity bid	0.3588*** (0.0524)		0.3621*** (0.0524)	
Depth beyond crossed ask		-0.3659* (0.2)		-0.3827* (0.1987)
Depth beyond crossed bid		-0.5092*** (0.1842)		-0.522*** (0.1828)
Depth crossed ask	-0.0375 (0.0961)	0.1862 (0.1451)	-0.0391 (0.0953)	0.1989 (0.1438)
Depth crossed bid	-0.0996 (0.1044)	0.2175 (0.1382)	-0.1013 (0.1031)	0.2233 (0.1365)
Fragmentation (value)			0.1909*** (0.0347)	0.1916*** (0.0346)
Fragmentation (volume)	0.1885*** (0.0347)	0.1889*** (0.0346)		
Traded value			0.1554*** (0.0567)	0.1472** (0.0576)
Traded volume	0.154** (0.0641)	0.1463** (0.066)		

Table 7.21: Multivariate contemporaneous models with daily data.

22. Multivariate contemporaneous models with 30-minute data

Variable	Model 1	Model 2	Model 3	Model 4
Constant	-0.5843*** (0.0668)	-0.7934*** (0.0617)	-0.832*** (0.0681)	-1.0484*** (0.0625)
Crossed bid-ask spread standardized	-0.1889*** (0.0272)	-0.2135*** (0.027)	-0.1676*** (0.0271)	-0.1902*** (0.0269)
Crossed midquote return volatility	0.3569*** (0.0153)	0.363*** (0.0153)	0.3292*** (0.0155)	0.3328*** (0.0155)
Crossed relative liquidity ask	0.0753*** (0.0189)		0.0786*** (0.0189)	
Crossed relative liquidity bid	0.0764*** (0.019)		0.0796*** (0.019)	
Depth beyond crossed ask		-0.1054 (0.0702)		-0.1461** (0.0692)
Depth beyond crossed bid		-0.3196*** (0.073)		-0.347*** (0.0721)
Depth crossed ask	-0.1731*** (0.0419)	-0.0482 (0.0541)	-0.1765*** (0.0415)	-0.0209 (0.0514)
Depth crossed bid	-0.2705*** (0.0439)	-0.0634 (0.0575)	-0.2675*** (0.0431)	-0.0334 (0.0556)
Fragmentation (value)			-0.1086*** (0.0158)	-0.1076*** (0.0158)
Fragmentation (volume)	-0.1141*** (0.0158)	-0.1148*** (0.0158)		
Traded value			0.1834*** (0.0139)	0.1817*** (0.0139)
Traded volume	0.1529*** (0.0144)	0.1461*** (0.0146)		

Table 7.22: Multivariate contemporaneous models with 30-minute data.

23. Depth as only regressor, univariate predictive models with daily data

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	1.9954*** (0.3925)	1.957*** (0.3927)	1.9298*** (0.3933)	1.9355*** (0.3913)	1.8651*** (0.3913)	1.8762*** (0.3912)
Depth ask	-0.3955*** (0.0903)					
Depth bid		-0.4174*** (0.0956)				
Depth best prices			-0.4955*** (0.1015)			
Depth beyond ask				-0.5564*** (0.0824)		
Depth beyond bid					-0.5962*** (0.0803)	
Depth beyond						-0.6227*** (0.0839)
Marginal effects						
Depth ask	-0.072***					
Depth bid		-0.076***				
Depth best prices			-0.0901***			
Depth beyond ask				-0.1009***		
Depth beyond bid					-0.108***	
Depth beyond						-0.1128***

Table 7.23: Depth as only regressor, univariate predictive models with daily data.

24. Depth as only regressor, univariate predictive models with 30-minute data

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	-2.0594*** (0.0733)	-2.076*** (0.0735)	-2.1606*** (0.0745)	-2.2623*** (0.0732)	-2.2881*** (0.0732)	-2.3108*** (0.0734)
Depth ask	-0.5596*** (0.0437)					
Depth bid		-0.5685*** (0.043)				
Depth best prices			-0.6809*** (0.0456)			
Depth beyond ask				-0.83*** (0.04)		
Depth beyond bid					-0.852*** (0.0393)	
Depth beyond						-0.875*** (0.0393)
Marginal effects						
Depth ask	-0.0203***					
Depth bid		-0.0206***				
Depth best prices			-0.0247***			
Depth beyond ask				-0.03***		
Depth beyond bid					-0.0308***	
Depth beyond						-0.0316***

Table 7.24: Depth as only regressor, univariate predictive models with 30-minute data.

25. Depth crossed as only regressor, univariate predictive models with daily data

Variable	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Constant	1.9866*** (0.3927)	1.9599*** (0.3916)	1.9164*** (0.3929)	1.9323*** (0.3911)	1.8768*** (0.3914)	1.8742*** (0.3913)
Depth crossed ask	-0.3946*** (0.0806)					
Depth crossed bid		-0.4208*** (0.0838)				
Depth best crossed prices			-0.5137*** (0.0909)			
Depth beyond crossed ask				-0.5317*** (0.0853)		
Depth beyond crossed bid					-0.5391*** (0.0835)	
Depth beyond crossed						-0.5892*** (0.0879)
Marginal effects						
Depth crossed ask	-0.0718***					
Depth crossed bid		-0.0765***				
Depth best crossed prices			-0.0933***			
Depth beyond crossed ask				-0.0965***		
Depth beyond crossed bid					-0.0978***	
Depth beyond crossed						-0.1069***

Table 7.25: Depth crossed as only regressor, univariate predictive models with daily data.

26. Depth crossed as only regressor, univariate predictive models with 30-minute data

Variable	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Constant	-2.0453*** (0.0722)	-2.0436*** (0.0722)	-2.1468*** (0.0733)	-2.1796*** (0.073)	-2.2017*** (0.073)	-2.2392*** (0.0733)
Depth crossed ask	-0.5653*** (0.0415)					
Depth crossed bid		-0.5328*** (0.0393)				
Depth best crossed prices			-0.6747*** (0.0424)			
Depth beyond crossed ask				-0.729*** (0.0407)		
Depth beyond crossed bid					-0.7403*** (0.0396)	
Depth beyond crossed						-0.7866*** (0.0401)
Marginal effects						
Depth crossed ask	-0.0205***					
Depth crossed bid		-0.0193***				
Depth best crossed prices			-0.0245***			
Depth beyond crossed ask				-0.0264***		
Depth beyond crossed bid					-0.0268***	
Depth beyond crossed						-0.0285***

Table 7.26: Depth crossed as only regressor, univariate predictive models with 30-minute data.

27. Bid-ask spread as only regressor, univariate predictive models with daily data

Variable	Model 13	Model 14	Model 15	Model 16
Constant	2.1644*** (0.39)	1.8635*** (0.3948)	2.1642*** (0.3902)	1.7042*** (0.3977)
Bid-ask spread	-0.3386*** (0.0905)			
Bid-ask spread standardized		-0.3838*** (0.0649)		
Crossed bid-ask spread			-0.415*** (0.11)	
Crossed bid-ask spread standardized				-0.4284*** (0.0618)
Marginal effects				
Bid-ask spread	-0.0617***			
Bid-ask spread standardized		-0.0697***		
Crossed bid-ask spread			-0.0756***	
Crossed bid-ask spread standardized				-0.0777***

Table 7.27: Bid-ask spread as only regressor, univariate predictive models with daily data.

28. Bid-ask spread as only regressor, univariate predictive models with 30-minute data

Variable	Model 13	Model 14	Model 15	Model 16
Constant	-1.5651*** (0.0716)	-1.6983*** (0.0676)	-1.5078*** (0.0714)	-1.7486*** (0.0679)
Bid-ask spread	-0.3132*** (0.0579)			
Bid-ask spread standardized		-0.1944*** (0.0444)		
Crossed bid-ask spread			-0.4441*** (0.0542)	
Crossed bid-ask spread standardized				-0.3774*** (0.0381)
Marginal effects				
Bid-ask spread	-0.0114***			
Bid-ask spread standardized		-0.0071***		
Crossed bid-ask spread			-0.0161***	
Crossed bid-ask spread standardized				-0.0137***

Table 7.28: Bid-ask spread as only regressor, univariate predictive models with 30-minute data.

29. Relative liquidity as only regressor, univariate predictive models with daily data

Variable	Model 17	Model 18	Model 19	Model 20	Model 21	Model 22
Constant	2.9293*** (0.3993)	3.0082*** (0.3963)	3.0946*** (0.3991)	2.9902*** (0.4029)	2.9839*** (0.3983)	3.241*** (0.4042)
Relative liquidity ask	0.6221*** (0.0477)					
Relative liquidity bid		0.6826*** (0.0473)				
Relative liquidity			0.7429*** (0.0505)			
Crossed relative liquidity ask				0.4739*** (0.0423)		
Crossed relative liquidity bid					0.5353*** (0.0422)	
Crossed relative liquidity						0.635*** (0.0465)
Marginal effects						
Relative liquidity ask	0.1112***					
Relative liquidity bid		0.1215***				
Relative liquidity			0.1321***			
Crossed relative liquidity ask				0.0852***		
Crossed relative liquidity bid					0.0958***	
Crossed relative liquidity						0.1134***

Table 7.29: Relative liquidity as only regressor, univariate predictive models with daily data.

30. Relative liquidity as only regressor, univariate predictive models with 30-minute data

Variable	Model 17	Model 18	Model 19	Model 20	Model 21	Model 22
Constant	-1.5431*** (0.0675)	-1.545*** (0.0675)	-1.5055*** (0.0675)	-1.4336*** (0.0688)	-1.4172*** (0.0689)	-1.3045*** (0.0693)
Relative liquidity ask	0.522*** (0.0212)					
Relative liquidity bid		0.5224*** (0.0211)				
Relative liquidity			0.6234*** (0.0226)			
Crossed relative liquidity ask				0.3389*** (0.0198)		
Crossed relative liquidity bid					0.3562*** (0.0198)	
Crossed relative liquidity						0.4515*** (0.0212)
Marginal effects						
Relative liquidity ask	0.0189***					
Relative liquidity bid		0.0189***				
Relative liquidity			0.0225***			
Crossed relative liquidity ask				0.0123***		
Crossed relative liquidity bid					0.0129***	
Crossed relative liquidity						0.0163***

Table 7.30: Relative liquidity as only regressor, univariate predictive models with 30-minute data.

31. Slope as only regressor, univariate predictive models with daily data

Variable	Model 17	Model 18	Model 19	Model 20	Model 21	Model 22
Constant	2.179*** (0.3891)	2.1735*** (0.389)	2.1724*** (0.3891)	2.1842*** (0.389)	2.1795*** (0.3893)	2.1802*** (0.3892)
Slope ask	-0.0226 (0.0372)					
Slope bid		-0.0581 (0.0547)				
Slope			-0.0556 (0.0483)			
Crossed slope ask				0.0407 (0.0399)		
Crossed slope bid					-0.067* (0.0384)	
Crossed slope						-0.0319 (0.0479)
Marginal effects						
Slope ask	-0.0041					
Slope bid		-0.0106				
Slope			-0.0101			
Crossed slope ask				0.0074		
Crossed slope bid					-0.0122*	
Crossed slope						-0.0058

Table 7.31: Slope as only regressor, univariate predictive models with daily data.

32. Slope as only regressor, univariate predictive models with 30-minute data

Variable	Model 17	Model 18	Model 19	Model 20	Model 21	Model 22
Constant	-1.7021*** (0.0676)	-1.7025*** (0.0676)	-1.7026*** (0.0676)	-1.7035*** (0.0676)	-1.7025*** (0.0676)	-1.7034*** (0.0676)
Slope ask	-0.003 (0.0278)					
Slope bid		0.003 (0.0145)				
Slope			0.0023 (0.0182)			
Crossed slope ask				0.0255 (0.019)		
Crossed slope bid					0.006 (0.0233)	
Crossed slope						0.0246 (0.0225)
Marginal effects						
Slope ask	-1e-04					
Slope bid		1e-04				
Slope			1e-04			
Crossed slope ask				9e-04		
Crossed slope bid					2e-04	
Crossed slope						9e-04

Table 7.32: Slope as only regressor, univariate predictive models with 30-minute data.

33. Volatility variables as only regressor, univariate predictive models with daily data

Variable	Model 23	Model 24	Model 25	Model 26	Model 27
Constant	2.1743*** (0.3892)	2.1359*** (0.3927)	2.1599*** (0.3894)	2.1427*** (0.39)	2.0073*** (0.3917)
Midquote return volatility	0.0874*** (0.0277)				
Volatility quintile 1		-0.2364*** (0.0879)			
Volatility quintile 2		-0.0839 (0.0791)			
Volatility quintile 4		0.0757 (0.0791)			
Volatility quintile 5		0.1937** (0.0917)			
High volatility 10%			0.1588* (0.0906)		
High volatility 25%				0.1668** (0.0716)	
High volatility 50%					0.2456*** (0.0607)
Marginal effects					
Midquote return volatility	0.0159***				
Volatility quintile 1		-0.0428***			
Volatility quintile 2		-0.0153			
Volatility quintile 4		0.0139			
Volatility quintile 5		0.0356**			
High volatility 10%			0.0291*		
High volatility 25%				0.0305**	
High volatility 50%					0.0448***

Table 7.33: Volatility variables as only regressor, univariate predictive models with daily data.

34. Volatility variables as only regressor, univariate predictive models with 30-minute data

Variable	Model 23	Model 24	Model 25	Model 26	Model 27
Constant	-1.9009*** (0.0715)	-1.8301*** (0.0731)	-1.7352*** (0.0675)	-1.81*** (0.0682)	-1.9399*** (0.0713)
Midquote return volatility	0.0961*** (0.0088)				
Volatility quintile 1		-0.2296*** (0.053)			
Volatility quintile 2		-0.0627 (0.0462)			
Volatility quintile 4		0.1904*** (0.0437)			
Volatility quintile 5		0.5497*** (0.0473)			
High volatility 10%			0.5627*** (0.045)		
High volatility 25%				0.4741*** (0.0353)	
High volatility 50%					0.36*** (0.0327)
Marginal effects					
Midquote return volatility	0.0035***				
Volatility quintile 1		-0.0068***			
Volatility quintile 2		-0.002			
Volatility quintile 4		0.0068***			
Volatility quintile 5		0.023***			
High volatility 10%			0.0248***		
High volatility 25%				0.0189***	
High volatility 50%					0.0129***

Table 7.34: Volatility variables as only regressor, univariate predictive models with 30-minute data.

35. Crossed volatility variables as only regressor, univariate predictive models with daily data

Variable	Model 28	Model 29	Model 30	Model 31	Model 32
Constant	2.1859*** (0.3889)	2.1313*** (0.3917)	2.1766*** (0.3893)	2.1579*** (0.3895)	2.0661*** (0.3914)
Crossed midquote return volatility	0.1547*** (0.0328)				
Crossed volatility quintile 1		-0.2388*** (0.0917)			
Crossed volatility quintile 2		-0.0651 (0.0807)			
Crossed volatility quintile 4		0.106 (0.0794)			
Crossed volatility quintile 5		0.15 (0.0986)			
Crossed high volatility 10%			0.2152** (0.0946)		
Crossed high volatility 25%				0.1473* (0.0772)	
Crossed high volatility 50%					0.1568** (0.0649)
Marginal effects					
Crossed midquote return volatility	0.0281***				
Crossed volatility quintile 1		-0.0432***			
Crossed volatility quintile 2		-0.0119			
Crossed volatility quintile 4		0.0194			
Crossed volatility quintile 5		0.0275			
Crossed high volatility 10%			0.0395**		
Crossed high volatility 25%				0.0269*	
Crossed high volatility 50%					0.0286**

Table 7.35: Crossed volatility variables as only regressor, univariate predictive models with daily data.

36. Crossed volatility variables as only regressor, univariate predictive models with 30-minute data

Variable	Model 28	Model 29	Model 30	Model 31	Model 32
Constant	-1.9903*** (0.0722)	-1.8443*** (0.0732)	-1.7276*** (0.0676)	-1.8339*** (0.0685)	-1.9623*** (0.0718)
Crossed midquote return volatility	0.1639*** (0.0126)				
Crossed volatility quintile 1		-0.2519*** (0.0551)			
Crossed volatility quintile 2		-0.0828* (0.0465)			
Crossed volatility quintile 4		0.2024*** (0.0437)			
Crossed volatility quintile 5		0.5622*** (0.0474)			
Crossed high volatility 10%			0.61*** (0.0445)		
Crossed high volatility 25%				0.5276*** (0.0354)	
Crossed high volatility 50%					0.3769*** (0.033)
Marginal effects					
Crossed midquote return volatility	0.0059***				
Crossed volatility quintile 1		-0.0073***			
Crossed volatility quintile 2		-0.0026*			
Crossed volatility quintile 4		0.0072***			
Crossed volatility quintile 5		0.0235***			
Crossed high volatility 10%			0.0272***		
Crossed high volatility 25%				0.0211***	
Crossed high volatility 50%					0.0134***

Table 7.36: Crossed volatility variables as only regressor, univariate predictive models with 30-minute data.

37. Trading activity variables as only regressor, univariate predictive models with daily data

Variable	Model 33	Model 34	Model 35	Model 36	Model 37	Model 38
Constant	2.2552*** (0.3894)	1.7992*** (0.3952)	0.4445 (0.4074)	0.4476 (0.4067)	2.2066*** (0.3889)	2.2067*** (0.3889)
Traded volume	0.2823*** (0.0511)					
Traded value		0.2677*** (0.049)				
Submitted limit buy orders			1.1432*** (0.0743)			
Submitted limit sell orders				1.1219*** (0.0735)		
Fragmentation (volume)					0.123*** (0.033)	
Fragmentation (value)						0.123*** (0.033)
Marginal effects						
Traded volume	0.0513***					
Traded value		0.0486***				
Submitted limit buy orders			0.2018***			
Submitted limit sell orders				0.1981***		
Fragmentation (volume)					0.0224***	
Fragmentation (value)						0.0224***

Table 7.37: Trading activity variables as only regressor, univariate predictive models with daily data.

38. Trading activity variables as only regressor, univariate predictive models with 30-minute data

Variable	Model 33	Model 34	Model 35	Model 36	Model 37	Model 38
Constant	-1.784*** (0.0685)	-2.173*** (0.0729)	-2.3568*** (0.0719)	-2.3595*** (0.072)	-1.7578*** (0.068)	-1.7578*** (0.068)
Traded volume	0.2541*** (0.0129)					
Traded value		0.2572*** (0.0125)				
Submitted limit buy orders			0.5409*** (0.0136)			
Submitted limit sell orders				0.5406*** (0.0135)		
Fragmentation (volume)					-0.1382*** (0.0186)	
Fragmentation (value)						-0.1382*** (0.0186)
Marginal effects						
Traded volume	0.0092***					
Traded value		0.0093***				
Submitted limit buy orders			0.019***			
Submitted limit sell orders				0.019***		
Fragmentation (volume)					-0.005***	
Fragmentation (value)						-0.005***

Table 7.38: Trading activity variables as only regressor, univariate predictive models with 30-minute data.

39. Multivariate predictive models with daily data

Variable	Model 1	Model 2	Model 3	Model 4
Constant	2.8219*** (0.4201)	1.5699*** (0.4004)	2.6702*** (0.4247)	1.3826*** (0.4029)
Crossed bid-ask spread standardized	-0.2459*** (0.0639)	-0.4012*** (0.0646)	-0.2391*** (0.0641)	-0.3954*** (0.0648)
Crossed midquote return volatility	0.1496*** (0.037)	0.1746*** (0.0405)	0.1435*** (0.0371)	0.1701*** (0.0406)
Crossed relative liquidity ask	0.2349*** (0.0514)		0.2361*** (0.0514)	
Crossed relative liquidity bid	0.3489*** (0.0515)		0.3503*** (0.0515)	
Depth beyond crossed ask		-0.3092 (0.1925)		-0.3259* (0.1913)
Depth beyond crossed bid		-0.2623 (0.1765)		-0.2782 (0.1759)
Depth crossed ask	-0.047 (0.0918)	0.0837 (0.1394)	-0.0484 (0.0916)	0.0952 (0.1383)
Depth crossed bid	-0.0681 (0.1004)	0.0528 (0.1329)	-0.0664 (0.1001)	0.0645 (0.1327)
Fragmentation (value)			0.1614*** (0.034)	0.1539*** (0.0339)
Fragmentation (volume)	0.1599*** (0.0339)	0.1516*** (0.0338)		
Traded value			0.0986* (0.053)	0.1125** (0.0525)
Traded volume	0.0825 (0.0554)	0.1076* (0.0563)		

Table 7.39: Multivariate predictive models with daily data.

40. Multivariate predictive models with 30-minute data

Variable	Model 1	Model 2	Model 3	Model 4
Constant	-2.0544*** (0.0827)	-2.5376*** (0.0797)	-2.2814*** (0.0837)	-2.778*** (0.0801)
Crossed bid-ask spread standardized	-0.3814*** (0.0423)	-0.4212*** (0.0422)	-0.3617*** (0.0421)	-0.3979*** (0.0419)
Crossed midquote return volatility	0.1514*** (0.0162)	0.1547*** (0.0162)	0.1339*** (0.0159)	0.1332*** (0.016)
Crossed relative liquidity ask	0.1823*** (0.0223)		0.1853*** (0.0223)	
Crossed relative liquidity bid	0.1996*** (0.0224)		0.2042*** (0.0224)	
Depth beyond crossed ask		-0.2311*** (0.076)		-0.2745*** (0.0746)
Depth beyond crossed bid		-0.425*** (0.0747)		-0.46*** (0.0736)
Depth crossed ask	-0.2276*** (0.0442)	-0.0401 (0.0562)	-0.2297*** (0.0439)	-0.0093 (0.0526)
Depth crossed bid	-0.1761*** (0.0429)	0.0623 (0.0502)	-0.1854*** (0.0424)	0.0832* (0.0482)
Fragmentation (value)			-0.1003*** (0.0192)	-0.1124*** (0.0191)
Fragmentation (volume)	-0.1056*** (0.0192)	-0.1201*** (0.0191)		
Traded value			0.1659*** (0.014)	0.168*** (0.0141)
Traded volume	0.1505*** (0.0144)	0.1445*** (0.0145)		

Table 7.40: Multivariate predictive models with 30-minute data.

41. Fragmentation and volatility interaction models without control variables, daily data

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	2.2066*** (0.3891)	2.1485*** (0.3928)	2.1727*** (0.3898)	2.1683*** (0.3898)	2.0317*** (0.3915)
Fragmentation	0.1361*** (0.0334)	0.1279* (0.0683)	0.1098*** (0.0354)	0.1389*** (0.0381)	0.1492*** (0.0435)
Midquote ret vol	0.09*** (0.0296)				
Fragmentation * Midquote ret vol	-0.0168 (0.0224)				
VolQ1		-0.2262** (0.0889)			
VolQ2		-0.0698 (0.0805)			
VolQ4		0.0884 (0.0806)			
VolQ5		0.227** (0.094)			
Fragmentation * VolQ1		0.1056 (0.0848)			
Fragmentation * VolQ2		-0.0724 (0.0846)			
Fragmentation * VolQ4		-0.0129 (0.0857)			
Fragmentation * VolQ5		0.0019 (0.0821)			
VolQ10			0.2238** (0.094)		
Fragmentation * VolQ10			0.1081 (0.0721)		
VolQ25				0.1863*** (0.0722)	
Fragmentation * VolQ25				-0.0256 (0.0559)	
VolQ50					0.258*** (0.0608)
Fragmentation * VolQ50					-0.0373 (0.0515)

Table 7.41: Fragmentation and volatility interaction models without control variables, daily data.

42. Fragmentation and volatility interaction models without control variables, 30-minute data

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	-1.9518*** (0.0718)	-1.8795*** (0.0737)	-1.7883*** (0.068)	-1.8627*** (0.0687)	-1.9902*** (0.0717)
Fragmentation	-0.1332*** (0.0186)	-0.1168*** (0.0416)	-0.1166*** (0.0199)	-0.1058*** (0.0219)	-0.1137*** (0.0262)
Midquote ret vol	0.1057*** (0.0117)				
Fragmentation * Midquote ret vol	0.0138 (0.0091)				
VolQ1		-0.2273*** (0.0534)			
VolQ2		-0.0665 (0.047)			
VolQ4		0.1846*** (0.0446)			
VolQ5		0.5339*** (0.0477)			
Fragmentation * VolQ1		-0.033 (0.0555)			
Fragmentation * VolQ2		0.0378 (0.0558)			
Fragmentation * VolQ4		0.0164 (0.0542)			
Fragmentation * VolQ5		-0.047 (0.0515)			
VolQ10			0.534*** (0.0458)		
Fragmentation * VolQ10			-0.0706 (0.0455)		
VolQ25				0.4641*** (0.0354)	
Fragmentation * VolQ25				-0.0683* (0.0351)	
VolQ50					0.3561*** (0.0328)
Fragmentation * VolQ50					-0.0337 (0.033)

Table 7.42: Fragmentation and volatility interaction models without control variables, 30-minute data.

43. Fragmentation and crossed volatility interaction models without control variables, daily data

Variable	Model 6	Model 7	Model 8	Model 9	Model 10
Constant	2.2286*** (0.3889)	2.2182*** (0.3946)	2.2052*** (0.3893)	2.1849*** (0.3891)	2.112*** (0.3919)
Fragmentation	0.1352*** (0.0334)	0.1537** (0.0694)	0.1435*** (0.0346)	0.1677*** (0.0369)	0.1837*** (0.0417)
Crossed midquote ret vol	0.099** (0.0441)				
Fragmentation * Crossed midquote ret vol	-0.0721** (0.0299)				
Crossed VolQ1		-0.1931** (0.0926)			
Crossed VolQ2		-0.0468 (0.082)			
Crossed VolQ4		0.1355* (0.0818)			
Crossed VolQ5		0.1657 (0.1009)			
Fragmentation * Crossed VolQ1		0.1795** (0.0819)			
Fragmentation * Crossed VolQ2		-0.1507* (0.0866)			
Fragmentation * Crossed VolQ4		-0.1174 (0.0894)			
Fragmentation * Crossed VolQ5		-0.1551* (0.0916)			
Crossed VolQ10			0.1808* (0.101)		
Fragmentation * Crossed VolQ10			-0.1449 (0.0906)		
Crossed VolQ25				0.1532** (0.0776)	
Fragmentation * Crossed VolQ25				-0.1646** (0.0645)	
Crossed VolQ50					0.1774*** (0.0654)
Fragmentation * Crossed VolQ50					-0.1312** (0.0554)

Table 7.43: Fragmentation and crossed volatility interaction models without control variables, daily data.

44. Fragmentation and crossed volatility interaction models without control variables, 30-minute data

Variable	Model 6	Model 7	Model 8	Model 9	Model 10
Constant	-2.0432*** (0.0719)	-1.8994*** (0.0739)	-1.7832*** (0.0681)	-1.8899*** (0.0689)	-2.0149*** (0.0722)
Fragmentation	-0.1268*** (0.0187)	-0.1006** (0.0418)	-0.1098*** (0.02)	-0.1008*** (0.0219)	-0.1078*** (0.0261)
Crossed midquote ret vol	0.213*** (0.0161)				
Fragmentation * Crossed midquote ret vol	0.0586*** (0.0114)				
Crossed VolQ1		-0.2507*** (0.0554)			
Crossed VolQ2		-0.0822* (0.0474)			
Crossed VolQ4		0.1993*** (0.0447)			
Crossed VolQ5		0.5515*** (0.0479)			
Fragmentation * Crossed VolQ1		-0.0194 (0.0554)			
Fragmentation * Crossed VolQ2		0.0113 (0.056)			
Fragmentation * Crossed VolQ4		0.0033 (0.0544)			
Fragmentation * Crossed VolQ5		-0.0999* (0.052)			
Crossed VolQ10			0.5791*** (0.0452)		
Fragmentation * Crossed VolQ10			-0.1162** (0.0464)		
Crossed VolQ25				0.5232*** (0.0355)	
Fragmentation * Crossed VolQ25				-0.0883** (0.0356)	
Crossed VolQ50					0.3763*** (0.0332)
Fragmentation * Crossed VolQ50					-0.0473 (0.0332)

Table 7.44: Fragmentation and crossed volatility interaction models without control variables, 30-minute data.

45. Fragmentation and volatility interaction models with all control variables, daily data

Variable	Model 11	Model 12	Model 13	Model 14	Model 15
Constant	3.1649*** (0.4086)	3.1509*** (0.4118)	3.1708*** (0.4096)	3.1662*** (0.4093)	3.0424*** (0.4114)
Fragmentation	0.14*** (0.0337)	0.1106 (0.0689)	0.1127*** (0.0359)	0.1386*** (0.0386)	0.1545*** (0.0441)
Midquote ret vol	0.0371 (0.0298)	0.0198 (0.0327)	0.0488 (0.0341)	0.0417 (0.0318)	0.0167 (0.0295)
Fragmentation * Midquote ret vol	-0.0077 (0.0225)				
VolQ1		-0.1846** (0.0907)			
VolQ2		-0.0517 (0.0818)			
VolQ4		0.0399 (0.0825)			
VolQ5		0.0327 (0.107)			
Fragmentation * VolQ1		0.1335 (0.0856)			
Fragmentation * VolQ2		-0.0511 (0.0852)			
Fragmentation * VolQ4		-0.0089 (0.0867)			
Fragmentation * VolQ5		0.0474 (0.0831)			
VolQ10			-0.0153 (0.1145)		
Fragmentation * VolQ10			0.1465** (0.0732)		
VolQ25				-0.0093 (0.0835)	
Fragmentation * VolQ25				3e-04 (0.0568)	
VolQ50					0.165** (0.0652)
Fragmentation * VolQ50					-0.0268 (0.0523)

Table 7.45: Fragmentation and volatility interaction models with all control variables, daily data.

46. Fragmentation and volatility interaction models with all control variables, 30-minute data

Variable	Model 11	Model 12	Model 13	Model 14	Model 15
Constant	-2.0314*** (0.0791)	-2.0698*** (0.0834)	-2.0197*** (0.0787)	-2.0688*** (0.079)	-2.1792*** (0.0812)
Fragmentation	-0.1077*** (0.0188)	-0.0971** (0.042)	-0.1021*** (0.0201)	-0.0975*** (0.022)	-0.1134*** (0.0262)
Midquote ret vol	0.0638*** (0.0131)	0.0334*** (0.0096)	0.0416*** (0.0095)	0.04*** (0.0094)	0.0465*** (0.0093)
Fragmentation * Midquote ret vol	0.0053 (0.01)				
VolQ1		-0.2126*** (0.0536)			
VolQ2		-0.0472 (0.0471)			
VolQ4		0.1587*** (0.0448)			
VolQ5		0.3697*** (0.0495)			
Fragmentation * VolQ1		-0.0549 (0.0559)			
Fragmentation * VolQ2		0.0181 (0.0561)			
Fragmentation * VolQ4		0.0173 (0.0548)			
Fragmentation * VolQ5		0.001 (0.0529)			
VolQ10			0.3189*** (0.0487)		
Fragmentation * VolQ10			-7e-04 (0.048)		
VolQ25				0.3239*** (0.037)	
Fragmentation * VolQ25				-0.0119 (0.0364)	
VolQ50					0.2718*** (0.0336)
Fragmentation * VolQ50					0.0181 (0.0337)

Table 7.46: Fragmentation and volatility interaction models with all control variables, 30-minute data.

47. Fragmentation and crossed volatility interaction models with all control variables, daily data

Variable	Model 16	Model 17	Model 18	Model 19	Model 20
Constant	3.2683*** (0.4143)	3.2731*** (0.4192)	3.2448*** (0.4138)	3.2481*** (0.4139)	3.2188*** (0.4172)
Fragmentation	0.158*** (0.0339)	0.1955*** (0.0705)	0.174*** (0.0351)	0.1957*** (0.0375)	0.2064*** (0.0424)
Crossed midquote ret vol	0.043 (0.046)	0.1093*** (0.0392)	0.1245*** (0.0386)	0.1239*** (0.038)	0.1111*** (0.0359)
Fragmentation * Crossed midquote ret vol	-0.0829*** (0.0299)				
Crossed VolQ1		-0.1559 (0.0954)			
Crossed VolQ2		-0.0128 (0.0835)			
Crossed VolQ4		0.0914 (0.0837)			
Crossed VolQ5		-0.0618 (0.1126)			
Fragmentation * Crossed VolQ1		0.1509* (0.0833)			
Fragmentation * Crossed VolQ2		-0.1533* (0.0876)			
Fragmentation * Crossed VolQ4		-0.1194 (0.0904)			
Fragmentation * Crossed VolQ5		-0.1329 (0.0933)			
Crossed VolQ10			-0.0504 (0.1143)		
Fragmentation * Crossed VolQ10			-0.087 (0.0925)		
Crossed VolQ25				-0.0425 (0.0871)	
Fragmentation * Crossed VolQ25				-0.1247* (0.0659)	
Crossed VolQ50					0.07 (0.0702)
Fragmentation * Crossed VolQ50					-0.0931* (0.0565)

Table 7.47: Fragmentation and crossed volatility interaction models with all control variables, daily data.

48. Fragmentation and crossed volatility interaction models with all control variables, 30-minute data

Variable	Model 16	Model 17	Model 18	Model 19	Model 20
Constant	-2.0143*** (0.0828)	-2.0113*** (0.087)	-1.9668*** (0.0827)	-2.0357*** (0.0828)	-2.1513*** (0.0849)
Fragmentation	-0.088*** (0.019)	-0.0678 (0.0422)	-0.0808*** (0.0202)	-0.0791*** (0.0222)	-0.0945*** (0.0264)
Crossed midquote ret vol	0.1492*** (0.0168)	0.0556*** (0.0112)	0.0706*** (0.0118)	0.0663*** (0.0115)	0.0791*** (0.0118)
Fragmentation * Crossed midquote ret vol	0.0436*** (0.0114)				
Crossed VolQ1		-0.2587*** (0.0559)			
Crossed VolQ2		-0.0662 (0.0475)			
Crossed VolQ4		0.1705*** (0.0449)			
Crossed VolQ5		0.4211*** (0.0498)			
Fragmentation * Crossed VolQ1		-0.042 (0.0561)			
Fragmentation * Crossed VolQ2		-0.007 (0.0565)			
Fragmentation * Crossed VolQ4		0.0122 (0.0551)			
Fragmentation * Crossed VolQ5		-0.0276 (0.0534)			
Crossed VolQ10			0.4096*** (0.0484)		
Fragmentation * Crossed VolQ10			-0.0057 (0.0491)		
Crossed VolQ25				0.403*** (0.0375)	
Fragmentation * Crossed VolQ25				-0.0096 (0.037)	
Crossed VolQ50					0.2971*** (0.0345)
Fragmentation * Crossed VolQ50					0.0192 (0.0341)

Table 7.48: Fragmentation and crossed volatility interaction models with all control variables, 30-minute data.

49. Fragmentation and volatility interaction models with one control variable, daily data

Variable	Model 21	Model 22	Model 23	Model 24	Model 25
Constant	2.2778*** (0.3894)	3.0256*** (0.3967)	2.9506*** (0.3994)	1.8914*** (0.3916)	1.9614*** (0.3914)
Fragmentation	0.1437*** (0.0334)	0.1337*** (0.0337)	0.1422*** (0.0336)	0.1412*** (0.0335)	0.1429*** (0.0335)
Midquote ret vol	0.0643** (0.0296)	0.0535* (0.0296)	0.059** (0.0295)	0.0851*** (0.0295)	0.0847*** (0.0295)
Fragmentation * Midquote ret vol	-0.0057 (0.0225)	-0.0143 (0.0223)	-0.0139 (0.0225)	-0.0166 (0.0224)	-0.0166 (0.0224)

Table 7.49: Fragmentation and volatility interaction models with one control variable, daily data.

50. Fragmentation and volatility interaction models with one control variable, 30-minute data

Variable	Model 21	Model 22	Model 23	Model 24	Model 25
Constant	-1.948*** (0.072)	-1.777*** (0.0718)	-1.7729*** (0.0717)	-2.5188*** (0.0767)	-2.4907*** (0.0766)
Fragmentation	-0.1107*** (0.0188)	-0.1284*** (0.0187)	-0.1265*** (0.0187)	-0.1251*** (0.0187)	-0.1264*** (0.0187)
Midquote ret vol	0.075*** (0.0126)	0.0912*** (0.0121)	0.0908*** (0.0121)	0.1043*** (0.0119)	0.1029*** (0.0119)
Fragmentation * Midquote ret vol	0.0145 (0.0097)	0.0074 (0.0094)	0.0076 (0.0094)	0.0126 (0.0093)	0.0121 (0.0093)

Table 7.50: Fragmentation and volatility interaction models with one control variable, 30-minute data.

51. Fragmentation and volatility interaction models with one control variable, daily data

Variable	Model 26	Model 27	Model 28	Model 29	Model 30
Constant	2.2066*** (0.3891)	2.2391*** (0.393)	2.9867*** (0.3997)	2.9215*** (0.4024)	1.8367*** (0.395)
Fragmentation	0.1361*** (0.0334)	0.1283* (0.0683)	0.1117 (0.0689)	0.1162* (0.0687)	0.1406** (0.0685)
Midquote ret vol	0.09*** (0.0296)				
Fragmentation * Midquote ret vol	-0.0168 (0.0224)				
VolQ1		-0.2207** (0.0889)	-0.1903** (0.0899)	-0.2104** (0.0898)	-0.2157** (0.089)
VolQ2		-0.0694 (0.0806)	-0.0531 (0.0816)	-0.0626 (0.0813)	-0.0644 (0.0807)
VolQ4		0.0728 (0.0808)	0.0641 (0.0817)	0.0517 (0.0815)	0.0942 (0.0809)
VolQ5		0.141 (0.0955)	0.1185 (0.0956)	0.1234 (0.0953)	0.2067** (0.0943)
Fragmentation * VolQ1		0.1092 (0.0848)	0.1244 (0.0856)	0.1347 (0.0856)	0.0984 (0.085)
Fragmentation * VolQ2		-0.0665 (0.0846)	-0.0591 (0.0853)	-0.0543 (0.0851)	-0.0772 (0.0847)
Fragmentation * VolQ4		-0.0061 (0.0858)	-0.0109 (0.0868)	-0.0196 (0.0863)	-0.0211 (0.086)
Fragmentation * VolQ5		0.0224 (0.0823)	0.0265 (0.0829)	0.0369 (0.0826)	-0.0121 (0.0824)

Table 7.51: Fragmentation and volatility interaction models with one control variable, daily data.

52. Fragmentation and volatility interaction models with one control variable, 30-minute data

Variable	Model 26	Model 27	Model 28	Model 29	Model 30
Constant	-1.9518*** (0.0718)	-1.9246*** (0.0743)	-1.7149*** (0.0737)	-1.7123*** (0.0737)	-2.436*** (0.0784)
Fragmentation	-0.1332*** (0.0186)	-0.1021** (0.0417)	-0.1121*** (0.0419)	-0.11*** (0.0419)	-0.1041** (0.0417)
Midquote ret vol	0.1057*** (0.0117)				
Fragmentation * Midquote ret vol	0.0138 (0.0091)				
VolQ1		-0.2161*** (0.0535)	-0.2405*** (0.0534)	-0.2368*** (0.0535)	-0.223*** (0.0534)
VolQ2		-0.0545 (0.0471)	-0.0651 (0.0471)	-0.0646 (0.0471)	-0.061 (0.047)
VolQ4		0.1754*** (0.0447)	0.1792*** (0.0447)	0.1785*** (0.0447)	0.1754*** (0.0447)
VolQ5		0.4436*** (0.0484)	0.4823*** (0.0478)	0.4801*** (0.0478)	0.5229*** (0.0478)
Fragmentation * VolQ1		-0.0457 (0.0555)	-0.0346 (0.0559)	-0.0362 (0.056)	-0.0502 (0.0555)
Fragmentation * VolQ2		0.0236 (0.0558)	0.0349 (0.0561)	0.0373 (0.0561)	0.0235 (0.0558)
Fragmentation * VolQ4		0.0162 (0.0543)	0.0168 (0.0546)	0.0152 (0.0546)	0.0216 (0.0544)
Fragmentation * VolQ5		-6e-04 (0.0521)	-0.0481 (0.0519)	-0.0466 (0.0519)	-0.0509 (0.0517)

Table 7.52: Fragmentation and volatility interaction models with one control variable, 30-minute data.

53. Fragmentation and volatility interaction models with one control variable, daily data

Variable	Model 31	Model 32	Model 33	Model 34	Model 35
Constant	1.9111*** (0.395)	2.1622*** (0.3927)	2.2611*** (0.3902)	3.0124*** (0.3976)	2.936*** (0.4004)
Fragmentation	0.1408** (0.0685)	0.1292* (0.0683)	0.1181*** (0.0355)	0.1041*** (0.0358)	0.1117*** (0.0357)
VolQ1	-0.222** (0.089)	-0.2082** (0.0896)			
VolQ2	-0.07 (0.0807)	-0.0612 (0.0807)			
VolQ4	0.0892 (0.0809)	0.0718 (0.0813)			
VolQ5	0.1988** (0.0943)	0.1534 (0.1048)			
Fragmentation * VolQ1	0.0991 (0.0849)	0.1065 (0.0848)			
Fragmentation * VolQ2	-0.0759 (0.0847)	-0.0721 (0.0846)			
Fragmentation * VolQ4	-0.0173 (0.086)	-0.0108 (0.0857)			
Fragmentation * VolQ5	-0.0111 (0.0823)	0.0045 (0.0821)			
Midquote ret vol		0.0509 (0.0321)			
VolQ10			0.1323 (0.0952)	0.1285 (0.0949)	0.1451 (0.0945)
Fragmentation * VolQ10			0.1226* (0.0726)	0.1349* (0.0729)	0.1394* (0.0725)

Table 7.53: Fragmentation and volatility interaction models with one control variable, daily data.

54. Fragmentation and volatility interaction models with one control variable, 30-minute data

Variable	Model 31	Model 32	Model 33	Model 34	Model 35
Constant	-2.4107*** (0.0784)	-1.9892*** (0.0766)	-1.8386*** (0.0687)	-1.6357*** (0.0681)	-1.6325*** (0.0681)
Fragmentation	-0.1044** (0.0417)	-0.1226*** (0.0416)	-0.1033*** (0.02)	-0.1142*** (0.02)	-0.1118*** (0.02)
VolQ1	-0.2216*** (0.0534)	-0.2081*** (0.0535)			
VolQ2	-0.0599 (0.047)	-0.0572 (0.047)			
VolQ4	0.1771*** (0.0447)	0.1698*** (0.0447)			
VolQ5	0.5202*** (0.0478)	0.4704*** (0.049)			
Fragmentation * VolQ1	-0.0506 (0.0555)	-0.0331 (0.0555)			
Fragmentation * VolQ2	0.0239 (0.0558)	0.0366 (0.0558)			
Fragmentation * VolQ4	0.0188 (0.0544)	0.0191 (0.0543)			
Fragmentation * VolQ5	-0.0526 (0.0517)	-0.0156 (0.0519)			
Midquote ret vol		0.058*** (0.0092)			
VolQ10			0.397*** (0.0471)	0.4774*** (0.0458)	0.4748*** (0.0457)
Fragmentation * VolQ10			-0.0189 (0.047)	-0.0607 (0.0461)	-0.0636 (0.0461)

Table 7.54: Fragmentation and volatility interaction models with one control variable, 30-minute data.

55. Fragmentation and volatility interaction models with one control variable, daily data

Variable	Model 36	Model 37	Model 38	Model 39	Model 40
Constant	1.8596*** (0.3922)	1.9302*** (0.392)	2.1909*** (0.3898)	2.2591*** (0.3901)	3.0118*** (0.3973)
Fragmentation	0.1161*** (0.0356)	0.1178*** (0.0356)	0.115*** (0.0355)	0.1435*** (0.0381)	0.1339*** (0.0385)
VolQ10	0.202** (0.0942)	0.1988** (0.0942)	0.0403 (0.1143)		
Fragmentation * VolQ10	0.1011 (0.0723)	0.1006 (0.0722)	0.1052 (0.0722)		
Midquote ret vol			0.0957*** (0.0342)		
VolQ25				0.1106 (0.0737)	0.0881 (0.0735)
Fragmentation * VolQ25				-0.0092 (0.0561)	-0.016 (0.0565)

Table 7.55: Fragmentation and volatility interaction models with one control variable, daily data.

56. Fragmentation and volatility interaction models with one control variable, 30-minute data

Variable	Model 36	Model 37	Model 38	Model 39	Model 40
Constant	-2.3532*** (0.0732)	-2.328*** (0.0733)	-1.9207*** (0.0716)	-1.9061*** (0.0693)	-1.7024*** (0.0687)
Fragmentation	-0.1087*** (0.02)	-0.1099*** (0.02)	-0.1211*** (0.0199)	-0.099*** (0.0219)	-0.1015*** (0.022)
VolQ10	0.5233*** (0.0461)	0.5202*** (0.0461)	0.4434*** (0.0479)		
Fragmentation * VolQ10	-0.0753 (0.0458)	-0.076* (0.0458)	-0.0203 (0.0466)		
Midquote ret vol			0.0653*** (0.0093)		
VolQ25				0.3888*** (0.036)	0.428*** (0.0355)
Fragmentation * VolQ25				-0.0179 (0.0358)	-0.0695** (0.0354)

Table 7.56: Fragmentation and volatility interaction models with one control variable, 30-minute data.

57. Fragmentation and volatility interaction models with one control variable, daily data

Variable	Model 41	Model 42	Model 43	Model 44	Model 45
Constant	2.9342*** (0.4)	1.8579*** (0.392)	1.9284*** (0.3919)	2.1856*** (0.3896)	2.1279*** (0.392)
Fragmentation	0.1393*** (0.0385)	0.1474*** (0.0382)	0.1491*** (0.0382)	0.142*** (0.0381)	0.1541*** (0.0436)
VolQ25	0.0995 (0.0733)	0.1644** (0.0726)	0.163** (0.0725)	0.0871 (0.082)	
Fragmentation * VolQ25	-0.0066 (0.0563)	-0.0366 (0.056)	-0.0365 (0.056)	-0.0207 (0.0559)	
Midquote ret vol				0.0803** (0.0314)	
VolQ50					0.224*** (0.0612)
Fragmentation * VolQ50					-0.0239 (0.0516)

Table 7.57: Fragmentation and volatility interaction models with one control variable, daily data.

58. Fragmentation and volatility interaction models with one control variable, 30-minute data

Variable	Model 41	Model 42	Model 43	Model 44	Model 45
Constant	-1.699*** (0.0687)	-2.4245*** (0.0738)	-2.3977*** (0.0738)	-1.988*** (0.0718)	-2.0279*** (0.0724)
Fragmentation	-0.0992*** (0.022)	-0.1003*** (0.0219)	-0.1014*** (0.0219)	-0.1125*** (0.0219)	-0.1096*** (0.0261)
VolQ25	0.4232*** (0.0355)	0.455*** (0.0354)	0.4521*** (0.0354)	0.4006*** (0.0367)	
Fragmentation * VolQ25	-0.0703** (0.0354)	-0.064* (0.0353)	-0.0642* (0.0353)	-0.0356 (0.0356)	
Midquote ret vol				0.0667*** (0.0091)	
VolQ50					0.3114*** (0.0331)
Fragmentation * VolQ50					0.0033 (0.0333)

Table 7.58: Fragmentation and volatility interaction models with one control variable, 30-minute data.

59. Fragmentation and volatility interaction models with one control variable, daily data

Variable	Model 46	Model 47	Model 48	Model 49	Model 50
Constant	2.8899*** (0.3992)	2.8107*** (0.4019)	1.7215*** (0.3938)	1.7913*** (0.3939)	2.0605*** (0.3917)
Fragmentation	0.148*** (0.044)	0.1597*** (0.044)	0.1562*** (0.0436)	0.1579*** (0.0436)	0.1512*** (0.0436)
VolQ50	0.1974*** (0.0617)	0.2083*** (0.0616)	0.2499*** (0.061)	0.2507*** (0.061)	0.2103*** (0.0644)
Fragmentation * VolQ50	-0.034 (0.0521)	-0.0395 (0.052)	-0.04 (0.0516)	-0.0399 (0.0516)	-0.0299 (0.0516)
Midquote ret vol					0.0653** (0.0289)

Table 7.59: Fragmentation and volatility interaction models with one control variable, daily data.

60. Fragmentation and volatility interaction models with one control variable, 30-minute data

Variable	Model 46	Model 47	Model 48	Model 49	Model 50
Constant	-1.8253*** (0.0716)	-1.8197*** (0.0716)	-2.5537*** (0.0766)	-2.5273*** (0.0766)	-2.1188*** (0.0744)
Fragmentation	-0.1116*** (0.0262)	-0.1094*** (0.0263)	-0.1146*** (0.0261)	-0.1153*** (0.0261)	-0.1226*** (0.0262)
VolQ50	0.3457*** (0.0329)	0.3422*** (0.0329)	0.3453*** (0.0328)	0.3441*** (0.0329)	0.3062*** (0.0335)
Fragmentation * VolQ50	-0.03 (0.0332)	-0.0303 (0.0332)	-0.0193 (0.0331)	-0.0203 (0.0331)	-0.0086 (0.0332)
Midquote ret vol					0.078*** (0.009)

Table 7.60: Fragmentation and volatility interaction models with one control variable, 30-minute data.

61. Fragmentation and crossed volatility interaction models with one control variable, daily data

Variable	Model 51	Model 52	Model 53	Model 54	Model 55
Constant	2.2909*** (0.3893)	3.0328*** (0.3988)	3.057*** (0.4029)	1.9279*** (0.3917)	1.9824*** (0.3912)
Fragmentation	0.1408*** (0.0334)	0.1442*** (0.0337)	0.1598*** (0.0338)	0.1395*** (0.0335)	0.1414*** (0.0335)
Crossed midquote ret vol	0.0464 (0.0454)	0.073* (0.0443)	0.0797* (0.0443)	0.0881** (0.0441)	0.0881** (0.0441)
Fragmentation * Crossed midquote ret vol	-0.0743** (0.0297)	-0.0784*** (0.0298)	-0.0783*** (0.0298)	-0.0744** (0.0298)	-0.0738** (0.0298)

Table 7.61: Fragmentation and crossed volatility interaction models with one control variable, daily data.

62. Fragmentation and crossed volatility interaction models with one control variable, 30-minute data

Variable	Model 51	Model 52	Model 53	Model 54	Model 55
Constant	-1.9928*** (0.0719)	-1.7556*** (0.0732)	-1.7716*** (0.0731)	-2.5071*** (0.0765)	-2.4825*** (0.0764)
Fragmentation	-0.1093*** (0.0188)	-0.1003*** (0.0188)	-0.1002*** (0.0189)	-0.1241*** (0.0187)	-0.1244*** (0.0187)
Crossed midquote ret vol	0.1392*** (0.0167)	0.2149*** (0.0161)	0.2152*** (0.0162)	0.1999*** (0.0161)	0.1986*** (0.016)
Fragmentation * Crossed midquote ret vol	0.0454*** (0.0113)	0.0579*** (0.0114)	0.0576*** (0.0115)	0.0534*** (0.0114)	0.0527*** (0.0114)

Table 7.62: Fragmentation and crossed volatility interaction models with one control variable, 30-minute data.

63. Fragmentation and crossed volatility interaction models with one control variable, daily data

Variable	Model 56	Model 57	Model 58	Model 59	Model 60
Constant	2.2286*** (0.3889)	2.301*** (0.3945)	3.0165*** (0.4048)	3.0489*** (0.408)	1.9091*** (0.3974)
Fragmentation	0.1352*** (0.0334)	0.1608** (0.0695)	0.1825*** (0.0703)	0.1842*** (0.0701)	0.1691** (0.0696)
Crossed midquote ret vol	0.099** (0.0441)				
Fragmentation * Crossed midquote ret vol	-0.0721** (0.0299)				
Crossed VolQ1		-0.1794* (0.0927)	-0.1938** (0.0937)	-0.2228** (0.0936)	-0.184** (0.0927)
Crossed VolQ2		-0.0341 (0.0821)	-0.0375 (0.083)	-0.045 (0.0828)	-0.0412 (0.0821)
Crossed VolQ4		0.1289 (0.0819)	0.1406* (0.0827)	0.1174 (0.0824)	0.1358* (0.0819)
Crossed VolQ5		0.067 (0.1026)	0.1295 (0.1019)	0.1305 (0.1017)	0.1442 (0.1011)
Fragmentation * Crossed VolQ1		0.178** (0.082)	0.1521* (0.083)	0.1633** (0.0829)	0.1687** (0.082)
Fragmentation * Crossed VolQ2		-0.1537* (0.0866)	-0.1589* (0.0874)	-0.1475* (0.0874)	-0.1629* (0.0867)
Fragmentation * Crossed VolQ4		-0.1141 (0.0894)	-0.1336 (0.0904)	-0.1139 (0.09)	-0.1171 (0.0896)
Fragmentation * Crossed VolQ5		-0.1326 (0.0918)	-0.1798* (0.0924)	-0.1552* (0.0921)	-0.1804** (0.0918)

Table 7.63: Fragmentation and crossed volatility interaction models with one control variable, daily data.

64. Fragmentation and crossed volatility interaction models with one control variable, 30-minute data

Variable	Model 56	Model 57	Model 58	Model 59	Model 60
Constant	-2.0432*** (0.0719)	-1.9433*** (0.0745)	-1.5998*** (0.0752)	-1.6154*** (0.0752)	-2.3842*** (0.0787)
Fragmentation	-0.1268*** (0.0187)	-0.096** (0.0417)	-0.0643 (0.0422)	-0.0702* (0.0422)	-0.0977** (0.0419)
Crossed midquote ret vol	0.213*** (0.0161)				
Fragmentation * Crossed midquote ret vol	0.0586*** (0.0114)				
Crossed VolQ1		-0.2474*** (0.0555)	-0.2899*** (0.0556)	-0.2867*** (0.0556)	-0.2488*** (0.0555)
Crossed VolQ2		-0.0723 (0.0474)	-0.0868* (0.0475)	-0.0893* (0.0474)	-0.0773 (0.0474)
Crossed VolQ4		0.1886*** (0.0447)	0.2047*** (0.0448)	0.201*** (0.0448)	0.1869*** (0.0447)
Crossed VolQ5		0.4615*** (0.0485)	0.5571*** (0.0479)	0.5548*** (0.0479)	0.5464*** (0.0478)
Fragmentation * Crossed VolQ1		-0.0179 (0.0554)	-0.0398 (0.0561)	-0.0314 (0.0561)	-0.0305 (0.0555)
Fragmentation * Crossed VolQ2		0.0048 (0.0559)	-9e-04 (0.0566)	0.009 (0.0566)	8e-04 (0.0561)
Fragmentation * Crossed VolQ4		0.009 (0.0544)	2e-04 (0.055)	0.0032 (0.055)	0.0127 (0.0546)
Fragmentation * Crossed VolQ5		-0.0335 (0.0526)	-0.1108** (0.0525)	-0.1028* (0.0525)	-0.0933* (0.0522)

Table 7.64: Fragmentation and crossed volatility interaction models with one control variable, 30-minute data.

65. Fragmentation and crossed volatility interaction models with one control variable, daily data

Variable	Model 61	Model 62	Model 63	Model 64	Model 65
Constant	1.9653*** (0.3968)	2.2545*** (0.394)	2.2829*** (0.3894)	3.0148*** (0.3989)	3.0385*** (0.403)
Fragmentation	0.168** (0.0696)	0.1549** (0.0695)	0.1509*** (0.0346)	0.1546*** (0.0349)	0.1694*** (0.035)
Crossed VolQ1	-0.1835** (0.0927)	-0.1293 (0.0941)			
Crossed VolQ2	-0.041 (0.0821)	-0.0181 (0.0824)			
Crossed VolQ4	0.1361* (0.0819)	0.089 (0.0827)			
Crossed VolQ5	0.1444 (0.1011)	-0.0201 (0.1116)			
Fragmentation * Crossed VolQ1	0.1706** (0.082)	0.1803** (0.0819)			
Fragmentation * Crossed VolQ2	-0.1592* (0.0867)	-0.1474* (0.0866)			
Fragmentation * Crossed VolQ4	-0.1121 (0.0895)	-0.1112 (0.0894)			
Fragmentation * Crossed VolQ5	-0.1765* (0.0918)	-0.1083 (0.0923)			
Crossed midquote ret vol		0.1477*** (0.0386)			
Crossed VolQ10			0.0474 (0.1042)	0.1584 (0.1018)	0.1852* (0.1012)
Fragmentation * Crossed VolQ10			-0.1187 (0.0911)	-0.1426 (0.0912)	-0.1376 (0.091)

Table 7.65: Fragmentation and crossed volatility interaction models with one control variable, daily data.

66. Fragmentation and crossed volatility interaction models with one control variable, 30-minute data

Variable	Model 61	Model 62	Model 63	Model 64	Model 65
Constant	-2.3611*** (0.0786)	-2.0519*** (0.0767)	-1.8357*** (0.0688)	-1.4938*** (0.0696)	-1.5101*** (0.0696)
Fragmentation	-0.0975** (0.0418)	-0.1065** (0.0418)	-0.1007*** (0.0199)	-0.0834*** (0.0201)	-0.0831*** (0.0201)
Crossed VolQ1	-0.2483*** (0.0555)	-0.2069*** (0.0557)			
Crossed VolQ2	-0.0766 (0.0474)	-0.0616 (0.0474)			
Crossed VolQ4	0.1874*** (0.0447)	0.171*** (0.0448)			
Crossed VolQ5	0.5438*** (0.0478)	0.4391*** (0.05)			
Fragmentation * Crossed VolQ1	-0.0308 (0.0555)	-0.0219 (0.0555)			
Fragmentation * Crossed VolQ2	0.0018 (0.0561)	0.0094 (0.056)			
Fragmentation * Crossed VolQ4	0.0114 (0.0546)	0.0052 (0.0545)			
Fragmentation * Crossed VolQ5	-0.0952* (0.0522)	-0.0489 (0.0526)			
Crossed midquote ret vol		0.0993*** (0.0117)			
Crossed VolQ10			0.4456*** (0.0465)	0.5894*** (0.0451)	0.5884*** (0.0451)
Fragmentation * Crossed VolQ10			-0.0354 (0.0479)	-0.1176** (0.0469)	-0.118** (0.0468)

Table 7.66: Fragmentation and crossed volatility interaction models with one control variable, 30-minute data.

67. Fragmentation and crossed volatility interaction models with one control variable, daily data

Variable	Model 66	Model 67	Model 68	Model 69	Model 70
Constant	1.8958*** (0.392)	1.9515*** (0.3916)	2.2164*** (0.3887)	2.2716*** (0.3894)	3.0021*** (0.3989)
Fragmentation	0.1506*** (0.0347)	0.1525*** (0.0347)	0.1494*** (0.0346)	0.1735*** (0.037)	0.1808*** (0.0373)
Crossed VolQ10	0.1762* (0.1015)	0.1767* (0.1014)	-0.0387 (0.1126)		
Fragmentation * Crossed VolQ10	-0.1601* (0.091)	-0.1598* (0.0909)	-0.0817 (0.0917)		
Crossed midquote ret vol			0.1656*** (0.0385)		
Crossed VolQ25				0.0799 (0.0788)	0.1226 (0.0783)
Fragmentation * Crossed VolQ25				-0.1418** (0.0647)	-0.1707*** (0.0649)

Table 7.67: Fragmentation and crossed volatility interaction models with one control variable, daily data.

68. Fragmentation and crossed volatility interaction models with one control variable, 30-minute data

Variable	Model 66	Model 67	Model 68	Model 69	Model 70
Constant	-2.2808*** (0.0735)	-2.2581*** (0.0734)	-1.9745*** (0.072)	-1.933*** (0.0696)	-1.6002*** (0.0704)
Fragmentation	-0.1075*** (0.02)	-0.1078*** (0.02)	-0.1149*** (0.02)	-0.0969*** (0.0219)	-0.0741*** (0.0221)
Crossed VolQ10	0.591*** (0.0452)	0.5863*** (0.0452)	0.4273*** (0.0482)		
Fragmentation * Crossed VolQ10	-0.1108** (0.0465)	-0.1106** (0.0465)	-0.0289 (0.048)		
Crossed midquote ret vol			0.1129*** (0.0124)		
Crossed VolQ25				0.4494*** (0.0361)	0.531*** (0.0355)
Fragmentation * Crossed VolQ25				-0.0282 (0.0363)	-0.0903** (0.0359)

Table 7.68: Fragmentation and crossed volatility interaction models with one control variable, 30-minute data.

69. Fragmentation and crossed volatility interaction models with one control variable, daily data

Variable	Model 71	Model 72	Model 73	Model 74	Model 75
Constant	3.0205*** (0.4029)	1.8783*** (0.3919)	1.9341*** (0.3913)	2.2184*** (0.3888)	2.2103*** (0.3924)
Fragmentation	0.1931*** (0.0374)	0.1765*** (0.037)	0.1778*** (0.037)	0.1695*** (0.0369)	0.1885*** (0.0418)
Crossed VolQ25	0.1282 (0.0782)	0.1358* (0.0777)	0.1349* (0.0777)	-0.0158 (0.0864)	
Fragmentation * Crossed VolQ25	-0.16** (0.0648)	-0.1796*** (0.0647)	-0.1771*** (0.0646)	-0.1169* (0.0653)	
Crossed midquote ret vol				0.1615*** (0.0373)	
Crossed VolQ50					0.1402** (0.0658)
Fragmentation * Crossed VolQ50					-0.1133** (0.0556)

Table 7.69: Fragmentation and crossed volatility interaction models with one control variable, daily data.

70. Fragmentation and crossed volatility interaction models with one control variable, 30-minute data

Variable	Model 71	Model 72	Model 73	Model 74	Model 75
Constant	-1.6179*** (0.0704)	-2.3764*** (0.0741)	-2.3532*** (0.074)	-2.0569*** (0.072)	-2.0528*** (0.0729)
Fragmentation	-0.0748*** (0.0221)	-0.102*** (0.0219)	-0.1018*** (0.0219)	-0.1089*** (0.022)	-0.1051*** (0.0261)
Crossed VolQ25	0.5287*** (0.0355)	0.5151*** (0.0354)	0.5139*** (0.0354)	0.4121*** (0.0377)	
Fragmentation * Crossed VolQ25	-0.0872** (0.0359)	-0.0771** (0.0357)	-0.0786** (0.0357)	-0.0362 (0.0363)	
Crossed midquote ret vol				0.1107*** (0.0119)	
Crossed VolQ50					0.334*** (0.0335)
Fragmentation * Crossed VolQ50					-0.0069 (0.0335)

Table 7.70: Fragmentation and crossed volatility interaction models with one control variable, 30-minute data.

71. Fragmentation and crossed volatility interaction models with one control variable, daily data

Variable	Model 76	Model 77	Model 78	Model 79	Model 80
Constant	2.9311*** (0.4019)	2.9522*** (0.4058)	1.8122*** (0.3947)	1.8664*** (0.3942)	2.1867*** (0.3923)
Fragmentation	0.1928*** (0.0421)	0.2064*** (0.0423)	0.1879*** (0.0418)	0.1892*** (0.0418)	0.1853*** (0.0418)
Crossed VolQ50	0.165** (0.0661)	0.1627** (0.0661)	0.1693*** (0.0655)	0.1696*** (0.0655)	0.0743 (0.0695)
Fragmentation * Crossed VolQ50	-0.1257** (0.056)	-0.122** (0.0559)	-0.1293** (0.0555)	-0.1278** (0.0555)	-0.1001* (0.0558)
Crossed midquote ret vol					0.1514*** (0.0347)

Table 7.71: Fragmentation and crossed volatility interaction models with one control variable, daily data.

72. Fragmentation and crossed volatility interaction models with one control variable, 30-minute data

Variable	Model 76	Model 77	Model 78	Model 79	Model 80
Constant	-1.7321*** (0.0734)	-1.7475*** (0.0734)	-2.4982*** (0.0771)	-2.4748*** (0.077)	-2.1799*** (0.0747)
Fragmentation	-0.0854*** (0.0264)	-0.084*** (0.0264)	-0.1141*** (0.0261)	-0.1138*** (0.0261)	-0.1185*** (0.0262)
Crossed VolQ50	0.3948*** (0.0333)	0.3913*** (0.0333)	0.3659*** (0.0332)	0.3647*** (0.0333)	0.283*** (0.0346)
Fragmentation * Crossed VolQ50	-0.0403 (0.0335)	-0.0434 (0.0335)	-0.0312 (0.0333)	-0.0324 (0.0333)	-0.008 (0.0336)
Crossed midquote ret vol					0.1304*** (0.0123)

Table 7.72: Fragmentation and crossed volatility interaction models with one control variable, 30-minute data.