SSE Riga Student Research Papers
2018 : 6 (205)

WHY DO SOME STOCK MARKETS FRAGMENT MORE THAN OTHERS? EVIDENCE FROM THE CHI-X TRADING PLATFORM

Authors: Hanna Bandarchyk
Artsiom Kazubouski

ISSN 1691-4643
ISBN 978-9984-822-

November 2018
Riga

Why Do Some Stock Markets Fragment More Than Others? Evidence from the Chi-X Trading Platform

Hanna Bandarchyk

and

Artsiom Kazubouski

Supervisor: Tālis J. Putniņš

November 2018
Riga

Table of Contents

Abstract

Over the past two decades, the structure of financial markets has changed dramatically due to technological development and liberating regulatory programs such as the Markets in Financial Instruments Directive (MiFID). These changes eliminated the monopolies that were enjoyed by traditional stock exchanges. Hence, increased competition was introduced among stock trading venues, ultimately leading equity trading to fragment across multiple venues. While the effects of fragmentation have been examined in a growing recent literature, little is known about why some markets fragment much more than others (some markets remaining unfragmented), and some stocks fragment much more than others even within the same market. We aim to fill this gap and provide evidence on the stock- and market-level drivers of fragmentation. We use Heckman two-stage selection model in which the first stage models the probability that a market or stock is added to a major competing equity venue (Chi-X) and the second stage models the determinants of the degree of fragmentation, accounting for the fact that the opportunity for fragmentation (the entry of Chi-X) is endogenous. We find that stocks or markets with higher market capitalization, larger minimum price increments (tick sizes), lower algorithmic trading activity and lower size of average trade tend to fragment more, while stocks with lower volatility, lower spread, higher traded volume, and index constituents are more likely to be added to the stocks traded on Chi-X.

Keywords: equity market fragmentation, alternative trading venue, multilateral trading facility, algorithmic trading, sample selection bias

Acknowledgements

We express our deep gratitude to our supervisor Tālis Putniņš for his invaluable support and suggestions throughout the whole process of writing our thesis. We are also thankful to Marta Khomyn for sharing her immersive knowledge and supporting us in our research endeavours.

1. Introduction

Over the past few decades, equity markets have undergone major changes due to technological development and liberating regulations. The second half of the 20th century was marked by the revolutionary shifts in the security trading, when traditional face-to-face or open outcry systems were gradually replaced with electronic trading platforms. Initially, these platforms were run by traditional stock exchanges. Further advancement in communication technology, which started in the 1990s, has brought new trading venues that became the main competitors to traditional stock exchanges. The volume of stocks, previously traded on stock exchanges, started to split across multiple venues due to the numerous advantages that new entrants offered, such as lower fees and transparency, higher speed, different execution strategies. Thus, competition over the order flow of stocks has led to an upturn in equity market fragmentation.

Technological development was accompanied by liberating regulations in the sphere of equity trading, such as the Regulation-National Market System (RegNMS) established in 2005 in the USA and the Markets in Financial Instruments Directive (MiFID) introduced in 2007 in the EU. Among all the changes brought by them, the one which impacted security trading and fragmentation was the abolition of "concentration rule", according to which all the trading in equities should be executed only on regulated markets. The goal was to foster the competitiveness of financial markets and stimulate efficient price formation. Consequently, it has also facilitated the development and expansion of new trading venues, such as Chi-X, BATS and Turquoise.

Therefore, nowadays markets are far from being consolidated and comprehensible, but rather consist of numerous entities such as traditional exchanges, electronic communication networks, alternative trading venues, and established financial companies, all of which try to capture the largest market share. Currently on the European equities market, London Stock Exchange Group and Euronext comprise the highest total market share of about 35%, while Cboe Europe Equities representing the alternative venues such as Chi-X and Bats Europe comprise the second largest share – more than 20% (Cboe Global Markets, 2017).

These dramatic changes in the equity markets stimulated research on this topic, which resulted in the emergence of a large body of literature and findings. While the effects of fragmentation have been examined in a growing recent literature (e.g.

Gomber, Sagade, Theissen, Weber, & Westheide, 2017), little is known about why some markets fragment much more than others (some markets remaining unfragmented), and some stocks fragment much more than others even within the same market. According to LiquidMetrix (2010), even stocks included in the same index vary in their trading volume across different venues. For example, in 2010, LSE had a market share in FTSE 100 stocks ranging from 50% to 80%, while the share of BATS in the same index varied from 2% to 15%. On the same note, looking at DAX stocks, it can be noted that Deutsche Borse Xetra had a market share of 37% for Siemens and 75% for Commerzbank in 2010. These examples demonstrate that there are some stock and market characteristics leading to higher degree of fragmentation for some stocks compared with others in one index. The reasons for such variability of stock fragmentation has been little studied and is of particular interest for trading venues, market participants, as well as for regulators. According to Gomber et al. (2017), market fragmentation leads to an increase in execution complexity and makes the choice of the appropriate regulatory policy in such conditions challenging.

We aim to fill the gap in the existing research by providing evidence on the stock- and market-level drivers of fragmentation. We study the European equity market during the period of 2007-2009, using the staggered roll-out of the Chi-X trading venue across 12 European countries. Chi-X Europe serves as a relevant case for the study of fragmentation in Europe because (i) it entered the European market 18 months before other alternative venues and was the only venue competing with traditional exchanges for that period of time, thus being the primary driver of fragmentation, (ii) it succeeded in capturing more than 13% of equity trading in Europe by the end of the second year of operations, thus significantly enhancing equity fragmentation, and (iii) its technological advancement and smart order routing systems facilitated the split of trade across numerous venues. Thus, to reach our research goals, we intend to answer the following research questions:

RQ1: What types of stocks are more prone to trade on Chi-X Europe?

RQ2: What determines the overall level of fragmentation in a stock?

To answer these questions, we use Heckman two-stage selection model in which the first stage models the probability that a stock is added to the Chi-X platform and the second stage models the determinants of the degree of fragmentation in a stock, accounting for the fact that the opportunity for fragmentation (the entry of Chi-X) is endogenous. While in a classical Heckman model only the second stage is of primary interest, we use information from both stages to answer each of the research questions. The first stage informs about market operator choices for selecting stocks to trade on Chi-X, thus providing evidence on characteristics of stocks that are more likely to be included in Chi-X. The second stage notifies about market participant's decisions on what and where to trade. Both market operator's and market participant's choices are required for fragmentation to occur.

This paper contributes to the existing body of literature in the following ways. Firstly, this study adds to the scarce evidence on the determinants of equity fragmentation by studying the case of the entrance of Chi-X to the European equity market. Secondly, this research accounts for endogeneity of a stock's introduction to Chi-X using Heckman two stage selection model, which was not used before in the studies on fragmentation. Our results show the presence of selection bias indicated by statistically significant Inverse Mills Ratio, suggesting that coefficients will be biased without Heckman correction procedure.

Our findings suggest that stocks with higher trading volume, lower spread, lower volatility and index constituents are more likely to enter Chi-X Europe. Regarding the overall level of fragmentation, we find that stocks with higher market capitalization, lower price, and larger minimum price increments (tick sizes) tend to fragment more. Moreover, we find that stocks traded more by retail traders and stocks with lower algorithmic trading have higher degree of fragmentation.

The results of our study may be valuable for market participants, market operators, regulators, as well as for founders and owners of stock exchanges and alternative trading facilities. Management of different trading venues and exchanges could modify their market structure or stock selecting strategy based on stock- and market-related factors that foster fragmentation. Thus, it would enable them to compete successfully against other rival venues. Moreover, this paper may be helpful for financial regulators in choosing an appropriate regulatory policy that affects market structure and competition.

The rest of the paper is structured as follows. Section 2 provides institutional and the European equity market overview as well as outlines the importance of the Chi-X case for the study on fragmentation. Section 3 presents the literature on the pros and cons of consolidation and fragmentation and reviews the main literature on the equity market fragmentation. Section 4 provides a description of sample construction approach and the list of data sources. Section 5 presents definitions of the variables used in the research and descriptive statistics. Section 6 describes the research design and specifications of the model. Section 7 provides regression results and their discussion. Section 8 includes robustness check of the obtained results. Section 9 concludes.

2. Institutional details

2.1 Overview of Alternative Trading Systems

Alternative trading systems (ATSs) are non-exchange trading venues such as Chi-X, Turquoise, NASDAQ OMX, and BATS. ATSs are classified into three main categories: electronic communication networks (ECNs), crossing networks, and dark pools (OECD, 2016). Electronic communication networks are a fully electronic category of ATSs, which match buyers and sellers by means of exchange. Crossing networks are similar to ECNs, but they don't use exchange as an intermediary to match buyers and sellers and perform it automatically (Conrad, Johnson, & Wahal, 2003). Dark pools are private exchanges or online forums, where buyers and sellers of securities can identify each other and perform a trade without a public announcement of their quotes.

Alternative Trading Systems have some features that distinguish them from traditional exchanges, while they also share certain similarities. Alternative venues are usually managed by private entities, such as investment banks, firms, while exchanges operate most commonly in the form of a public company. Only traditional exchanges offer listing services for the companies, whereas ATSs provide an opportunity for traders to perform operations with securities that are already listed on one or several exchanges (Di Febo & Angelini, 2015). Thus, ATSs organize trading of securities between the parties similar to exchanges, but they cannot affirm the admission to trading for these securities. In addition to that, different rules and procedures for authorizing transactions are applied for ATSs and regulated markets. Another important feature of ATSs is their lack of transparency: they are not obliged to report their trading services, operations, and fees, which also creates an advantage for traders compared with traditional exchanges (OECD, 2016).

2.2 Overview of the European equity market

In Europe, ATSs are called Multilateral Trading Facilities (MTFs), which have grown in the popularity after the adoption of MiFID in 2007. MTFs are defined by MiFID as venues which connect the retail and other multiple parties to trade in the system under non-discretionary rules (Official Journal of the European Union, 2014). Nowadays, Europe experiences the flow of stock trading volume from traditional stock exchanges such as London Stock Exchange or Euronext to MTFs.

Currently, the largest MTFs by market share and value traded in Europe are BATS Europe and Chi-X Europe held by a parent company Cboe Global Markets. Currently, these MTFs capture more than 20% of European equity market (Figure 1).

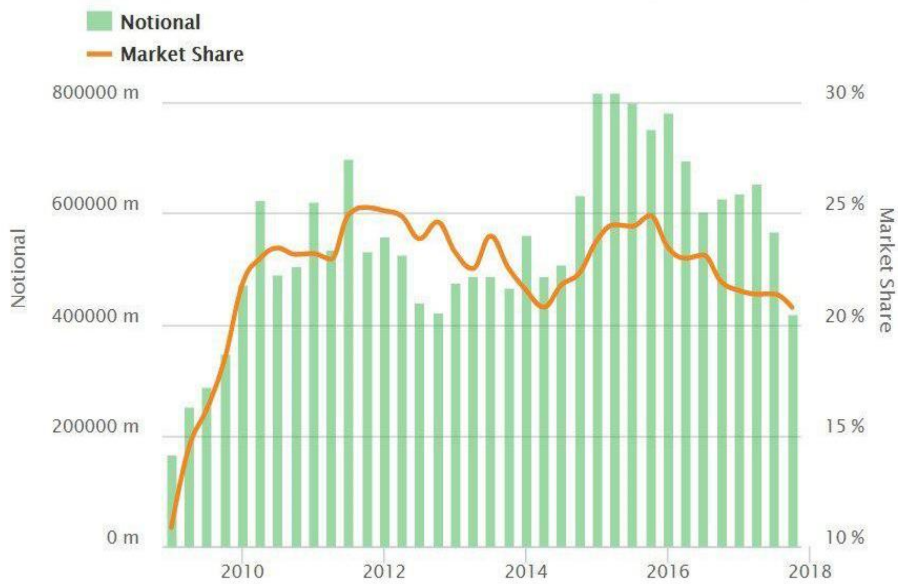

Figure 1. Traded volume on Cboe in EUR and its market share in Europe based on traded volume. Source: Cboe Global Markets, 2017.

Cboe Global Markets is an exchange holding company, which operates across the USA and Europe and offers trading in a wide range of products covering various asset classes and geographies. Its European part, Cboe European equities, is presented by two trading venues: BATS Europe and Chi-X Europe, with more than 70% of volume trading on the latter (Cboe Global Markets, 2017). During the last decade, Chi-X Europe has succeeded in attracting order flow in equities and outperformed traditional exchanges, such as LSE and Euronext, which dominated the market for a long time. On the Figure 2, it can be seen that the market shares of LSE Group and Euronext have been steadily decreasing since 2009 and dropped by around 8% for the former and 7% for the latter. At the same time, Chi-X was confidently strengthening its position on the European equity market and managed to increase its market share by almost 20%.

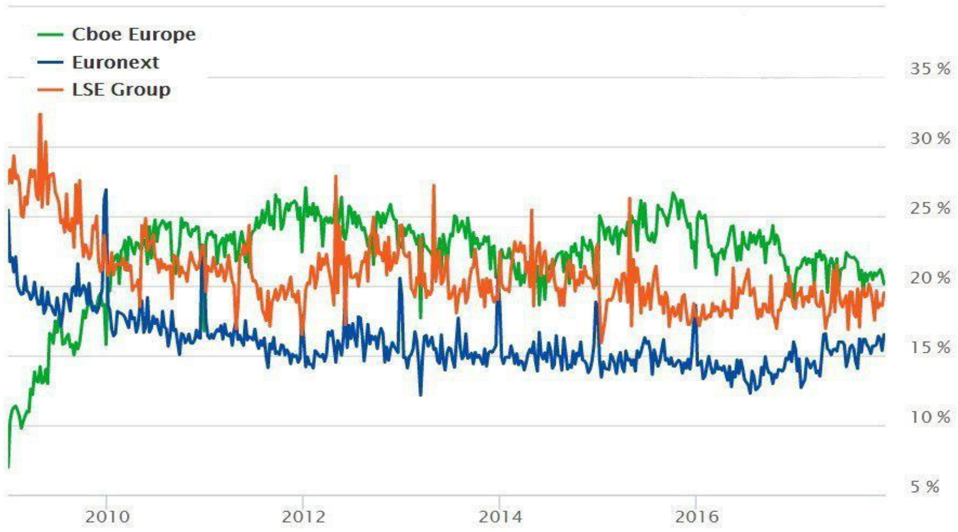

Figure 2. Market shares of the main trading venues in Europe (LSE Group, Euronext, and Cboe Europe) in the period of 2008-2018. Source: Cboe Global Markets, 2017.

All this suggests that alternative trading venues can successfully compete with established traditional exchanges and impact equity markets. Therefore, Chi-X Europe is of particular interest for our research, as it is a prominent example of the venue that managed to become the leading trading platform in equities in such a short period.

2.3 Chi-X Europe

Chi-X is an alternative trading venue which has successfully competed with traditional exchanges across Europe and worldwide. There are several factors that make this venue the most representative case for this study. Chi-X Europe was the first successful MTF in the European region. It was launched in March 2007, seven months prior to MiFID (November 2007) and eighteen months prior to other MTFs launch in Europe (Chi-X Europe, 2011a). For that period, it was the only venue competing with traditional exchanges, thus being the primary driver of fragmentation. Furthermore, this favourable timing gave an advantage for Chi-X to start capturing market share in equity trading. Then due to its competitive advantages such as lower trading costs, faster execution, and new instruments and products added continuously, Chi-X Europe became a leader among MTFs in Europe.

Chi-X is a prominent example of a successful alternative trading venue, which was able to become the largest MTF in Europe and gain a market share close to the ones of the leading exchanges during the first years of its expansion. From the beginning of its operation, Chi-X Europe gradually started capturing order flows gradually while entering new European markets. Initially, the venue was based on a platform that traded only Dutch and German stocks. However, during a two-year period from its launch, the trading venue spread across 13 European countries: Austria, Belgium, Denmark, France, Finland, Germany, Italy, the Netherlands, Norway, Spain, Sweden, Switzerland, and the UK (the illustration of the Chi-X entry in the European markets is provided in Appendix A). After the first year of its operations, Chi-X Europe captured almost 5% of all trades in Europe, having particularly high share in Dutch stocks (13.6%) (Menkveld, 2013). Thus, the entry of Chi-X shattered the monopoly of traditional exchanges and considerably stimulated equity fragmentation by attracting order flow to the platform.

Chi-X was competing with traditional exchanges using a business model of low fees and high speed. When the platform was launched in 2007, the main European stock exchanges charged trading fees exceeding 0.5 bps for each side of the trade. On the contrary, Chi-X Europe employed a maker-taker fee structure: an order that is executed immediately and takes out liquidity from the market was charged 0.3 bps, while a passive order, which increases liquidity, received a rebate of 0.2 bps (He, Jarnecic, & Liu, 2015). Thus, the platform provided an incentive for liquidity providers to enter the market. In addition to favorable fee structure, Chi-X Europe was advantageous due to employing systems supporting high speed transactions. In 2008, Chi-X announced that the speed on the platform is up to 10 times higher than on the fastest European primary exchange (Menkveld, 2013). Moreover, Chi-X Europe employed smart order routing system, which enabled traders to route their orders to the destination with the best execution. Such system also facilitated a split of trading across multiple venues.

Due to these factors, Chi-X reached the second position by volume traded among equity exchanges in Europe in 2010 with the value traded exceeding EUR 1.58 billion (Chi-X Europe, 2011b). These successes led to the negotiations with BATS Europe throughout 2010 and later in February 2011, Chi-X Europe was acquired by BATS. In 2017, after a new round of negotiations BATS was sold to Cboe Global Markets (Cboe Global Markets, n.d.) and currently Chi-X Europe operates under the Cboe Europe Equities brand.

3. Literature Review and Hypotheses

In this section, we present the literature on the debate around fragmentation and consolidation, showing advantages and disadvantages of both forms of market organization, review the major studies on equity market fragmentation, as well as outline factors that can cause equity market to fragment.

3.1 Fragmentation and consolidation: a comparative review

Researchers express opposing views regarding the most beneficial way of trading: fragmented across multiple venues or consolidated on a single one. According to O'Hara and Ye (2011), the factors usually taken into account in this analysis are related to trading design (particularly, fee structure of the markets and network effects) and the effects of competition on the market. On the one hand, consolidation is beneficial due to an opportunity to decrease significant set-up and operational costs of the venue by splitting them on a large number of shares, so that costs per share are much smaller for large trading venues than for smaller ones. In addition to that, large venues provide positive network externality in the form of greater ability to match buyers and sellers, which consequently leads to further reduction in trading costs. On the other hand, consolidation may result in the creation of monopolies that harm market participants by acting non-competitively. In that case, fragmentation is favored for increasing competition, and thus stimulating trading costs to decline.

The debate around the advantages and disadvantages of consolidation and fragmentation is reflected in numerous research papers. For a long time, traditional stock exchanges were the only means to trade and were considered to have monopolistic power due to economies of scale and network externalities. For example, Stigler (1964) argues that economies of scale on traditional exchanges result in efficient price setting of securities. He claims that if a large number of transactions in trading of a particular equity are concentrated on one stock exchange, the price of this equity is set solely on this exchange. Similarly, Bloch and Schwartz (1978) argue that although monopoly power of traditional stock exchanges harm investors through high fees, it provides an advantage in the form of price efficiency. One more argument in favor of consolidation is developed by Pagano (1989), Chowdhry, and Nanda (1991). They assert that in the presence of two markets, orders naturally concentrate on one, more liquid market, due to the fact that investors benefit from higher liquidity, as it results in better trading terms. High level of liquidity of the trading venue stimulates other investors to join it, which further fosters liquidity and increases benefits for investors. Consequently, it leads to the situation when all investors are concentrated on a single trading venue. Madhavan (1995) also points out disadvantages of fragmentation in the form of higher price volatility and other price distortions. He also claims that trade disclosure leads to consolidation and fragmentation occurs if disclosure rules are not mandatory. According to his model, benefits are distributed to less competitive dealers and investors who hide their trades.

Although consolidation of trading has several positive effects, the growing number of recent studies favor fragmentation and point out its advantages. In one of such studies, Economides (1996) asserts that benefits of monopolistic trading venue, such as network externalities, do not outweigh losses it imposes on the market participants and suggested that welfare improvement can be attained by switching to fragmentation. Harris (1993) argues that fragmentation results from different needs of investors and problems they are facing. Hendershott and Mendelson (2000) point out the benefits of fragmentation in the form of lower inventory risk of individual dealers. Moreover, it is considered that new trading venues are becoming globally consolidated nowadays, since they can employ sophisticated order routing technologies and consolidated tape which prevents liquidity to split over these venues, but rather creates a virtually consolidated market with multiple points of entry (O'Hara & Ye, 2011).

In order to test theoretical hypotheses regarding benefits and drawbacks of consolidation and fragmentation, numerous empirical studies are conducted. The focus of these studies is placed on the effects of fragmentation on trading costs, liquidity, and market efficiency. According to the research done by Gomber, Gsell, and Lutat (2011), equity market fragmentation is positively affecting the level of liquidity on the European market. In addition to that, the study by Fioravanti and Gentile (2011) supports the conclusion on the positive effect on liquidity. The authors consider that such effect can be explained by the proliferation of high frequency traders across the venues, which could lead to creation and attracting of trade flows. Battalio (1997) proves the increase in the liquidity on NYSE caused by a new dealer entering the market. The same effect on liquidity is found by Boehmer and Boehmer (2003) in their research on the effects of NYSE initiated trading ETFs listed on the American Stock

Exchange. The research by Foucault and Menkveld (2008) also supports the beneficial effect of fragmentation on liquidity by concluding that the increase of competition for Dutch stock led to higher liquidity measured by depth. In addition to the studies on liquidity, Fong, Madhavan, and Swan (2001) research the effects of fragmentation on trading costs. They show that large Australian stocks face lower transaction costs in case of off-exchange execution.

However, some empirical findings arrive at opposing conclusions on the effects of fragmentation. The study on the Australian stock market executed by Frino (2012) shows that fragmentation of stocks in the ASX 200 index resulted in the decrease of trading by around 10%, thus leading to the contraction of liquidity. Gajewski and Gresse (2007) in their study of European equity market's fragmentation identify that trading costs are lower on the consolidated market than on the market of numerous competing venues. Similar outcome is reached by Bennett and Wei (2006), who researched stocks moving between two venues: one being more consolidated (NYSE), and the other more fragmented (NASDAQ). They show that execution costs for a stock fell in case of switching from NASDAQ to NYSE. Amihud, Lauterbach, and Mendelson (2003) demonstrate on the basis of a warrant exercise that consolidation is more beneficial compared to fragmentation providing the evidence of its positive effects on liquidity. Thus, it can be stated that the existing theoretical and empirical evidence does not provide a clear answer on whether consolidated or fragmented market leads to higher market quality.

Overall, this comparative review between advantages and disadvantages of market fragmentation and consolidation allows us to frame our further research and identify potential variables for our empirical study. Moreover, it helps us to distinguish areas that are worth focusing on while looking for determinants of equity fragmentation and the characteristics of stocks which are more prone to fragment. Considering above mentioned literature, there are several important factors that need further attention, such as liquidity, transaction costs, price volatility, as well as price efficiency and discovery. In the next section, we continue to review literature on equity fragmentation focusing on these and some other factors in more detail.

3.2. Factors of equity market fragmentation

In this subsection, we review the studies connected to market fragmentation, market structure, and alternative trading. It allows us to frame hypotheses on which factors can influence stock inclusion on Chi-X and what determines the overall level of equity market fragmentation.

Since market fragmentation is one of major changes on financial markets in recent times, it has attracted much attention of researchers. The largest part of these studies is focused on the effect of market fragmentation on market quality. The summary of these findings can be found in a survey by Gomber et al. (2017). However, little is known about why some markets fragment more than others and some remain consolidated, as well as why some stocks fragment much more than others even within the same market. We use some findings from the studies on effects and reasons of market fragmentation in our research to identify the main factors that may cause stocks and equity markets to fragment. The summary of these findings can be found in Appendix B.

The main reasons of market fragmentation are claimed to be heterogeneous preferences of traders towards the characteristics of exchanges: trading fees, tick size, liquidity level, order sizes, etc. (Harris, 1993). In their research, Fong, Madhavan, and Swan (2001) study market fragmentation on the example of off-market trading and conclude that market fragmentation is driven by institutional trading interest (trading volume, indexation) and liquidity (bid-ask spread and market depth). The importance of liquidity as a factor of fragmentation is also suggested by Bennet and Wei (2006), who conclude that less liquid stocks benefit more from switching from a fragmented market to consolidated. In addition, they show that trading volume, the number of market makers, the exchange industry concentration index, and the daily return explain stocks switching decisions from more fragmented exchanges to a consolidated one.

Among other factors that may influence stock fragmentation is the size of a company**.** On the one hand, large companies are more likely to experience higher interest from the side of investors and traders. As a result, stocks with higher market capitalization tend to be more liquid and therefore, they are more prone to the relocation of their ample liquidity on other trading venues. On the other hand, as determined in the study by O’Hara and Ye (2011), fragmentation is more beneficial for small firms than for large ones. For example, they show that liquidity improves for small cap stocks with fragmentation, while it has no effect on large cap stocks. However, fragmentation causes execution speed of stocks with large market capitalization to decline, while leaves execution speed of small cap stocks unaffected.

In addition, Gajewski and Gresse (2007) in their study on trading costs on fragmented and consolidated markets identify several characteristics of stocks traded on each of the markets. Firstly, they conclude that price volatility is significantly higher for stocks on fragmented markets than on consolidated. It can be explained by lower liquidity and thus, larger spreads on fragmented markets. When volatility of the stock is high, it might be more beneficial for it to stay on the primary market with concentrated liquidity and order flow. On the other hand, increase in volatility for a stock can lead to price inefficiencies, and thus, create opportunities for traders to exploit them with cross-venue arbitrage, which positively affects the number of venues the stock is traded on. Secondly, Gajewski and Gresse (2007) show that the average size of transactions is larger for fragmented markets, which can be explained by the fact that institutional traders are subject to lower transaction costs for their large volume trades on alternative venues.

Besides that, HFT might be one of the important factors stimulating fragmentation, as proliferation of algorithmic trading (AT) gives an opportunity to employ strategies using different trading platforms (Hendershott, Jones, & Menkveld, 2011). Nowadays, equity trading is highly impacted by algorithms, which are used to exploit trading opportunities by constantly searching for them across a vast range of securities and trading venues. Moreover, one of the important conditions for the operation of trading algorithms is the ability of a trading venue to handle orders with high speed. Due to that, AT tends to concentrate on alternative trading venues, as they offer higher speed than traditional exchanges (Hendershott et al., 2011). In addition to that, presence of algorithms may influence the amount and type of traders a venue can attract. While institutional traders might prefer alternative trading venues for their speed of execution and anonymity, they may be reluctant to shift from the primary exchange, which offers more natural liquidity and order flow.

Composition of traders in the trading activity of a certain stock is also one of the factors that might stimulate certain stocks to fragment more. According to Lin, Michayluk, Oppenheimer, and Sabherwal (2009), retail investors are more sensitive to higher costs of trading in particular stocks and trading venues, while institutional investors are to a certain extent indifferent about which trading venue to choose.

Moreover, as minimal size of the trade for institutional investors is 10,000 shares, they tend to choose markets with higher liquidity to be able to execute large orders. In the study on the impact of decimalization on NYSE on the composition of traders, Lin et al. (2009) show that lower depth on the market stimulates institutional investors to switch to more liquid trading venues, where a large order can be executed without significant impact on the stock price. Therefore, it can be concluded that stocks traded by institutional investors are less likely to fragment, compared to the ones traded mostly by retail investors. However, the study by Barclay, Hendershott, and McCormick (2003) on the competition between trading venues in the USA shows that alternative trading venues attract institutional traders due to higher speed of execution and higher level of anonymity.

In addition, tick size can be an important indicator for market fragmentation. Tick size is the smallest price increment an equity or other security can make on an exchange. According to Verousis, Perotti, and Sermpinis (2018), tick size has an effect on market quality and market structure. They also show that regulations of tick size schedules have a tight link with the changes in high frequency trading activity. Moreover, Verousis et al. (2018) argue that the smaller tick sizes result in the narrower difference of bid-ask spread. Besides that, Hameed and Terry (1994), Anderson and Peng (2014) show that decrease in tick size have positive effect on liquidity of heavily traded stocks. The studies also demonstrate that when a tick size decreases, then both spreads and depths of trades follow it (Harris, 1994; Alampieski & Lepone, 2009; Van Ness, Van Ness & Pruitt, 2000).

Tick size schedule differed across trading venues in Europe in 2007, with the maximum number of 25 different schedules at one point. After MiFID was implemented, it led to the increase of competition in the financial markets. One of the drivers of venues' competition was lower tick sizes of new entrants in comparison to incumbent exchanges, which drove trading volume to alternative trading venues and fostered fragmentation (Meling & Ødegaard, 2017). This period is also characterized by proliferation of smart order routing, which shifted trades execution from incumbent exchanges with higher tick sizes to the MTFs with lower tick sizes. This fact also contributed to the rise of equity fragmentation. Therefore, tick size can be considered as one of the potential explanatory variables of the degree of fragmentation.

Considering all the theoretical and empirical evidence on the potential factors of equity fragmentation, we form hypotheses on the predicted effects of these variables. The list of hypotheses reflecting the expected effect of the variables on probability on trading on Chi-X and fragmentation can be found in Table 1.

Table 1. Hypotheses

This table provides definitions of eight hypotheses. The first part of the table states three hypotheses for the first research questions: what are the determinants of the probability to be traded on the Chi-X trading platform. The second part of the table includes five hypotheses for the second research question: what are the determinants of the overall level of fragmentation in a stock.

Variables	*Hypotheses*
	First stage
1. Liquidity	Less liquid stocks have lower probability to be traded on Chi-X.
2. Trading volume	Stocks with higher trading volume are more likely to be included in Chi-X.
3. Volatility	Stocks with higher volatility are less likely to trade on Chi- X.
	Second stage
4. Market capitalization	Stocks with higher market capitalization are more likely to fragment.
5. AT	Stocks attracting more AT activity are more likely to fragment.
6. Composition of traders	Stocks traded by institutional traders are less likely to fragment.
7. Tick size	Larger tick size on the home exchange will positively affect the degree of fragmentation.
8. Price	Higher stock prices result in lower degree of fragmentation.

4. Data and sample description

The sample period spans from February 1, 2007 till February 28, 2009, covering the entry of Chi-X in 13 European countries. It starts two months prior to the Chi-X entry in the European market and extends till the two months after Chi-X entered the Spanish market. Figure 3 (Appendix A) shows the sequential entry of Chi-X to each country in our sample. The trading on this venue launched with the opening of the platform in Germany and the Netherlands (April, 2007). The number of stocks traded significantly increased with the entry to the UK market and continued its growth till the end of our sample period. At the end of the timespan, 741 stocks were traded on this platform.

Our sample is structured in the following way. We include all stocks which are traded on the Chi-X trading venue during our sample period. Then, for each market we add top 75 stocks (for the United Kingdom we include top 150 stocks) with the largest aggregate trading volume in our sample period following the sample construction approach of Malceniece, Malcenieks, and Putniņš (2018). If stocks are also traded on the Chi-X platform, they are not added for the second time. The final sample includes 32,233 stock-months observations for 1,311 companies in 12 European countries: Austria, Belgium, Denmark, France, Finland, Germany, Italy, the Netherlands, Norway, Spain, Sweden and the UK (the list of countries is based on Chi-X trading platform's expansion during 2007-2009). Switzerland is omitted due to data deficiency. We exclude the effects of the financial crisis of 2008 (which is included in our sample period) by implementing time fixed effects approach.

We use Thomson Reuters Datastream as the primary source to compile our dataset. The Datastream provides us with monthly bid and ask quotes, number of trades, trading volume, and market capitalization for the stocks. To construct high frequency trading variables, we adopt the Data appendix to the paper by Malceniece et al. (2018). Historical tick sizes are derived using daily closing stock prices, and then verified by the data from public sources such as trading venues web-sites (e.g. Euronext, Nasdaq, Cboe Global Markets). The trading data is obtained for three markets: primary listing of a stock (home market), Chi-X trading venue (if this stock is traded there), and consolidated values across all markets. To construct consolidated records of trading activity for each stock, we combine the data from all the venues where this stock is traded. The data on different markets are used for the construction of fragmentation measures.

We also account for the difference in currency across the countries in our sample and convert all the necessary values to EUR. Weekends and national holidays are excluded from our sample.

5. Measures and descriptive statistics

In this section, we provide definitions and formulas of the major measures. Further, descriptive statistics for these measures and other variables in the sample is provided. We conclude this section by highlighting differences between Chi-X and non-Chi-X stocks based on descriptive statistics for each group.

5.1 Market fragmentation measures

As suggested by Malceniece et al. (2018), several measures of stock fragmentation are calculated to get robust results. The first measure, $Frag1_{i,t}$ reflects the number of venues, on which stock i was traded during the month t.

The second fragmentation measure is used by Bennet and Wei (2006), Degryse, de Jong, and van Kervel (2015) to analyze fragmentation in European equity market. This measure is represented by the Herfindahl-Hirschman Index (HHI) and is calculated as follows:

$$Frag2_{i,t} = 1 - \sum_{j} \left(\frac{dvol_{i,j,t}}{\sum_t dvol_{i,j,t}} \right)^2,$$

where $dvol_{i,j,t}$ is dollar traded volume of stock i on market j in month t.

$FRAG3_{i,t}$ is also based on the Herfindahl-Hirschman Index (HHI), but differs from $FRAG2_{i,t}$ by using the number of trades instead of trading volume.

$$Frag3_{i,t} = 1 - \sum_{j} \left(\frac{ntrades_{i,j,t}}{\sum_t ntrades_{i,j,t}} \right)^2,$$

where $ntrades_{i,j,t}$ is number of trades of stock i on market j in month t.

$FRAG4_{i,t}$ is calculated as the volume market share of all venues except for the home exchange. We perform our baseline test using the first measure, and other measures are used in the robustness check.

5.2 Liquidity measures

As the main liquidity measure, we use relative quoted bid-ask spread, that was frequently used in studies on fragmentation (Boneva, Linton, & Vogt, 2013; Bennet & Wei, 2006; He et al., 2015). This measure is expressed in bps and calculated as follows:

$$spread_{i,t} = 10{,}000\frac{ask_{i,t}-bid_{i,t}}{\frac{ask_{i,t}+bid_{i,t}}{2}},$$

where $ask_{i,t}$ and $bid_{i,t}$ are the ask and bid quotes in month t for stock i.

As an alternative measure of liquidity, we use illiquidity measure proposed by Amihud (2002), but we reverse it so as to get a measure of liquidity, $liq_{i,t}$ (similar to studies by Boneva et al. (2013); Malceniece et al. (2018)). This measure is an efficient proxy for price impact of trading volume and is highly correlated with other liquidity measures such as bid-ask spread and depth (Boneva et al., 2013). Liquidity measure is calculated as follows:

$$liq_{i,t} = -\log(1 + 1{,}000\frac{|r_{i,t}|}{dvol_{i,t}})),$$

where $|r_{i,t}|$ is the absolute midquote return in bps per 1,000 EUR traded volume of stock i in month t. We take the log value of the variable so as to make the distribution of the liquidity measure closer to normal. Moreover, minus sign reverses the interpretation of the estimate from illiquidity to liquidity. This measure is used in a robustness check of the validity of our results.

5.3 AT measures

Trading data does not allow to distinguish between an order placed by a person or through a computer algorithm. Therefore, to estimate the effect of HFT on equity fragmentation, we use a proxy for algorithmic trading similar to the one used by Hendershott et al. (2011) and Boehmer, Fong, and Wu (2015). The logic behind the first proxy is based on the fact that AT employs algorithms, which submit large amount of orders so as to exploit trading opportunities. Therefore, the activity of algorithms can be tracked by spotting abnormal message traffic. Messages sent to the trading venue include submissions, modifications, and cancelations of orders. However, it should be noted that stocks with higher trading activity attract higher flow of order messages. Therefore, we normalize message traffic proxy by the number of trades:

$$AT_{i,t} = \frac{messages_{i,t}}{trades_{i,t}},$$

where $messages_{i,t}$ is the number of messages for stock i in month t, and $trades_{i,t}$ is the number of trades for stock i in month t. We expect to get positive effect of AT measures on equity fragmentation.

5.4 Composition of traders

Composition of traders can be estimated using proprietary data, which indicates the type of investor, or by classifying trades by their size. We use the second option and proxy the prevalence of either type of investors by calculating the average trade size for each stock, similarly to Oppenheimer and Sabherwal (2003), Lin et al. (2009). Higher average trade size indicates larger proportion of institutional investors relative to retail investors, as it is commonly regarded that institutional traders post orders of 10,000 shares and more, while retail traders typically trade less than 500 shares (Oppenheimer & Sabherwal, 2003). Therefore, our measure of composition of traders reflects the average dollar size of the trade in each stock and calculated as following:

$$avg\_trade_{i,t} = \frac{dvol_{i,t}}{trades_{i,t}},$$

where $dvol_{i,t}$ is the consolidated dollar trading volume for each stock i in month t, and $trades_{i,t}$ is the number of trades executed for each stock i in month t.

5.5 Descriptive statistics

In this section, we present the main statistics on the data used in the study. Moreover, we provide explanations for all variables which are not yet defined. Afterwards, we compare two groups of stocks in our sample, Chi-X and non-Chi-X stocks. Table 2 presents descriptive statistics for the complete dataset.

Table 2. Descriptive statistics

This table contains descriptive statistics for the dataset used in regression analysis. It encompasses the period between February 1, 2007 and February 28, 2009. The dataset contains 32,233 stock-month observations for 1,311 stocks.

Variable	*Mean*	*Standard Deviation*	*25th percentile*	*50th percentile*	*75th percentile*
$Frag1_{i,t}$	3.537	3.291	1.105	2.304	4.762
$Frag2_{i,t}$	0.071	0.111	0.000	0.017	0.098
$Frag3_{i,t}$	0.068	0.123	0.000	0.007	0.072
$Frag4_{i,t}$	0.087	0.200	0.000	0.012	0.080
$Chix_{i,t}$	0.572	0.495	0.000	1.000	1.000
$mcap_{i,t}$	5.227	14.029	0.188	0.911	3.160
$dvol_{i,t}$	34.183	100.373	0.360	3.559	19.822
$price_{i,t}$	28.247	56.266	4.697	11.951	29.605
$volatility_{i,t}$	80.980	324.012	38.069	54.939	81.605
$high\_low_{i,t}$	350.256	320.200	171.511	286.803	454.545
$spread_{i,t}$	124.756	569.845	13.641	33.969	98.396
$liq_{i,t}$	1.559	3.171	-0.479	1.798	3.778
$home\_tick\_size_{i,t}$	0.068	0.080	0.0100	0.0200	0.100
$home\_tick\_to\_price_{i,t}$	0.003	0.019	0.000	0.000	0.001
$AT_{i,t}$	35.800	465.668	4.393	7.093	13.474
$messages_{i,t}$	17.528	54.388	0.303	3.108	11.674
$trades_{i,t}$	1.469	2.832	0.047	0.452	1.596
$avg\_trade_{i,t}$	22.946	272.367	5.184	9.268	16.842
$in\_index_{i,t}$	0.248	0.432	0.000	0.000	1.000

(1) $Chix_{i,t}$ is the dummy variable specifying whether stock *i* is traded on the Chi-X platform in month *t*. As the mean is above 0.5, we have more stocks which are included in Chi-X at some point in time comparing to the cases when they are not included.

(2) $mcap_{i,t}$ is the market capitalization (EUR billion) of stock *i* in month *t*. The market capitalization of an average stock in our sample is 5.23 billion EUR.

(3) $dvol_{i,t}$ is the consolidated value of executed trades (EUR million) for stock *i* in month *t*. The value of average monthly trading volume is 34.18 million EUR.

(4) $price_{i,t}$ is the midquote price (EUR) of stock *i* in month *t*.

(5) $volatility_{i,t}$ is the standard deviation of monthly midquote returns in bps of stock *i* in month *t* for the consolidated order book.

(6) $high\_low_{i,t}$ is an alternative measure of volatility in bps for stock *i* in month *t* for robustness check. We calculate it as follows: $\frac{2*(high_{i,t}-low_{i,t})}{high_{i,t}+low_{i,t}}$, where $high_{i,t}$ and $low_{i,t}$ are high and low prices of a respective stock i in month t.

(7) $home\_tick\_to\_price_{i,t}$ is an effective measure of tick size in bps accounting for changes in price levels. The variable is calculated as the ratio of home tick size to closing price of stock *i* in month *t*. Home exchange tick size is derived from daily closing prices of stock *i* in month *t* and proven by historical tick size schedules available at public sources.

(8) $messages_{i,t}$ is the monthly number of messages (thousands) for stock i in month *t* consolidated across all markets. This variable is used to calculate proxies for AT activity. An average stock gets 17.5 thousand quote messages in a month.

(9) $trades_{i,t}$ is the monthly number of executed trades (thousands) for stock *i* in month *t* in consolidated order book. The average number of trades per month is almost 1.5 thousand.

(10) $in\_index_{i,t}$ is the dummy variable which reflects whether stock i is included in the major European indices (see the list in Appendix C) in month t.

5.6 Descriptive statistics comparison for Chi-X and non-Chi-X stocks

We provide comparison of statistics for the main variables in our dataset between Chi-X and non-Chi-X stocks in Table 3. It allows us to get an insight about their distribution and differences across markets.

Table 3. Descriptive statistics for Chi-X and non-Chi-X stocks

This table contains descriptive statistics for Chi-X (1) and non-Chi-X stocks (2). It encompasses the period between February 1, 2007 and February 28, 2009. The dataset contains 18,425 stock-month observations for Chi-X stocks and 13,808 stock-month observations for non-Chi-X stocks. The table provides difference between means of Chi-X and non-Chi-X stocks: $(1) - (2)$. Relative difference compares the difference between Chi-X and non-Chi-X means to non-Chi-X values: $\frac{(1)-(2)}{(2)}$.

Variable	*Mean for Chi-X stocks*	*Mean for non-Chi-X stocks*	*Difference*	*Relative Difference*
$Frag1_{i,t}$	4.829	1.814	3.015	1.662
$Frag2_{i,t}$	0.107	0.022	0.085	3.864
$Frag3_{i,t}$	0.105	0.018	0.087	4.833
$Frag4_{i,t}$	0.134	0.024	0.110	4.583
$mcap_{i,t}$	8.449	0.927	7.522	8.114
$dvol_{i,t}$	58.239	2.084	56.155	26.946
$price_{i,t}$	23.212	35.135	-11.923	-0.339
$volatility_{i,t}$	73.203	91.589	-18.386	-0.201
$high\_low_{i,t}$	361.441	333.183	28.258	0.085

$spread_{i,t}$	36.591	242.932	-206.341	-0.849
$liq_{i,t}$	19.284	15.472	3.812	0.246
$home\_tick\_size_{i,t}$	0.072	0.062	0.010	0.161
$home\_tick\_to\_pric$	0.001	0.006	-0.005	-0.833
$AT_{i,t}$	50.198	16.574	33.624	2.029
$messages_{i,t}$	29.606	1.413	28.193	19.953
$trades_{i,t}$	2.388	0.244	2.144	8.787
$avg\_trade_{i,t}$	30.412	12.976	17.436	1.344
$in\_index_{i,t}$	0.408	0.033	0.375	11.364

The fragmentation measures for Chi-X stocks has significantly larger values than for non-Chi-X stocks. Thus, it can be concluded that Chi-X stocks are on average more fragmented than the ones not traded on Chi-X. One of the most noticeable differences between the two groups of stocks refer to their market capitalization and trading volume. Chi-X stocks have more than 9 times higher average market capitalization and more than 25 times higher trading volume. The reason for that can be the fact that when Chi-X enters a new market, it starts with including the most traded and large cap stocks, while consequently adding other types of stocks. Moreover, Chi-X stocks have higher average size of trade, which suggests higher interest from institutional traders towards these stocks or the Chi-X platform. One more sizable difference between Chi-X and non-Chi-X stocks is the values of spread. Chi-X stocks have on average 6 times lower spread than stocks not traded on Chi-X. It can be explained by Chi-X favoring more liquid stocks with higher trading volumes. Moreover, Chi-X stocks have lower volatility, which can also stem from higher liquidity and trading volume. Home exchange tick size values are lower for an average non-Chi-X stock comparing to Chi-X stocks. It may suggest that Chi-X attracts stocks with higher tick size on their home exchange. One more notable difference refers to the number of messages and number of trades. These variables are considerably higher for Chi-X stocks, and consequently, their ratio (AT), which proxies algorithmic trading, is 3 times larger for stocks traded on Chi-X. This evidence is in line with the theory that alternative trading venues attract larger number of high frequency traders than traditional exchanges.

6. Research Design

6.1 Addressing the endogeneity issue

Research on fragmentation can suffer from an endogeneity problem, which affects the estimation in several ways. Firstly, to get an opportunity to fragment, stocks should be chosen to trade on an alternative trading venue. As the selection procedure is not random, the estimated effect on fragmentation is affected by selection bias. Moreover, these stocks share some common characteristics that enabled them to be chosen for trading on this platform. With simple OLS regression, these characteristics are unobservable factors that lead to biased coefficients (Briggs, 2004). Therefore, to solve the problem of sample selection and unobservable factors, we employ Heckman two-stage selection model. In the first stage, we use all observations (for stocks traded and not traded on Chi-X) to determine the factors that influence their selection to the Chi-X platform. From this regression, Inverse Mills Ratio (IMR) is calculated for each observation in the sample. In the second stage, we use the sample of stocks that were traded on Chi-X and estimate the effects of stock and market characteristics on the degree of fragmentation. IMR is added to this stage to correct the selection bias and obtain consistent estimates of the effects on fragmentation.

In addition to that, to deal with the issue of unobservable factors that affect both the probability of a stock being traded on Chi-X and the degree of fragmentation, we use instrumental variable. In the Heckman model, such variable should be a strong predictor of the selection (inclusion to Chi-X), but not the outcome (the degree of fragmentation) (Briggs, 2004). In our case, we use a dummy variable which captures stock inclusion into major market index. The list of major market indices for our sample is provided in the Appendix C. When Chi-X enters new markets, it starts off by rolling out trading in stocks that belong to major indices (Malceniece et al., 2018). At the same time, market participants don’t necessarily choose the stocks which are included in major indices when they choose what and where to trade. Thus, index inclusion can be used as an instrument in our model, as it affects the Chi-X decision to select the stock, but not market participant’s decisions, that affect fragmentation, other than through the Chi-X decision.

One more issue of endogeneity that should be considered is the problem of simultaneity or reverse causality, meaning that the dependent variable has an impact on independent variables, and vice versa. In our case, we use stock and market characteristics to estimate probability of inclusion to Chi-X and the degree of fragmentation. At the same time, these characteristics can be influenced by trading on Chi-X or fragmentation. Thus, for the independent variables that are likely to be impacted by the dependent variables in our model, we use 3-month lagged values to account for endogeneity and the problem of reverse causality.

6.2 Two-stage selection model

We use Heckman two-stage selection model (Heckman, 1979) in which the first stage models the probability that a stock is added to the Chi-X platform and the second stage models the determinants of the degree of fragmentation in a stock, accounting for the fact that the opportunity for fragmentation (the entry of Chi-X) is endogenous. Unlike a typical application of Heckman selection model where the second stage is of primary interest and the first stage is employed merely to produce an unbiased second stage, in our application both stages are of interest. The first stage informs about market operator choices and the second about market participant's trading decisions, both of which are required for fragmentation to occur.

The choice of variables for the first and second stage of Specification 1 is based on the papers discussed in the literature review and follow the hypotheses presented in Table 1. We also apply an iterative approach to test different combinations of variables in the model. We choose the final combination of variables for the baseline equation based on (i) economic reasoning and (ii) highest significance and explanatory power.

In the first stage of Heckman model, we use probit regression to estimate the probability of a stock being included to the Chi-X platform. We perform log-conversion of several variables to make their distribution close to normal. The regression equation for the first stage is the following:

$$Chi - X_{i,t} = \beta_0 + \beta_1\, ln(dvol)_{i,t-3} + \beta_2\, ln(volatility)_{i,t-3} + \beta_3 ln(spread)_{i,t-3} + \beta_4 in\_index_{i,t} + \varepsilon_{i,t},$$

where $ln(dvol)_{i,t-3}$ is 3-month lagged value of the EUR trading volume of stock i in month t; $ln(volatility)_{i,t-3}$ is 3-month lagged value of volatility for stock i in month t; $ln(spread)_{i,t-3}$ is 3-month lagged value of relative quoted spread in bps for stock i in month t; $in\_index_{i,t}$ is the dummy variable reflecting stock's i inclusion in major European indices in month t.

The second stage estimates the effect on the degree of market fragmentation. We employ the following regression to identify the link between stock and market characteristics and the degree of market fragmentation:

$$Frag1_{i,t} = \beta_0 + \beta_1 mcap_{i,t-3} + \beta_2 price_{i,t-3} + \beta_3 home\_tick\_to\_price_{i,t} + \beta_4 AT_{i,t-3} + \beta_5 avg\_trade_{i,t} + \beta_6 IMR_{i,t} + \varepsilon_{i,t},$$

where $Frag1_{i,t}$ reflects the number of venues, on which stock i was traded during the month t, $mcap_{i,t-3}$ is 3-month lagged value of market capitalization of stock i in month t; $price_{i,t-3}$ is 3-month lagged value of midquote price of stock i in month t; $home\_tick\_to\_price_{i,t}$ is the ratio of the tick size on the home exchange to the closing price of stock i in month t; $AT_{i,t-3}$ is 3-month lagged proxy for algorithmic trading, calculated as the number of messages per trade; $avg\_trade_{i,t}$ shows the average traded dollar volume per trade of stock i in month t. $IMR_{i,t}$ is estimated as $IMR_{i,t} = \frac{\varphi(p_i)}{\Phi(p_i)}$, where $p_{i,t}$ is the predicted probability of stock i to be traded on Chi-X in month t from the first stage of a probit model, φ is the standard normal density function (pdf), and Φ is the cumulative normal distribution function (cdf).

In both equations, we employ double-clustered standard errors, clustered by stocks and year-months. This step corrects estimation of standard errors and eliminates the effects of correlation between entities and individual time series in the panel data. In addition to that, we standardize all the variables, except dummy and dependent variables, to be able to derive conclusions about the economic effects of these factors and to compare their effects.

6.3 Country- and time-fixed effects model specification

As we use financial panel data in our research, we face an issue that our results can be affected by country- and time-specific effects. For example, countries differ in their economic conditions, size and quality of markets, regulations, etc. Moreover, stock and market characteristics can vary in time and be affected by certain periodic economic and financial conditions, which also influences the coefficients. To account for it, we bring country- and time-fixed effects to our model. Our Specification 2 with country-fixed effects includes 12 market dummies ($Country_i$), apart all the other variables present in Specification 1. Specification 3 of the model contains 25 time-fixed effect dummy variables at monthly intervals ($YearMonth_t$). We employ both country- and time-fixed effects in Specification 4 of the model. Controlling for these variables, we are able to get the refined coefficients for our main variables excluding unobservable country- and time-effects. We provide regression equations for Specification 4 and use this specification in our analysis as a baseline for robustness check:

$$1\ stage: Chi - X_{i,t} = \beta_0 + \beta_1\, ln(dvol)_{i,t-3} + \beta_2\, ln(volatility)_{i,t-3} + \beta_3 ln(spread)_{i,t-3} + \beta_4 in\_index_{i,t} + Country_i + YearMonth_t + \varepsilon_{i,t}$$

$$2nd\ stage: Frag1_{i,t} = \beta_0 + \beta_1 mcap_{i,t-3} + \beta_2 price_{i,t-3} + \beta_3 home\_tick\_to\_price_{i,t} + \beta_4 AT_{i,t-3} + \beta_5 avg\_trade_{i,t} + \beta_6 IMR_{i,t} + Country_i + YearMonth_t + \varepsilon_{i,t}$$

Specification 2 and 3 are structured the same way, but they include only $Country_i$ and $YearMonth_t$ respectively.

7. Regression results and discussion

In this section, we present the results of our model and provide the discussion of the findings. Table 4 shows the results for Specifications 1-4 of our model with standardized variables.

Table 4. Regression results for Heckman two stage selection model

This table contains results for the specification of two stage Heckman model with country, time, and country and time fixed effects. We control for cross-country variations by adding dummy variables ($Country_i$) for each of 12 countries and for time variation by adding dummy variables ($YearMonth_t$) for each of 25 time periods. All the independent variables are standardized to have mean of 0 and standard deviation of 1. We provide regression equations for Specification 4. The equations for Specifications 1-3 are identical, but include no or only one type of fixed effects.
$1st\ stage$: $Chix_{i,t} = ln(dvol)_{i,t-3} + ln(volatility)_{i,t-3} + \ln(spread)_{i,t-3} + in_{index_{i,t}} + Country_i + YearMonth_t$;
$2nd\ stage$: $Frag1_{i,t} = mcap_{i,t-3} + price_{i,t-3} + home\_tick\_to\_price_{i,t} + AT_{i,t-3} + avg\_trade_{i,t} + Country_i + YearMonth_t + IMR_{i,t}$.
The number of stock-month observations is 31,436. ***, **, and * refer to statistical significance at 1%, 5%, and 10% levels, respectively. T-statistics are provided in parentheses.

	No FE	*Country FE*	*Time FE*	*Country and Time FE*
First Stage	$Chix_{i,t}$	$Chix_{i,t}$	$Chix_{i,t}$	$Chix_{i,t}$
$ln(dvol)_{i,t-3}$	0.80***	0.69***	0.84***	0.71***
	(7.32)	(5.43)	(7.38)	(5.57)
$ln(volatility)_{i,t-3}$	-0.04	-0.05*	-0.11**	-0.11**
	(-1.04)	(-1.83)	(-2.22)	(-2.35)
$ln(spread)_{i,t-3}$	-0.24***	-0.32**	-0.24***	-0.35***
	(-3.24)	(-2.81)	(-3.00)	(-2.73)
$in\_index_{i,t}$	1.25***	2.92***	1.22***	2.90***
	(7.75)	(11.15)	(7.61)	(11.09)
Second stage	$Frag1_{i,t}$	$Frag1_{i,t}$	$Frag1_{i,t}$	$Frag1_{i,t}$
$mcap_{i,t-3}$	1.01***	1.12***	1.04***	1.14***
	(9.89)	(10.75)	(9.83)	(10.69)
$price_{i,t-3}$	0.51**	-0.70***	0.81***	-0.29***
	(2.37)	(-3.43)	(3.97)	(-2.91)
$home\_tick\_to\_price_{i,t}$	0.79***	0.26*	0.24*	-0.23
	(3.87)	(1.68)	(1.77)	(-1.13)
$AT_{i,t-3}$	-0.03	-0.06**	-0.05**	-0.08**
	(-1.39)	(-2.47)	(-2.19)	(-2.51)
$avg\_trade_{i,t}$	-0.12***	-0.16***	-0.10**	-0.14***
	(-3.97)	(-4.94)	(-4.56)	(-7.03)
$IMR_{i,t}$	-2.48***	-1.94***	-2.53***	-2.04***
	(-6.42)	(-5.99)	(-6.80)	(-6.33)

1st stage: R-squared	37%	64%	38%	65%
2nd stage: R-squared	31%	47%	53%	67%

We provide the results with non-standardized variables for Specifications 1-4 in Appendix D. We employ double clustered standard errors by stock and month in all the regressions.

To begin with, we review and discuss the results of our Specification 1 without fixed effects. Most of the obtained results are in line with our hypotheses. The coefficient for volatility is not significant at 10% level, while other variables are significant at 1% significance level. In the first stage, where we model the probability to be included in Chi-X ($Chix_{i,t}$), we find that the inclusion of the stock in the major market index is one of the most important determinants of being selected to trade on Chi-X. Such effect is predictable as index inclusion was chosen as an instrumental variable that is a strong predictor of the probability to enter Chi-X. It proves that Chi-X was adding the stocks from major market indices at the beginning of its expansion to the European market.

Moreover, the results suggest that stocks with wider spread are less likely to trade on Chi-X (supported Hypothesis 1). This effect supports the finding of Bennet and Wei (2006) that less liquid stocks tend to stay on their primary market. The evidence on spread suggests that market operators are more likely to choose more liquid stocks to trade on their platforms, as they have abundant liquidity that can migrate to other venues. Moreover, trading in more liquid stocks is likely to attract traders to the venue, enhancing its liquidity. Our results are also in line with Hypothesis 2, and show that stocks' trading volume is positively associated with the probability to trade on Chi-X. This fact also suggests that the venue tries to choose stocks that are characterized by high levels of traders' interest and high liquidity. Similar conclusion was reached by Fong, Madhavan, and Swan (2001), who find that high trading volume is the most important factor for a stock to start trading off-market.

The second stage of our model provides results on the effects of stock and exchange related characteristics on fragmentation measure ($Frag1_{i,t}$), which refers to the number of venues stock i trades in month t. The signs of the obtained coefficients are in line with our hypotheses. Market capitalization of the stock shows the most pronounced effect on fragmentation: an increase in one standard deviation of market capitalization results in the rise of number of venues a stock trades by 1.01. This finding is coherent with the fact that large cap stocks attract higher interest from investors, and therefore, are more likely to split their trading across multiple venues. In line with our predictions, we find positive effect of home tick to price ratio on fragmentation, meaning that the larger is the ratio of the tick size on the primary exchange to the price of the stock, the higher is the fragmentation of this stock. The coefficient suggests that an increase in tick size relative to its price on the home exchange by one standard deviation results in higher fragmentation measure by 0.79. This means that the larger the tick size on the home exchange is, the higher is the incentive of traders to route their orders to the venues with smaller tick sizes so as to decrease queueing of the orders on the exchange and fasten their execution. In this specification, our results do not support Hypothesis 5 that stocks with higher AT activity are more likely to fragment, as this coefficient is not significant at 10% significance level.

In line with Hypothesis 6, average trade size negatively influences fragmentation of the stock. The results suggest that stocks traded by institutional investors are less likely to fragment, namely an increase in average trade size by one standard deviation causes fragmentation measure to fall by 0.12. In line with findings by Lin et al. (2009), it can be explained by the fact that retail investors are more sensitive to higher costs of trading in particular stocks and trading venues, and therefore, their brokers choose alternative trading venues, which offer lower fees and faster execution. Moreover, high levels of algorithmic trading on alternative trading venues may lead to the average size of the trade to fall, as HFT relies on small orders. Institutional investors are also likely to use algorithms for splitting their large orders into smaller ones so as to eliminate significant movements in stock price.

In all specifications, the coefficient of IMR is negative and significant at 1%. The significance of IMR shows that our model corrects the endogeneity issue of sample selection bias. Without the correction, the regression coefficients would be downward biased or underestimated. Using Heckman model, we get more accurate estimates while eliminating sample selection bias.

Further, we discuss the results of the regression model with both country- and time-fixed effects (Specification 4). Signs of the coefficients in the first stage remain the same and are in line with our hypotheses. The coefficients for volatility and AT are significant at 5% level, for all the other variables – at 1% significance level. However, home tick to price ratio becomes insignificant in this specification.

When controlling for country- and time-fixed effects, the results support Hypothesis 3 that stocks with higher volatility are less likely to trade on Chi-X. It can be explained by the fact that stocks with high volatility benefit more from staying on the home exchange due to higher liquidity and stability. This finding is supported by the conclusion of Ye (2011) on a dark pool market, which loses its market share during periods of high stock volatility. Furthermore, considering the conclusion of Gajewski and Gresse (2007) that stocks have higher volatility on fragmented markets, we can state that a highly volatile stock is more likely to stay on the primary market with concentrated liquidity and order flow and not be traded on the alternative trading venue such as Chi-X. Other variables in the first stage do not experience the change in signs and significance compared with Specification 1.

The second stage results for market capitalization remain the same compared to Specification 1 (no fixed effects) and its coefficient increases marginally. The coefficient for price changes its sign and becomes negative, which is in line with our Hypothesis 8. The reason for this change may stem from the variations in price levels across countries. Furthermore, home tick to price coefficient becomes insignificant in this specification. The significance level of this variable can be affected by different tick size regimes as well as changes in local regulations across the countries in our sample. At the same time, the coefficient for AT variable becomes significant at 5% level and indicates negative effect on fragmentation. It rejects our Hypothesis 5 that stocks attracting more AT activity are more likely to fragment. It can be explained by the fact that investors prefer to stay on the home exchange, which provides natural and accessible liquidity, comparing to ATS with high levels of algorithmic trading. Moreover, lower message-to-trade ratio enables investors to execute their trades faster and with higher probability. This finding is also in line with the conclusions of Conrad et al. (2003), that traders, especially institutional, prefer venues with lower trading costs and faster execution process.

The results of specifications with country- (Specification 2) and time- (Specification 3) separately are mostly in line with the previous results of Specification 1 and 4, except for some differences. The coefficients on the first stage do not change their signs and remain statistically significant, meaning that country and time-related variations have no significant influence on our specification without fixed effects.

The results from the second stage show that all the variables retain their signs, except for price which has negative coefficient in Specification 2 and positive in Specification 3. This change may be explained by the elimination of price differences between countries after controlling for country-fixed effects. Therefore, we observe the change of its sign in both Specifications 2 and 4 where we have country fixed effects. Home tick size to price ratio, average trade, and IMR retains the same sign as in Specification 1. AT coefficient have the same sign and significance level as in Specification 4, however it differs from Specification 1. It suggests that controlling for both country- and time-fixed effects provide the opportunity to eliminate both country and time variations and see the "clean" effects.

Our study exploits a certain period of time and sample of countries in order to determine the drivers of fragmentation. However, the results obtained could be also extrapolated for other countries with similar market conditions and different time periods. This can be illustrated by comparing our results with the findings of similar papers on market fragmentation and microstructure. For example, the effect of liquidity on fragmentation is supported by the study of Bennet and Wei (2006) that was conducted for the US market in the period of 2002-2003. Our conclusion on the influence of trading volume is in line with the research by Fong et al. (2001), who studied Australian market in 1993-1998. Obtaining results consistent with other studies on similar markets suggests that the results of this research could also be applicable to other countries (e.g. with similar level of liquidity in equity markets). However, as this study of fragmentation is novel in several dimensions, conducting similar research on other markets and time periods would be beneficial for further consideration.

Moreover, the case of Chi-X is especially prominent for this type of research because it can serve as an example for other markets in similar conditions. This case is exemplary for the reasons outlined previously in the paper: it was the first and only operating alternative trading platform during the sample period, and it gathered significant trading volume in the first years of operations. At the same time, this time period is characterised by two different market conditions: more calm period of 2007 and more volatile crisis period of 2008-2009. However, we alleviate this concern by implementing time-fixed effects which helps us to derive conclusions irrespective of conditions of our study period. Therefore, our results provide determinants of market fragmentation which are not only specific for this case but rather these results can be extended for other geographic markets in similar conditions and in different time periods.

As the next step, we check the validity of the results using several robustness checks. For that purpose, we employ Specification 4 of the Heckman model, which eliminates country- and time- specific effects on the variables of interest.

8. Robustness check

In this section, we present the robustness checks using alternative measures for fragmentation and independent variables to check the validity of our results. First, we run our specification with country- and time-fixed effects using all fragmentation measures. The results of these regressions with standardized variables are provided in the Table 5. The version with non-standardized variables can be found in Appendix E.

Table 5. Robustness check using alternative fragmentation measures

This table contains results for specification of two stage Heckman model with both country- and time-fixed effects. At the second stage, alternative measures of fragmentation ($Frag2_{i,t}$, $Frag3_{i,t}$, $Frag4_{i,t}$) are used as dependent variables. All the independent variables are standardized to have mean of 0 and standard deviation of 1. The number of stock-month observations is 31,436. ***, **, and * refer to statistical significance at 1%, 5%, and 10% levels. T-statistics is provided in parentheses.

First Stage	$Chix_{i,t}$			
$ln(dvol)_{i,t-3}$	0.71***			
	(5.57)			
$ln(volatility)_{i,t-3}$	-0.11**			
	(-2.35)			
$ln(spread)_{i,t-3}$	-0.35***			
	(-2.73)			
$in\_index_{i,t}$	2.90***			
	(11.09)			
Second stage	$Frag1_{i,t}$	$Frag2_{i,t}$	$Frag3_{i,t}$	$Frag4_{i,t}$
$mcap_{i,t-3}$	1.14***	0.02***	0.02***	0.02***
	(10.69)	(4.24)	(4.26)	(2.83)
$price_{i,t-3}$	-0.29**	-0.01	-0.01**	-0.005
	(-2.91)	(-1.42)	(-2.42)	(-0.55)
$home\_tick\_to\_price_{i,t}$	-0.23	-0.01	-0.01	0.06*
	(-1.13)	(-0.91)	(-0.49)	(1.73)
$AT_{i,t-3}$	-0.08**	-0.001	0.0002	0.01
	(-2.51)	(-0.60)	(0.07)	(0.97)
$avg\_trade_{i,t}$	-0.14***	-0.002***	0.002**	0.02**
	(-7.03)	(-4.17)	(2.56)	(2.25)
$IMR_{i,t}$	-2.04***	-0.03***	-0.03*	0.12***
	(-6.33)	(-2.77)	(-1.92)	(4.14)
1st stage: R-squared	65%	65%	65%	65%
2nd stage: R-squared	67%	47%	41%	10%

The first stage is the same across all specifications of robustness check with alternative fragmentation measures. Therefore, the results for the first stage of robustness check repeat our results of Specification 4 (probit with country- and time-fixed effects). The version of the second stage with $Frag1_{i,t}$ is provided for the purpose of comparison with the results of specifications with alternative measures.

Using different fragmentation measures, most of the obtained results are in line with the baseline equation for all significant variables. However, some variables in the second stage change their signs and significance level. The reason for that may lie in the way how fragmentation measures are estimated. $Frag2_{i,t}$ and $Frag3_{i,t}$ are calculated using trading activity measures, such as dollar trading volume and the number of trades. Furthermore, $Frag4_{i,t}$ reflects the dollar volume market share of all venues other than home market. As our study period includes the first years of the expansion of the alternative trading platforms, trading activity on them are much less compared with the home exchanges. In that case, the changes in trading volume or the number of trades on these platforms are not effectively reflected by fragmentation measures based on the trading activity, even if a stock does fragment. Therefore, $Frag1_{i,t}$, which refers to the number of venues a stock trades on, is the most effective measure of fragmentation for our research. Moreover, the specification with that fragmentation measure shows the highest value of R-squared compared with specifications using other proxies.

Further we check whether the obtained results for the both stages are robust to the choice of our measures used in the selection equation. For this purpose, we use alternative measures for liquidity: $liq_{i,t}$ (following Amihud, 2002), and volatility: $high\_low_{i,t}$. The detailed results for standardized variables are provided in the Table 6. The results for non-standardized variables are reported in Appendix F.

We omit trading volume variable from the first stage due to its inclusion in the calculations of $liq_{i,t}$, which leads to high correlation. The results obtained with alternative liquidity measure are in line with our baseline specification. The positive coefficient of the liquidity proxy suggests that more liquid stock have higher probability to be traded on Chi-X which supports our Hypothesis 1. The same results were obtained previously using spread as a proxy for liquidity. The alternative measure of volatility, which is calculated using monthly high and low prices, also suggests the negative effect of this variable on the probability to be traded on Chi-X. These results confirm our previous findings with the other measure of volatility. Therefore, we can conclude that our results for the first stage are robust to the choice of the measures for the model.

Table 6. Robustness check using alternative measures for liquidity and volatility

This table contains results for specification of two stage Heckman model with both country- and time-fixed effects. On the first stage, $ln(liq)_{i,t-3}$ is used as an alternative measure for liquidity instead of $ln(spread)_{i,t-3}$ and $ln(high\_low)_{i,t-3}$ is an alternative measure for $ln(volatility)_{i,t-3}$. All the independent variables are standardized to have mean of 0 and standard deviation of 1. The number of stock-month observations is 24,466. ***, **, and * refer to statistical significance at 1%, 5%, and 10% levels.

First Stage	$Chix_{i,t}$			
$ln(high\_low)_{i,t-3}$	-0.06**			
	(-1.95)			
$ln(liq)_{i,t-3}$	0.94***			
	(6.77)			
$in\_index_{i,t}$	2.91***			
	(11.82)			
Second stage	$Frag1_{i,t}$	$Frag2_{i,t}$	$Frag3_{i,t}$	$Frag4_{i,t}$
$mcap_{i,t-3}$	1.18***	0.02***	0.02***	0.02***
	(10.12)	(4.03)	(4.13)	(2.87)
$price_{i,t-3}$	-0.33***	-0.01**	-0.01***	-0.01
	(-3.20)	(-2.07)	(-2.63)	(-0.86)
$home\_tick\_to\_price_{i,t}$	-0.36	-0.02	-0.02	0.08**
	(-1.28)	(-1.26)	(-1.17)	(2.22)
$AT_{i,t-3}$	-0.08**	0.0001	0.002	0.01
	(-2.43)	(0.07)	(0.64)	(1.05)
$avg\_trade_{i,t}$	-0.12***	-0.002***	0.01**	0.01**
	(-5.21)	(-6.11)	(2.57)	(2.13)
$IMR_{i,t}$	-1.98***	-0.03***	-0.02*	0.07***
	(-7.49)	(-2.82)	(-1.79)	(3.11)
1st stage: R-squared	63%	63%	63%	63%
2nd stage: R-squared	69%	49%	45%	9%

Moreover, as the selection equation is used to calculate IMR, which corrects the effects on the second stage, we can check whether the choice of the measures for our first stage impacts the results obtained in the second. Comparing the results from Table 5 and 6, it can be concluded that the second stage is not influenced by using alternative measures for liquidity and volatility, as all the variables preserve their signs. Moreover, IMR remains significant, meaning that the model still corrects selection bias.

9. Conclusion

Our study aims to advance the current understanding of stock- and market-level drivers of fragmentation. For that purpose, we exploit the expansion of the Chi-X trading platform, the first and most prominent Multilateral Trading Facility in Europe, to 12 European equity markets. Fragmentation occurs conditional on trading decisions made by market operators and market participants. We study the impact on both of them using Heckman two-stage selection model.

Our empirical results show that index inclusion and larger trading volume are the most critical factors for a stock to be traded on Chi-X. Furthermore, stock liquidity, proxied by spread and Amihud's liquidity measure, positively influences market operator's decision to include a stock on the Chi-X platform. After controlling for country- and time-fixed effects, the results suggest that lower volatility increases the probability of a stock to enter Chi-X. Robustness of this finding is verified using high-low measure of volatility.

Among the factors impacting overall fragmentation in a stock, market capitalization shows the most pronounced effect. A one standard deviation increase in market capitalization leads to increase in number of venues by 1.14 of its mean. Larger tick to price ratio on the home exchange shows economically meaningful effect on fragmentation in the studied period, being the second most important factor. However, controlling for country- and time-fixed effects, it has no statistical significance. Contrasting to the recent rise in algorithmic trading, we find that it has a negative impact on fragmentation. It shows that market participants prefer more natural sources of stock liquidity, faster execution and thus, choose platforms with lower AT activity. Price and average trade size have a negative impact on the number of venues trading a stock, suggesting that stocks traded by retail investors are more likely to fragment.

Our study points out the issue of endogeneity in the research on fragmentation. We document the presence of sample selection bias, which is suggested by statistically significant IMR ratio. Using Heckman selection model, we are able to minimize the influence of the bias and get more robust results. Our study reiterates the importance of correcting bias of non-randomly chosen samples and illustrates its potential presence in studies on market fragmentation.

While our results provide evidence on some of stock- and market-level drivers of fragmentation, further research on the topic of the drivers of equity fragmentation is needed. Latency, depth, fees structure are only some examples of factors which can also affect fragmentation. In addition, our results need to be explored in the extended period. We partly alleviate the issue of time-related variance by controlling for time-fixed effects. However, our sample contains two quite different periods: the calmer before-crisis period (2007) and the more volatile period of financial crisis (2008-2009). Thus, longer time frames can be considered for further research. As for now, the obtained set of factors can be used by market operators in altering their market structure or stock selecting strategy, as well as by market participants in making their order routing decisions. Furthermore, the findings may be used by financial regulators in forming their policy decisions.

10. References

Alampieski, K., Lepone, A. (2009). Impact of a tick size reduction on liquidity: Evidence from the Sydney futures exchange. *Accounting & Finance, 49,* 1–20.

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

Amihud, Y., Lauterbach, B., & Mendelson, H. (2003). The value of trading consolidation: Evidence from the exercise of warrants. *Journal of Financial and Quantitative Analysis, 38*(4), 829–846.

Anderson, H. D., & Peng, Y. (2014). From cents to half-cents and its liquidity impact. *Pacific Accounting Review, 26*, 160–176.

Barclay, M. J., Hendershott, T., & McCormick, D. (2003). Competition among trading venues: Information and trading on electronic communications networks. *Journal of Finance, 58*(6), 2637-2666.

Battalio, R. H. (1997). Third market broker-dealers: cost competitors or cream skimmers? *Journal of Finance, 52*(1), 341–352.

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

Bloch, E., & Schwartz, R. A. (1978). The great debate over NYSE Rule 390. *Journal of Portfolio Management, 5*(1), 5–8.

Boehmer, B., & Boehmer, E. (2003). Trading your neighbor's ETFs: competition or fragmentation? *Journal of Banking and Finance, 27*(9), 1667–1703.

Boehmer, E., Fong, K., & Wu, J. (2015). International evidence on algorithmic trading. *Working paper*.

Boneva, L., Linton, O., & Vogt, M. (2015). The effect of fragmentation in trading on market quality in the UK equity market. *Journal of Applied Econometrics, 31*(1), 192-213.

Briggs, D. C. (2004). Causal Inference and the Heckman Model. *Journal of Educational and Behavioral Statistics, 29*(4), 397-420.

Cboe Global Markets (n.d.). Market Statistics. Retrieved from http://markets.cboe.com/europe/equities/market_statistics/.

Cboe Global Markets. (n.d.). European Equities Overview. Retrieved from http://markets.cboe.com/europe/equities/market_statistics/.

Chi-X Europe. (2011a). Company Description. Retrieved from https://web.archive.org/web/20110221092314/http://www.chi-xeurope.com:80/home/company.asp.

Chi-X Europe. (2011b). Chi-X Europe breaks above €1.5trn mark in 2010 [Press-Release]. Retrieved from https://web.archive.org/web/20110220054944/http://www.chi-xeurope.com:80/chi-x-press-releases/chi-x-europe-q4-2010.pdf.

Chowdhry, B., & Nanda, V. (1991). Multimarket trading and market liquidity. *Review of Financial Studies, 4*(3), 483–511.

Conrad, J., Johnson, K. M., & Wahal, S. (2003). Institutional trading and alternative trading systems. *Journal of Financial Economics, 70,* 99-134.

Degryse, H., de Jong, F., & van Kervel, V. (2015). The impact of dark trading and visible fragmentation on market quality. *Review of Finance, 19*(4), 1-23.

Di Febo, E., & Angelini, E. (2015). The launch of alternative trading venues and their evolution in the European Union: Spanish exchanges. *Current Politics & Economics of Europe*. *26*(3), 267-297.

Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EU Text with EEA relevance. (2014). *Official Journal of the European Union, L173,* 349-496.

Economides, N. (1996). The economics of networks. *International Journal of Industrial Organization, 14,* 673–699.

Fioravanti, S., & Gentile, M. (2011). The impact of market fragmentation on European stock exchanges. *Working Paper.*

Fong, K., Madhavan, A., & Swan, P. L. (2001). Why Do Markets Fragment? A Panel-Data Analysis of Off-Exchange Trading. *Working Paper.*

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

Gajewski, J., & Gresse, C. (2007). Centralised order books versus hybrid order books: A paired comparison of trading costs on NSC (Euronext Paris) and SETS (London Stock Exchange). *Journal of Banking & Finance, 31*(9), 2906-2924.

Gomber, P., Sagade, S., Theissen, E., Weber, M. C., & Westheide, C. (2017). Competition between equity markets: A review of the consolidation versus fragmentation debate. *Journal of Economic Surveys, 31*(3), 792-814.

Gomber, P., Gsell, M. & Lutat, M. (2011). The quality of European equity markets after MiFID. *Trends in the European Securities Industry,* 179–200.

Hameed, A., & Terry, E. (1998). The effect of tick size on price clustering and trading volume. *Journal of Business Finance and Accounting, 25,* 849–867.

Harris, L. E. (1993). Consolidation, fragmentation, segmentation and regulation. *Financial Markets, Institutions & Instruments, 2*(5), 1–28.

Harris, L. E. (1994). Minimum price variations, discrete bid–ask spreads, and quotation sizes. *The Review of Financial Studies, 7,* 149–178.

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

Heckman, J. (1979). Sample selection biases as a specification error. *Econometrica, 47,* 153–162.

Hendershott, T., & Mendelson, H. (2000). Crossing networks and dealer markets: competition and performance. *Journal of Finance, 55*(5), 2071–2115.

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

Lin, B., Michayluk, D., Oppenheimer, H. R., & Sabherwal, S. (2009). French and U.S. trading of cross-listed stocks around the period of U.S. decimalization: Volume, spreads, and depth effects. *International Review of Financial Analysis, 18*(5), 223-231.

LiquidMetrix (2010). What kind of stocks fragment the most? *Working Paper.*

Madhavan, A. (1995). Consolidation, fragmentation, and the disclosure of trading information. *Review of Financial Studies, 8*(3), 579–603.

Malceniece, L., Malcenieks, K., & Putniņš, T. (2018). High frequency trading and co-movement in financial markets. *Journal of Financial Economics*, Forthcoming.

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

Meling, T., & Ødegaard, B. A. (2017). Tick size wars, high frequency trading, and market quality. *Working Paper.*

OECD (2016). Changing business models of stock exchanges and stock market fragmentation. *OECD Business and Finance Outlook 2016*. Paris: OECD Publishing.

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

Oppenheimer, H. R., & Sabherwal, S. (2003). The competitive effects of US decimalization: Evidence from the US-listed Canadian stocks. *Journal of Banking & Finance, 27*(9), 1883-1910.

Pagano, M. (1989). Trading volume and asset liquidity. *Quarterly Journal of Economics, 104*(2), 255–274.

Stigler, G. J. (1964). Public regulation of the securities markets. *Business Lawyer, 19*(3), 721–753.

Van Ness, B., Van Ness, R., & Pruitt, S. (2000). The impact of the reduction in tick increments in major U.S. markets on spreads, depth, and volatility. *Review of Quantitative Finance and Accounting, 15,* 153–167.

Verousis, T., Perotti, P., & Sermpinis, G. (2018). One size fits all? High frequency trading, tick size changes and the implications for exchanges: Market quality and market structure considerations. *Review of Quantitative Finance and Accounting, 50*(2), 353-392.

Ye, M. (2011). A Glimpse into the dark: Price formation, transaction cost and market share of the crossing network. *Working Paper.*

Appendices

Appendix A. Staggered entrance of Chi-X on the European markets

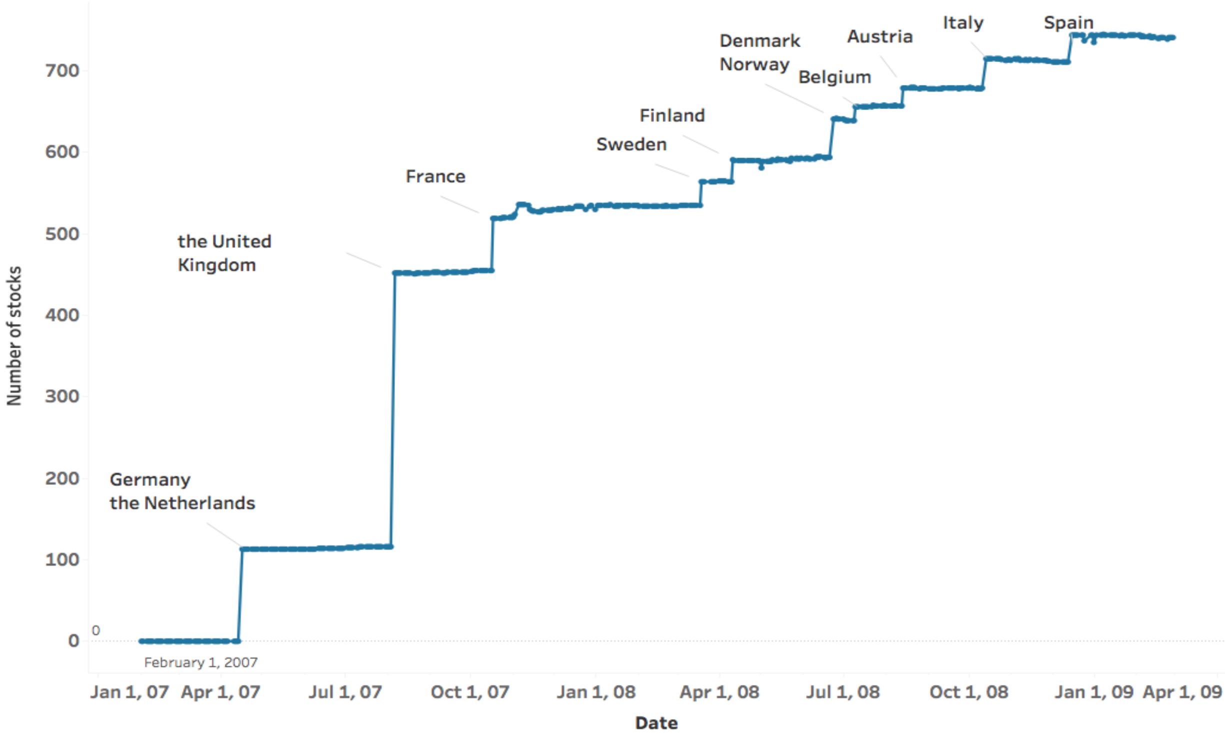

Figure 3. Number of stocks traded on Chi-X during the sample period. The graph shows the staggered entry of Chi-X Europe to European equity market during the sample period from February 1, 2007 to February 28, 2009. The vertical axis shows the number of stocks in our sample traded on Chi-X Europe on the daily basis during the studied timeframe.

Appendix B. Summary of empirical findings on the drivers of equity fragmentation

Table 7. Summary of empirical findings on equity fragmentation

The table presents the main empirical finding on the drivers of equity fragmentation. These findings were used to form hypotheses on the effect of stock- and market- characteristics on fragmentation.

Determinant	**Finding**	**Paper**
Liquidity	(1) market fragmentation is higher for more liquid stocks	Fong, Madhavan, & Swan (2001)
	(2) less liquid stocks benefit more from switching from fragmented market to consolidated	Bennet & Wei (2006)
Size of the company	liquidity improves for small cap stocks with fragmentation	O'Hara & Ye (2011)
Volatility	price volatility is significantly higher for stocks on fragmented markets than on consolidated	Gajewski & Gresse (2007)
Institutional vs retail traders	(1) average size of transactions is larger for fragmented markets	Gajewski & Gresse (2007)
	(2) retail investors are more sensitive to trading costs; lower depth on the market stimulates institutional investors to switch to more liquid trading venues	Lin, Michayluk, Oppenheimer, & Sabherwal (2009)
	(3) alternative trading venues attract institutional traders due to higher speed of execution and higher level of anonymity	Barclay, Hendershott, & McCormick (2003)

Appendix C. List of major European indices

Table 8. List of major European indices

The table contains the major European indices by countries, which are used to construct the instrumental dummy variable on index inclusion of a stock $in\_index_{i,t}$. This variable takes the value of 1 if the stock *i* was included in the respective index on month *t*, and 0 otherwise.

Index name	Country
AMX	Austria
BEL 20	Belgium
OMXC20	Denmark
OMXH25	Finland
CAC 40	France
DAX	Germany
Euro Stoxx 50	Germany
FTSE MIB	Italy
AEX	Netherlands
OSEBX	Norway
Oslo OBX	Norway
IBEX 35	Spain
OMXS30	Sweden
FTSE 100	UK

Appendix D. Regression results with non-standardized variables

Table 9. Regression results with non-standardized variables.

This table contains results for the specification of two stage Heckman model with country, time, and country and time fixed effects. We control for cross-country variations by adding dummy variables ($Country_i$) for each of 12 countries and for time variation by adding dummy variables ($YearMonth_t$) for each of 25 time periods. We exclude one $Country_i$ and $YearMonth_t$ variable to avoid dummy variable trap. The variables are not standardized. Our regression equations are as follows:

$1st\ stage$: $Chix_{i,t} = ln(dvol)_{i,t-3} + ln(volatility)_{i,t-3} + \ln(spread)_{i,t-3} + in_{index_{i,t}} + Country_i + YearMonth_t$;

$2nd\ stage$: $Frag1_{i,t} = mcap_{i,t-3} + price_{i,t-3} + home\_tick\_to\_price_{i,t} + AT_{i,t-3} + avg\_trade_{i,t} + Country_i + YearMonth_t + IMR_{i,t}$.

The number of stock-month observations is 31,436. ***, **, and * refer to statistical significance at 1%, 5%, and 10% levels, respectively. T-statistics are provided in parentheses.

	No FE	*Country FE*	*Time FE*	*Country and Time FE*
First Stage	$Chix_{i,t}$	$Chix_{i,t}$	$Chix_{i,t}$	$Chix_{i,t}$
$ln(dvol)_{i,t-3}$	0.27***	0.23***	0.28***	0.24***
	(7.32)	(5.43)	(7.38)	(5.57)
$ln(volatility)_{i,t-3}$	-0.04	-0.07*	-0.13**	-0.14**
	(-1.04)	(-1.83)	(-2.22)	(-2.35)
$ln(spread)_{i,t-3}$	-0.15***	-0.20***	-0.15***	-0.22***
	(-3.24)	(-2.81)	(-3.00)	(-2.73)
$in\_index_{i,t}$	1.25***	2.92***	1.22***	2.90***
	(7.75)	(11.15)	(7.61)	(11.09)
Second Stage	$Frag1_{i,t}$	$Frag1_{i,t}$	$Frag1_{i,t}$	$Frag1_{i,t}$
$mcap_{i,t-3}$	0.07***	0.08***	0.07***	0.08***
	(9.89)	(10.75)	(9.83)	(10.69)
$price_{i,t-3}$	0.009**	-0.01***	0.01***	-0.01***
	(2.37)	(-3.43)	(3.97)	(-2.91)
$home\_tick\_to\_price_{i,t}$	42.11***	13.84*	12.92*	-12.10
	(3.87)	(1.68)	(1.77)	(-1.13)
$AT_{i,t-3}$	-0.0001	-0.0001**	-0.0001**	-0.0002**
	(-1.39)	(-2.47)	(-2.19)	(-2.51)
$avg\_trade_{i,t}$	-0.0004***	-0.001***	-0.0004***	-0.0005***
	(-3.97)	(-4.94)	(-4.56)	(-7.03)
$IMR_{i,t}$	-2.48***	-1.94***	-2.53***	-2.04***
	(-6.42)	(-5.99)	(-6.80)	(-6.33)
1st stage: R-squared	37%	64%	38%	65%
2nd stage: R-squared	31%	47%	53%	67%

Appendix E. Robustness check using alternative fragmentation measures with non-standardized variables

Table 10. Robustness check using alternative fragmentation measures with non-standardized variables

This table contains results for specification of two stage Heckman model with both country- and time-fixed effects. At the second stage, alternative measures of fragmentation ($Frag2_{i,t}$, $Frag3_{i,t}$, $Frag4_{i,t}$) are used as dependent variables. All the variables are not standardized. ***, **, and * refer to statistical significance at 1%, 5%, and 10% levels. The number of stock-month observations is 31,436. T-statistics are provided in parentheses.

First Stage		$Chix_{i,t}$		
$ln(dvol)_{i,t-3}$		0.24***		
		(5.57)		
$ln(volatility)_{i,t-3}$		-0.14**		
		(-2.35)		
$ln(spread)_{i,t-3}$		-0.22***		
		(-2.73)		
$in\_index_{i,t}$		2.90***		
		(11.09)		
Second stage	$Frag1_{i,t}$	$Frag2_{i,t}$	$Frag3_{i,t}$	$Frag4_{i,t}$
$mcap_{i,t-3}$	0.08***	0.001***	0.001***	0.001***
	(10.69)	(4.24)	(4.26)	(2.83)
$price_{i,t-3}$	-0.01**	-0.0001	-0.0002**	-0.0001
	(-2.91)	(-1.42)	(-2.42)	(-0.55)
$home\_tick\_to\_price_{i,t}$	-12.10	-0.50	-0.36	3.27*
	(-1.13)	(-0.91)	(-0.49)	(1.73)
$AT_{i,t-3}$	-0.0002**	-2.29e-06	3.43e-07	0.00001
	(-2.51)	(-0.60)	(0.07)	(0.97)
$avg\_trade_{i,t}$	-0.001***	-0.000008***	0.000008**	0.0001**
	(-7.03)	(-4.17)	(2.56)	(2.25)
$IMR_{i,t}$	-2.04***	-0.03***	-0.03*	0.12***
	(-6.33)	(-2.77)	(-1.92)	(4.14)
1st stage: R-squared	65%	65%	65%	65%
2nd stage: R-squared	67%	47%	41%	10%

Appendix F. Robustness check using alternative measures for liquidity and volatility with non-standardized variables

Table 11. Robustness check using alternative measures for liquidity and algorithmic trading with non-standardized variables

This table contains results for specification of two stage Heckman model with both country- and time-fixed effects. On the first stage, $liq_{i,t}$ is used as an alternative measure for liquidity instead of $spread_{i,t}$. On the second stage, algorithmic trading is proxied by the ration of dollar trading volume to number of messages instead of the ratio of messages to trades used previously. The variables are not standardized. The number of stock-month observations is 24,466. ***, **, and * refer to statistical significance at 1%, 5%, and 10% levels.

First Stage	$Chix_{i,t}$			
$ln(high\_low)_{i,t-3}$	-0.08**			
	(-1.95)			
$ln(liq)_{i,t-3}$	0.30***			
	(6.77)			
$in\_index_{i,t}$	2.91***			
	(11.82)			
Second stage	$Frag1_{i,t}$	$Frag2_{i,t}$	$Frag3_{i,t}$	$Frag4_{i,t}$
$mcap_{i,t-3}$	0.08***	0.001***	0.001***	0.001***
	(10.12)	(4.03)	(4.13)	(2.87)
$price_{i,t-3}$	-0.01***	-0.0002**	-0.0002***	-0.0001
	(-3.20)	(-2.07)	(-2.63)	(-0.86)
$home\_tick\_to\_price_{i,t}$	-19.30	-0.89	-0.91	4.09**
	(-1.28)	(-1.26)	(-1.17)	(2.22)
$AT_{i,t-3}$	-0.0001**	3.31e-07	4.15e-06	0.00001
	(-2.43)	(0.07)	(0.64)	(1.05)
$avg\_trade_{i,t}$	-0.0004***	-0.000008***	0.000005**	0.0001**
	(-5.21)	(-6.11)	(2.57)	(2.13)
$IMR_{i,t}$	-1.98***	-0.03***	-0.02*	0.07***
	(-7.49)	(-2.82)	(-1.79)	(3.11)
1st stage: R-squared	63%	63%	63%	63%
2nd stage: R-squared	69%	49%	45%	9%