

Multi-market Trading: Simulations and Empirical Analysis

Andrew E. Todd

December 3, 2016

Contents

1	Introduction	11
2	Methodology	18
2.1	Abstract	18
2.2	Introduction	19
2.2.1	What is agent-based modeling?	19
2.2.2	Why agent-based modeling for the study of financial markets?	20
2.2.3	What is the research strategy?	24
2.3	Modeling financial markets	30
2.3.1	Agents	30
2.3.2	Environment and Topology	36
2.4	Conclusion	40
3	Crossed and Locked Quotes	41
3.1	Abstract	41
3.2	Introduction	42
3.2.1	Limit Order Book	45
3.2.2	Market Fragmentation	48
3.3	Methods	51
3.3.1	Single-Market Model	52
3.3.2	Multi-Market Model	53
3.4	Results and Discussion	59
3.4.1	Sensitivity Analysis	66
3.5	Conclusions	70
4	Correlations, Arbitrage and Trade-throughs	72
4.1	Abstract	72
4.2	Introduction	73
4.2.1	Market Structure and Multi-market Trading	73

4.2.2	Data and Sample Period	77
4.2.3	Organization of the chapter	84
4.3	Correlations	84
4.4	Arbitrage opportunities	90
4.5	Trade-throughs	96
4.6	Conclusion	102
5	Conclusion	103
6	Appendix	105
6.1	Appendix A: Nordic Market Supporting Figures	105
6.1.1	Volume	105
6.1.2	Market Updates	113
6.1.3	Spread	117
6.1.4	Depth	121
6.1.5	Mid-quote correlation	125
6.1.6	Trade-through Volume	129
6.1.7	Trade-through Costs	133

List of Figures

1.1	Cost of computation. Cost of a million computations in 2006 USD as calculated by William Nordhaus (2001).	13
2.1	Diagram of the literature. How does agent-based modeling fit in with existing economics literature? This figure depicts the relevant literature along two axes. The vertical axis reflects the level of institutional detail that is explicitly included in the model. The horizontal axis represents the solution concept. In the bottom left of the model, we have neoclassical models, which represent the most general class of models that have to do with markets. Moving vertically, we add more and more institutional detail. The models on the left-hand side are solved analytically, whereas models on the right-hand side are generally solved numerically through simulation.	22
2.2	Schelling's model. In Schelling's model of segregation, households are distributed on a grid. Despite the majority of households favoring integration, the resulting neighborhoods are segregated. The figure shows the initial random distribution of households on the left. The right-hand side shows the outcome, which is an example of emergence.	23
2.3	Starting in the middle. Illustration of the relationship between traditional agent-based models and the models of econophysics. A traditional agent-based model of a financial market starts with a model of strategic behavior and demonstrates the emergence of certain statistical characteristics of prices. The models of econophysics start with a model of order flow and do not explicitly account for the strategic origin of those orders.	33
2.4	Neighborhoods. Illustration of neighborhoods.	36

2.5	Limit order book. Example of three events in a limit order book: 1) a new limit order, 2) a new market order is placed, 3) an existing order is canceled. The diagram also depicts the best bid and ask prices, and the mid-quote, which are important measures of price. Depth measures the aggregate quantity or resting orders in the book given a particular price or side.	39
3.1	Limit order book diagram. A new buy limit order arrives at price $b(t)$ increasing the depth from 2 to 3. A new buy market order executes at the best ask increasing the spread and changing the mid-quote. . .	47
3.2	Crossed and locked quotes. Market 1 and Market 2 are crossed since $a_1(t) < b_2(t)$. Market 3 and Market 4 are locked since $a_3(t) = b_4(t)$. . .	50
3.3	Order routing policy. A new buy limit order at price $\hat{p} = b_2(t) = b_1(t) + 1$ arrives and is routed to market 1 since it will have better priority at that price, i.e. $d_1^b(\hat{p}, t) < d_2^b(\hat{p}, t)$. In market 1, a new limit order to buy at \hat{p} is first in the queue with no depth ahead toward the best ask. In market 2, the same order is third in the queue. A new market order to sell is routed to market 2 since $b_2(t) > b_1(t)$	56
3.4	Welch's graphical method for identifying the warm-up period for the simulation.	59
3.5	Percentage of time markets are crossed.	61
3.6	Percentage of time markets are locked.	62
3.7	Percentage of time markets are normal (neither crossed nor locked).	63
3.8	Mean spread in the consolidated book.	64
3.9	Volatility as measured by the variance of the difference in transaction prices.	65
3.10	Sensitivity of crossed state duration. Proportion of time market is in crossed state as a function of limit and market order arrival rates for $\alpha = 1$ and $\beta = 0$	67
3.11	Sensitivity of locked state duration. Proportion of time market is in locked state as a function limit and market order arrival rates for $\alpha = 1$ and $\beta = 0$	68
3.12	Sensitivity analysis of the consolidated spread. Mean consolidated spread as market and limit order arrival rates are varied for $\alpha = 1$ and $\beta = 1$	69
3.13	Sensitivity analysis of volatility. Volatility of transaction prices as a function limit and market order arrival rates for $\alpha = 1$ and $\beta = 1$	70

4.1	Monthly volume. Monthly volume for each stock in each venue. . .	80
4.2	Market updates. Monthly market updates for each stock in each venue.	81
4.3	Daily average spread. Spreads in this figure are averaged over time and given as a percentage of price.	82
4.4	Daily average depth. Depths are calculated from best prices and averaged over time.	83
4.5	Nordea mid-quotes. Mid-quotes are shown in a common currency, the euro. Quotes in SEK are converted to euros using the mid-quote of the spot exchange rate.	87
4.6	Correlation range for miq-quote returns in native currency. This figure is based on the average daily correlation for Nokia, Nordea Group, Stora Enso and Telia. For each stock, we calculate the daily correlation of the mid-quote returns. We then take the average for each year, 2009 through 2013. The filled curves show the range of those averages across the four stocks for 2009 and 2013. Data for all 5 years is available in the appendix.	88
4.7	Time-evolution of millisecond correlations. Daily correlations based on a 16ms resample.	89
4.8	Cumulative number of arbitrage opportunities. cumulative number of arbitrage opportunities based on our 0.005 euro threshold. The calculation aggregates the number of arbitrage opportunities in Nokia, Nordea Group, Stora Enso and Telia Sonera over the 5 year sample period.	92
4.9	Number of arbitrage opportunities per stock.	94
4.10	Length of arbitrage opportunities. Histogram of the length of the arbitrage opportunities across all stocks in the sample for the five year time period.	95
4.11	Value of arbitrage opportunities. Histogram of the value of the arbitrage opportunities across all stocks in the sample for the five year time period. We define the value of the arbitrage opportunity to be the profit from the transaction if it were to be perfectly timed in a hypothetical market in which one could transact in both stocks in euros.	96
4.12	Volume of trade-throughs. Percentage of volume that trades through better quotes	100
4.13	Cost of trade-throughs. Difference between two available quotes at the time of the trade-through averaged over volume.	101

6.1	Euro to Swedish Kronor exchange rate over the sample period. . . .	106
6.2	Volume for Ericsson.	107
6.3	Volume for Nordea Group.	108
6.4	Volume for Nordea Group.	109
6.5	Volume for Stora Enso.	110
6.6	Volume for Telia Sonera.	111
6.7	Volume for Tieto.	112
6.8	Number of market updates for Nokia.	113
6.9	Number of market updates for Nordea Group.	114
6.10	Number of market updates for Stora Enso.	115
6.11	Number of market updates for Telia Sonera.	116
6.12	Average spread for Nokia.	117
6.13	Average spread for Nordea Group.	118
6.14	Average spread for Stora Enso.	119
6.15	Average spread for Telia Sonera.	120
6.16	Average depth at top of book for Nokia.	121
6.17	Average depth at top of book for Nordea Group.	122
6.18	Average depth at top of book for Stora Enso.	123
6.19	Average depth at top of book for Tela Sonera.	124
6.20	Mid-quote correlations for Nokia.	125
6.21	Mid-quote correlations for Nordea Group.	126
6.22	Mid-quote correlations for Stora Enso.	127
6.23	Mid-quote correlations for Telia Sonera.	128
6.24	Volume of trade-throughs for Nokia as a percentage of total volume. .	129
6.25	Volume of trade-throughs for Nordea as a percentage of total volume.	130
6.26	Volume of trade-throughs for Stora Enso as a percentage of total volume.	131
6.27	Volume of trade-throughs for Telia Sonera as a percentage of total volume.	132
6.28	Cost of trade-throughs in Nokia.	133
6.29	Cost of trade-throughs in Nordea Group.	134
6.30	Cost of trade-throughs in Stora Enso.	135
6.31	Cost of trade-throughs in Telia Sonera.	136

List of Tables

3.1	Aggregate arrival rates for the multi-market model	53
3.2	Market order routing	54
3.3	Limit order classification	55
3.4	Non-improving priority routing mechanism	56
3.5	Limit order routing for improving orders	57
3.6	Consolidated spread and volatility results	66
4.1	Reuters Identification Codes	77
4.2	Annual volume. Volume share in each currency within NASDAQ OMX where HE is NASDAQ OMX Helsinki, ST is NASDAQ OMX Stockholm and CO is NASDAQ OMX Copenhagen. Ericsson volume in HE is approximately 10^{-5} percent.	78
4.3	Annual number of arbitrage opportunities based on 0.005 euro threshold.	91
6.1	Reuters Identification Codes for the 6 cross-listed stocks.	105

Acknowledgements

I would like to thank Peter Beling, William Scherer, Patrik Sandås, Gerard Learmonth, Sr., Jamey Thompson, Steve Patek, Alfredo Garcia, Randy Cogill, Michael Smith, Yacov Haimen, Barry Horowitz, Steve Yang, Andrei Kirilenko, Mark Paddrik, Roy Hayes, Matthew Burkett and Faraz Dadgostari. I'd also like to thank my family and my wife, Garah.

Executive Summary

Careful construction of properly scoped agent-based models can offer meaningful insights about financial markets and lead to improved decision aides for financial regulators and other stakeholders. However, it is of critical importance to back these models with sound empirical analysis and careful consideration of market institutions. Analysis of financial data is increasingly difficult due to the sheer volume of data, but also due to complex market structures. Open out-cry markets have been replaced with electronic distributed-access systems, and trading strategies are highly automated, no longer subject to human response times. This dissertation presents a new agent-based model for analyzing fragmented markets, and provides an empirical analysis that demonstrates the sea change in financial markets.

In the first part of this dissertation, we survey the literature on constructing agent-based models of financial markets. The survey provides a comprehensive review of a number of topics regarding the design, implementation and analysis of agent-based financial markets. The survey also contributes to the discussion of the methodological distinction between several classes of models that are commonly referred to as agent-based models.

Next, we introduce a model designed to analyze strategic behavior in a multi-market setting. In our model, the prevalence of order routing is an experimental variable. The model allows us to study the causal relationship between certain types of strategic behavior and the statistical properties of a multi-market system. The unique approach to the specification of strategic behavior in a multi-market system is a methodological contribution to the literature on agent-based financial markets.

Finally, we provide an empirical analysis of the fragmented market for a sample of Nordic equities. The institutional setting is ideal in that the platforms are identical and centrally located, but securities trade in different currencies, which provides enough of an economic friction for us to observe the evolution of market inefficiencies as market structure changes throughout the sample period. Correlations, arbitrage opportunities and trade throughs are analyzed. We document the drastic transformation in markets with respect to the speed of exchanges and the response times of market participants. The analysis highlights the need for models that explicitly account for the mechanics and structure of markets.

In summary, this dissertation contributes to the literature on agent-based modeling, multi-market systems and market efficiency with a new model of multi-market trading and an empirical analysis of a fragmented market for equities.

Chapter 1

Introduction

NASDAQ introduced the first electronic market in 1971. At that time, market participants in over 400 cities could receive quotes in “5 or 6 seconds,” and the system handled “more than a million messages” on a busy, high volume day [1]. NASDAQ’s new computerized market was a resounding success prompting one author to worry that the “high level of automation” might cause operators to become “extremely bored” and unable to support the complex system that they created, and that was in 1972 [2]! But, it was not until around 1998, after the passage of Regulation ATS (Alternative Trading Systems), that electronic trading really began to take off [3]. The introduction new electronic trading networks led to an increase in financial market fragmentation. Not much later, in 2001, came the decimalization of U.S. equities and Reg National Market System, both of which put pressure on exchanges and participants to develop more sophisticated electronic platforms [4].

Needless to say, computerization and automation of the financial systems con-

tinues. The reduction in computing costs (illustrated in Figure 1.1) has led to the increased adoption of computing throughout every aspect of the business [5]. Traders have all but abandoned open outcry markets for fully automated systems. Nearly every aspect of market structure, including access, anonymity, products, protocols, and transparency are affected by the computerization of markets. Simply put, computing allows market participants to do more: offer new products, open new exchanges, trade more markets, transact with counter-parties around the world, and, ultimately, do more business. The cost of storing data has also decreased affording traders, exchanges, governments and other market participants the opportunity to record details the of their activity. All of the above reminds us that the financial market is a complex system, and with a paper trail!

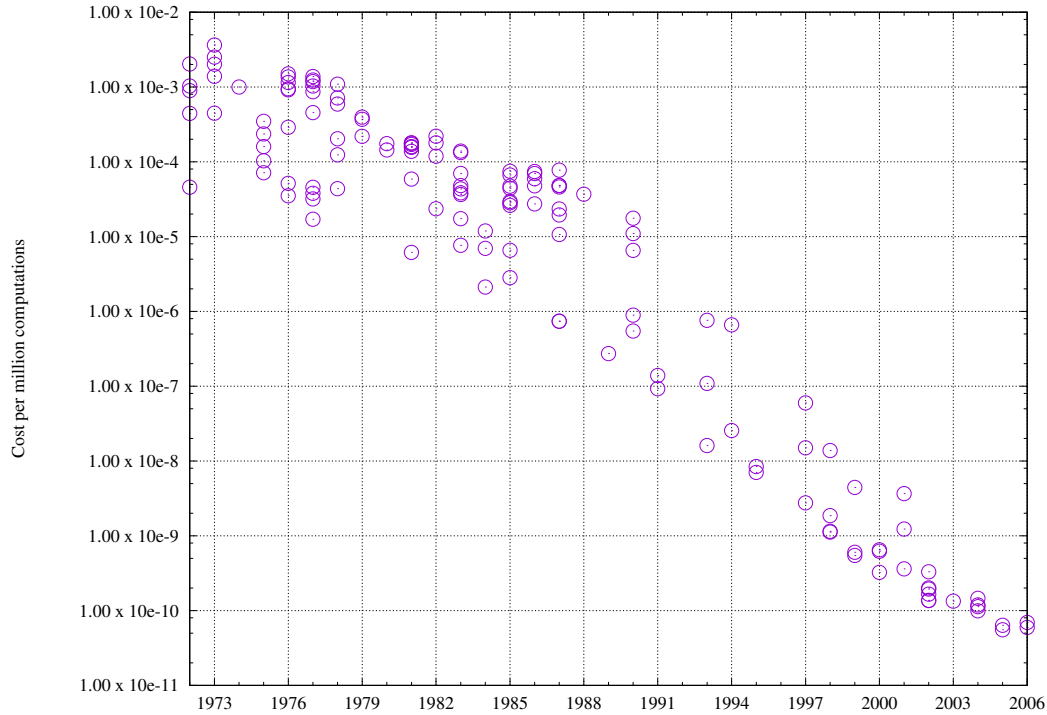


Figure 1.1: **Cost of computation.** Cost of a million computations in 2006 USD as calculated by William Nordhaus (2001).

Traders already take advantage of the data recorded about the economy by building models to evaluate the risk and reward of individual trades, but efforts at the same level of sophistication by regulators to understand the economy as a whole are largely absent [6]. Economists have long viewed markets and the economy as a complex system, but models from that perspective capable of informing policy decisions are still a work in progress. Farmer (2009) strongly advocates for a holistic approach to the agent-based modeling of economy, but concedes that it is an “ambitious” project [6]. Cliff and Northrop (2012) also advocate for agent-based models

of financial markets, but point out that such an undertaking is a “major endeavor . . . requiring national-scale levels of investment” [7]. Levin and Lo (2015) even argue that the economy has “crossed a threshold of complexity where the system is evolving faster than regulators and regulations can keep pace” [8]. Where does this leave aspiring modelers?

Agent-based modeling of the economy is a daunting prospect, but properly scoped, empirically validated, theoretically sound agent-based models that respect the real institutions of markets already contribute to understanding of financial markets. In the remainder of this introduction, we highlight a number of existing models to provide context for the work presented in this dissertation.

Complexity Economics and the Santa Fe Model

The fundamental motivation for agent-based modeling of economies and financial markets is complexity economics, which recognizes that the economy or the market is not in equilibrium, but rather in a constant state of flux in which all participants are constantly reacting to changes in global or high-level patterns that they in aggregate have created through their their local interactions [9]. The Santa Fe Artificial Stock Market is an early and important agent-based model that is still under active use and study. A recent contribution is that of Norman Ehrentreich in which the model is used to demonstrate that evolutionary mechanisms alone can not guarantee that only optimal trading rules will survive [10].

Decimalization

Darley and Outkin (2007) present a model of the NASDAQ market built to explore the consequences of decimalization [11]. At the time, NASDAQ was facing a mandate from the SEC to implement the move from one sixteenth of a dollar to \$0.01 tick sizes in the equity markets. Darley and Outkin were commissioned to build a simulation model to predict unintended consequences. (See Harris (1997) for an outline of the arguments for and against decimalization as they stood at the time [12].) Darley and Outkin used their agent-based financial model to make detailed predictions about price discovery, volume, wealth effects, spreads and trading strategies. We do not attempt to evaluate the quality of the predictions here, but rather highlight that the model was successfully applied to the evaluation of an important policy decision prior to the availability of empirical data.

Mortgage prepayments

Geanakoplos et al. (2012) describe a agent-based model developed to address the mortgage prepayment problem, and later used to understand the housing crash and financial crisis of 2008. Correct prediction of prepayments is critical to proper valuation of mortgage-back securities. Conventional approaches use “exogenously specified functional forms to describe aggregate behavior” whereas the model developed by Geanakoplos and his associates directly model the heterogeneous population of borrowers and their prepayment decisions [13]. The model, found to be “generally superior to aggregate modeling,” is another successful example of agent-based modeling applied in the private sector to evaluate financial decisions.

Market design

The design of the U.S. equity market comes up frequently in debate. Budish, Crampton and Shim (2015), for example, argue that financial exchanges should use frequent batch auctions in lieu of operating continuous limit order books [14]. Wah and Wellman (2013) use an agent-based simulation to evaluate the allocative efficiency of a fragmented market under two different market designs, the continuous limit order book and frequent batch auctions [15]. The model is an important contribution to the literature since it demonstrates the value and need to explicitly model the real mechanics of markets, and is another example of an agent-based model that can directly contribute to policy discussions.

Flash crashes

The market event of May 6th, 2010, now well known as the Flash Crash, is an example of how fast automated markets can move. The flash crash may have been a perfect storm that hit a relatively immature new electronic market structure, but there are still lessons to be learned. Events like the flash crash are intimately related to market structure, and can not be explained by fundamental information or the rational behavior of market participants. Debate about the causes of the flash crash are ongoing, but it is clear that the event is closely tied to the highly automated and distributed market environment [16]. Agent-based models provide a powerful framework for analyzing alternative hypothesis about the flash crash, as well as better understanding the mechanics of the event. Paddrik et al (2012) explore the hypothesis of a large order flow imbalance and the reaction of different classes of

traders [17]. Vuorenmaa and Wang (2014) also explore the mechanics of an order flow imbalance coupled with automated market making strategies [18]. Leal et al (2016) study the interaction of low- and high-frequency traders to better understand how flash crash dynamics [19]. These works demonstrate the power and convenience of agent-based approaches to understanding the dynamics associated with increasingly complicated and distributed market structures.

Summary

The above selection of agent-based models illustrate the numerous decision problems and market dynamics that can be analyzed with agent-based models. With these models in mind, we embark on a thorough review of the methodology, followed by presentation of a new agent-based model of a distributed market structure, and finally, an empirical analysis that highlights the evolution of financial market data and the need for creative new approaches to analyzing financial markets.

Chapter 2

Methodology

2.1 Abstract

An agent-based model is a computer simulation driven by the individual decisions of programmed agents. Such models provide a promising alternative to traditional economic modeling in that they can fully capture the diversity of agents and the institutional detail of the underlying an economic system. In this chapter, we provide a methodological review of the agent-based approach to modeling financial markets. We review the research strategy, which is organized into a discussion of formulation, implementation, verification and validation. We conclude the paper with a review of the domain focusing on modeling market participants and market structure.

2.2 Introduction

Agent-based modeling provides a powerful framework for modeling a heterogeneous population of economic actors, market mechanisms, and institutions that are often absent from traditional theoretical models. In this paper, we introduce agent-based modeling by answering several basic questions about the method. We then provide a survey of agent-based modeling of financial markets organized around major design points.

2.2.1 What is agent-based modeling?

We begin this survey with the obvious question: what is agent-based modeling? Agent-based modeling is a simulation-based methodology that has found applications in a number of fields including anthropology (Epstein and Axelrod (1996), ecology (Grimm and Railsback, 2005), economics (Tesfatsion and Judd, 2005), political science (Marchi and Page, 2014), and others [20, 21, 22, 23]. Agent-based modeling involves the simulated interaction of a heterogeneous population of autonomous agents. For general introductions to agent-based modeling, see the monographs by Miller and Page (2007) and Railsback and Grimm (2012), Wilensky and Rand (2015), and Namatame and Chen (2016) [24, 25, 26, 27].

The specification of an agent-based model includes a set of agents, a topology and an environment [28]. The agents have individual rules of behavior and internal states. The topology provides a spatial orientation for the agents where “nearby agents are more likely to interact and better able to influence each other” [29]. Examples

include a grid, a network or euclidean space. The environment consists of non-agent entities that may influence the decisions of agents. Models are typically analyzed with numerical simulations.

2.2.2 Why agent-based modeling for the study of financial markets?

Why choose agent-based modeling for the study of financial markets? Bankes (2002), in a survey of research presented at a colloquium of the National Academy of Sciences, identifies three basic reasons researchers cite for their selection of agent-based modeling as a methodology [30].

The first reason that Bankes cites is the “unsuitability of competing modeling formalisms.” Traditional approaches such as differential equations and statistical modeling lead to restrictive assumptions such as “linearity, homogeneity, normality, and stationarity.” For example, in economics, the “analytical convenience” of the representative agent is criticized as being neither “descriptive of actual behavior nor useful as a normative model” [31, 32]. Computational methods, such as agent-based simulation, allow such assumptions to be relaxed.

Simulation modeling, however, has not been wholly embraced by the economics community. According to Lehtinen and Kuorikoski (2007) this is due to the fact that simulation “does not fit the conception of understanding inherent in mainstream economics – economists’ understanding is constituted by analytical derivation from a set of fundamental economics axioms” [33]. Simulation results may be rejected on the grounds that the “deductive links between the assumptions and the consequences are

not transparent” and therefore do not enhance economic understanding, which raises the question: are agent-based models suitable for developing economic theory? Is a computational methodology a substitute for the “assumption-theorem-proof” approach, or is it merely complimentary [34]? Figure 2.1 orients agent-based modeling within the existing literature on financial markets. The figure is a simplification, and only provides some basic intuition for the structure of the literature.

The second reason Bankes cites is that the agent-based approach provides a natural representational formalism. Consider the canonical example of a flock of birds [35]. Reynolds finds that three simple rules, collision avoidance, velocity matching, and centering, lead to convincing simulation of flocking. It seems natural to model a flock by specifying the individual rules of behavior of a single bird, but quite daunting to come up with an equation of their flight trajectories. Access to modeling individual behavior directly and explicitly is a strength of agent-based modeling, especially given the centrality of economic agents in financial markets. The agent-based approach also allows for explicit specification of market institutions and market structure.

The final reason is emergence: the materialization of “novel and coherent structures, patterns and properties during the process of self-organization in complex systems” [36]. In other words, agent-based models allow us to study how micro-motives lead to macro-behaviors [37]. Schelling’s pioneering model of segregation is an early example of emergence in the economics literature. Households situated on a grid take turns following simple relocation rules according to their preferences. The procedure continues until each household is satisfied. Figure 2.2 illustrates one of

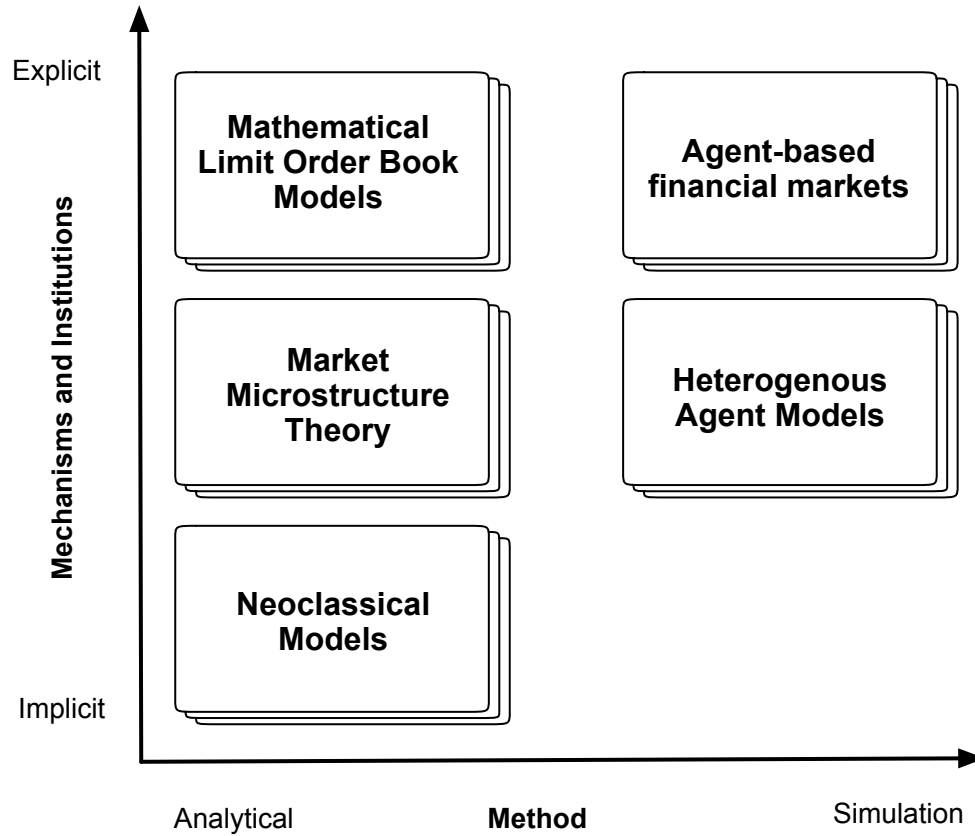


Figure 2.1: **Diagram of the literature.** How does agent-based modeling fit in with existing economics literature? This figure depicts the relevant literature along two axes. The vertical axis reflects the level of institutional detail that is explicitly included in the model. The horizontal axis represents the solution concept. In the bottom left of the model, we have neoclassical models, which represent the most general class of models that have to do with markets. Moving vertically, we add more and more institutional detail. The models on the left-hand side are solved analytically, whereas models on the right-hand side are generally solved numerically through simulation.

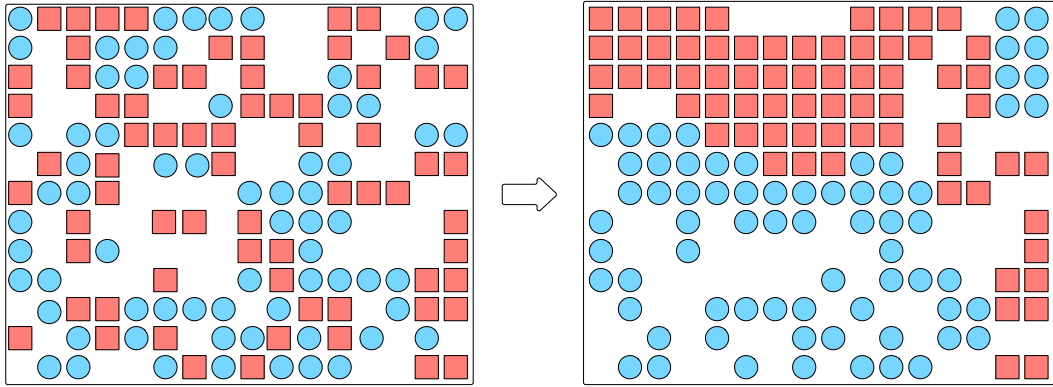


Figure 2.2: **Schelling's model.** In Schelling's model of segregation, households are distributed on a grid. Despite the majority of households favoring integration, the resulting neighborhoods are segregated. The figure shows the initial random distribution of housesholds on the left. The right-hand side shows the outcome, which is an example of emergence.

Schelling’s original experiments in which a random distribution of households eventually organizes into segregated neighborhoods despite the fact that the majority of households prefer integration. In the context of financial markets, we might ask how flash crashes in financial markets arise from the trading strategies of market participants in a particular market structure [17]. Bankes points out the difficulties with identifying emergent phenomena, reminding the reader that most studies “rely on human observers to declare emergence to have occurred based on graphical computer outputs.” As an aside, it is important that emergent behavior is rigorously quantified.

In summary, agent-based modeling is a preferred framework for the following reasons:

1. Unsuitability of competing approaches
2. Natural representational formalism
3. Evolutionary modeling and emergent behavior

2.2.3 What is the research strategy?

The goal of agent-based models is to provide superior understanding of complex systems and provide explanations of emergent phenomena. The goal of an individual model is usually the provision of an explanation for phenomena that are not adequately accounted for by existing theories. Below we provide a rough outline of an agent-based research strategy based on the survey by Hedstrom et al [38].

- Identify phenomena lacking a satisfactory explanation

- Formulate hypothesis of micro-level mechanisms
- Translate hypotheses into a computational model
- Simulate the model to derive the macro-level results
- Compare the simulations results with empirical data

Stylized facts

The phenomena of interest in the agent-based modeling of financial markets are the stylized facts. The stylized facts are statistical regularities “common across a wide range of instruments, markets and time periods” [39]. Pagan (1996) establishes a number of stylized facts of financial market time-series, and concludes that “statistical approaches to the modeling of financial series have possibly reached the limits of their usefulness,” a comment which one might interpret as motivation for a generative, agent-based approach [40]. Stylized facts range from statistical concepts, such as the distribution of returns, spread, depth and volume to more abstract notions such as stock market bubbles and crashes, and leverage effects [41, 42]. As the data quality and availability improves, we expect the stylized facts of financial markets to converge to a stable set of broadly observable features. To see what has been accomplished in this area, see the work of Chen, Chang and Du (2012), which identifies the stylized facts that are reproducible by different classes of agent-based models [43].

Generating hypotheses

The second step is to formulate hypotheses about the relevant “micro-level mechanisms,” and develop a conceptual model. In this stage, we need to identify entities and state variables and design the agents. The design of agents raises a number of practical questions: Do the agents observe the environment or learn? Do they make predictions? How are the agents initialized? Does the population evolve? Are there stochastic elements of the model? Random shocks, for example, such as a hurricane or economic depression. In this stage, we specify the model on a conceptual level and determine an experimental design. It is helpful to follow an established protocol, such as the OOD protocol to ensure that all aspects of the conceptual model are well defined [44].

Implementing the model

Next is the translation of the conceptual model into a computational model: an exercise in software engineering. A number software packages are available for agent-based simulation, but many simulations are necessarily developed from scratch [45, 46]. Rossiter (2015) outlines three basic requirements of a simulation implementation [47]:

- automated reproducibility
- cohesive, loosely-coupled design
- testability

Goodman, Fanelli and Ioannidis (2016) identify “methods reproducibility” as “the ability to implement, as exactly as possible, the experimental and computational procedures, with the same data and tools, to obtain the same results” [48]. Automated reproducibility takes this notion a step further and requires automated recreation of the computing environment, which can be of critical importance since computer architecture and underlying libraries may affect results.

Cohesive, loosely-coupled design simply refers to the application of software best practices, for example, the use of design patterns. Good design will lead to “modeling flexibility” — “the ability to model a system whose operating procedures can have any amount of complexity” (Law 2007, Ch. 3.4.1) [49]. Modeling flexibility is of critical importance in simulation since we are often comparing alternative system designs. In agent-based modeling of financial markets, we may be comparing alternative market mechanisms and market institutions as well as a variety of agent specifications.

Testability requires simulations to provide a suite of unit tests or regression tests that automate the traditional process of verification, which is the essential process of ensuring that the implementation or coding of the computational model is consistent with the conceptual model [50]. Verification is not to be conflated with validation, which refers to examination of the validity of modeling assumptions with respect to existing theoretical and empirical knowledge.

At the core of simulation models is random number generation. Random number generation (RGN) is usually available out of the box in most programming languages. However, being that it is a critical component of that analysis, it is worth taking the

time to ensure that is of sufficient quality for simulation analysis. Random number generation provided in some early programming languages is known to be unusable for monte-carlo style analysis. L’Ecuyer and Simard (2007) provide a comprehensive review of RGN algorithms and provide a package (TestU01) to evaluate random number generators [51]. We now review a few common structural components of agent-based models that that require RNG.

Agent activation

How are the agents activated, i.e., how do they take turns? It is a subtle consideration, but one that may significantly affect outcomes [52]. Radax and Rengs (2010) survey the literature on activation schemes for agent-based models [53]. Their survey points out that the topic is somewhat neglected in the agent-based modeling literature, and they recommend drawing on the cellular automata literature. There are two basic modes of activation: synchronous and asynchronous. Classical cellular automata update synchronously [54]. Variations of the update scheme, however, are found to affect the dynamics [55]. Activation in agent-based modeling is often asynchronous (sequential). A simple choice for activation is to select an agent uniformly at random. Ultimately, we’d prefer simulation results to be robust across activation regimes, and it is recommended that such design choices are varied as part of a sensitivity analysis.

Agent initialization

Another decision point for agent-based modelers is agent initialization. For example, how much of a particular resource is an agent endowed with? Where does an agent start in a spatial sense? Consider again the Schelling model: each location on the grid is initialized randomly as either a type of household or empty. Aspects of the topology, such as the lattice on which households exist in Schellings model, are generated randomly. However, initialization based on empirical distributions can sometimes prove critical to an analysis.

For example, Gabix et al (2003) highlight the importance of considering the distribution of the size of market participants in explanations of financial market fluctuations [56]. Toth et al (2011) and Mastromatteo, Toth, and Bouchaud (2014) employ similar ideas in agent-based models which account for the size-distribution of “meta-orders” [57, 58] Large parent orders are often broken up into smaller child orders, which may be submitted and executed over a a few minutes, several days or even months. The size distribution of such orders can explain certain statistical features of the order flow.

Simulation analysis

Simulations are generally classified as finite-horizon simulations or steady-state simulations. A steady-state simulation, i.e. a non-terminating simulation, may have steady-state parameters that reach an equilibrium value or periodicity [59]. The type of simulation governs the method of output analysis. There are well-known statistical methods for analyzing both finite-horizon and steady state systems [60].

A finite-horizon model, such as the model of Schelling, requires replications. When simulating steady-state systems one may decide to proceed with independent replications or batch observations in a single run [61]. Batching “ameliorates the effects of initialization bias”, but may produce correlated batch means, where as replication “yields independent sample means, but may suffer from initialization bias.” Agent-based modelers should also be wary of typical problems with steady-state simulation such as initialization bias, i.e. steady-state simulations often require a “warm up” period. Welch’s method is a practical approach to dealing with the problem of initial transience [62].

2.3 Modeling financial markets

We now turn to the domain specific details of the agent-based modeling of financial markets. We consider each component, agents, topology, and environment, from a financial market perspective.

2.3.1 Agents

An artificial agent in the most general sense is an entity that perceives its environment through sensors and acts upon the observations through actuators [63]. In the context of financial markets, agents observe the market process and fundamental information about the securities being traded and initiate trades based on their observations. According to LeBaron (2001), an agent’s directive is to “digest the large amounts of time series information generated during a market simulation, and convert this into

portfolio decisions” [64]. However, as LeBaron notes, there are many ways to process that data, and thus many ways to define agents. Furthermore, not every agent is identically motivated to trade.

According to a taxonomy of traders provided by Harris (2001), there are utilitarian traders, speculators, dealers, and even futile traders [65]. According to Harris, utilitarian traders include, asset exchangers, hedgers, gamblers and fledglings (among others). Hedgers trade to manage risk. Asset exchangers convert from one type of holding to a holding of more immediate value. Gamblers trade for the thrill of it. Fledglings trade to determine if trading may be profitable. Speculators process information to predict future prices. Speculators may trade on fundamental information about the value of an asset or on technical information related to the trading process, or even a combination of both. Dealers exist to intermediate markets. Futile traders are irrational or victimized market participants that do not profit from their activity.

The variety of financial market participants presents a tough modeling challenge. Traditionally, economics has solved this problem by working with a perfectly rational representative agent. So, how do designers of artificial markets approach the design of agents? Let’s assume we abandon the perfectly rational agent and attempt to respect the heterogeneity of participants and their cognitive and informational limitations. It is a tall order; why? Holland and Miller (1991) put the challenge this way: “Usually, there is only one way to be fully rational, but there are many ways to be less rational” [66]. In other words, when we abandon the rational agent model, we’re met with a modeling problem with many degrees of freedom. Developing agents based on evolutionary principles is one particular approach to addressing the problem [67].

Clearly, agents vary widely with respect to their motivations and how they observe and learn from their environments, and modeling them is a difficult and open problem. So, where to start? Consider the extreme case of an agent that does not observe, does not learn and selects from its space of actions randomly. Here, you have what is known as a zero-intelligence agent. On the other end of the spectrum you have an agent that learns from its observations, makes predictions based on internal models, and evolves its behavior accordingly, i.e., an agent that relies on artificial intelligence. While the models on the two extremes of this spectrum are, broadly speaking, agent-based models, they represent distinct methodological approaches [68].

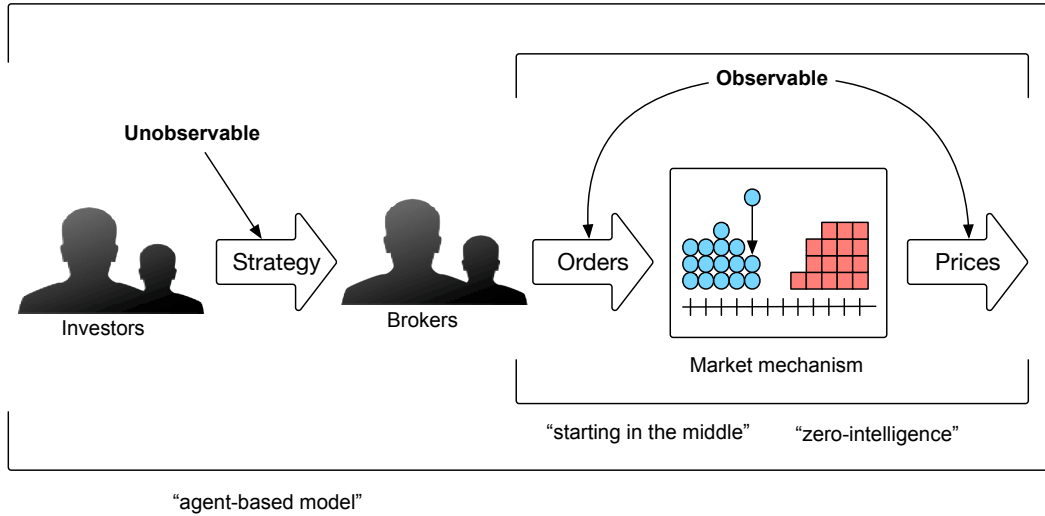


Figure 2.3: **Starting in the middle.** Illustration of the relationship between traditional agent-based models and the models of econophysics. A traditional agent-based model of a financial market starts with a model or models of strategic behavior and demonstrates the emergence of certain statistical characteristics of prices. The models of econophysics start with a model of order flow and do not explicitly account for the strategic origin of those orders.

As Schinckus points out, strictly statistical methods and agent-based methods are often associated in the literature, but represent very different approaches to modeling agents. What is the link between these two methodologies? Let us first consider the work of Mike and Farmer (2008). In their paper, they develop a rather sophisticated model for order placement and order cancelation. The order flow parameters are estimated empirically, and the statistical properties of the resulting price process are compared to those of the actual price series. The methodology is motivated by the fact that market participants are incredibly diverse, their motivations and strategies

are unobservable, and all in all, their behavior hard to model. The approach relies only on observable quantities in the system: the input — the order activity, and the output — the price process. Consider an explanation of the approach given in their own words:

The methodological approach that we have taken here can be viewed as a divide and conquer strategy. We have tackled the problem of price formation by starting in the middle. Rather than trying to immediately derive a model based on strategic motivations, we have empirically characterized behavioral regularities in order flow. From here one can work in two directions, either working forward to understand the relation between order flow and price formation, or working backward to understand the strategic motivations that give rise to the regularities in the first place.

Models of this type have the advantage of avoiding the need to directly model human (agent) behavior. Clearly, this type of model is not an agent-based models in the traditional sense. It is a purely statistical models, often referred to as an econophysics models, which adopts the approach of modeling aggregate order flow statistically without considering the individual strategic origin of that order flow. However, we can still provide an agent interpretation. Simply put, agents exist for a fleeting moment in which they submit a single order.

Importantly, these statistical models are still able to include whatever degree of institutional and mechanical detail the modeler deems necessary or correct. Models of this type also naturally lend themselves to the traditional process of model development and evaluation employed in predictive modeling, i.e. develop the model on

a subset of data and evaluate it on a hold-out set of data.

Now, on to the left side of Figure 2.3. How do you go about designing a population of agents? How do these models deal with heterogeneity of market participants and the “many degrees of freedom” problem? An early approach to the problem is to only consider a limited number of classes of agents, i.e. the few-type model [?]. The canonical example of an n-type model is the class of models that pit “fundamentalists” versus “chartists” [69]. Fundamentalists base their trading activity on analysis of economic fundamentals, e.g. a firm’s annual report. Chartists base their trading activity on historical prices and “technical analysis.” [70] for example, study a model in which the fraction of agents that believe in the mean-reversion of prices to fundamental values varies dynamically in a market where the remainder of the agents base their expectations on historical trends.

Alternatively, one might consider the evolutionary approach to agent construction. Agents are initialized with simple internal models, but are endowed with the ability to learn and adapt those rules through an evolutionary mechanism [71]. Early works used genetic programming to evolve double auction strategies that ultimately out-performed hand-coded strategies [72]. Evolving agents with genetic programming is well-known in the agent-based modeling community due to its application in the Santa Fe Artificial Stock Market [73, 10]. Genetic programming has also been employed as a mechanism for agents to improve strategies based on technical analysis [74].

2.3.2 Environment and Topology

Where do agents live? Do they move around? Who are their neighbors? What can the agents observe? A topology and environment, defined by the modeler, answers these questions. Traditionally, agents are oriented on a grid or network as shown in Figure 2.4. In the case of financial market modeling, topology also defines aspects of the market structure. We may be tempted to think of network of traders that transact directly with each other, but a more common scenario is a centralized exchange. In taking a closer look at the topology of agent-based financial markets, we review market structure concepts and the market mechanisms themselves.

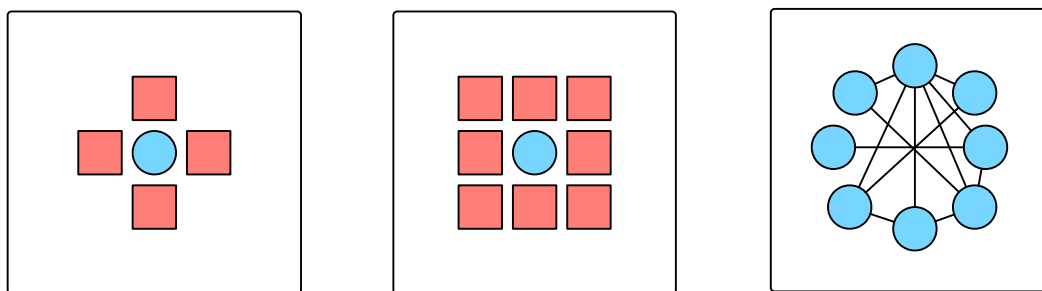


Figure 2.4: **Neighborhoods.** Illustration of neighborhoods.

Market Structure

Market structure, in the most general sense, refers to the competitive organization of an industry. However, when the term is used specifically in the context of financial markets it often refers to the financial firms, exchange, products, rules and systems that make up that market. In our discussion of market structure, we review two important concepts for agent-based modelers: market access and market

fragmentation.

Market Access

Market access is an important aspect of market structure. It defines the venues at which agents can trade. Markets are no longer “physically convened” (for the most part), and are instead “distributed access systems” [65]. In a distributed access system, a market participant does not necessarily need to be located on the premises of an exchange, and interaction with markets takes place electronically over computer networks. While some participants may invest heavily in computer infrastructure colocated with the exchange, others may use brokers or other less costly means. Market access has important consequences in terms of the strategic behavior, and is therefore an important consideration when developing an agent-based model.

Market Fragmentation

Markets may also be fragmented, i.e. the same or nearly the same security may trade in on a number of different venues. Market fragmentation increases the complexity of modeling agents since now trading decisions involve not only the how to trade, but also where to trade. Moreover, the existence of more than one market introduces the possibility that the properties of those markets may diverge. For example, a security may trade on a continuous limit order book at two different exchanges, but those exchanges could have different fees or margin requirements. Markets may also have different trading rules and protocols. For example, U.S. equities trade on open limit order book, but also trade on dark pools [75]. The United States has a notoriously

complex equity market structure. Reg NMS places specific requirements on where and how U.S. equities can be traded. Agent-based modeling offers researchers a flexible platform for evaluating alternative market structures and regulations.

Market Mechanisms

In an agent-based model of financial markets, a market mechanism plays a defining role in how agents interact. A typical approach is to mimic a dealer market by modeling price formation as function of excess demand, i.e. $\Delta P = f(D(P))$ where $D(P)$ is the excess demand at price P and f is an increasing function [76, 77, 69]. Financial markets, however, have largely transitioned from being quote-driven markets to order-driven markets, and such an approach eliminates the possibility of modeling continuous markets. More recent work explicitly models the limit order book mechanism, which underlies the majority of financial trading globally.

In a limit order book, illustrated in Figure 2.5, agents may continuously place orders, modify those orders, and transact against the resting orders of other participants. (See Gould (2013) for a detailed review of limit order books [78].) Adopting a limit order book mechanism has important modeling implications since now our agents need to know how to trade in addition to making a decision to buy or sell. However, from a statistical point of view, we can now observe more data related to an agent's strategy. Limit order books are either level-based or order-based. In an order-based book, the placement, modification and cancelation of each individual order is observable. In a level-based book, only the total number of orders and total number of participants at each price is available. Both types of data can be used to

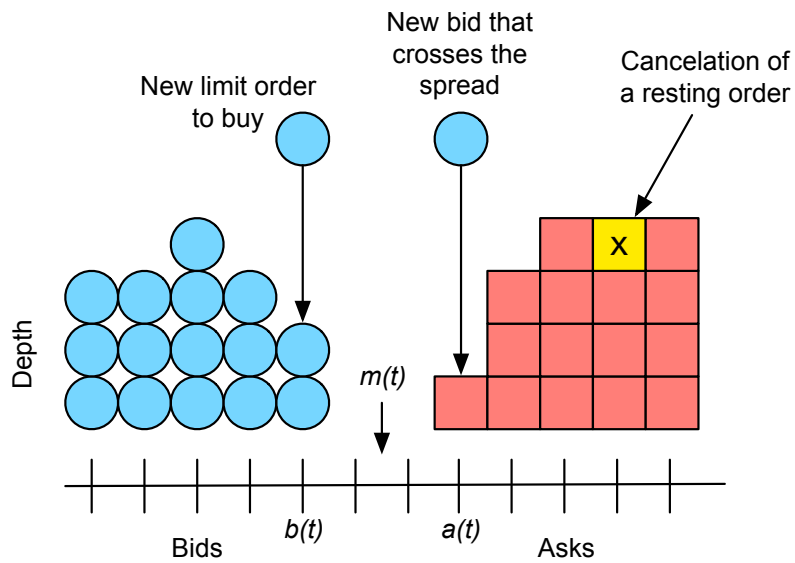


Figure 2.5: **Limit order book.** Example of three events in a limit order book: 1) a new limit order, 2) a new market order is placed, 3) an existing order is canceled. The diagram also depicts the best bid and ask prices, and the mid-quote, which are important measures of price. Depth measures the aggregate quantity or resting orders in the book given a particular price or side.

empirically validate artificial agents.

2.4 Conclusion

Our introduction to agent-based modeling of financial markets provides a review of the methodology, as well as detailed discussion of major design points including the design of agents and the design of the markets (environment and topology). Agent-based modeling is still in its infancy when it comes to producing predictive models and decision-aides for regulators and other stakeholders. However, it provides a realistic experimental platform for the evaluation of new rules, structures and regulations, and we expect it to be an area of continuing growth and interest.

Chapter 3

Crossed and Locked Quotes

3.1 Abstract

Financial markets are often fragmented, introducing the possibility that quotes in identical securities may become crossed or locked. There are a number of theoretical explanations for the existence of crossed and locked quotes, including competition, simultaneous actions, inattentiveness, fee structure and market access. In this paper, we perform a simulation experiment designed to examine the effect of simple order routing procedures on the properties of a fragmented market consisting of a single security trading in two independent limit order books. The quotes in the two markets are connected solely by the routing decision of the market participants. We report on the health of the consolidated market as measured by the duration of crossed and locked states, as well as the spread and the volatility of transaction prices in the consolidated market. We aim to quantify exactly how the prevalence of order

routing among a population of market participants affects properties of the consolidated market. Our model contributes to the zero-intelligence literature by treating order routing as an experimental variable. Additionally, we introduce a parsimonious heuristic for limit order routing, allowing us to study the effects of both market order routing and limit order routing. Our model refines intuition for the sometimes subtle relationships between the prevalence of order routing and various market measures. Our model also provides a benchmark for more complex agent-based models.

3.2 Introduction

Financial market participants face a competitive and diverse multi-venue financial system. Brokers and traders monitor a multitude of market feeds, which they rely upon to manage their trading operations. The number and variety of venues increases the complexity of trading decisions. Orders may be routed to a large number of exchanges, which may vary widely with respect to the trading platform, fees and rebates, liquidity and transparency.

Despite the complexity of the trading landscape, prices across venues should largely agree; the law of one price in financial markets should hold so far as traders can identify and act on arbitrage opportunities [79]. However, arbitrage is not the only type of strategic behavior that aligns quotes and prices across venues. Market participants with access to multiple trading venues route new orders and revise existing orders according to the state of the market as a whole.

Understanding the interaction of competing markets is of critical importance

for policy-makers. Zero-intelligence (ZI) models have proven useful in the analysis of individual limit order book markets, but have not been widely applied to study multi-market systems. We propose an extension to the standard ZI model to a multi-market setting by introducing order routing as an experimental variable. Prior to discussing the model in detail, we first discuss the explanatory power of ZI models, and discuss how our model in particular contributes to the existing literature.

ZI models provide explanations by focusing on the deterministic mechanisms at play within a given system. The approach differs from the usual methods of economic theorizing, where the focus is on the strategic behavior of agents. Not only do ZI models differ from the standard game-theoretic models of economics, but they also differ from agent-based models. While agent-based models often do include realistic implementations of market mechanisms, their focus is usually on local interactions and the evolution of strategies or agent compositions. While ZI models may be viewed as simple agent based models, their mode of explanation is fundamentally different.

ZI models provide a mechanism-based explanation, with the idea that “proper explanations should detail the cogs and wheels of the causal process through which the outcome to be explained was brought about” [80]. In a ZI model of financial markets, however, we admit defeat with respect to the possibility of modeling all of the possible strategies at play. Instead, we provide a simple random model of aggregate order flow. The order flow model is a maximum entropy distribution across allowable actions within the market. With a model of order flow in place, we can then isolate the effects of the mechanisms, which in our case, are the limit order

book and the order routing procedures.

ZI models have a long history, dating back to work by Becker [81]. More recently, ZI models have demonstrated remarkable predictive power when order flow parameters are fit to data [82]. ZI models have also been used to theorize about fundamental properties of markets such as liquidity and price impact [83]. ZI models have also proven useful for the evaluation of competing realize variance estimators [84]. In addition, ZI models have been used to compare the allocative efficiency of alternative market structures and latencies [85].

In this work, we present a two-market model where order routing is the only mechanism that integrates prices across markets. Our experimental design varies the prevalence of limit order and market order routing. Agents either have access to both markets, in which case they may route their orders, or have access to only one market, in which case their orders may trade through better quotes. The primary contribution is the study of the prevalence of order routing on various market measures, including spread, volatility and crossed or locked states. Prior work does not focus on order routing as the experimental variable. We present new results based on a unique experimental design. Along these lines, the contribution of the paper includes development of intuition with regards to the above relationships. Additionally, we introduce a parsimonious heuristic for limit order routing, which allows us to study the effects of both market order routing and limit order routing. Finally, our model provides baseline results with which to compare more complex models.

3.2.1 Limit Order Book

A limit order book is a continuously evolving record of outstanding orders to buy or sell a financial product. When a limit order is submitted to an exchange it is either matched with another outstanding order, queued in the limit order book, or rejected. Limit orders have a price, quantity and side, which are the primary attributes that govern how they are processed. The price and quantity of a limit order are restricted to multiples of the tick size and lot size, respectively. The tick size is the minimum price increment (e.g. a penny). The lot size is the minimum quantity that may be traded in a single transaction (e.g. 100 shares). A limit order may be rejected if it does not adhere to these restrictions. A limit order only executes if there exists an appropriately priced order on the opposite side of the market. Limit orders are queued in the book if they cannot be executed. Queued orders may be modified or canceled.

The queued orders establish the best bid and ask prices. Market participants seeking to trade immediately will buy at the best ask price and sell at the best bid price. The queued orders are executed according to rules of precedence. The primary rule of precedence is price. The secondary rule of precedence is time. The limit order book is essentially a FIFO queue in which orders are executed based on price and origination time. However, orders may also be matched on a pro-rata basis, and matching algorithms, in general, may vary across exchanges and products.

Generally speaking, exchanges offer limit order book functionality beyond the basics required to operate a continuous double auction. Limit orders may have additional attributes that govern specific aspects of their display and execution. For

example, some exchanges offer the ability to place hidden orders. Other attributes might control details of execution (e.g. immediate-or-cancel, fill-or-kill or all-or-none). In the United States, exchanges have introduced order types that are tied to specifics of the market structure (e.g. orders pegged to the national best bid and offer).

For modeling purposes, we choose to focus on only the core functionality required for the operation of a continuous double auction. Going forward, we make a few simplifying assumptions. First, the only order types will be market orders or limit orders with unit quantity. Unit order quantities eliminate the need for logic to handle partial executions. Instructions may be placed for the creation of a new order or the cancelation of an existing order.

Fig 3.1 depicts the evolution of the limit order book as a new limit order to buy and a market order to sell arrive. The best bid and ask at time t are denoted $b(t)$ and $a(t)$, respectively. In the figure, a buy limit order increases the depth at the best bid. Depth is the aggregate quantity available at a particular price. A market order to buy consumes the remaining depth at $a(t)$, subsequently increasing both the spread, $a(t) - b(t)$, and mid-quote, $(a(t) + b(t)) / 2$.

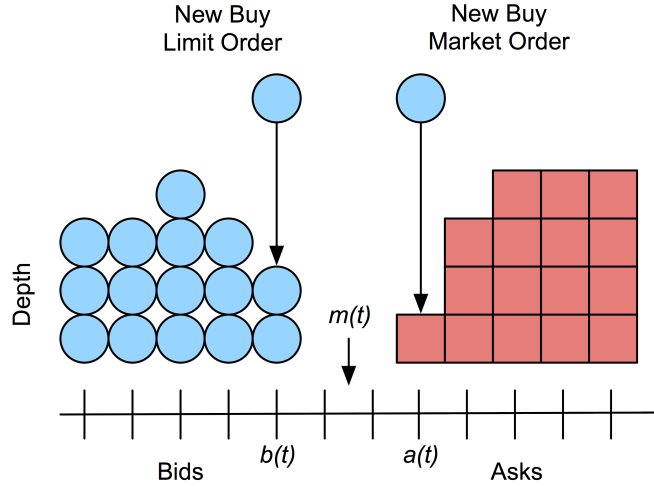


Figure 3.1: **Limit order book diagram.** A new buy limit order arrives at price $b(t)$ increasing the depth from 2 to 3. A new buy market order executes at the best ask increasing the spread and changing the mid-quote.

Limit order books, and financial time-series more generally, may exhibit a number of statistical features, commonly referred to as stylized fact [78]. The stylized facts of an order book vary across securities and markets, and may be effected by external factors, such as wider market volatility. For the purposes of our analysis, the primary focus is on the interaction of bid and ask quotes in a two-market system. However, we

also examine transaction price volatility. While there are a number of other stylized facts of financial markets, we feel that we have selected those most relevant to the goal of the analysis.

3.2.2 Market Fragmentation

Many countries have opened the door for competition in their financial markets. In the United States, equity trades are distributed across 11 exchanges and over 40 alternative trading systems. There are also over 200 broker-dealers who internalize client orders [86]. In Europe, the Markets in Financial Instruments Directive (MiFID) also increased competition in equity trading [87]. Securities also trade in other countries and in different currencies. Competition among exchanges can benefit investors, but market fragmentation underlies a number of controversial practices related to high frequency trading [88]. Understanding how individual markets are linked together is critical for our ability to assess the quality and function of the overall market [89].

In a fragmented market, traders must determine how attractive each venue is for trading. When a market participant makes an investment decision that decision is translated into a stream of orders, which may be routed to a number of available exchanges. Market participants may also trade on alternative venues, such as dark pools. Decision-makers must weigh a number of factors, such as price, liquidity, fees and rebates. Orders that are not appropriately routed may trade through better quotes.

A trade-through is a transaction that executes at quotes that are inferior to quotes simultaneously available in another market. For example, if $b_1(t) > b_2(t)$

and a market sell order executes in market 2, a trade-through occurs. Given the opportunity, a rational market participant would prefer to sell at a higher price and route the order to market 1 since it has superior quotes. The United States has regulations that are designed to protect retail order flow. For example, Reg NMS (Rule 610.d and Rule 611) establishes specific rules regarding trade-throughs and the display of crossed or locked quotes. Many markets, however, have no such restrictions and quotes across markets are connected only by the behavior of market participants.

Market participants may also seek to exploit arbitrage opportunities. An opportunity for instantaneous profit exists if a bid price in one market is greater than an ask price in another. The size of the arbitrage opportunity is constrained by the depth available at each price. In a two-market setting, markets are crossed if $b_1(t) > a_2(t)$ or $b_2(t) > a_1(t)$. Markets are locked if $b_1(t) = a_2(t)$ or $b_2(t) = a_1(t)$ (Fig 3.2). Crossed markets represent a theoretical arbitrage opportunity, which may be difficult to exploit in practice.

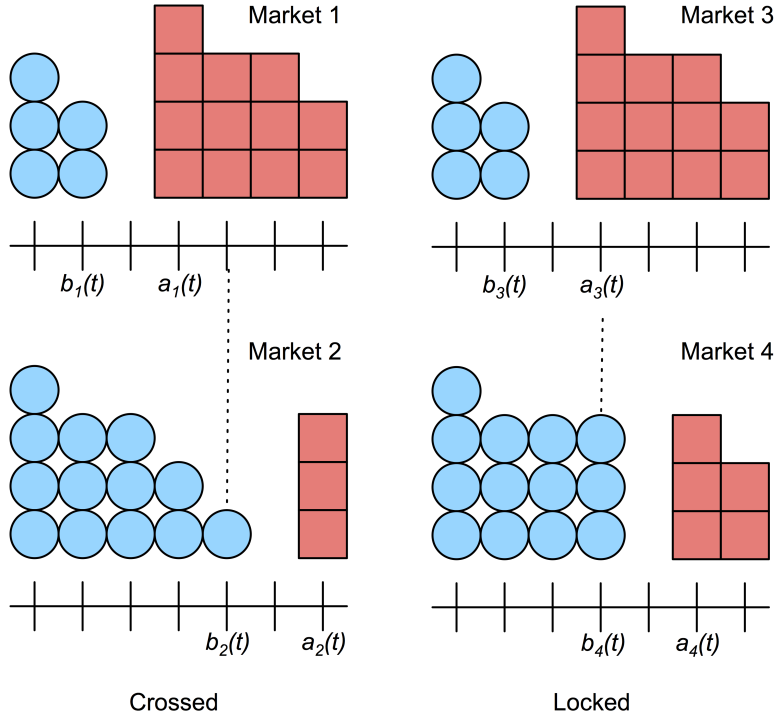


Figure 3.2: **Crossed and locked quotes.** Market 1 and Market 2 are crossed since $a_1(t) < b_2(t)$. Market 3 and Market 4 are locked since $a_3(t) = b_4(t)$.

Crossed and locked markets arise naturally in fragmented markets. Cao, Ghysels and Hatheway (2000) find locked and crossed markets in the pre-opening period are used by dealers to signal information [90]. Shkilko, Van Ness and Van Ness (2008) find that the national best bid and offer is locked or crossed roughly 10.5% of the time for NASDAQ stocks and approximately 4% of the time for NYSE stocks and cite a number of potential causes, including simultaneous actions, lack of coordination,

infrastructure issues, as well as differential fees and rebates [91]. Garvey and Murphy (2006) find markets locked 14% of the time and crossed 0.5% of the time for a sample of 20 NASDAQ stocks and conclude that institutional traders should be able to exploit arbitrage opportunities presented by crossed markets [92].

In this paper, we quantify the relationship between order routing and the time markets spend in cross or locked states. We also report the effect of order routing on other statistical properties of the consolidated limit order book.

3.3 Methods

Our methodology is motivated by the zero-intelligence approach popularized by Gode and Sunder (1993) and applied widely in both economics and finance [93]. The key notion is that the mechanics of the system, in our case the limit order book and the order routing algorithms, are of utmost importance, and potentially dominate specifications of strategic behavior. From an agent-based modeling perspective, the model is populated with simple randomly behaving agents that interact within a given market structure.

The approach has a number of origins and motivations. In our case, we seek to avoid making detailed assumptions regarding the strategic behavior of agents (beyond routing) since that behavior varies widely and is difficult to model. More generally, we seek a balance between the idea of “keeping it simple” and “keeping it descriptive” [94, 95]. Specifically, our model builds on a line of work that models order arrivals and cancellations as a random process [96, 97, 98, 99, 100, 101]. We

adopt a model of order flow based on Smith et al (2003).

3.3.1 Single-Market Model

In the single market setting, the limit order book is modeled as a multi-dimensional Markov process (Eq. (3.1)), where negative components represent bid depth and positive components represent ask depth.

$$X_i(t) = \{X_i^1(t), \dots, X_i^n(t)\} \in \mathbb{Z}^n \quad (3.1)$$

The best bid and ask prices at time t in market i , as mentioned previously, are denoted $b_i(t)$ and $a_i(t)$, respectively. If the bid side of the market is cleared out the best bid is set to 0. If the ask side of the market is cleared out, the best ask is set to n . Any given state transition consists of the addition or subtraction of a vector e^j , which is the j th row of an identity matrix of size n , i.e. for state $x \in \mathbb{Z}^n$, all transitions are for the form $x \rightarrow x \pm e^j$. Each transition represents the addition or cancelation of a limit order or the execution of a market order.

For convenience, the process is defined by the arrival rates of limit orders, market orders and cancellations. Limit orders arrive uniformly at random relative to the opposite quotes according to a band ν . Limit orders arrive at a rate of ℓ per price per unit time. The total limit order arrival rate is $2\nu\ell$. Buy limit orders are uniformly distributed across the prices $P_i^b = \{a_i(t) - \nu, a_i(t) - 1\}$. Sell limit orders are uniformly distributed across the prices $P_i^a = \{b_i(t) + 1, b_i(t) + \nu\}$. Market orders arrive at rate m per side per unit time making the total market order arrival rate $2m$. Cancellations occur at a rate c per unit depth per unit time. The arrival rate of cancellations at

price \hat{p} is $c|X_i^{\hat{p}}(t)|$. The total arrival rate of cancellations at time t is $c(\sum_{p=1}^n |X_i^p(t)|)$.

3.3.2 Multi-Market Model

The multi-market model is the process $Y(t)$ defined in Eq. (3.2), which models the simultaneous evolution of two limit order books. The index of price p in market 2 in the joint process is $p+n$. Events originate in each book according to the same rates as the single-market model, but events originating in one book may be routed to another based on the markets' combined state. Market orders are routed with probability α . Limit orders are routed with probability β . Orders that are not routed execute or accrue in their home markets. The arrival rates for the multi-market model are summarized in Table 3.1.

$$Y(t) = (X_1(t) X_2(t)) = \{Y^1(t), \dots, Y^{2n}(t)\} \in \mathbb{Z}^{2n} \quad (3.2)$$

Table 3.1: Aggregate arrival rates for the multi-market model

Order	Side	Route	Rate
Limit	Bid	No	$\nu(1 - \beta)\ell$
Limit	Bid	Yes	$\nu\beta\ell$
Limit	Ask	No	$\nu(1 - \beta)\ell$
Limit	Ask	Yes	$\nu\beta\ell$
Market	Bid	No	$(1 - \alpha)m$
Market	Bid	Yes	αm
Market	Ask	No	$(1 - \alpha)m$
Market	Ask	Yes	αm
Cancellation	n/a	n/a	$c \sum_{i=1}^{2n} Y^i(t) $

The arrival rates for events in the multi-market model.

With arrival rates defined, we must connect order arrivals to updates of the state of the limit order book. Market orders decrease depth at the best bid and ask prices in their target markets. Limit orders increase depth in their target market at their specified price. Recall that bid depth is represented by a negative number, so a new limit order to buy at price \hat{p} in market 2, which is not being routed, results in the state transition $y \rightarrow y - e^{\hat{p}+n}$, where e^j is the j th row of an identity matrix of size $2n$.

The target markets are defined by the routing mechanisms which are as follows. Market orders are simply routed to the markets with the best opposite quote. If the opposite quotes are identical, the order remains in its originating market. The market order routing mechanism is described formally in Table 3.2. Limit order routing is more involved. Limit orders are classified into three groups: non-improving, improving and marketable. The routing of each type is treated differently. The classification and routing procedure for limit orders is provided in Table 3.3.

Table 3.2: Market order routing

Side	Condition	Market
Bid	$a_1(t) < a_2(t)$	1
Bid	$a_1(t) > a_2(t)$	2
Bid	$a_1(t) = a_2(t)$	h
Ask	$b_1(t) > b_2(t)$	1
Ask	$b_1(t) < b_2(t)$	2
Askd	$b_1(t) = b_2(t)$	h

Market orders are simply routed to the market with the most favorable quote.

Non-improving limit orders are routed according to priority, which is defined as the depth between a new bid price $\hat{p} \in P_i^b(t)$ and the best ask in market i as defined

Table 3.3: Limit order classification

Side	Price	Classification	Routing
Bid	$\hat{p} \leq \max\{b_1(t), b_2(t)\}$	Non-improving	Priority
Bid	$\max\{b_1(t), b_2(t)\} < \hat{p} < \min\{a_1(t), a_2(t)\}$	Improving	Ticks to opposite quote
Bid	$\min\{a_1(t), a_2(t)\} \leq \hat{p}$	Executable	Executing market
Ask	$\min\{a_1(t), a_2(t)\} \leq \hat{p}$	Non-improving	Priority
Ask	$\max\{b_1(t), b_2(t)\} < \hat{p} < \min\{a_1(t), a_2(t)\}$	Improving	Ticks to opposite quote
Ask	$\hat{p} \leq \max\{b_1(t), b_2(t)\}$	Executable	Executing market

Classification of limit orders for routing procedures.

by Eq. (3.3). Analogously, the depth between a new ask price $\hat{p} \in P_i^a(t)$ and the best bid in market i is defined by Eq. (3.4). A non-improving limit order at price \hat{p} is routed to the market with the least depth between the new order and the opposite quote. The routing for non-improving limit orders is given in Table 3.4. The routing of a non-improving limit order is illustrated in Fig 3.3.

$$d_i^b(\hat{p}, t) = \sum_{p=\hat{p}}^{a_i(t)} |X_i^p(t)| \quad (3.3)$$

$$d_i^a(\hat{p}, t) = \sum_{p=b_i(t)}^{\hat{p}} |X_i^p(t)| \quad (3.4)$$

Table 3.4: Non-improving priority routing mechanism

Side	Condition	Destination
Bid	$d_1^b(\hat{p}, t) < d_2^b(\hat{p}, t)$	1
Bid	$d_1^b(\hat{p}, t) > d_2^b(\hat{p}, t)$	2
Bid	$d_1^b(\hat{p}, t) = d_2^b(\hat{p}, t)$	h
Ask	$d_1^a(\hat{p}, t) < d_2^a(\hat{p}, t)$	1
Ask	$d_1^a(\hat{p}, t) > d_2^a(\hat{p}, t)$	2
Ask	$d_1^a(\hat{p}, t) = d_2^a(\hat{p}, t)$	h

Routing procedures for non-improving limit orders. Limit orders originating with price \hat{p} are routed according to the conditions in the table where h indicates the originating market.

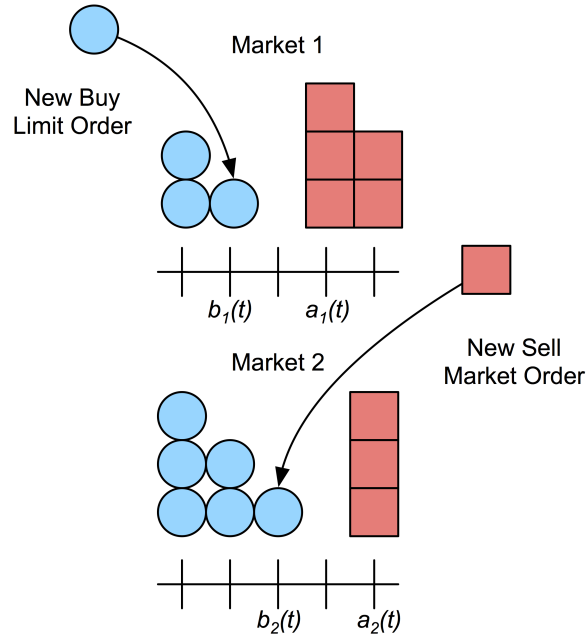


Figure 3.3: **Order routing policy.** A new buy limit order at price $\hat{p} = b_2(t) = b_1(t) + 1$ arrives and is routed to market 1 since it will have better priority at that price, i.e. $d_1^b(\hat{p}, t) < d_2^b(\hat{p}, t)$. In market 1, a new limit order to buy at \hat{p} is first in the queue with no depth ahead toward the best ask. In market 2, the same order is third in the queue. A new market order to sell is routed to market 2 since $b_2(t) > b_1(t)$.

Improving limit orders are routed according to the distance in ticks to the opposite quote. The routing of improving limit orders is provided in Table 3.5. If the distance to the opposite quote is 0 in one of the two markets, then the order is marketable and will execute in that market.

Table 3.5: Limit order routing for improving orders

Side	Condition	Destination
Bid	$a_1(t) - \hat{p} < a_2(t) - \hat{p}$	1
Bid	$a_1(t) - \hat{p} > a_2(t) - \hat{p}$	2
Bid	$a_1(t) - \hat{p} = a_2(t) - \hat{p}$	h
Ask	$\hat{p} - b_1(t) < \hat{p} - b_1(t)$	1
Ask	$\hat{p} - b_1(t) > \hat{p} - b_1(t)$	2
Ask	$\hat{p} - b_1(t) = \hat{p} - b_1(t)$	h

Classification of limit orders for routing procedures.

Simulation Procedure

The procedure for simulating the model is based on the standard stochastic simulation algorithm, also known as the Gillespie algorithm [102]. An outline of the algorithm is as follows.

1. Initialize the state $Y(t)$ at $t = 0$ and select a duration T for the simulation to run.
2. Denote the arrival rates $\lambda_i(t)$ for all possible events $i = 1 \dots m$.
3. Calculate the sum of the transition rates $\lambda(t) = \sum_i^m \lambda_i(t)$.

4. Simulate the time, τ , until the next transition by drawing from an exponential distribution with mean $1 / \lambda(t)$.
5. Choose the transition type by drawing from a discrete distribution where the probability of event i is $\lambda_i(t) / \lambda(t)$.
6. Update the process according to the interarrival time and transition type.
7. Repeat steps 2-6 until $t > T$.

The model is initialized with a depth of 5 lots at the best ν prices on each side of the market. The best bid and best ask in each market is set to $\lfloor n/2 \rfloor$ and $\lfloor n/2 \rfloor + 1$, respectively. For all replications, the model is warmed up to avoid initialization bias. We use Welch’s graphical method to determine the warm-up period [59]. In brief, the method consists of averaging over replications and then examining a moving average of the various output quantities. Fig 3.4 displays the moving average of several output quantities for $\alpha = 0.05$ and $\beta = 0.05$. The output is based on 10 replications of approximately 100,000 epochs. The values are averaged at each epoch and the moving average is calculated with respect to the last 1000 observations. Using this approach, we end the initial warm up period at $t_s = 650$, which is equivalent to approximately 100,000 epochs.

The embedded Markov chain associated with $Y(t)$ is simulated by normalizing the arrival rates of individual events by $\lambda(t)$ given in equation Eq. (3.5), which is the sum of the rates given in Table 3.1. At each discrete epoch, the clock is advanced according by τ , which is drawn from an exponential distribution with mean $1 / \lambda(t)$,

and the state is updated according to the model. The model terminates when $t > T$. Statistics are calculated for $t_s < t < T$ where $T = 1650$.

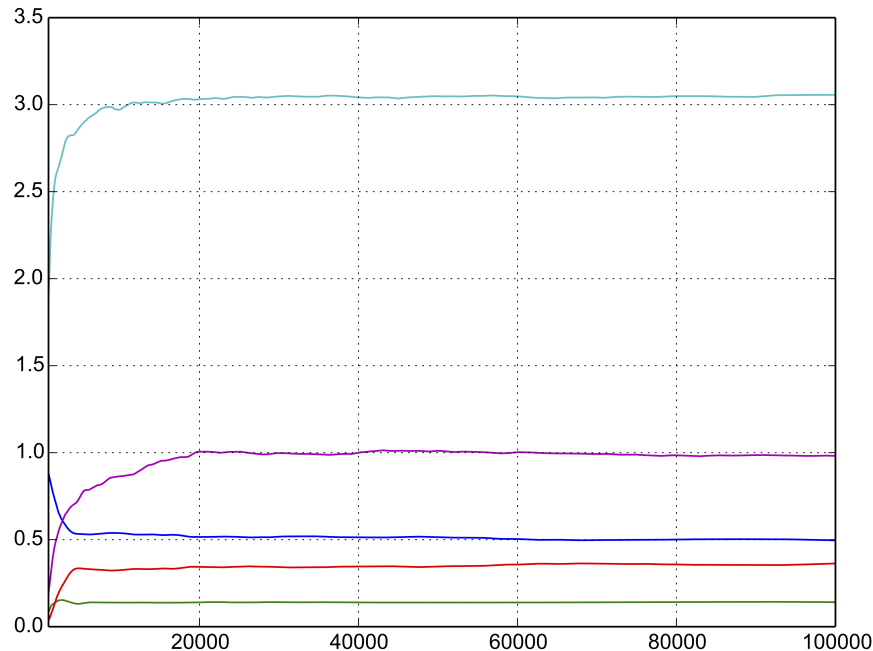


Figure 3.4: **Welch's graphical method for identifying the warm-up period for the simulation.**

$$\lambda(t) = 4v\ell + 4m + c \sum_{p=1}^{2n} |Y^p(t)| \quad (3.5)$$

3.4 Results and Discussion

The main results are summarized by response surfaces for crossed and locked states as α and β are varied between 0 and 1. We also report on mean spread and volatility of the market as defined by the consolidated books. The results are based on 30

replications of length $T - t_s = 1000$ for each parameterization of α and β . We also report on the sensitivity of the main results to 10% changes in the arrival of market and limit order parameters ℓ and m . We expect some sensitivity as these two parameters drive the spread and volatility of the single market model.

Gatheral and Oomen (2010) use the single market model to evaluate alternative procedures for computing realized variance. They use parameters $\ell = 1$, $m = 5$, $c = 0.2$, $\nu = 10$, which they find results in a microstructure noise ratio similar to that of stocks in the DJ30. Our main results are estimated with parameters $\ell = 1$, $m = 5$, $c = 0.2$, $\nu = 20$, which results in an average depth in each market of 150 lots. Increasing the band ν from 10 to 20 results in a higher mean depth across the markets, which also results in additional stability. With $\nu = 20$, neither side of the book emptied out across all of our replications. Earlier simulations with $\nu = 10$ were not stable for some values of α and β , which resulted in extreme price moves when the book emptied out on either the buy or sell side. The model, in general, is stable over a wide range of the parameters. Our goal is not to replicate the stylized facts of any particular market, but to study the effects of order routing under reasonable assumptions of order flow, spread and volatility.

The size of the price grid is $n = 10,000$, which is large enough that results are not affected by the bounds of the grid. We present the main results as response surfaces and then report on a sensitivity analysis of order flow parameters. Specific parameters, replications and other simulation details are given alongside the results.

Order Routing Results

The probability of market routing and limit routing, as defined earlier, is varied between 0 and 1 by increments of 0.05. At any given time, the state of the market is classified as either crossed, locked or normal. The average portion of time that the market spends in a crossed, locked or normal state, as a function of market and limit order routing, is reported in Fig 3.5, Fig 3.6 and Fig 3.7, respectively. Each point in the lattice of the wireframe represents the mean of the outputs across 30 replications. Crossed and locked market durations fall rapidly as routing is increased. Markets are locked 7.2% of the time and crossed 7.1% of the time when $\alpha = 1$ and $\beta = 0$. When $\beta = 1$, all limit orders are routed eliminating the possibility of crossed or locked markets.

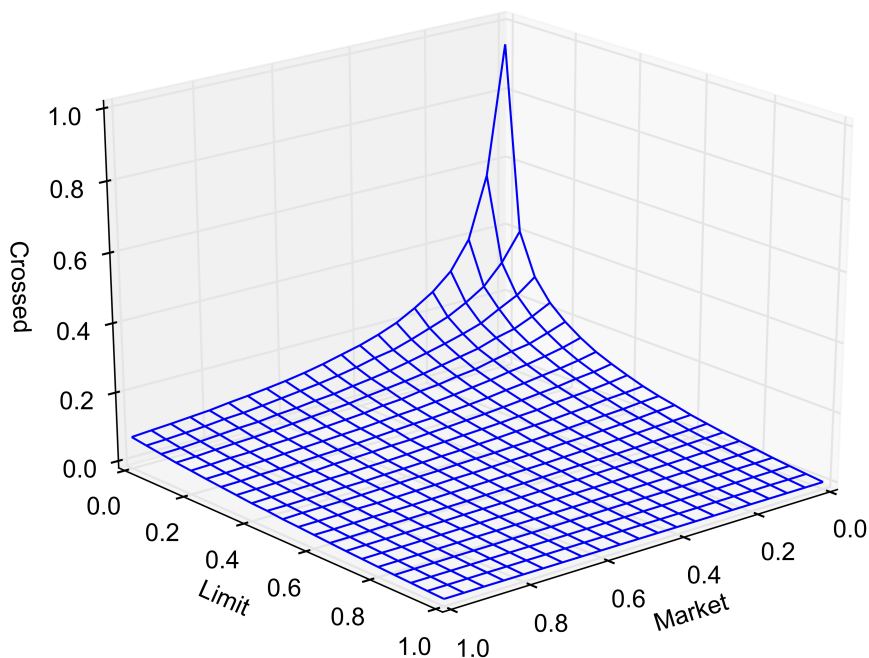


Figure 3.5: **Percentage of time markets are crossed.**

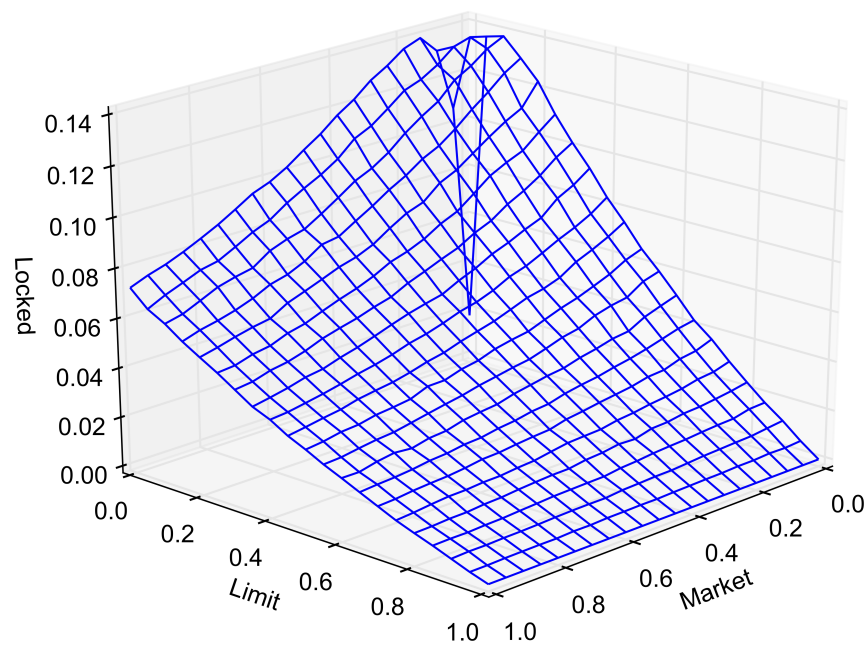


Figure 3.6: **Percentage of time markets are locked.**

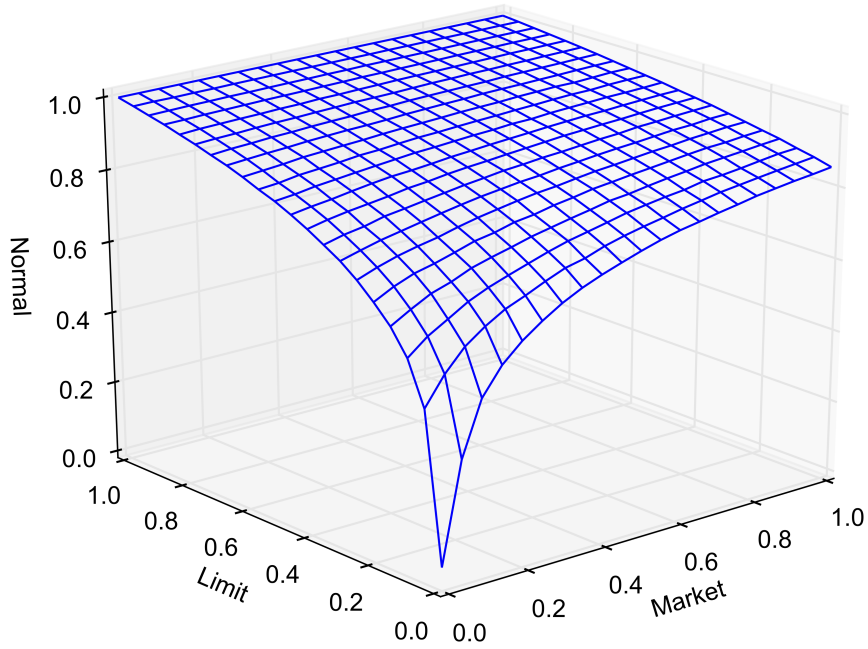


Figure 3.7: **Percentage of time markets are normal (neither crossed nor locked).**

Fig 3.8 reports the mean spread of the consolidated order book, i.e. $\min\{a_1(t), a_2(t)\} - \max\{b_1(t), b_2(t)\}$. Note that the spread in the consolidated book can be negative when $\beta \neq 1$. Order routing increases the mean consolidated spread by eliminating episodes where it is non-positive. Consider, however, the case when $\beta = 1$, where the consolidated spread is strictly positive. The spread is increasing with α since routed market orders are more likely to affect the best quotes.

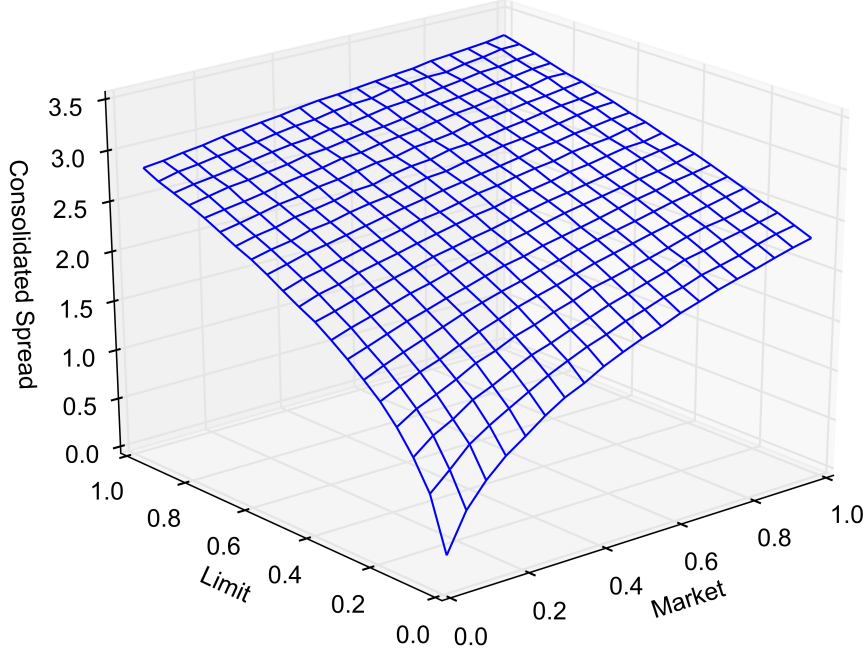


Figure 3.8: **Mean spread in the consolidated book.**

Volatility, like the other statistics we report, is calculated online. We define volatility to be the variance of the differences in transaction prices. The variance is computed using the well-known recursion given in Eq. 3.6 and Eq. 3.7 where $\mu_1 = x_1$ and $\phi = 0$ [103]. The k th transaction price is denoted x_k . When the simulation terminates the variance is given as $\phi / (K - 1)$ where K is the total number of transactions. Market order routing decreases volatility as expected. Routed market orders execute at the best bid and ask in the consolidated book, which is by definition a smaller price range. When all market orders are routed to the best quotes ($\alpha = 1$), there appear to be two competing effects when limit order routing is increased. With relatively few routed orders, an increase draws the market nearer, whereas later,

more routing increases the spread, resulting in higher price volatility. Fig 3.9 reports volatility as a function of α and β . Spread and volatility results for several discrete cases are reported in Table 3.6.

$$\mu_k = \mu_{k-1} + (x_k - \mu_{k-1}) / k \quad (3.6)$$

$$\phi_k = \phi_{k-1} + (x_k - \mu_{k-1})(x_k - \mu_k) \quad (3.7)$$

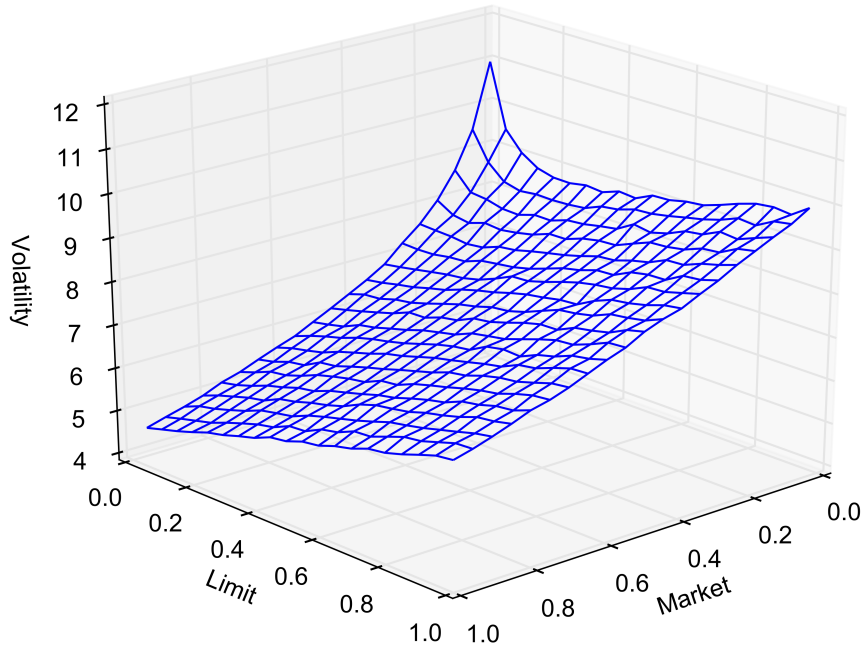


Figure 3.9: Volatility as measured by the variance of the difference in transaction prices.

Table 3.6: Consolidated spread and volatility results

Routing		Spread		Volatility	
α	β	Mean	Std. Dev.	Mean	Std. Dev.
1	1	3.25	0.04	6.8	0.20
0	1	2.76	0.03	10.3	0.23
1	0	2.12	0.02	4.6	0.13
0.5	0.5	2.46	0.03	7.3	0.16

Replication results based on 30 replications with $\ell = 1, m = 5, c = 0.2, v = 20, n = 10,000$. The spread is reported for the consolidated order book.

3.4.1 Sensitivity Analysis

The crossed and locked state durations, as well as spread and volatility, show some sensitivity to the main parameters of the market. This is to be expected as limit and market order arrival rates drive the characteristics of the single market model. For each of the model outputs reported in the main results, we provide a response surface as the market and limit arrival rates are varied by $\pm 10\%$. Fig 3.10 and Fig 3.11 report on crossed and locked state durations for $\alpha = 1$ and $\beta = 0$. Fig 3.12 and Fig 3.13 report the sensitivity of spread and volatility to changes in market and order arrival rates for $\alpha = 1$ and $\beta = 1$.

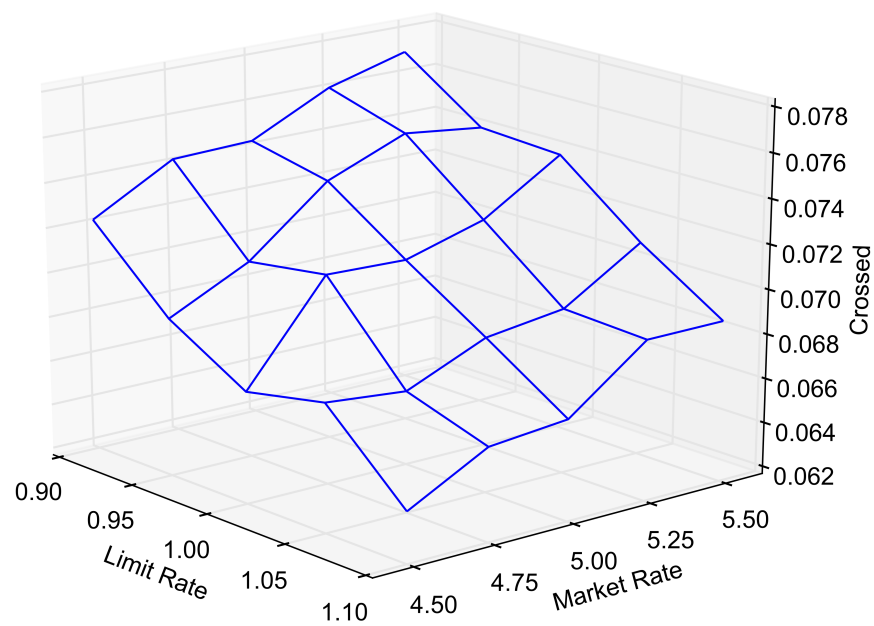


Figure 3.10: **Sensitivity of crossed state duration.** Proportion of time market is in crossed state as a function of limit and market order arrival rates for $\alpha = 1$ and $\beta = 0$.

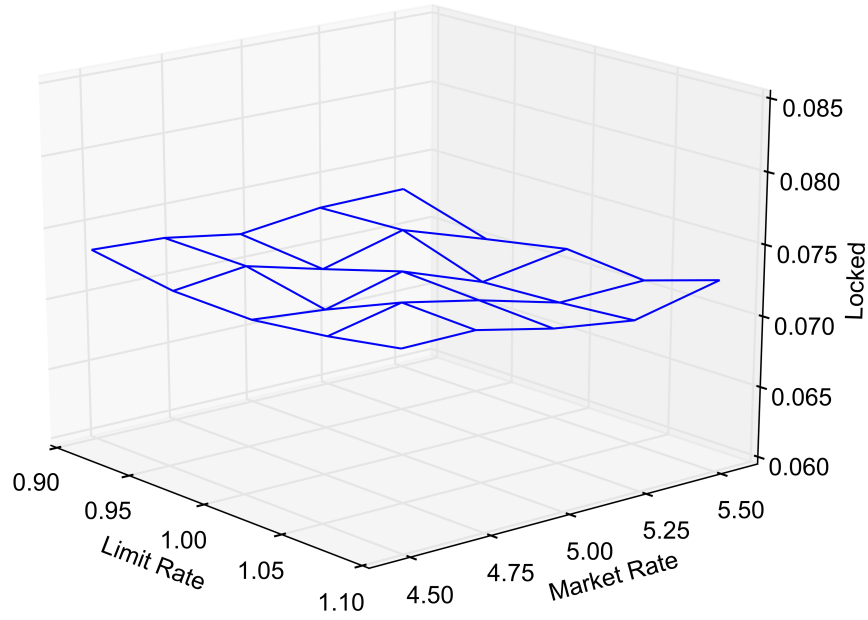


Figure 3.11: **Sensitivity of locked state duration.** Proportion of time market is in locked state as a function limit and market order arrival rates for $\alpha = 1$ and $\beta = 0$.

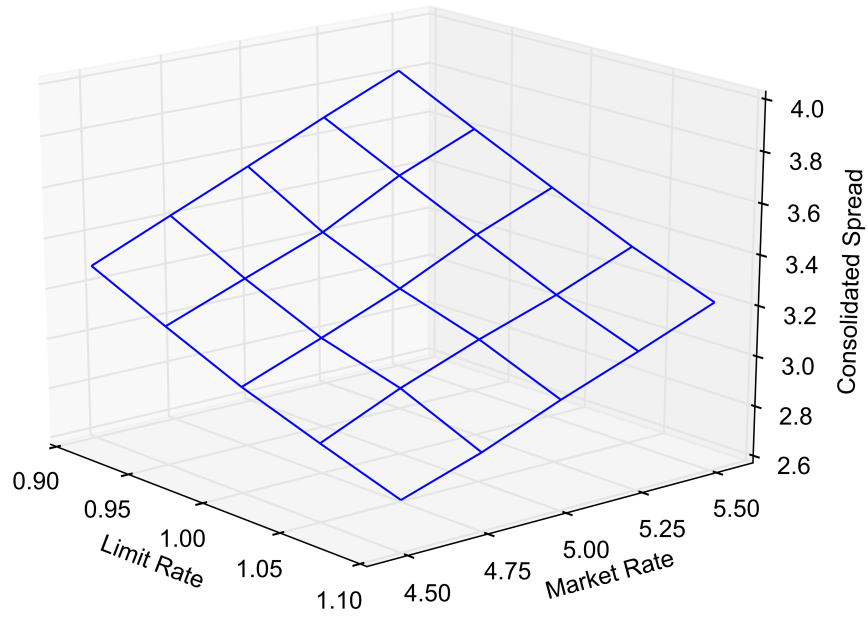


Figure 3.12: **Sensitivity analysis of the consolidated spread.** Mean consolidated spread as market and limit order arrival rates are varied for $\alpha = 1$ and $\beta = 1$.

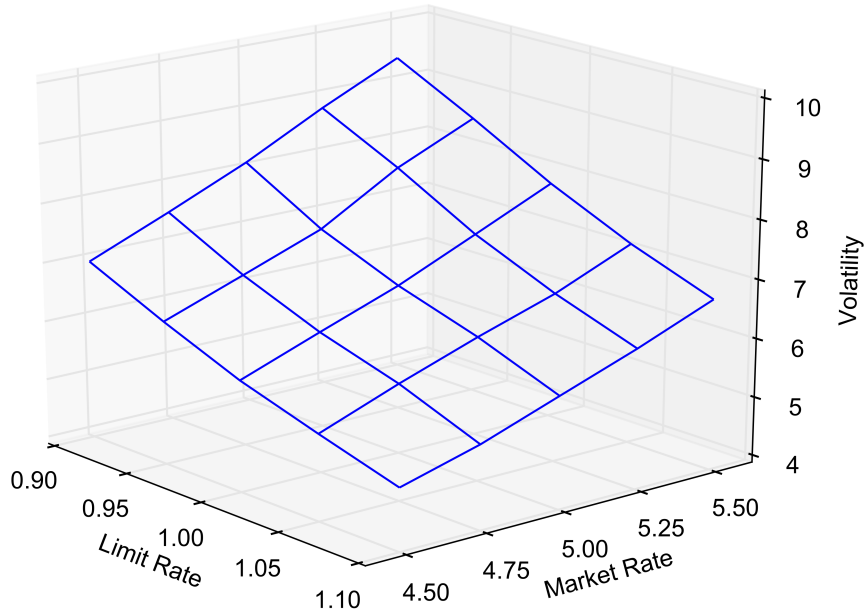


Figure 3.13: **Sensitivity analysis of volatility.** Volatility of transaction prices as a function limit and market order arrival rates for $\alpha = 1$ and $\beta = 1$.

3.5 Conclusions

The zero-intelligence approach is a powerful modeling framework for examination of fragmented markets. By modeling order routing as part of the market mechanism, we maintain a parsimonious model of order-driven markets that has the potential to explain stylized facts beyond crossed and locked markets. The model helps develop intuition for the effects of simple order routing procedures on the statistical properties of a multi-market system. The results also provide a benchmark against which more complex agent-based models may be compared.

Empirical validation is also a possibility since all model parameters, including α

and β , could potentially be estimated from data. The model has an interpretation in terms of market demographics as well. For example, parameterizations of our model can be interpreted as a markets consisting of a certain percentage of participants with superior market access. We believe models of this type are uniquely capable of examining complex market structures while simultaneously accounting for market demographics. The model has the flexibility to be extended in a number of directions, including examination of markets with asymmetric trading volume and more complex market structures.

Chapter 4

Correlations, Arbitrage and Trade-throughs

4.1 Abstract

We study correlations, arbitrage opportunities, and trade-throughs in an action-packed period for Nordic market structure during the 2009 through 2013 period. The introduction of a new trading platform at Nasdaq, centralized clearing and an tick size harmonization initiative are just a few of the many events expected to improve Nordic equity markets. The period also includes increased competition with new exchanges entering the market. We study cross-listed Nordic equities traded on the NASDAQ OMX platform and characterize changes in correlations, arbitrage opportunities and trade-throughs.

4.2 Introduction

In a fragmented financial market, prices and quotes no longer obey the normal mechanical restrictions of a limit order book. For example, the bid in one market may exceed the ask, introducing the possibility of an arbitrage trade. In a single market, such a bid would induce a transaction. In the same vein, a transaction to buy or sell may occur at a quote that is less than optimal with respect to quotes available in other markets. The law of one price suggests that arbitrage opportunities and suboptimal transactions should be rare or even non-existent. But, there are “limits to arbitrage,” and despite the best intentions of all market participants to maximize profits and minimize costs, inefficiencies persist [79, 104, 105]. In this paper, we undertake a study of market efficiency by examining correlations, arbitrage opportunities, and trade-throughs in a selection of cross-listed Nordic equities. We document the change in these quantities as markets evolve from 2009 to 2013.

4.2.1 Market Structure and Multi-market Trading

Markets are consolidated when all trade for a particular product occurs in the same place, or, in the case of financial markets, on the same exchange. There is, in fact, a natural tendency for markets to consolidate. The basic notion is that of a network effect, in which the benefit from the use of a system grows with the number of users (think Facebook). The more market participants, the more liquidity, and more liquidity attracts more traders [65]. Historically, some exchanges have enjoyed monopoly power, but those monopolies, such as the New York Stock Exchange, have

been disrupted through a combination of regulation, technology and competition [106]. While some exchanges still enjoy monopoly power on certain products, the majority of financial markets are fragmented. Numerous exchanges and platforms compete for order flow.

A market is fragmented when the same product or very similar products are traded on more than one exchange. Harris (2002) outlines a number of causes of market fragmentation including unequal size of trades, asymmetric information, unequal patience, unequal access and varying levels of creditworthiness [65]. The heterogeneity of market participants makes the optimality of any single market mechanism unlikely. While some market participants may prefer continuous trading in a limit order book, others may prefer an upstairs-style market in which they simply negotiate a price. New exchanges enter the market and cater to the variety of demands arising from the diverse population of market participants. In our case, historical mergers have resulted in populations of investors with demands to trade in a number of different currencies.

Broadly speaking, fragmented markets are linked by various institutions, regulations and strategic behavior. The National Market System in the United States is an example of a regulation that provides explicit rules on how equity markets are integrated. Exchanges are required to execute trades at the best available quotes [107]. (See Reg NMS Rule 611(a)(1).) In Europe, MiFiD also provides a best execution directive that requires firms to take “all reasonable steps to obtain the best possible result for their clients, taking into account price, costs, speed, likelihood of execution and settlement, size, nature or any other consideration relevant to order

execution,” but offers firms a “considerable degree of flexibility” on how to meet the requirement [108].

Naturally, this leads us to consider the role of the individual behavior in the integration of markets. Obviously, if participants have the required access to markets they will attempt to exploit arbitrage opportunities. However, trading that aims specifically to exploit mis-pricing is not the only type of behavior that links markets. Many brokerages make order routing technology available to their clients in order to improve execution quality [109]. Any type of strategic behavior that links markets relies on access to markets. There are, however, various costs and difficulties associated with accessing markets, which can lead to market segmentation

A market is considered segmented if a subset of market participants do not have access to all available venues. Consider a hypothetical two-market system in which a stock trades at two different locations. One group of investors trades at both locations, the other group only trades at the exchange nearest their office. The market is fragmented since the same shares trade on two exchanges, and it is partially segmented since some market participants only access one of the two exchanges. In fact, markets tend to be partially segmented where a minority of sophisticated participants invest in access to more markets than others.

National borders are a source of market segmentation. For example, traders may be captive to their local exchange if accessing a market in another country or currency proves to be too costly or too much of a hassle. Trading across national borders also requires traders to understand new regulations and market institutions. In our case, we study Nordic equities where the primary difficulty for investors is

currency. Equities in the Stockholm market trade in Swedish Kronor, whereas the same equities listed in the Helsinki market trade in euros.

There is a significant microstructure literature dedicated to the study of multi-market trading. A well known approach is that of Hasbrouck (1995), which studies the contribution of various venues to price discovery in U.S. equities [110]. Similar approaches have been applied to international cross-listings [111]. There is also a large literature that examines arbitrage opportunities in internationally cross-listed stocks. De Jong, Rosenthal and Van Dijk (2009) study the returns of an arbitrage strategy applied to dual listed companies [112]. Another recent examples is the study by Gagnon and Karolyi (2010) of a large sample of stocks that are cross-listed in the United States in the form of depository receipts [113].

The work is also related to the burgeoning literature on the speed of trading. Menkveld and Zoican (2016) find that speeding up exchanges does not necessarily improve liquidity [114]. Pagnotta and Philippon (2015) analyze a number of issues around trading speed and fragmentation, and find that it is not optimal to limit speed, and that it may be optimal to push slow venues to upgrade [115]. Menkveld (2016) surveys a number of relevant issues related to the speed of trading [116]. In our study, we draw on a number of different literatures related to the study of correlations, arbitrage and trade-throughs, and refer to relevant literature in each section.

4.2.2 Data and Sample Period

Our data consists of order book and trade data for a selection of cross-listed stocks trading on the Nasdaq OMX Nordic. Nasdaq Nordic refers to Nasdaq operated exchanges for Helsinki, Copenhagen, Stockholm, Iceland, Riga, Tallinn and Vilnius. We identify six cross-listed securities, each of which are listed on the Stockholm and Helsinki exchanges. Nordea Group is the lone security in our sample that is listed on three exchanges, the third being Copenhagen. Our trade and quote data, with microsecond granularity, was acquired from Reuters. Table 4.1 provides the currency and Reuters market abbreviation for each of the three markets.

Exchange	RIC suffix	Currency
Nasdaq OMX Nordic Stockholm	ST	SEK
Nasdaq OMX Nordic Helsinki	HE	EUR
Nasdaq OMX Nordic Copenhagen	CO	DKK

Table 4.1: Reuters Identification Codes

In selecting our sample, we eliminate two stocks, Ericsson and Tieto, due to either low absolute volume or low relative volume. We consider Nordea in Helsinki and Stockholm only in order to avoid the additional complexity of a third currency. (See Table 4.2 and Figure 6.2 and Figure 6.7 in the appendix.) The remaining stocks are Nokia, Nordea Group, Stora Enso and Telia Sonera.

Nokia is the well known telecommunications company headquartered in the Helsinki metropolitan area. Nordea Group is a Swedish bank that has grown through a number of mergers with Danish, Finnish, Norwegian and Swedish banks, which accounts for its cross listing on three exchanges. Stora Enso is a paper company headquartered in Helsinki. The company is a result of a merger the Swedish company Stora and

	Ericsson		Nokia		Nordea			Stora Enso		Telia Sonera		Tieto	
Year	HE	ST	HE	ST	HE	ST	CO	HE	ST	HE	ST	HE	ST
2009	0.00	1.00	0.95	0.05	0.11	0.77	0.12	0.82	0.18	0.11	0.89	0.89	0.11
2010	0.00	1.00	0.92	0.08	0.11	0.82	0.07	0.78	0.22	0.10	0.90	0.89	0.11
2011	0.00	1.00	0.94	0.06	0.08	0.88	0.04	0.81	0.19	0.10	0.90	0.89	0.11
2012	0.00	1.00	0.82	0.18	0.10	0.83	0.07	0.77	0.23	0.10	0.90	0.89	0.11
2013	0.00	1.00	0.76	0.24	0.09	0.83	0.08	0.73	0.27	0.08	0.92	0.95	0.05

Table 4.2: Annual volume. Volume share in each currency within NASDAQ OMX where HE is NASDAQ OMX Helsinki, ST is NASDAQ OMX Stockholm and CO is NASDAQ OMX Copenhagen. Ericsson volume in HE is approximately 10^{-5} percent.

the Finnish company Enso Oyj, explaining the cross-listing. Finally, Telia Sonera, a telecommunications company resulting from the merger of Swedish company Telia and Finnish company Sonera.

Figure 4.1 shows the relative volume on each market for the selected stocks over the sample period. In the case of each stock, there is a dominant market with the majority of volume, and a satellite market with a much smaller, but not inconsequential share of volume. Nasdaq OMX Helsinki is the dominant market with the majority of volume for Nokia and Store Enso. Nasdaq OMX Stockholm is the dominant market for Nordea Group and Telia Sonera.

Interestingly, there is near parity in the number of messages in each venue as shown in figure 4.2. Despite an imbalance in volumes, market makers still need to keep quotes current across markets. Spreads and depths, as seen in figures 4.3 and 4.4, also appear to be relatively competitive in each of the four stocks, with spreads appearing more competitive later in the period.

Prices and quotes in Swedish Kronor (SEK) are converted to euros (EUR) using exchange rate data acquired from Olsen Associates. The foreign exchange data

is provided in millisecond granularity. We use the contemporaneous mid-quote to convert the bid and ask in stocks on NASDAQ OMX Stockholm to euro. Fig 6.1 in the appendix shows the EUR to SEK exchange rate for the sample period.

Our sample period consists of years 2009 through 2013 and contains a number of important changes to market structure. One of the most important changes is the introduction of the INET platform by NASDAQ [117, 118]. At the time Bjørn Sibbern of NASDAQ noted that it may be “the biggest change in 20 years” for Nordic markets [119]. An updated platform is attractive for a number of reasons, and should, in general, increase volume.

Central counter-party (CCP) clearing was also introduced in 2009 [120]. According to the press release from NASDAQ, the introduction of CCP enables netting of trades, decreases margin requirements, requires fewer settlements, and overall, reduces costs. At the time, the president of NASDAQ OMX hailed it as “the biggest structural change for the Nordic equity markets since trading became electronic in the early 1990s,” and predicted a similar increase in volume.

Another important change in the market structure during the sample period is the effort to harmonize tick sizes. NASDAQ OMX implemented a new tick size regime in Stockholm on October 26, 2009 and in Helsinki and Copenhagen on January 4, 2010. NADAQ predicted that it would “lower the barrier for cross border trading” [121]. Kjell (2015) found that the implementation of new tick sizes improved market quality and encouraged trade [122].

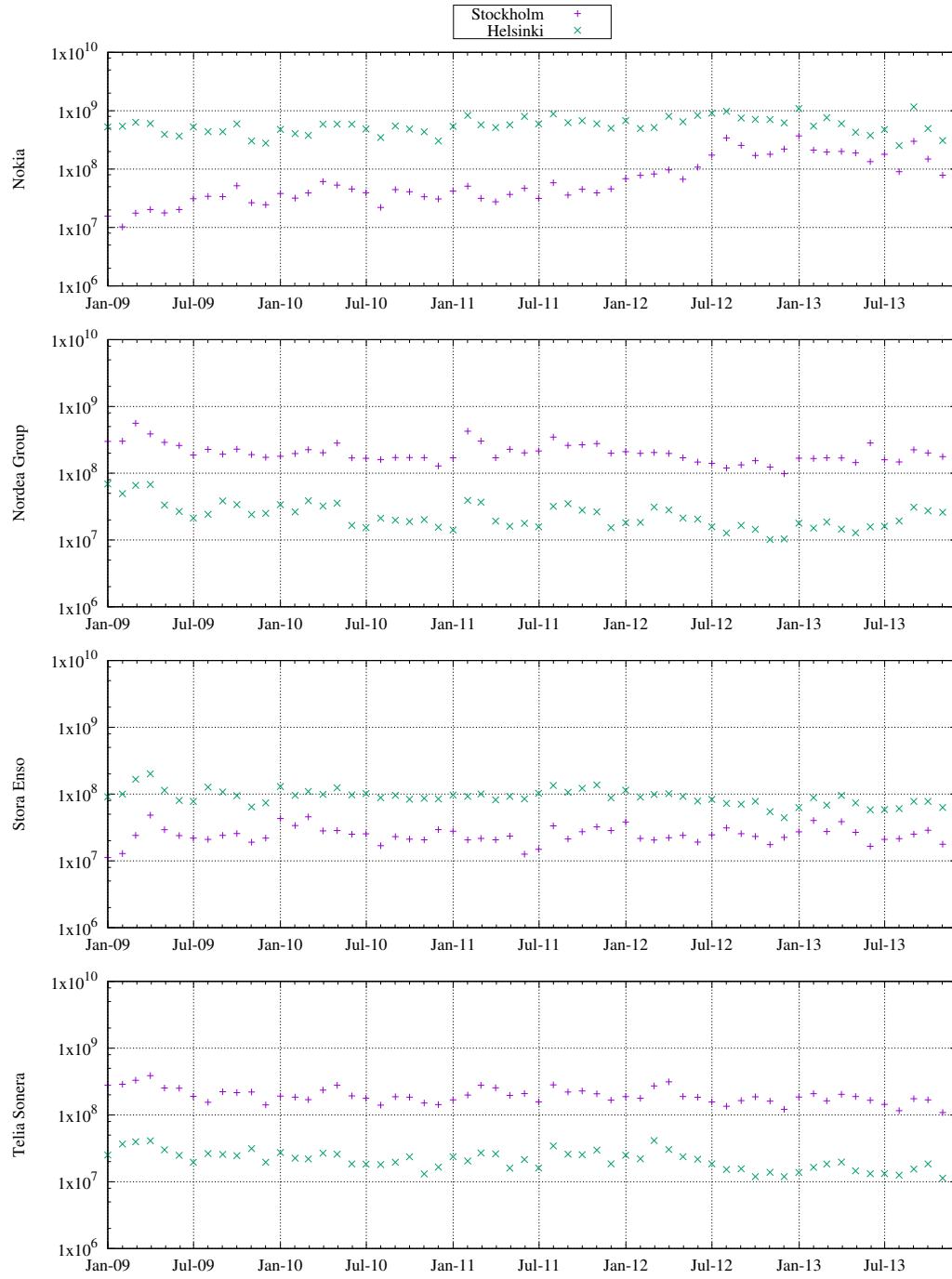


Figure 4.1: **Monthly volume.** Monthly volume for each stock in each venue.

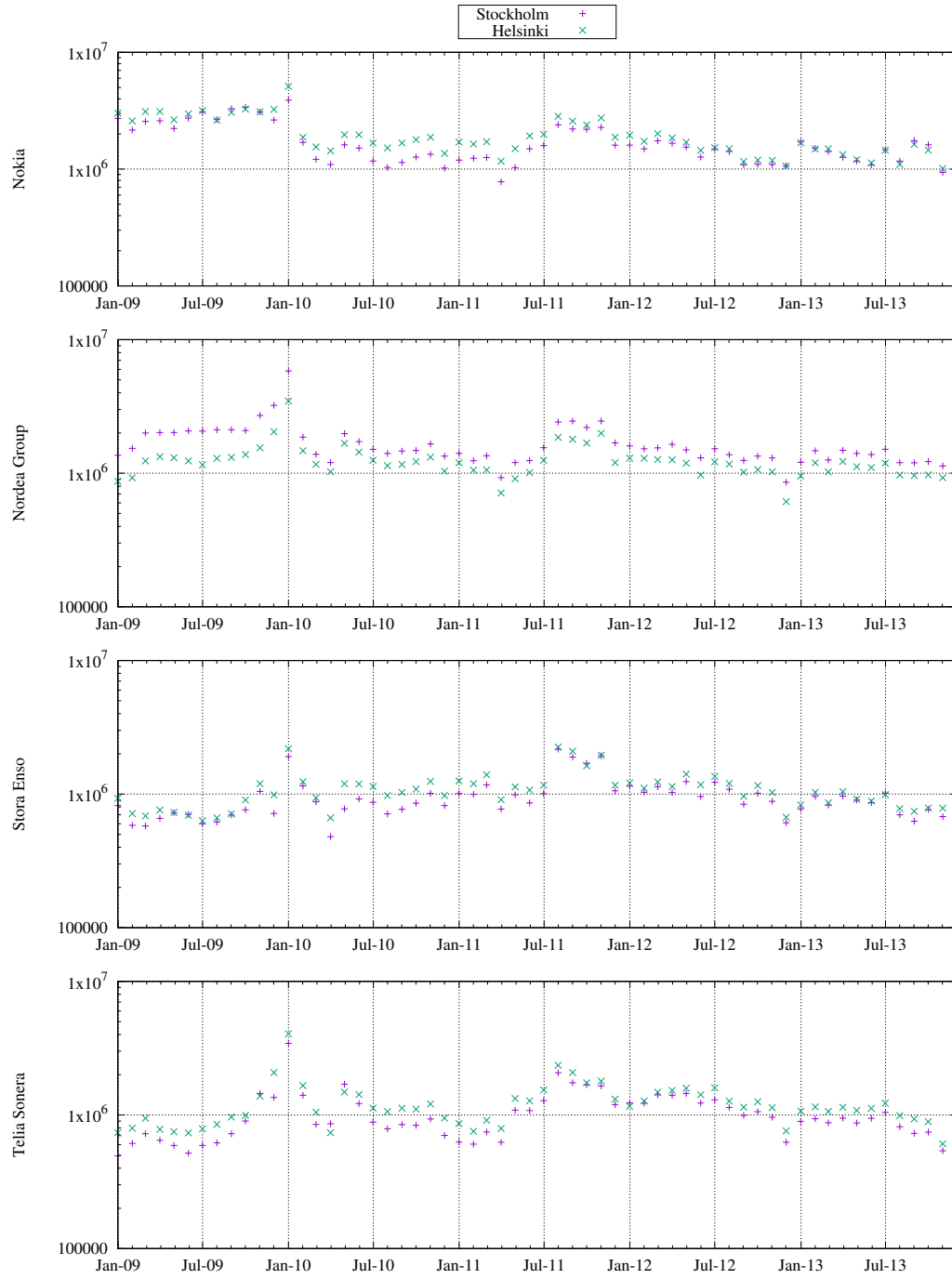


Figure 4.2: **Market updates.** Monthly market updates for each stock in each venue.

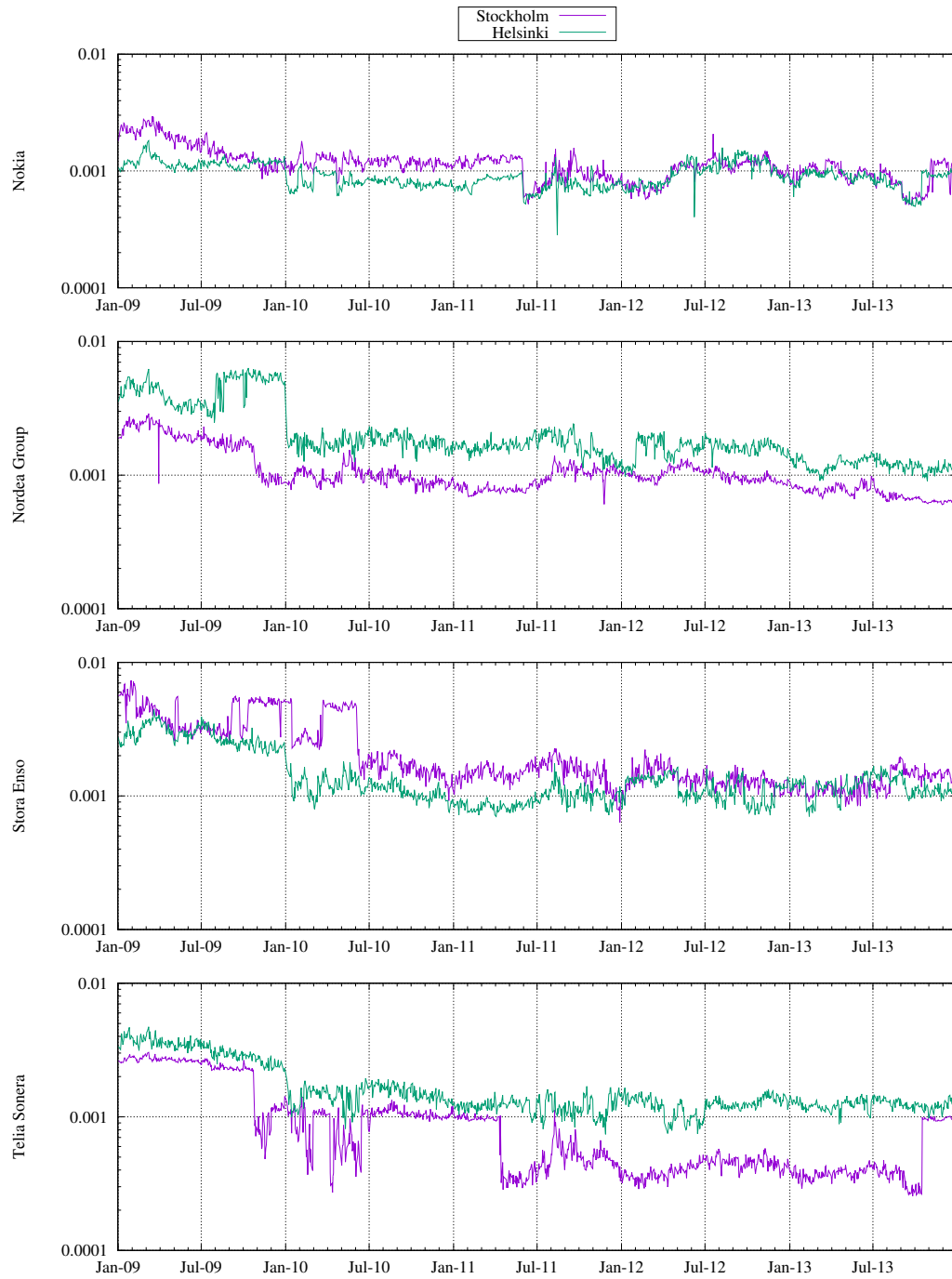


Figure 4.3: **Daily average spread.** Spreads in this figure are averaged over time and given as a percentage of price.

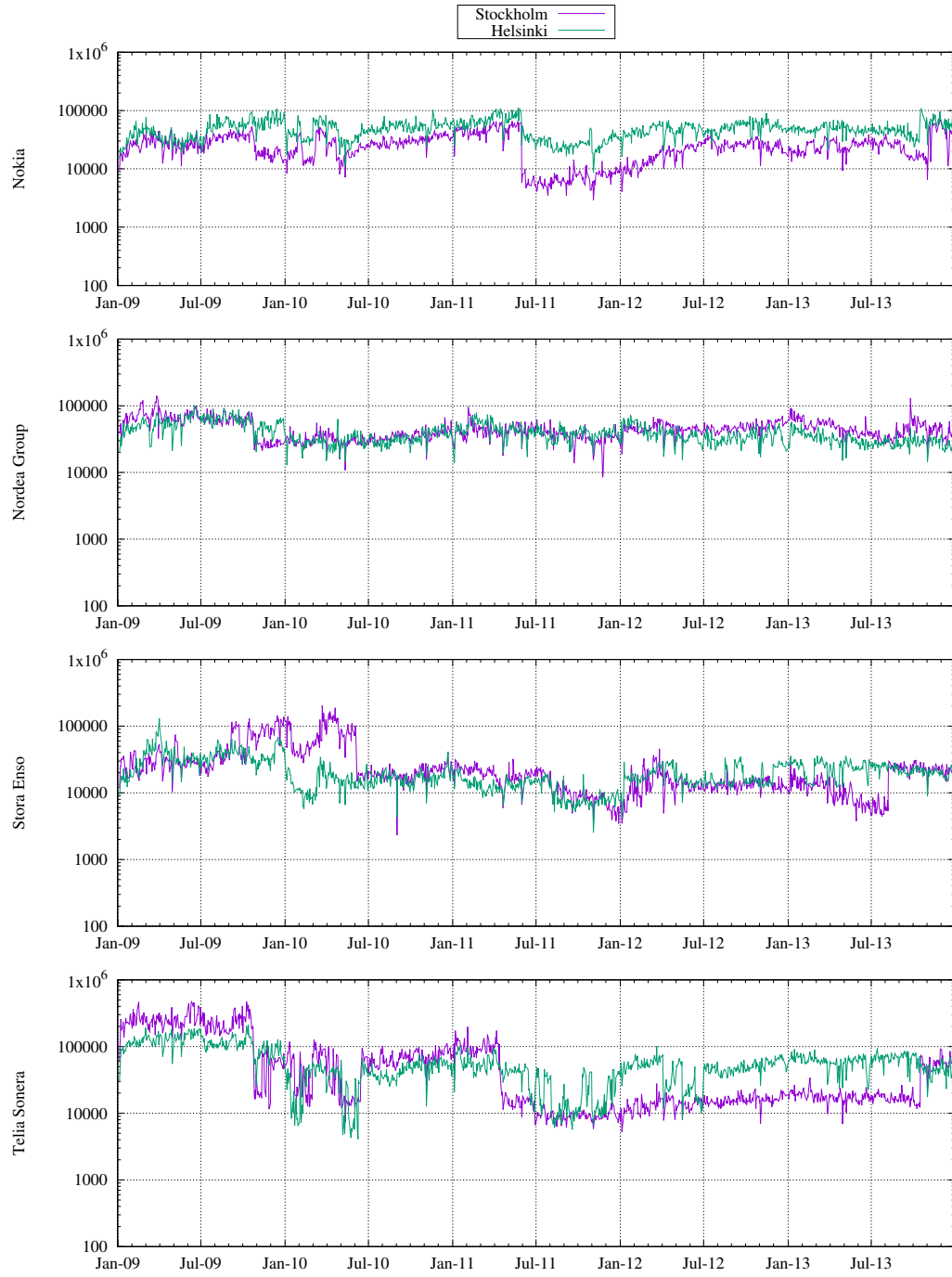


Figure 4.4: **Daily average depth.** Depths are calculated from best prices and averaged over time.

4.2.3 Organization of the chapter

We motivate this work by posing a few basic questions about the speed and efficiency of markets. First, are faster markets also more efficient markets in an informational sense? Do arbitrage opportunities become more rare? Do the properties of arbitrage opportunities, such as size and duration change in any significant way? Are trade-throughs less frequent or less costly?

To answer these questions we organize our inquiry around three basic measurements. The first is mid-quote return correlations at different sampling frequencies. Second, we characterize arbitrage opportunities based on a 0.005 euro threshold and a conversion of quotes to euros based on the EUR-SEK spot rate. Finally, we characterize the frequency and cost of trade-throughs.

4.3 Correlations

Epps first documented that stock correlations decrease in sampling frequency [123]. In a recent paper, Budish, Crampton and Shim (2015) also point to the “break down” of correlations at high-frequency, and note that “nothing in the current market architecture that would allow different securities’ prices to move at exactly the same time” [124]. The asynchronicity of market activity, which has long been a concern of economists, is the primary working explanation for the so-called Epps effect [125, 126, 127]. The phenomenon is visible in Figure 4.5, which shows the mid-quote for Nordea in euros in both Helsinki and Stockholm. Correlations appear to break down as we zoom in on the price chart.

Tóth and Kertész study the time evolution of the Epps effect and suggest that “the phenomenon is related to the reaction time of market participants” [128, 129]. The correlation structure, in other words, can give us an idea about the speed with which market participants react to new information and new quotes. During our sample period, the reaction of market participants is most certainly improved by the introduction of INET. For this reason, we expect an increase in correlations across the board.

Correlations are calculated on a daily basis at intervals ranging from 1 millisecond to approximately 2 minutes. To compute correlations, first, we calculate the mid-quote for each stock in a pair. Quotes outside of Helsinki are converted to euro using the contemporaneous spot FX mid-quote. We then resample the mid-quotes every 2^p milliseconds for $p = 0 \dots 17$. The lowest frequency sample is 2^{17} milliseconds, which is around 131 seconds, or roughly 2 minutes. Finally, we calculate the correlations of the log returns in the standard way.

Figure 4.6 shows the range of the mean correlations for mid-quote returns on an annual basis across the four cross listed securities. There is a clear upward shift with correlations increasing clearly during the 2009 to 2013 period. Figures 6.20, 6.21, 6.22, 6.23 in the appendix show the individual annual correlation curves for each of the 5 years. In our case, since we are studying cross-listed securities, we can reasonably expect correlations to be near 1 at a reasonable sampling frequency.

Most interesting is the increase in correlations at very granular times in the millisecond range. Figure 4.7 shows the increase in correlations for a 16 millisecond resample. This is mostly likely an effect related to improvements in the technology

infrastructure of the exchange, such as the introduction of INET. There appear to be two regimes, a low correlation regime and a high correlation regime starting at the beginning of 2010, which corresponds to the introduction of INET. Correlations in Nokia increase almost five fold, whereas increases in correlations in the other three stocks are more subdued. Our next question is, naturally, what happens to arbitrage opportunities?

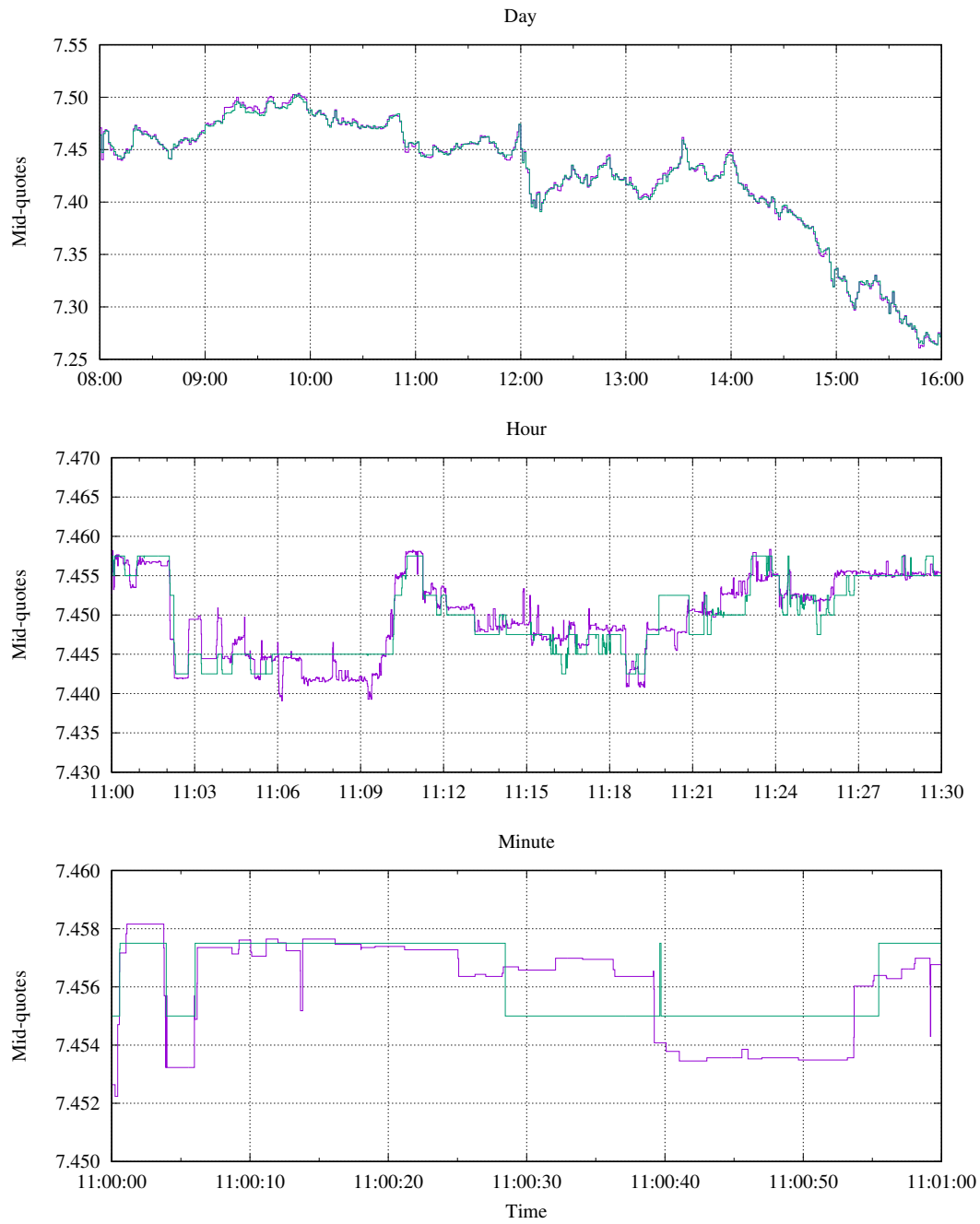


Figure 4.5: **Nordea mid-quotes.** Mid-quotes are shown in a common currency, the euro. Quotes in SEK are converted to euros using the mid-quote of the spot exchange rate.

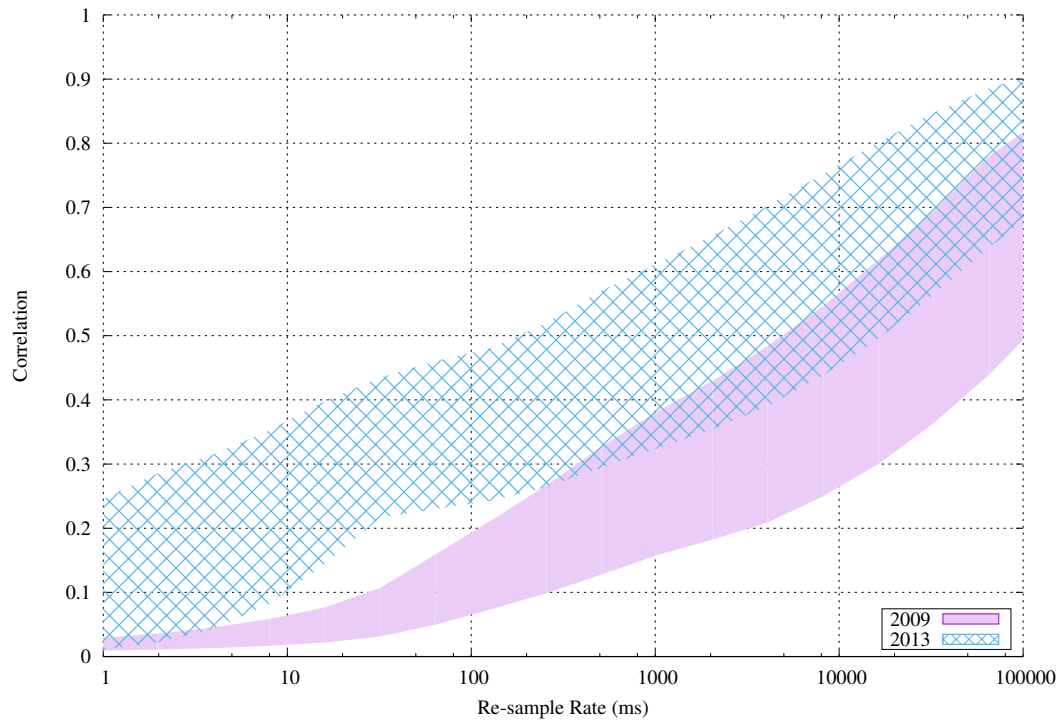


Figure 4.6: **Correlation range for miq-quote returns in native currency.** This figure is based on the average daily correlation for Nokia, Nordea Group, Stora Enso and Telia. For each stock, we calculate the daily correlation of the mid-quote returns. We then take the average for each year, 2009 through 2013. The filled curves show the range of those averages across the four stocks for 2009 and 2013. Data for all 5 years is available in the appendix.

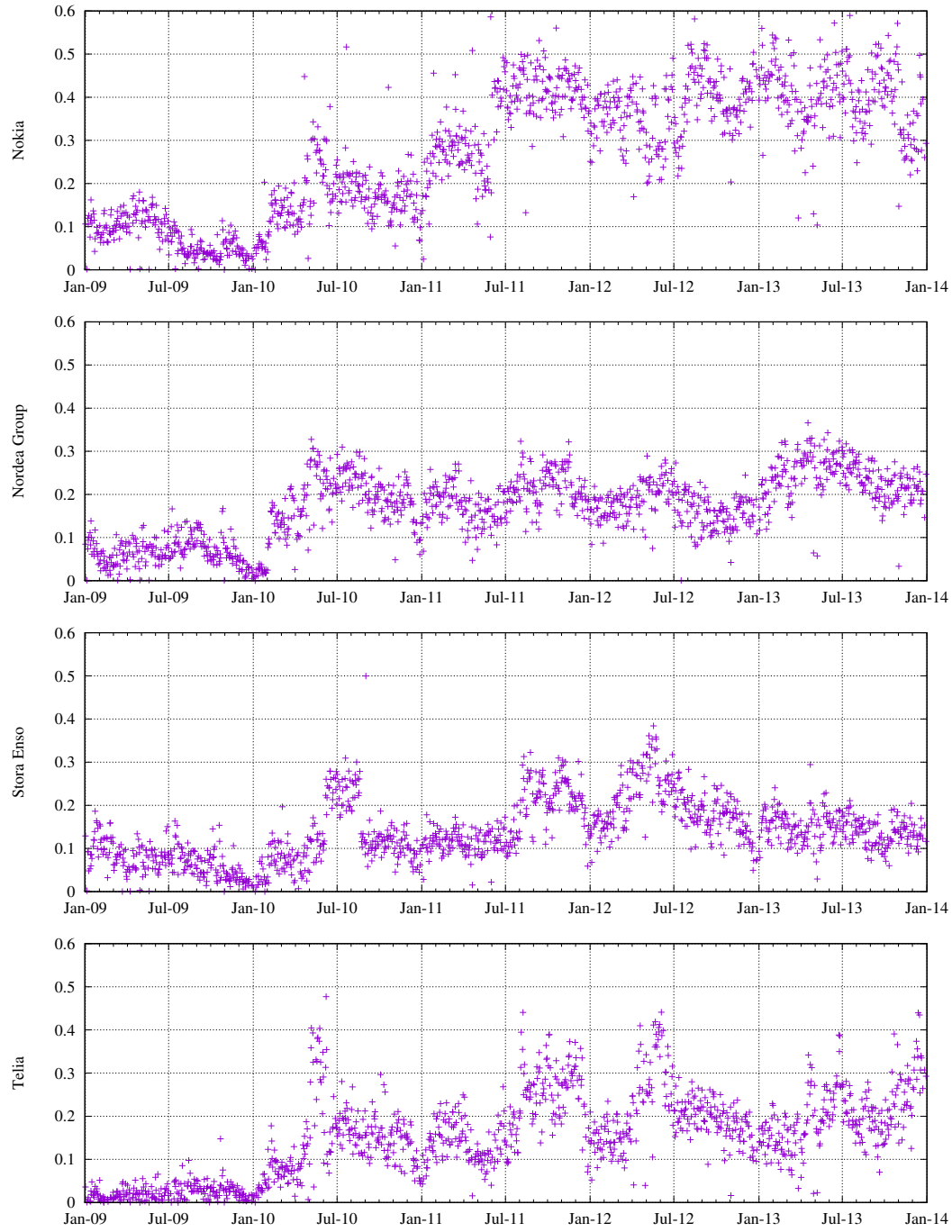


Figure 4.7: **Time-evolution of millisecond correlations.** Daily correlations based on a 16ms resample.

4.4 Arbitrage opportunities

Arbitrage is a fundamental concept in finance, and plays a crucial role in the efficient markets hypothesis (EMH). As summarized by Shleifer, the EMH relies on three basic arguments [104]. First, market participants are assumed to be rational, and therefore rationally value securities. Second, irrational participants, should they exist, trade in a way that is essentially random. Their behavior cancels out and does not affect prices. Finally, if irrational participants do happen to exist and manage to act in unison, their affect on prices is mitigated or eliminated by arbitrageurs. Therein lies the importance of arbitrage in financial markets.

In our study, we suspect arbitrage opportunities arise, again, due to the asynchronous activity of traders in a fragmented market. As mentioned in the prior section, our markets are mechanically independent. Even in a market where all participants are doing their best to act rationally, the physics of the market and the asynchronicity of activity leads to a system in which arbitrage opportunities are likely to arise. Based on our analysis of correlations, we suspect that mis-pricings exist at short time-horizons. Furthermore, we expect that these potential arbitrage opportunities, should they exist, either decrease in duration or number as markets become faster and more connected.

Practically speaking, the actual mechanics of exploiting an arbitrage trade are too difficult to simulate. The details, including trading costs, taxes, and, in our case, foreign exchange, are quite complicated. We seek to avoid these troubles by developing a measure that gives an upper bound estimate on the value an arbitrage opportunity. We define an arbitrage opportunity as a continuous period of time in

Year	Nokia	Nordea	Stora	Telia
2009	18086	7729	3676	857
2010	1877	2103	896	243
2011	187	910	1428	122
2012	45	126	134	24
2013	110	330	463	155

Table 4.3: Annual number of arbitrage opportunities based on 0.005 euro threshold.

which the difference between crossed bid and ask prices exceed a threshold of 0.005 euros. With quotes converted to euros using the spot exchange rate, we consider an arbitrage opportunity to exist if $b_h - a_s \geq 0.005$ euros or if $b_s - a_h \geq 0.05$ euros where b_h and a_h are the bid and ask in the Helsinki market and b_s and a_s are the bid and ask in euros in the Stockholm market. We calculate the value of the each discrete arbitrage opportunity as either $\max\{d(t)(b_h(t) - a_s(t))\}$ or $\max\{d(t)(b_s(t) - a_h(t))\}$ where $d(t)$ represents the available depth at the time of the transaction, which is the minimum of the depth available at the bid and the depth at the ask.

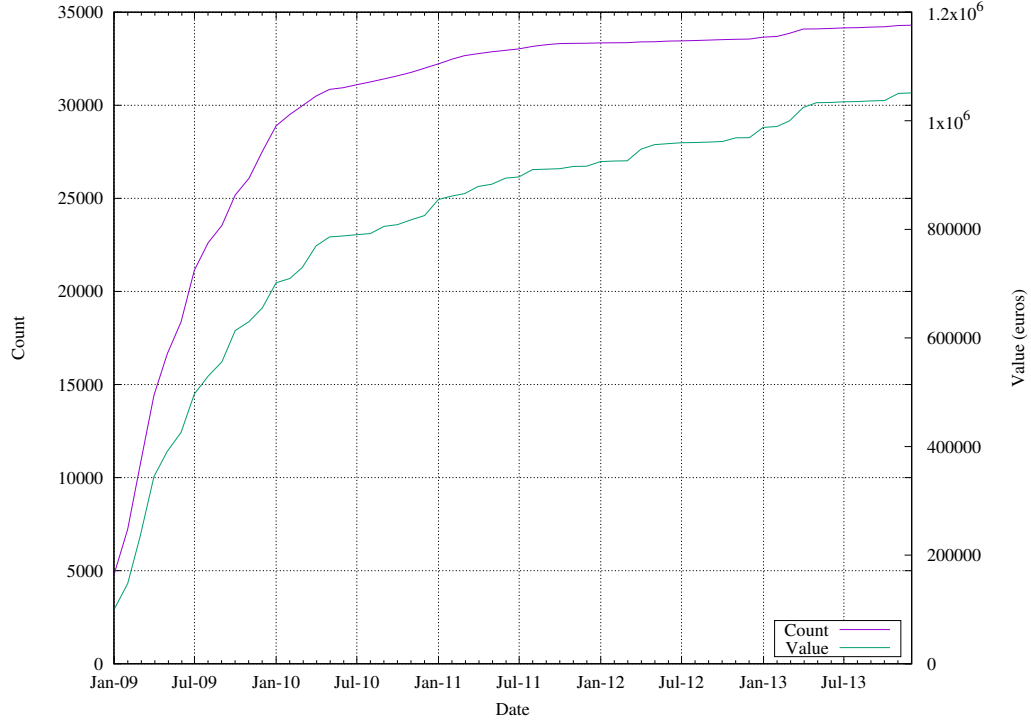


Figure 4.8: **Cumulative number of arbitrage opportunities.** cumulative number of arbitrage opportunities based on our 0.005 euro threshold. The calculation aggregates the number of arbitrage opportunities in Nokia, Nordea Group, Stora Enso and Telia Sonera over the 5 year sample period.

Figure 4.8 shows the cumulative number of arbitrage opportunities over our sample period taking into account all four stocks in the sample. Table 4.3 shows the number of each individual stock. Clearly, the number of opportunities, based on our definition and the 0.005 euro threshold, decreases significantly over the sample period. Figure 4.9 shows the number of opportunities in the four individual stocks. Later, we address a number of reasons why the observed number of arbitrage oppor-

tunities are decreasing.

Figure 4.10 provides a distribution of the length of the arbitrage opportunities in aggregate across all four stocks for the entire sample period. The obvious fact to note about the distribution is that longer arbitrage opportunities are less frequent. The vast majority of arbitrage opportunities exist for less than 10 seconds. There are a number of reasons why such opportunities persist. First, despite our 0.005 euro threshold, they may be too small to exploit. Second, since the value of an arbitrage opportunity may change during its lifetime (e.g. depth increases at the active price), it may be exploited later in its life when the value exceeds a threshold which is deemed profitable to the arbitrageur. Finally, what appear to be arbitrage opportunities in our data may not reflect “the truth on the ground” for market participants, a point we’ll return to later.

Figure 4.11 provides a distribution of the values of the arbitrage opportunities as defined previously. Again, less profitable opportunities are less frequent. In fact, only a handful of arbitrage opportunities exceeding 100 euros exist in the entire five year period across all four stocks. Our results point to remarkable efficiency of the market considering our calculations are based on the spot rate of FX, which may or may not be the effective rate for a trader.

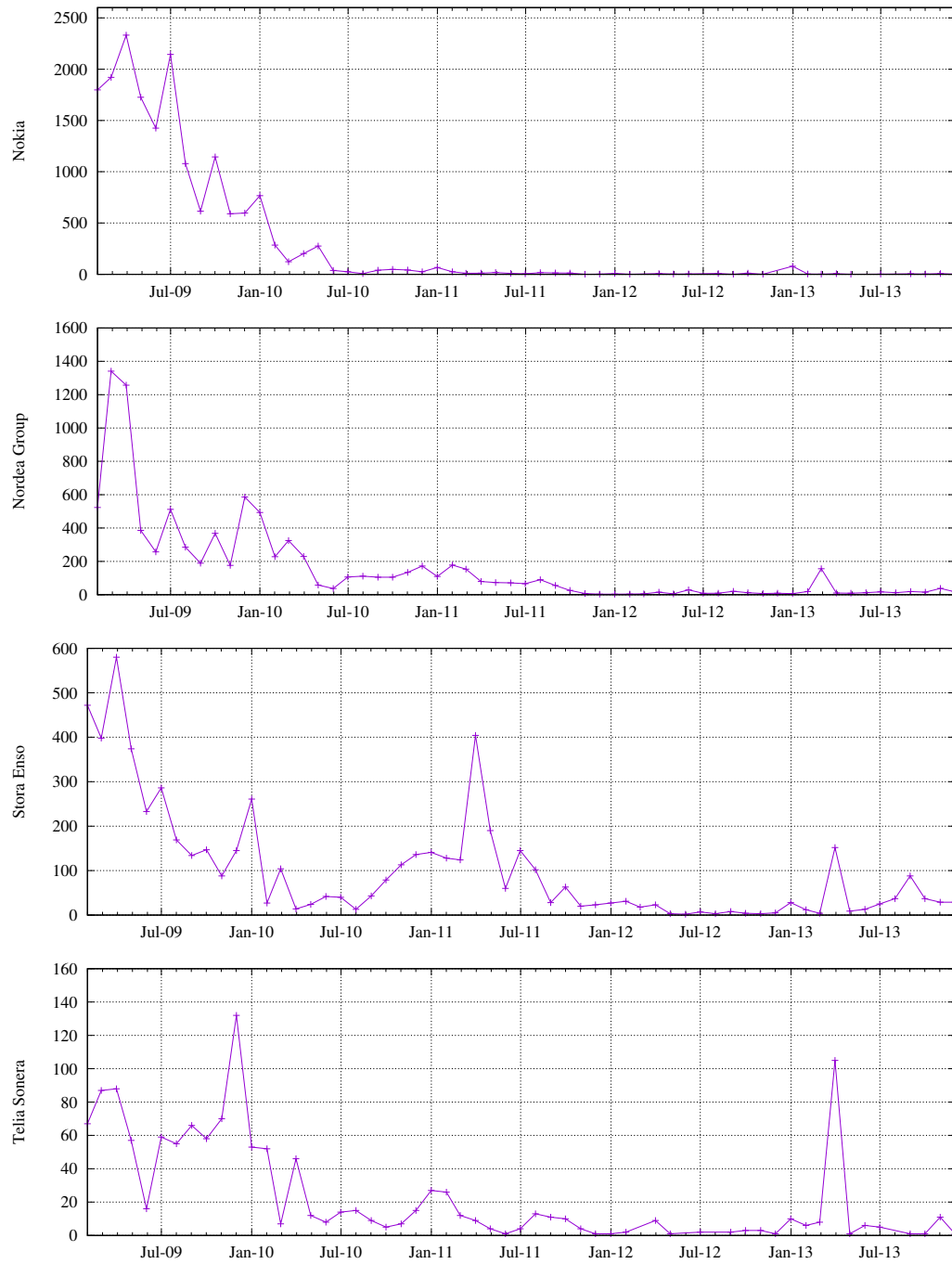


Figure 4.9: Number of arbitrage opportunities per stock.

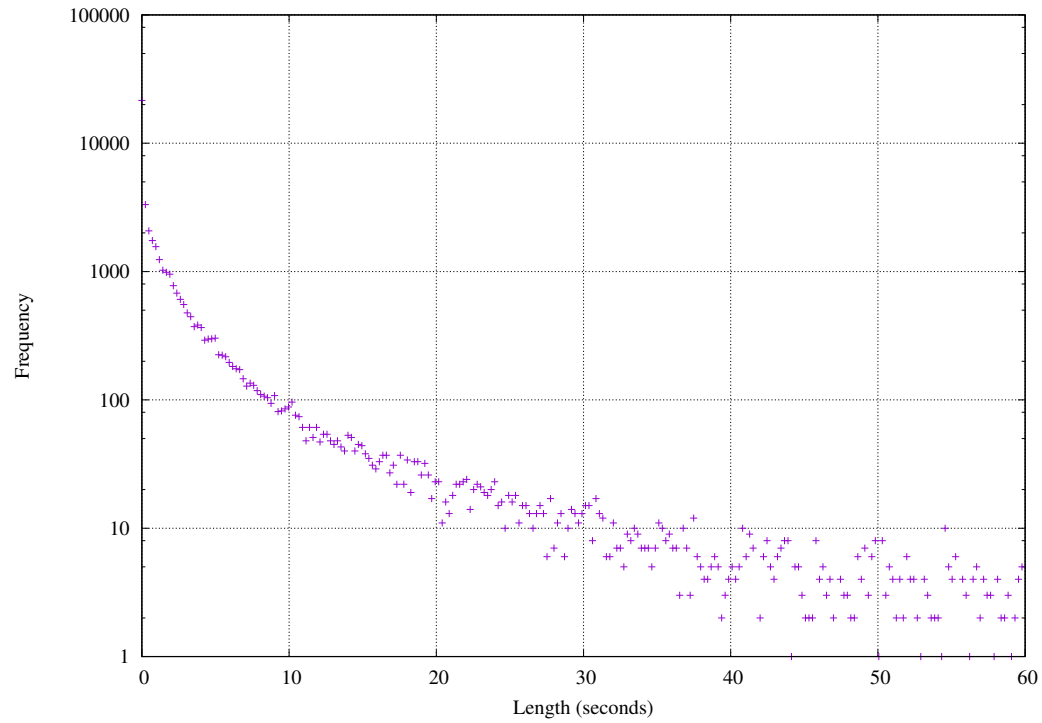


Figure 4.10: **Length of arbitrage opportunities.** Histogram of the length of the arbitrage opportunities across all stocks in the sample for the five year time period.

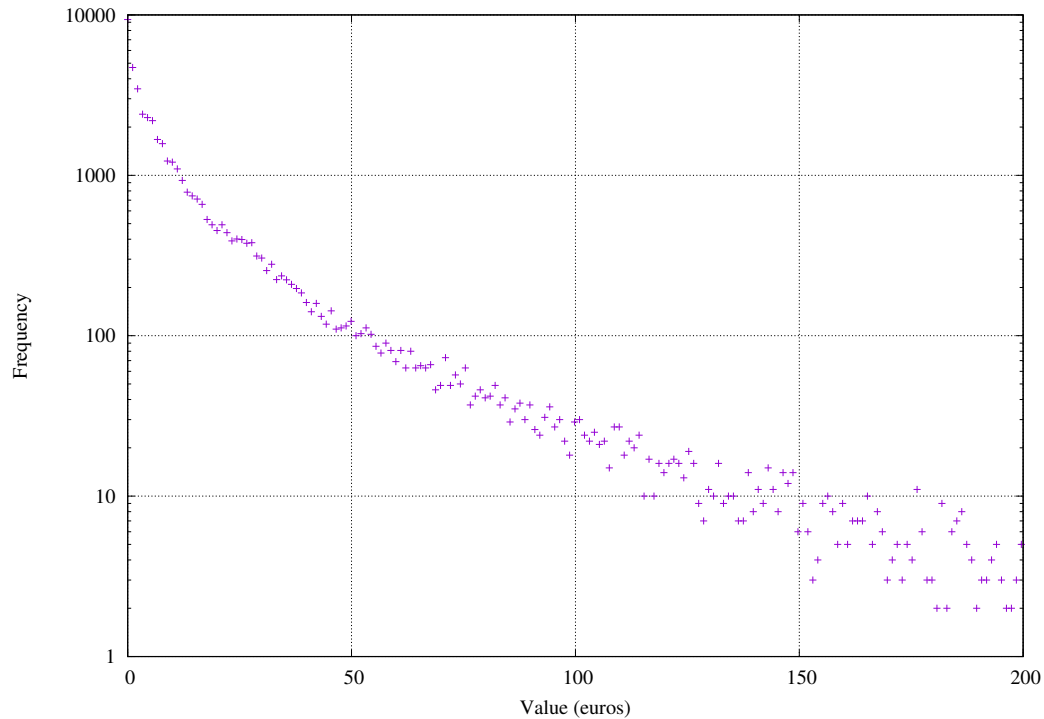


Figure 4.11: **Value of arbitrage opportunities.** Histogram of the value of the arbitrage opportunities across all stocks in the sample for the five year time period. We define the value of the arbitrage opportunity to be the profit from the transaction if it were to be perfectly timed in a hypothetical market in which one could transact in both stocks in euros.

4.5 Trade-throughs

A trade through (TT) is a transaction at a quote that is inferior to other simultaneously available quotes. In the United States, regulations restrict transactions that result in trade-throughs. In Europe, however, there is no such restriction. Ende,

Gomber and Lutat (2010) find that around 6.71% of transactions traded-through superior quotes in a sample of EURO STOXX 50 securities [109]. Similarly, Storkenmaier and Wagener (2011) find that trade-throughs account for 5.2% to 8.7% of volume in a 2009 observation period and for 4.7% to 6.9% of volume in a 2010 observation period using a sample of FTSE 100 stocks [130]. Kohler and von Wyss find similar numbers for a sample of Swiss stocks examined over the time period of 2008 to 2010 [131].

In our analysis, we study trade-throughs by again using the EURSEK spot rate to convert prices and quotes to euros. Denote the bid and ask prices in Helsinki as $b_h(t)$ and $a_h(t)$, respectively. Recall that prices in Helsinki are quoted in euros. Now, denote the bid prices and ask price in Stockholm as $b_s(t)$ and $a_s(t)$, where the prices are quoted in Swedish Kronor. We classify a transaction at time t at the ask in Helsinki when $a_h(t) > (a_s(t)/x(t))$ as a trade-through, where $x(t)$ denotes the EURSEK spot rate. Likewise, a transaction at the bid in Helsinki when $b_h(t) < (b_s(t)/x(t))$ is classified as a trade-through. Trade-throughs in Stockholm are defined analogously.

Volume of trade-throughs

Figure 4.12 gives the percentage of volume that trades through better quotes in each of the markets for the five stocks in our sample. The volumes of trade throughs we present are not comparable to those we quoted from earlier studies simply due to the fact that the tick size of the exchange rates results in a much more granular price grid. The price in one market may improve on a price in the other by fractions of a

euro cent.

The first and obvious thing to note about the volume of trade-throughs is that more transactions trade through superior quotes in the smaller markets by volume. This is to be somewhat expected since we found spreads to be higher on average in those markets. Qualitatively, the number of trade throughs appear to be converging later in the sample period, with the exception of Telia Sonera. This fact, along with spreads, suggests that the venues are becoming more competitive over the sample period. Volume of trade throughs can only tell us so much, and in this case, the volume primarily reflects differences in the spread. More interesting, and more economically important, are the costs of trade throughs.

Cost of trade-throughs

Figure 4.13 gives the volume-average cost of the trade throughs for each side of the market in each venue for the four stocks in the sample. The costs of trade throughs are calculated as follows. Let $v_i(t)$ be the volume of the i th trade-through at the bid in a given month in a given market at time t . For a total number of N trade throughs at the bid in that market, the total volume of those trade-throughs is $V = \sum_{i=1}^N v_i(t)$. We then define the average cost of trade throughs at the bid in Stockholm as

$$\frac{\sum_{i=1}^N v_i(t)((b_s(t)/x(t)) - b_h(t))}{V}$$

Trade throughs as the ask and in Helsinki are defined analogously. The cost of trade throughs appears to decrease throughout the sample period for Nokia. The same can not be said about the other stocks in the sample. Nokia, is a well known

stock internationally with an ADR listed on the New York Stock Exchange, which may have resulted in more attention from arbitrageurs. Nokia, however, had lower volume over the period than Nordea, which saw trade through costs remain even for the period. Overall, the result of our trade through costs is ambiguous, but total costs reflect a fairly efficient market.

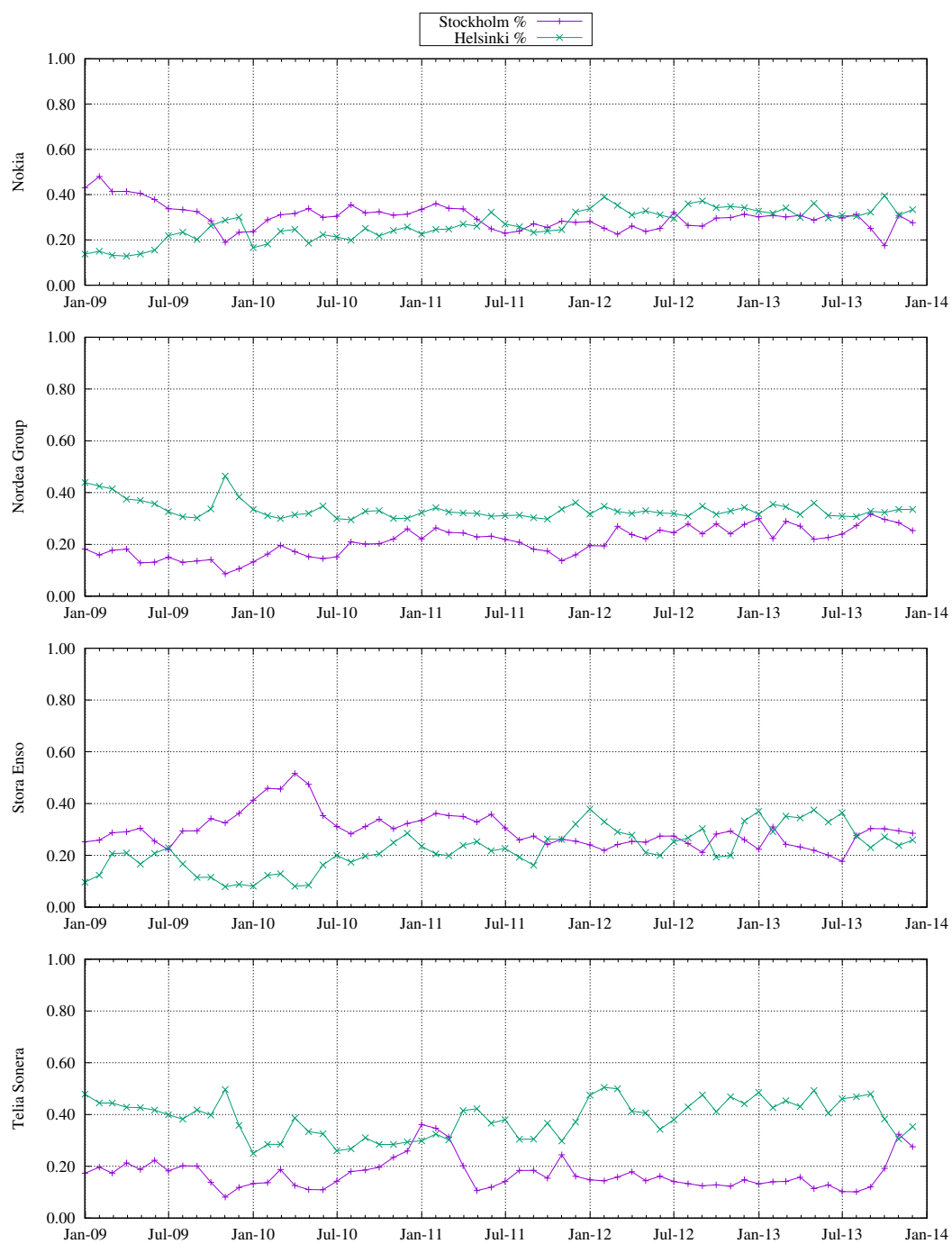


Figure 4.12: **Volume of trade-throughs.** Percentage of volume that trades through better quotes

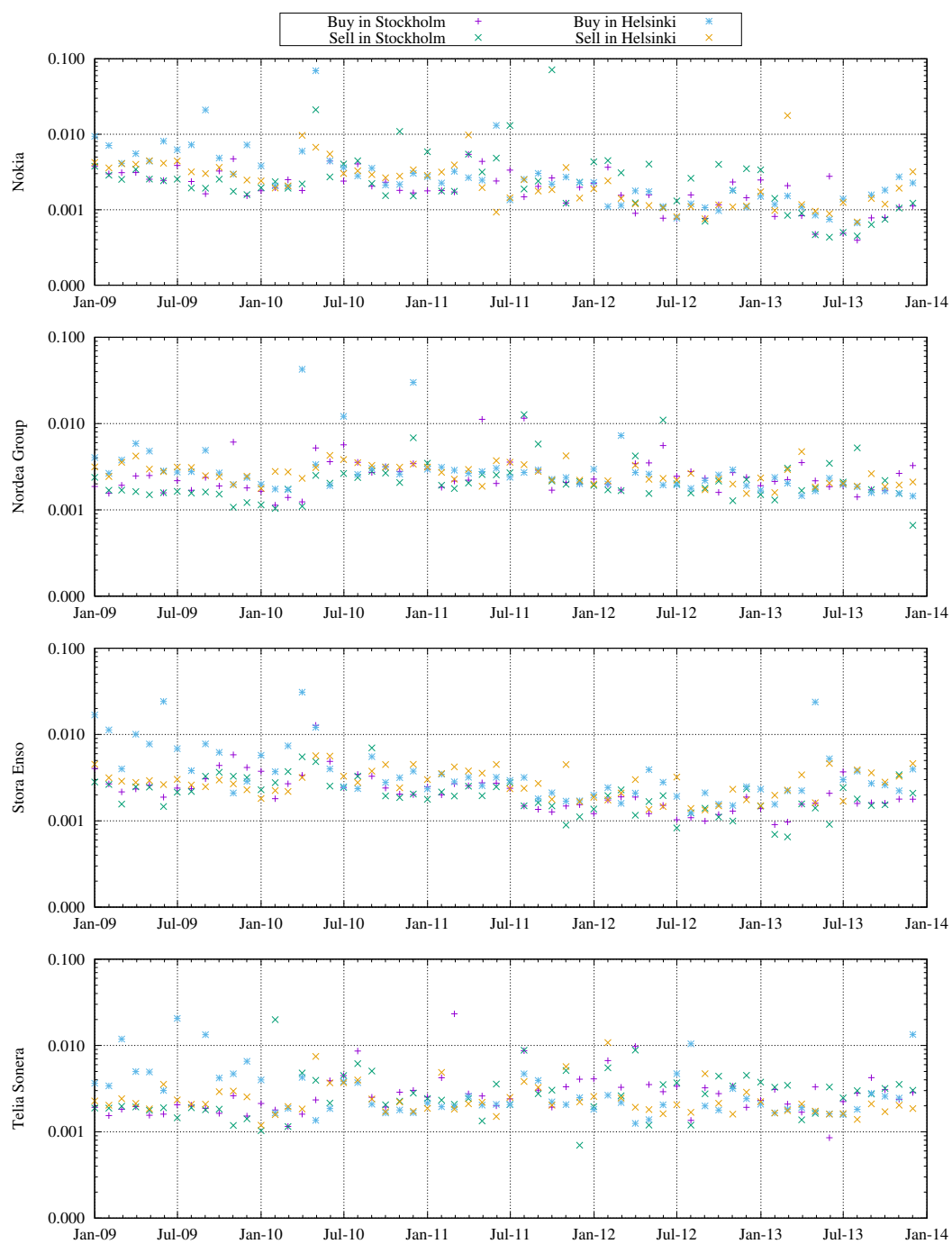


Figure 4.13: **Cost of trade-throughs.** Difference between two available quotes at the time of the trade-through averaged over volume.

4.6 Conclusion

In this chapter, we've provided an analysis of correlations, arbitrage opportunities and trade throughs in a sample of cross-listed Nordic equities. The sample period, from 2009 to 2013, consisted of at least three major improvements to market structure. Correlations at the millisecond level increased remarkably indicating an improvement in the response time of market participants. Arbitrage opportunities, based on a currency conversion using the spot FX rate, all but disappeared, suggesting entrants that may be able to exploit those opportunities. Finally, we found a significant decrease in the cost of trade-throughs in Nokia. Cost results for other stocks in the sample were ambiguous.

Chapter 5

Conclusion

Financial markets are evolving at a pace that will continue to challenge regulators and policy makers. In an ideal world, such parties should have a platform to evaluate new policies and ideas. In this dissertation, we present a simulation-based analysis of a multi-market system alongside an empirical analysis of a fragmented market for equities. In both cases, we were able to learn about the dynamics of distributed markets, but the way forward for purely empirical approaches will be increasingly difficult.

Researchers often turn to agent-based models when empirical analysis is expensive or impossible. The global financial market is one of the largest and most complex systems in the world, but it also has one of the most detailed paper trails. Financial data are everywhere and vendors compete to sell that information to consumers. So, why do we need agent-based models? Entire industries are built on the analysis of financial data, but purely statistical approaches to the analysis of markets is compli-

cated by the daunting complexity of the market structure, and a purely statistical approach often discounts the importance of institutions and mechanics.

Agent-based models provide a platform to include real institutions and real market mechanisms, and those models need not be entirely built from the ground up. Statistically constructed agents allow us to avoid the extreme difficulty of modeling an entire heterogeneous population of economic actors. Instead, we can characterize the population and the output of their strategic behavior statistically. This approach has numerous advantages. Statistical agents allow more naturally for experimental designs. Models can also be fit to data when strategic behavior is either unknown or too difficult to model. Finally, statistical agents can be intermingled with deterministic, rules-based agents, providing additional flexibility and value for policy evaluation.

In this dissertation, we contribute to a young and rapidly evolving field. Agent-based models for analysis of financial markets and the economy at the level of national policy is still a work in progress. Agent-based modeling, however, will become more and more compelling as an alternative to traditional methods as computing power and modeling platforms continue to improve. Empirical finance also needs to evolve in order to keep up with the volume of data and complexity of market structure. Ultimately, the analysis of financial markets will by necessity be an interdisciplinary field. Financial economists will be well served to partner with computer scientists and engineers to build models of financial markets and the economy.

Chapter 6

Appendix

6.1 Appendix A: Nordic Market Supporting Figures

6.1.1 Volume

In this section of the appendix, we provide volume figures for all six of the cross-listed securities considered for our sample. Table 6.1 gives the Reuters Identification Codes for each of the securities. Figures 6.2, 6.3, 6.4, 6.5, 6.6, and 6.7 show monthly and cumulative volume.

Ericsson	Nordea Group	Nokia	Stora Enso	Telia Sonera	Tieto
ERICb.ST	NDA.ST	NOKI.ST	STEr.ST	TLSN.ST	TIEN.ST
ERIBR.HE	NDA1V.HE	NOK1V.HE	STERV.HE	TLS1V.HE	TIE1V.HE
	NDA.CO				

Table 6.1: Reuters Identification Codes for the 6 cross-listed stocks.



Figure 6.1: Euro to Swedish Kronor exchange rate over the sample period.

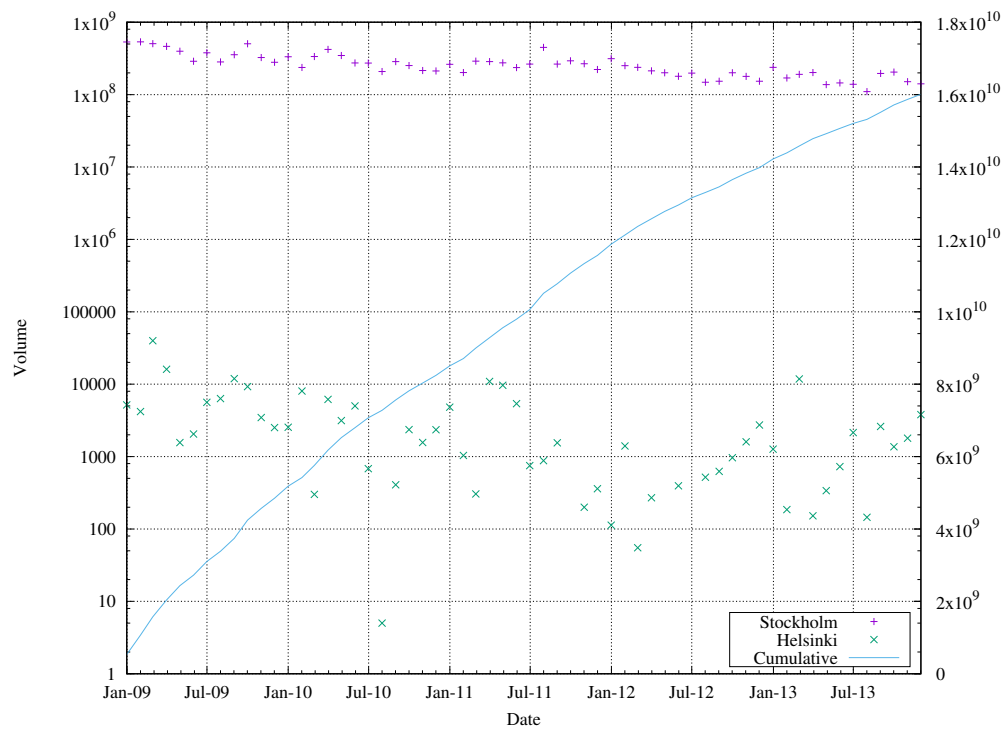


Figure 6.2: Volume for Ericsson.

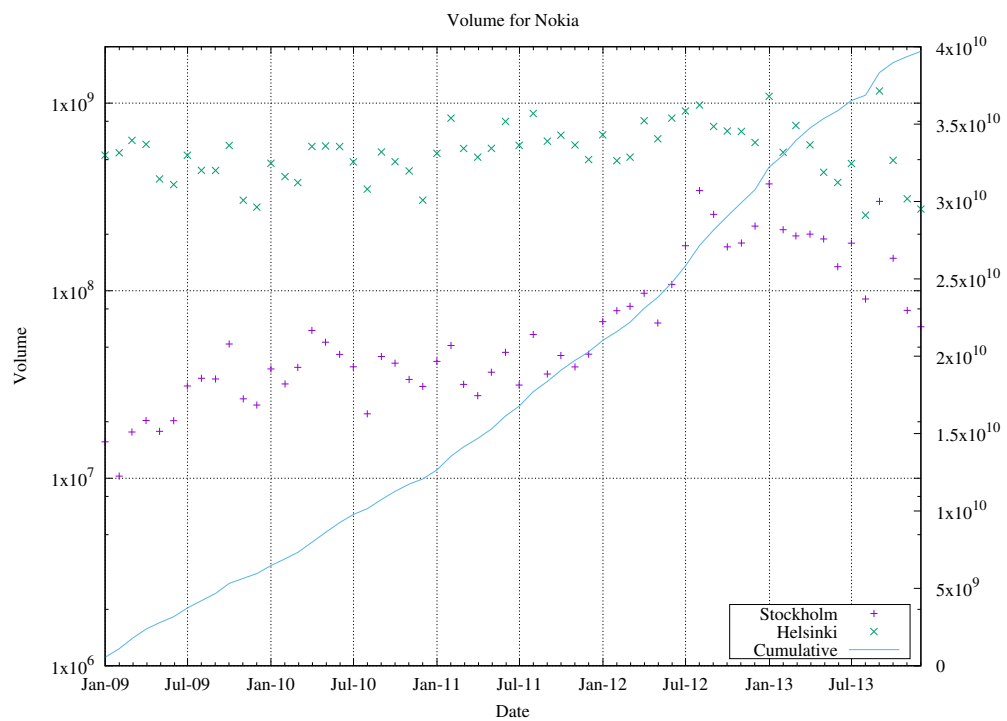


Figure 6.3: Volume for Nordea Group.

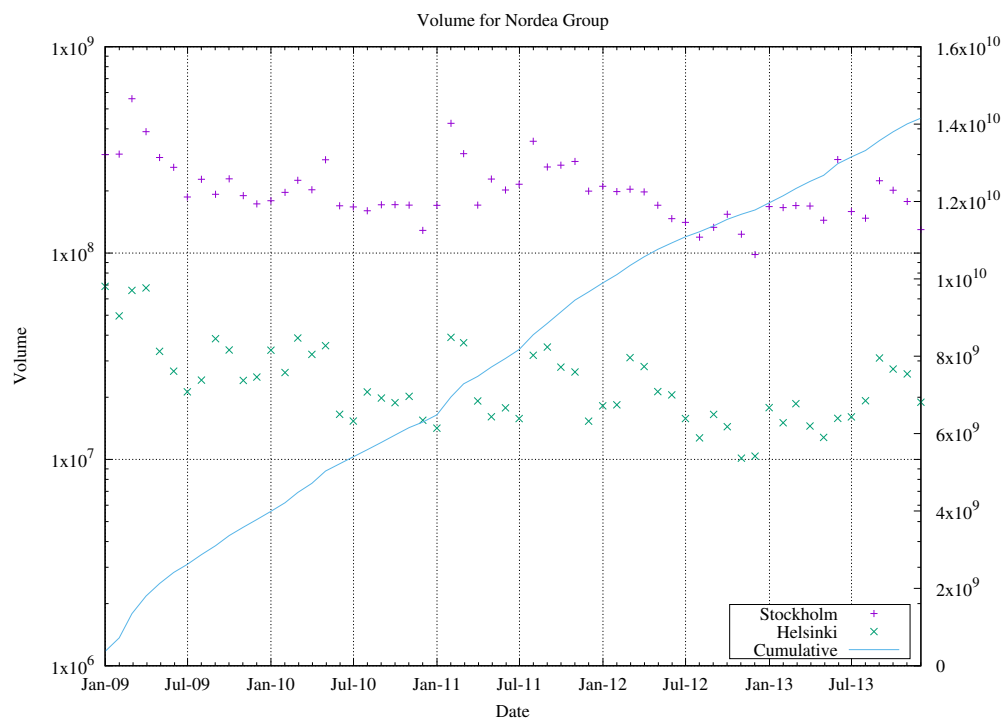


Figure 6.4: Volume for Nordea Group.

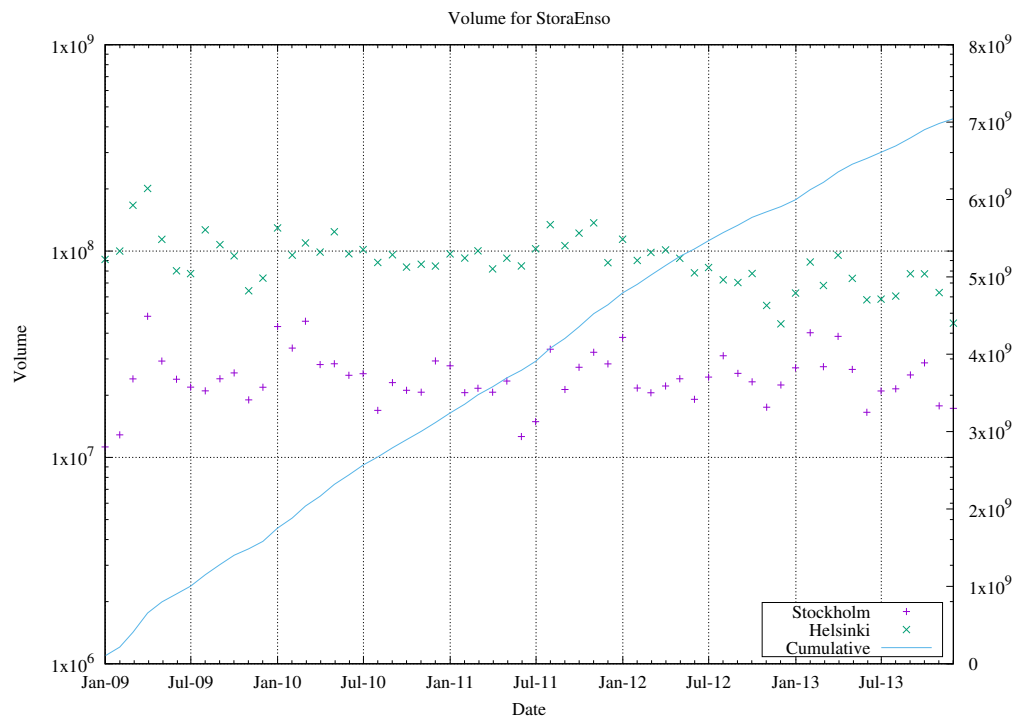


Figure 6.5: Volume for Stora Enso.

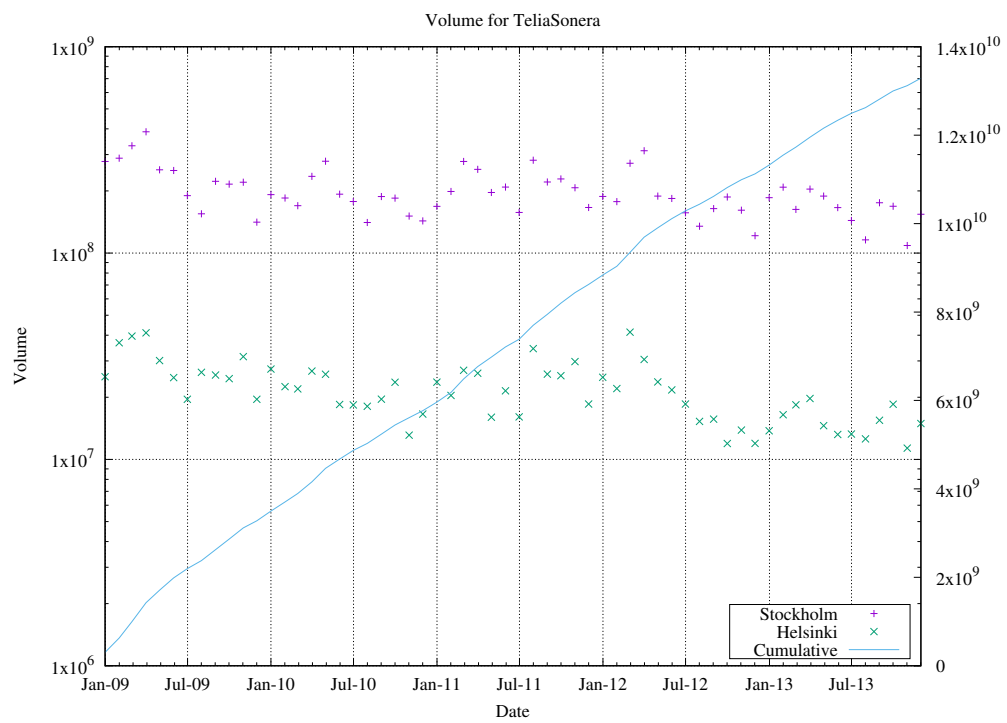


Figure 6.6: Volume for Telia Sonera.

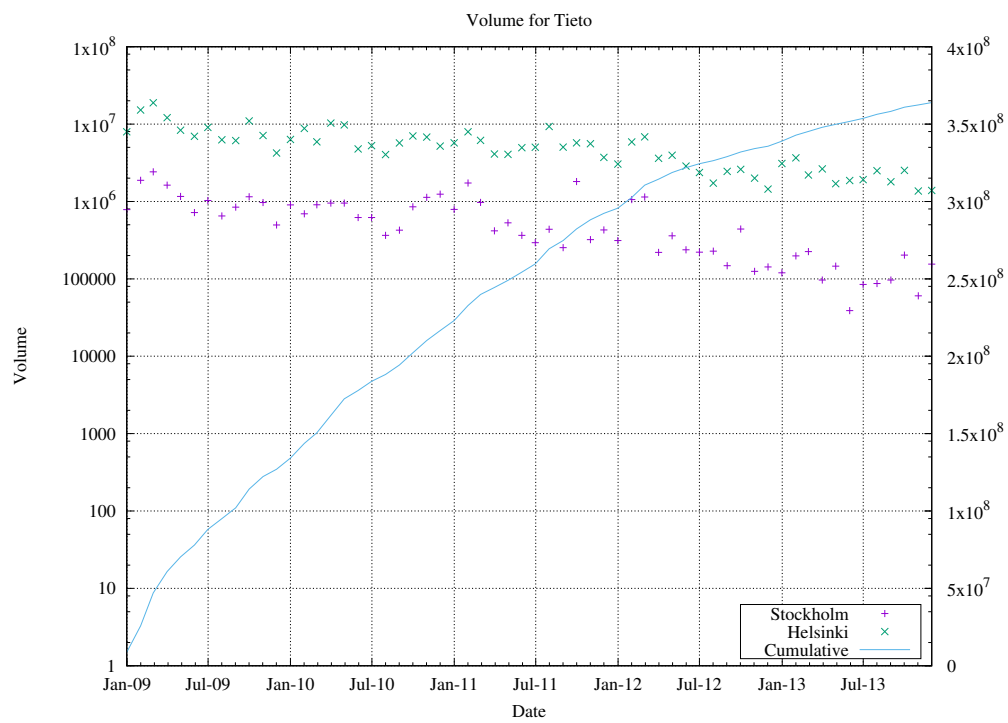


Figure 6.7: Volume for Tieto.

6.1.2 Market Updates

In this section of the appendix, we provide market update figures for the four cross-listed securities considered in our sample. Figures 6.8, 6.9, 6.10 and 6.11 show monthly and cumulative market updates.

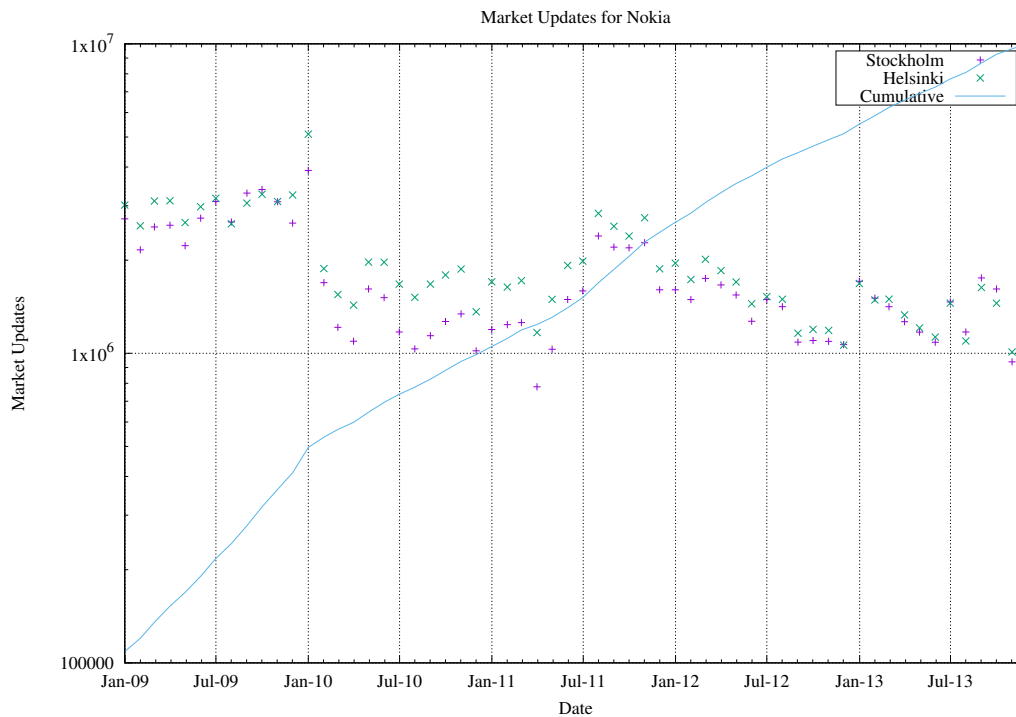


Figure 6.8: Number of market updates for Nokia.

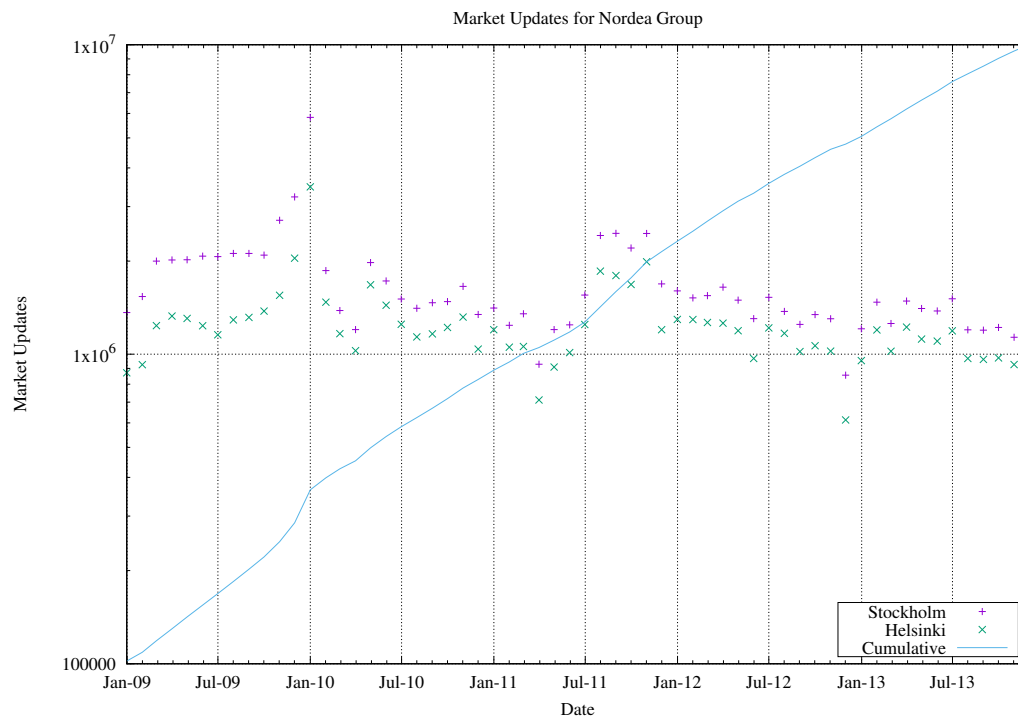


Figure 6.9: Number of market updates for Nordea Group.

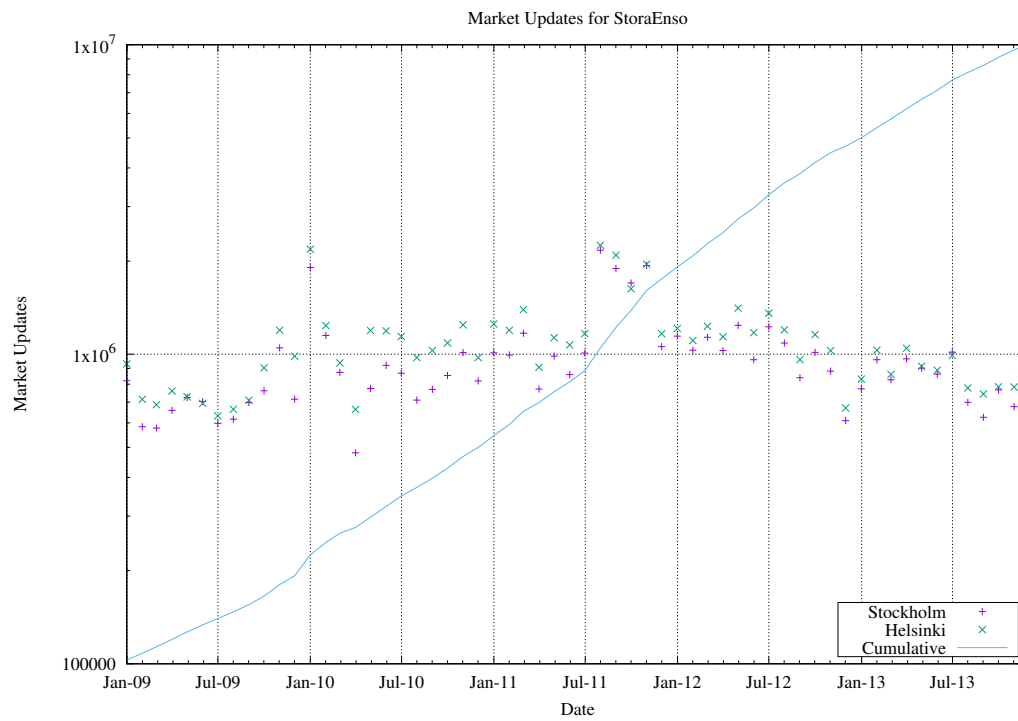


Figure 6.10: Number of market updates for Stora Enso.

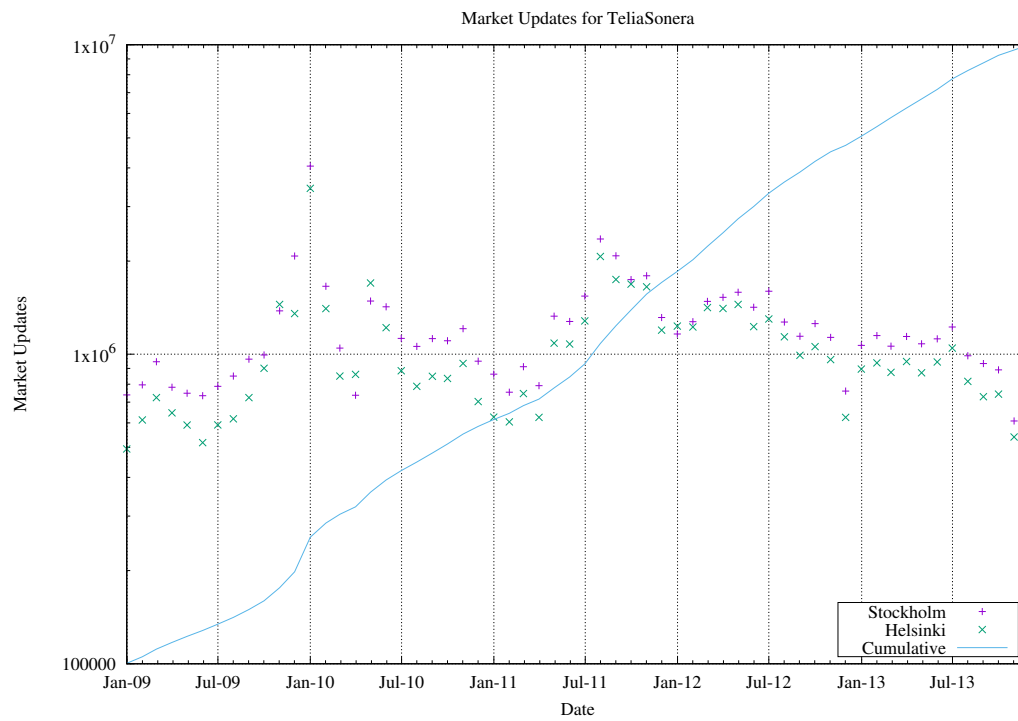


Figure 6.11: Number of market updates for Telia Sonera.

6.1.3 Spread

In this section of the appendix, we provide average daily spread figures for the four cross-listed securities considered in our sample. Figures 6.12, 6.13, 6.14 and 6.15 show monthly and cumulative market updates.

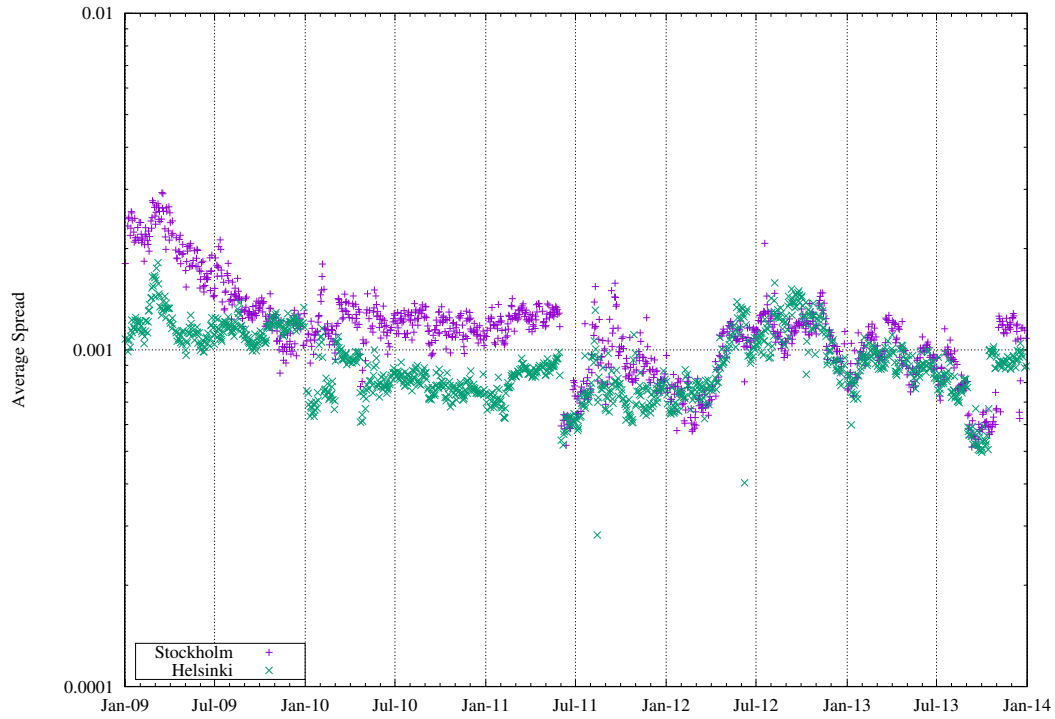


Figure 6.12: Average spread for Nokia.

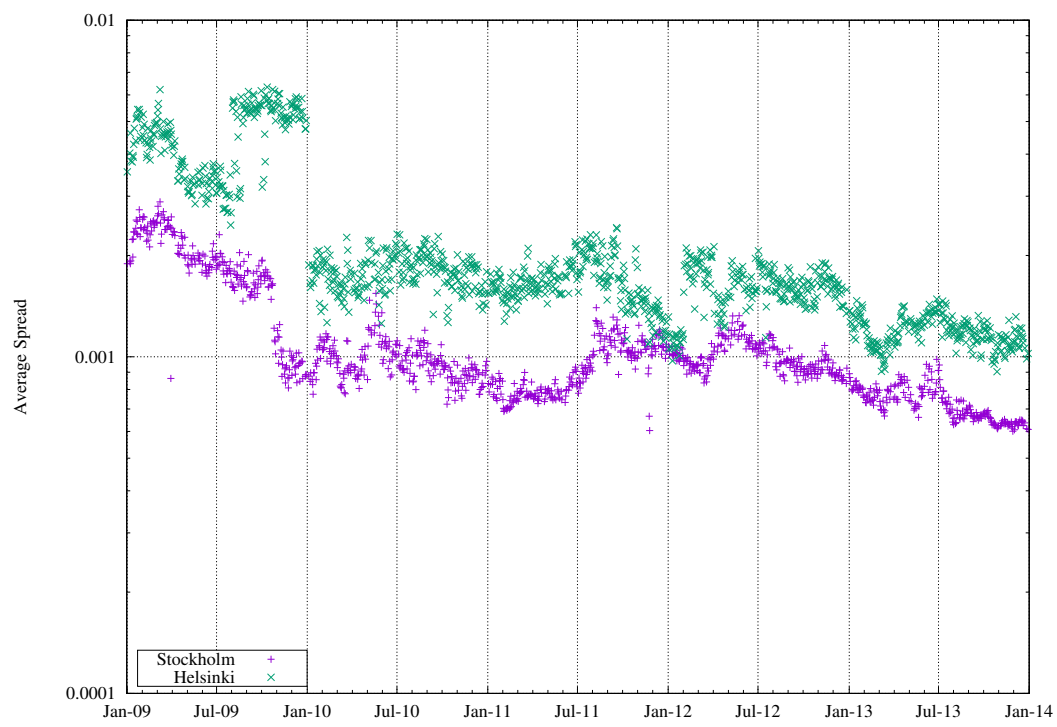


Figure 6.13: Average spread for Nordea Group.

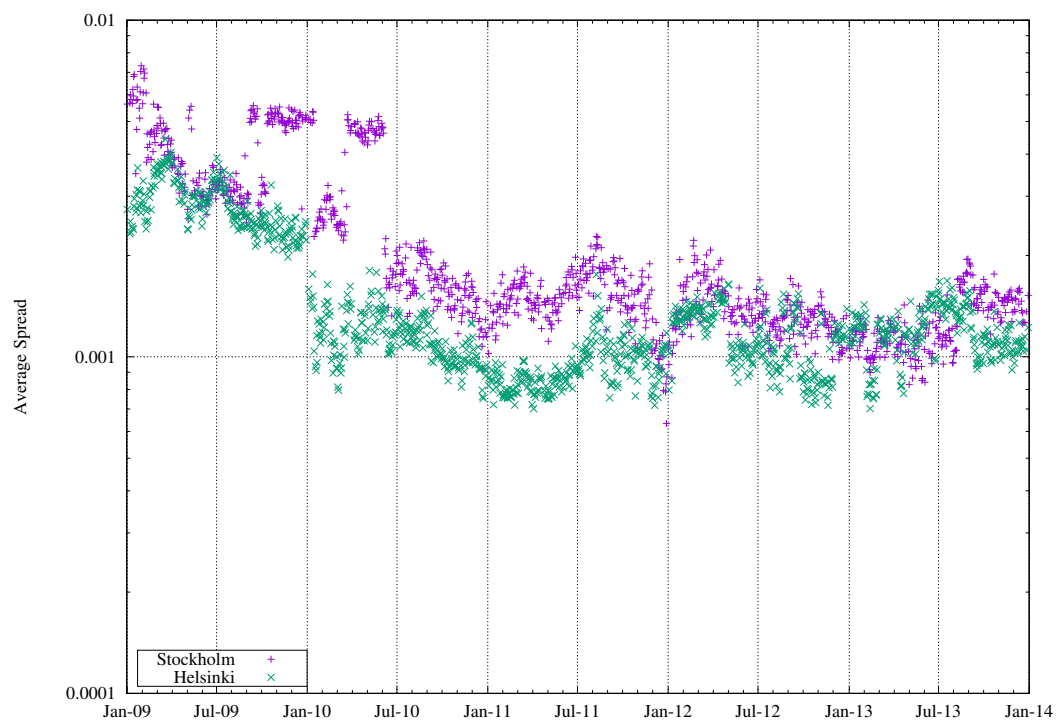


Figure 6.14: Average spread for Stora Enso.

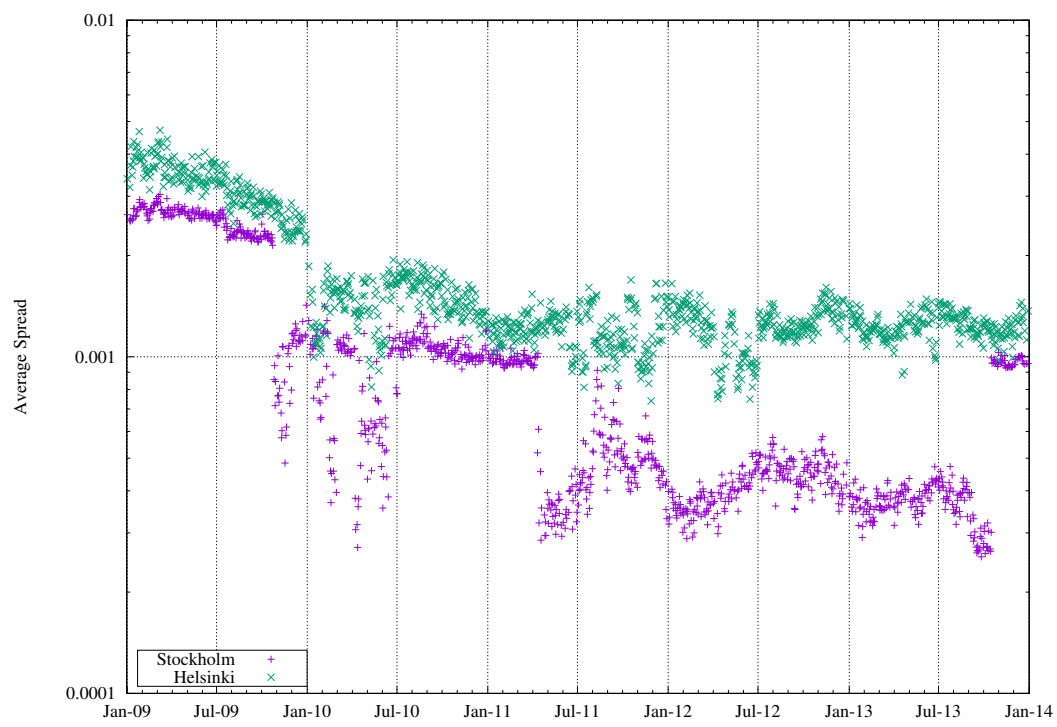


Figure 6.15: Average spread for Telia Sonera.

6.1.4 Depth

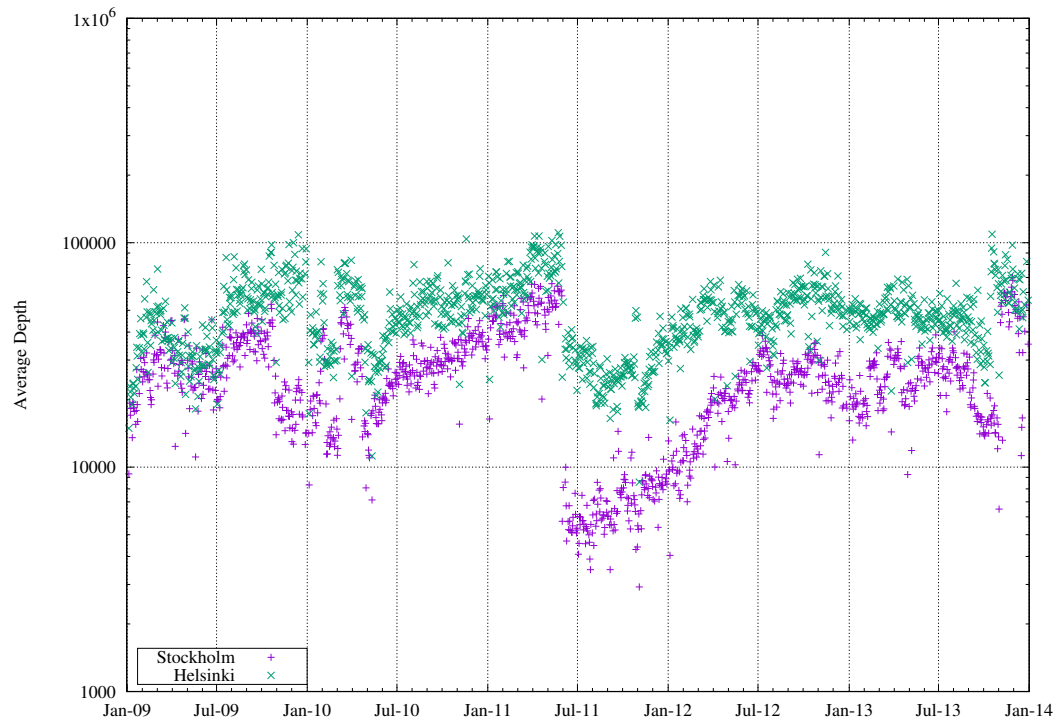


Figure 6.16: Average depth at top of book for Nokia.

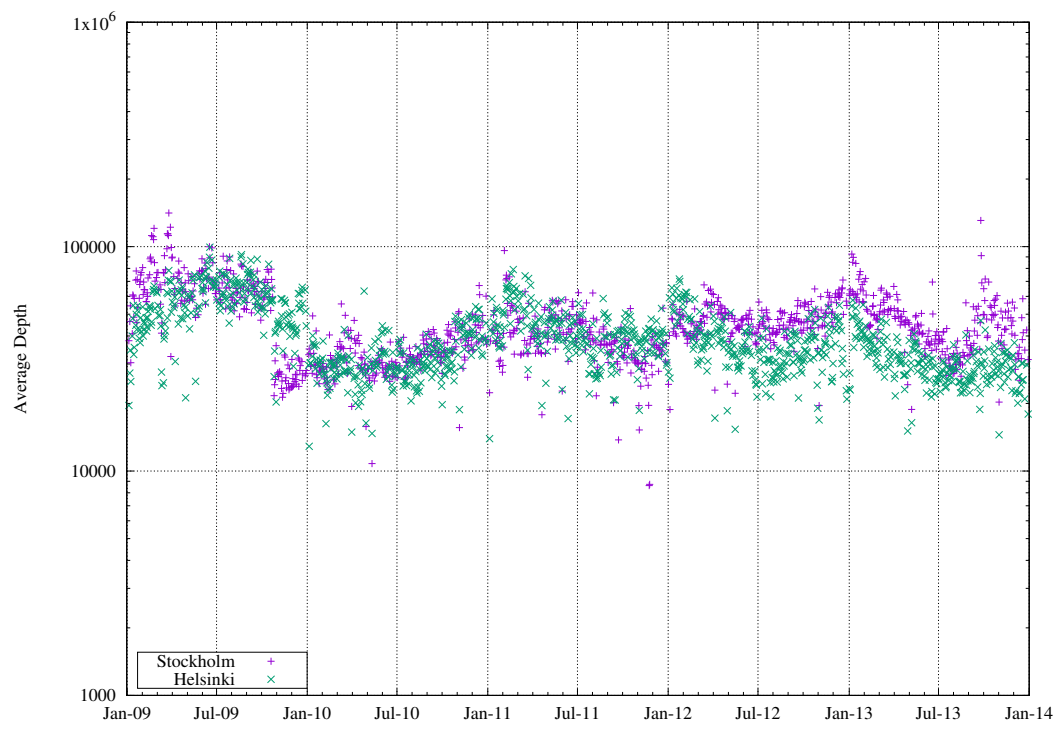


Figure 6.17: Average depth at top of book for Nordea Group.

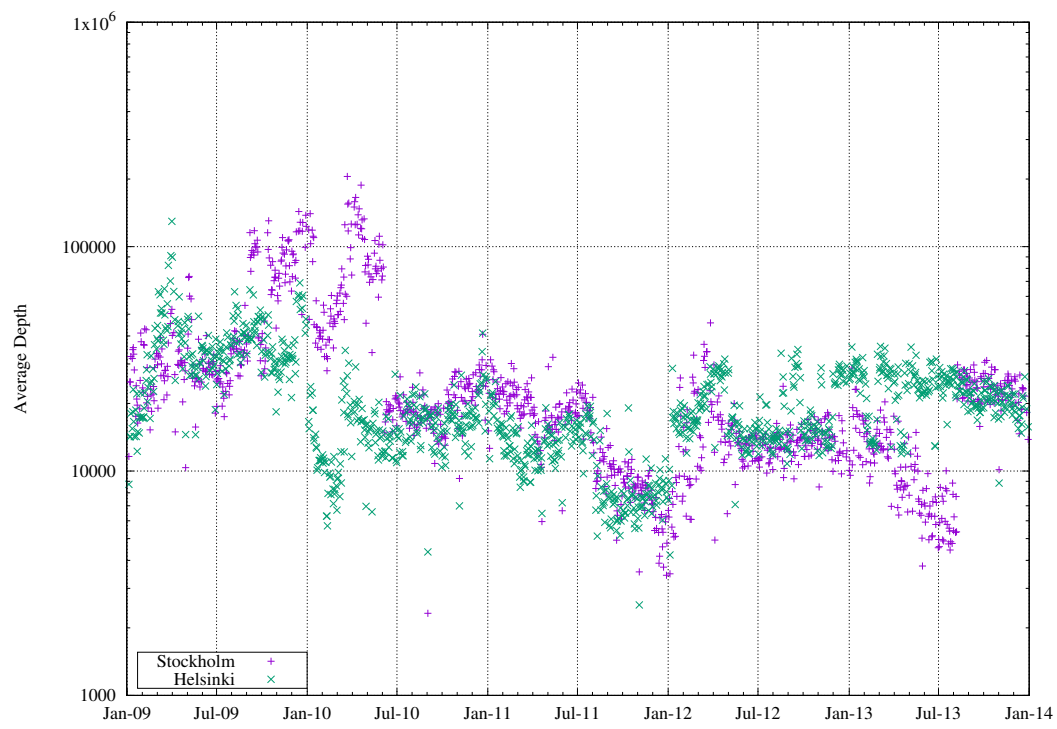


Figure 6.18: Average depth at top of book for Stora Enso.

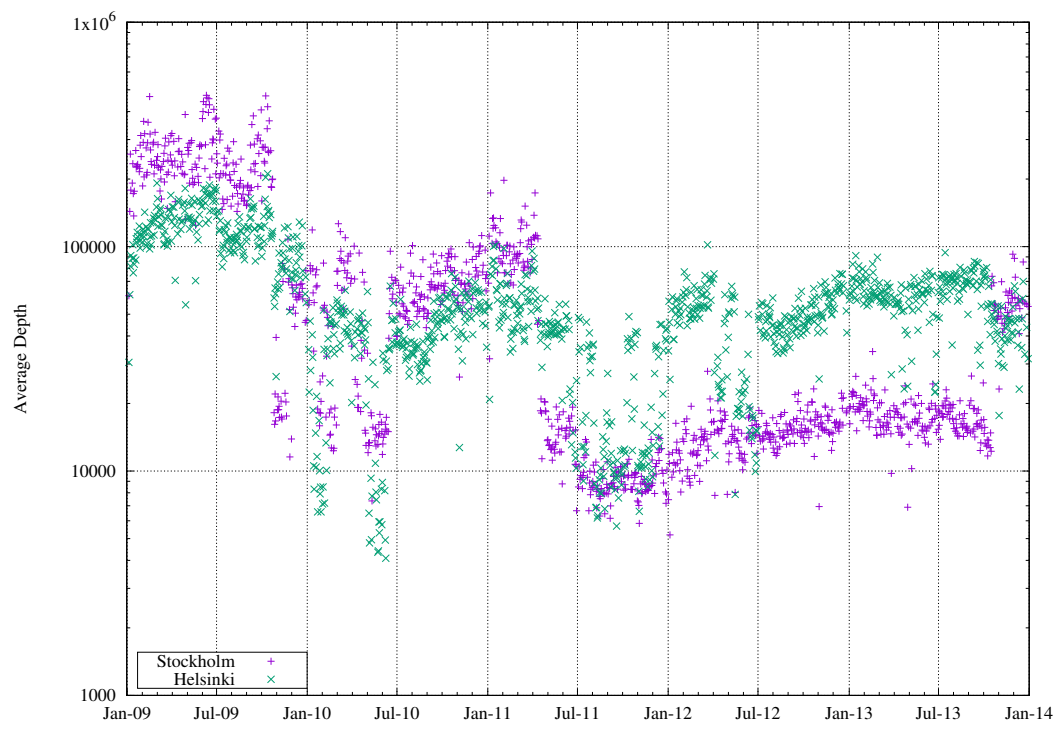


Figure 6.19: Average depth at top of book for Tela Sonera.

6.1.5 Mid-quote correlation

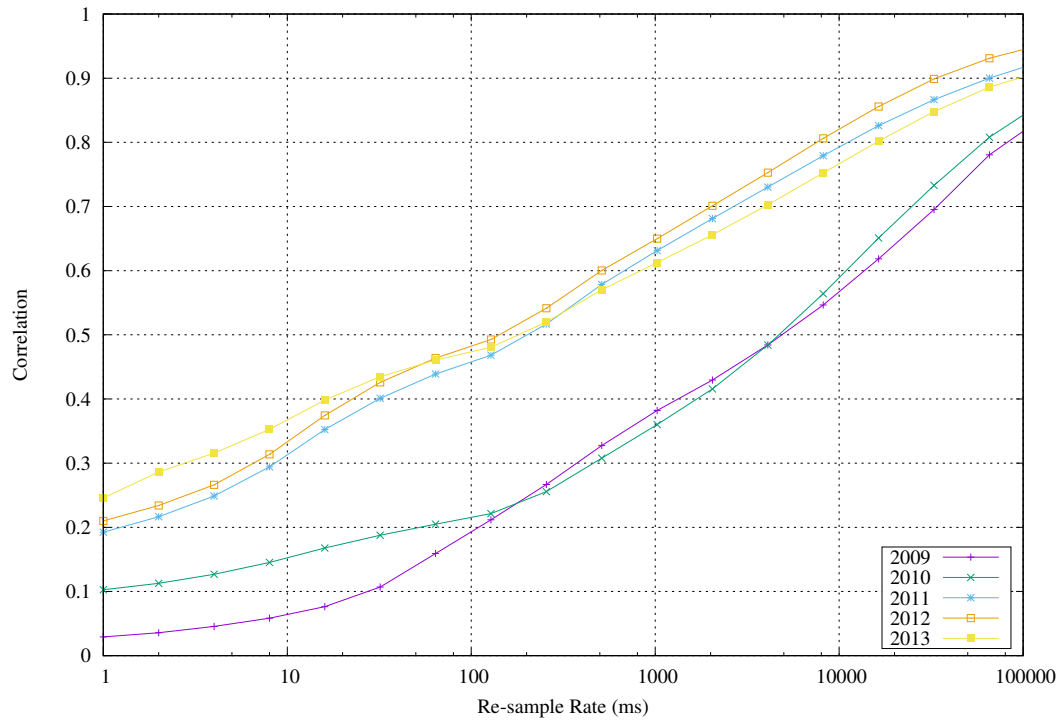


Figure 6.20: Mid-quote correlations for Nokia.

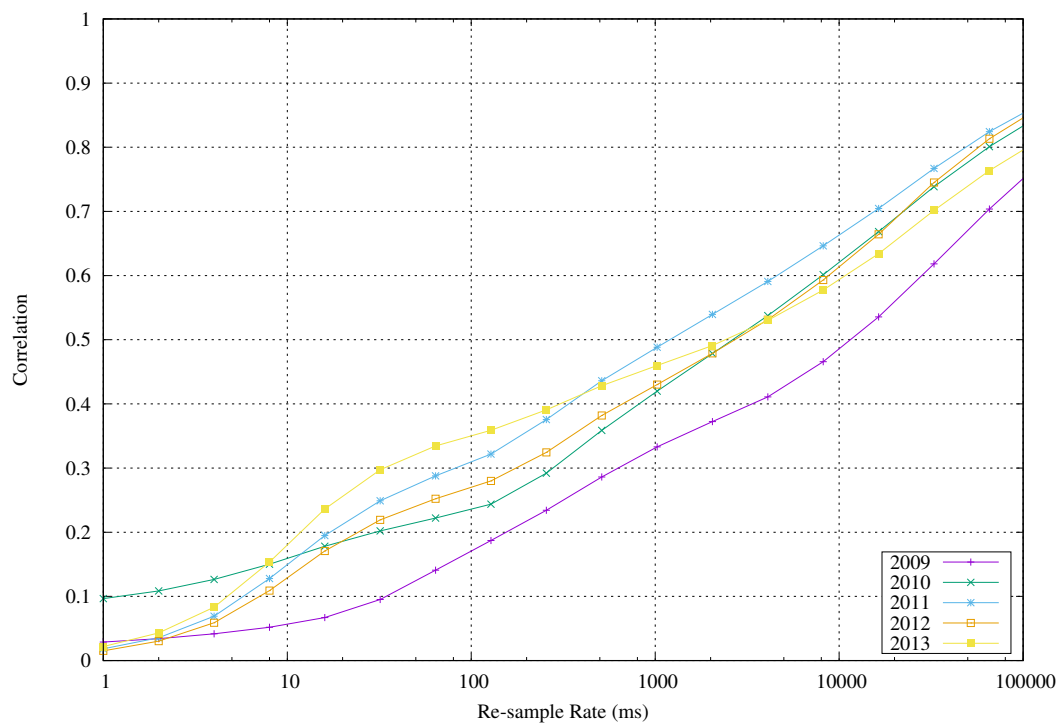


Figure 6.21: Mid-quote correlations for Nordea Group.

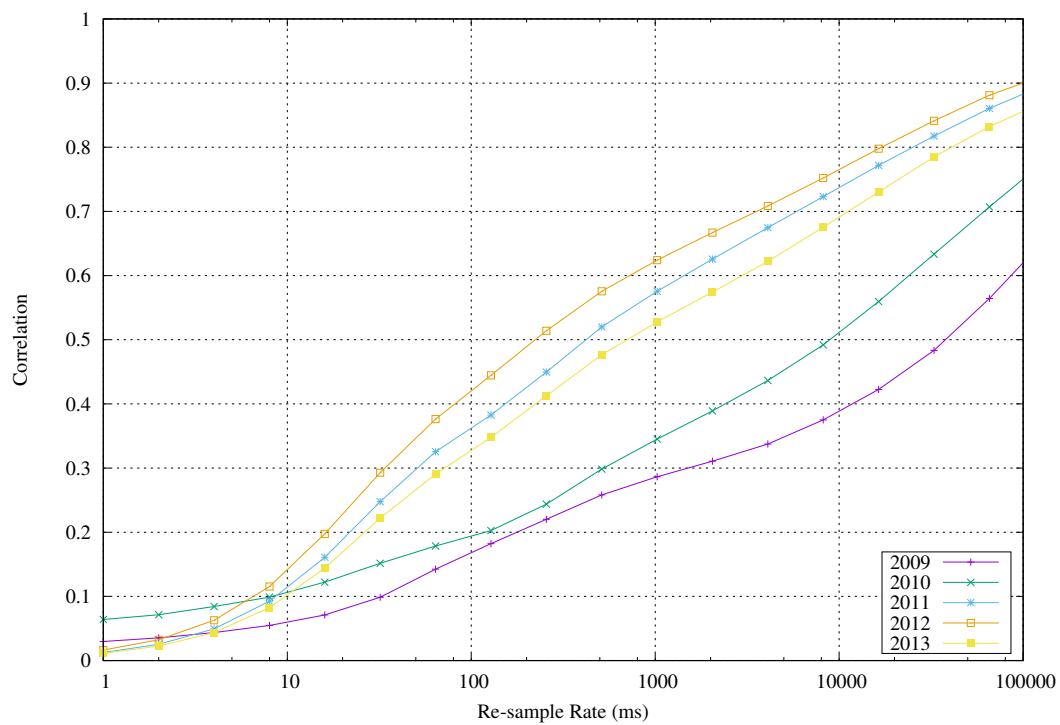


Figure 6.22: Mid-quote correlations for Stora Enso.

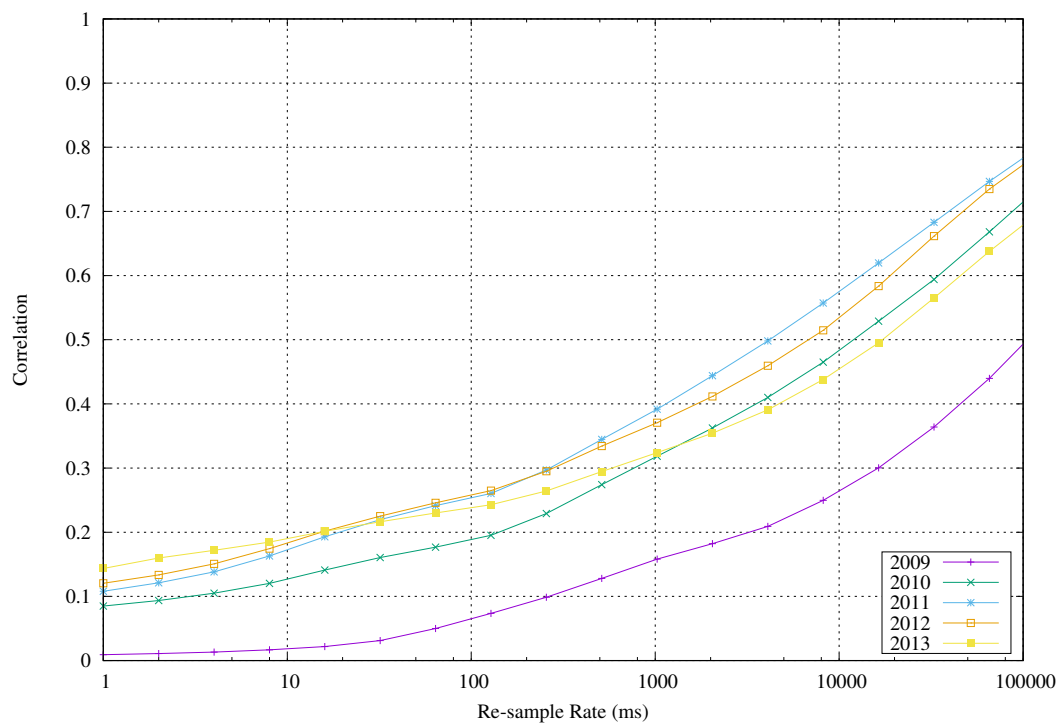


Figure 6.23: Mid-quote correlations for Telia Sonera.

6.1.6 Trade-through Volume

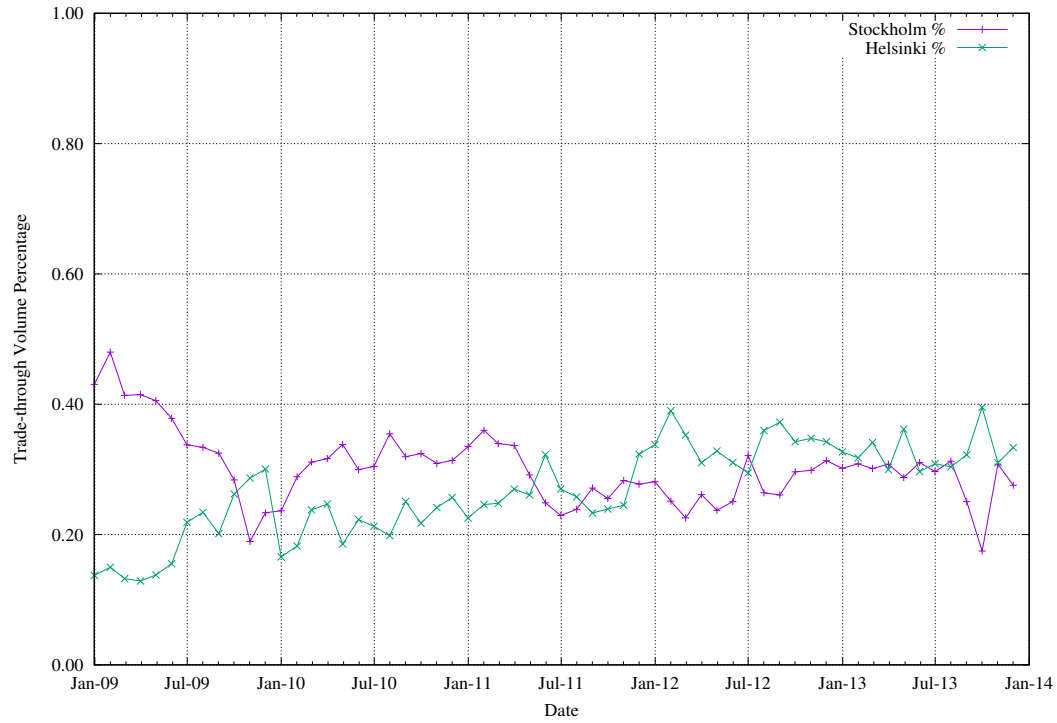


Figure 6.24: Volume of trade-throughs for Nokia as a percentage of total volume.

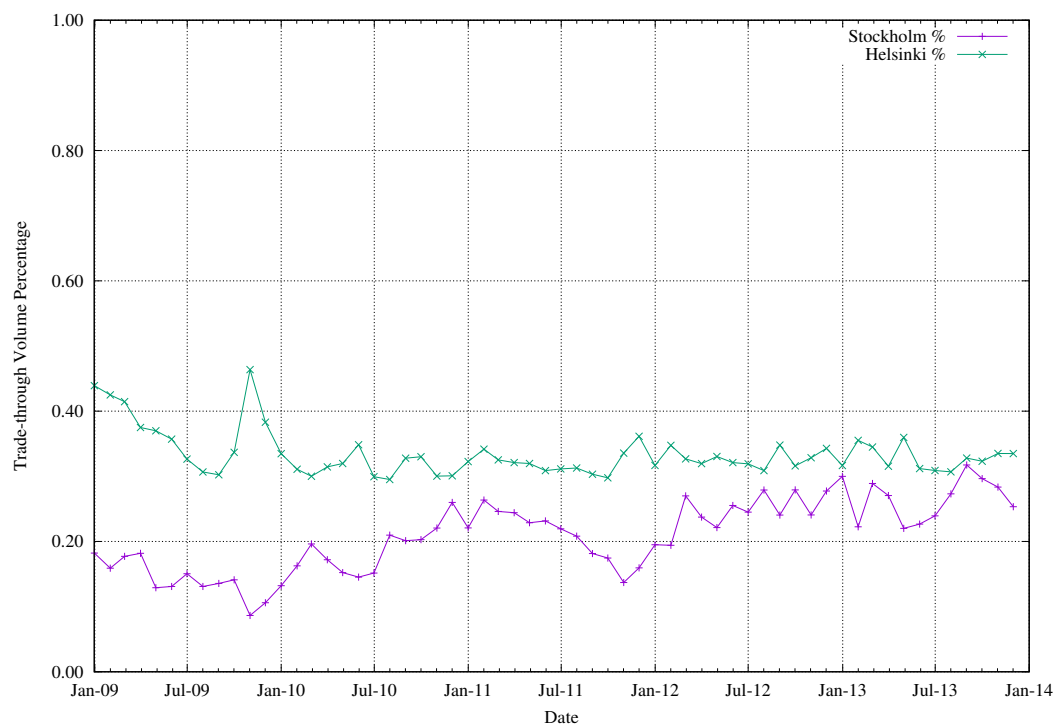


Figure 6.25: Volume of trade-throughs for Nordea as a percentage of total volume.

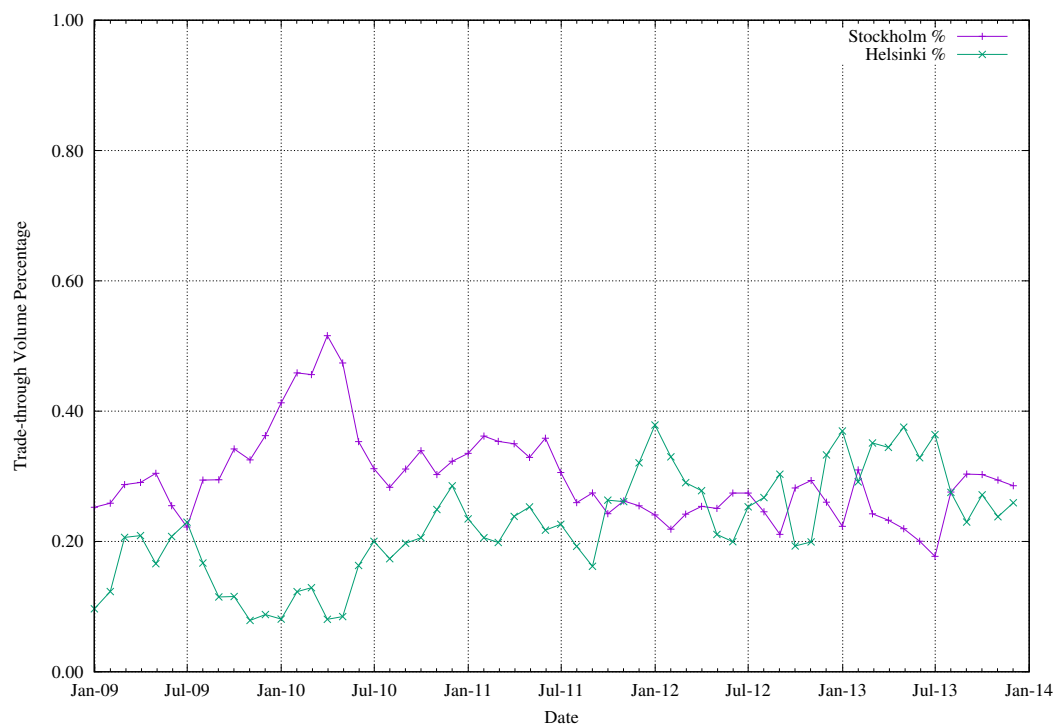


Figure 6.26: Volume of trade-throughs for Stora Enso as a percentage of total volume.

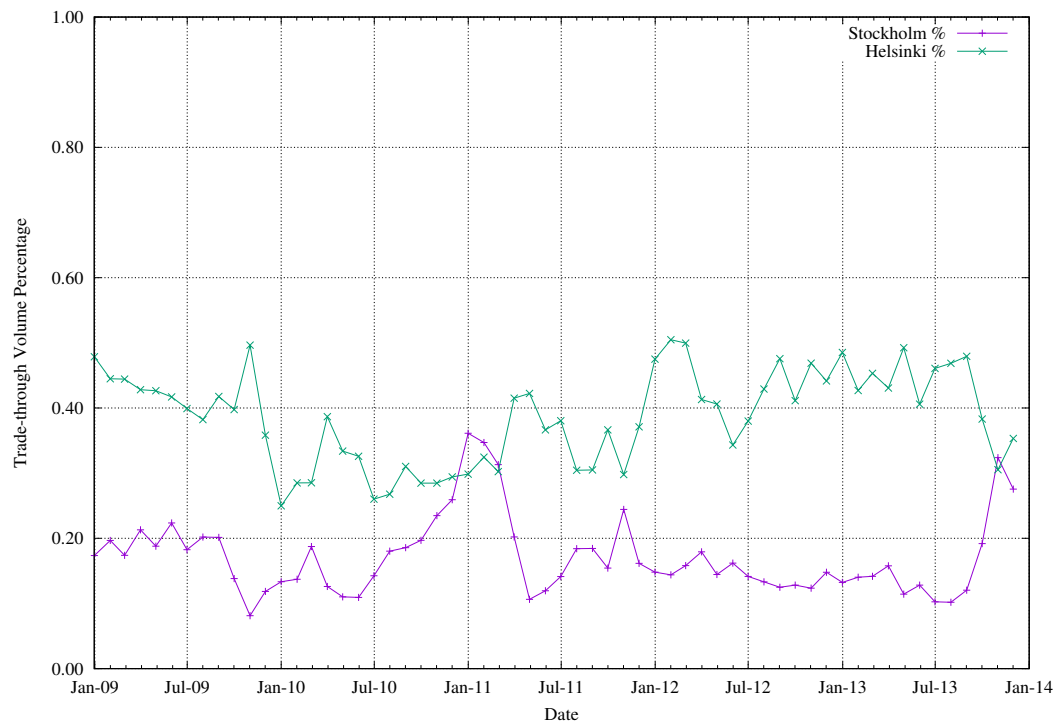


Figure 6.27: Volume of trade-throughs for Telia Sonera as a percentage of total volume.

6.1.7 Trade-through Costs

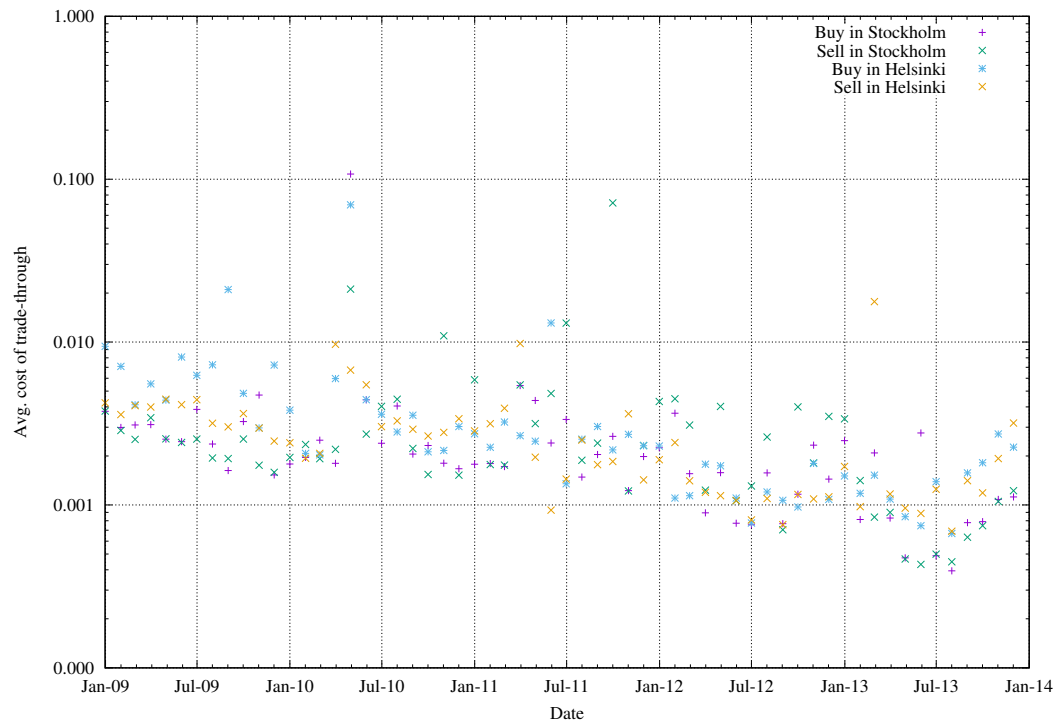


Figure 6.28: Cost of trade-throughs in Nokia.

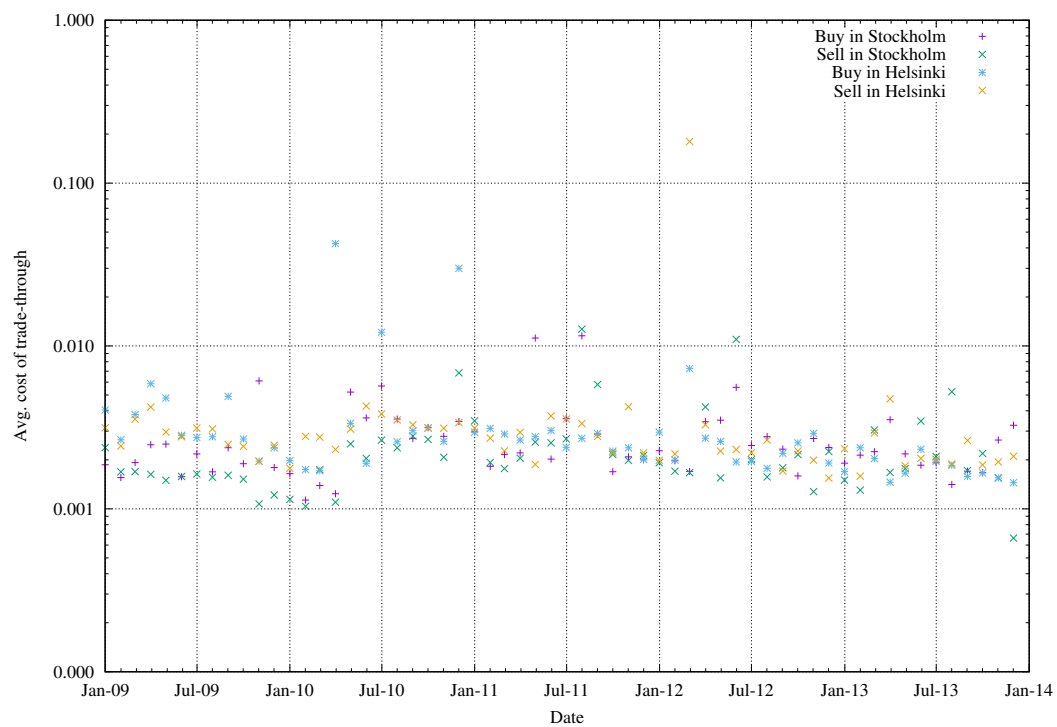


Figure 6.29: Cost of trade-throughs in Nordea Group.

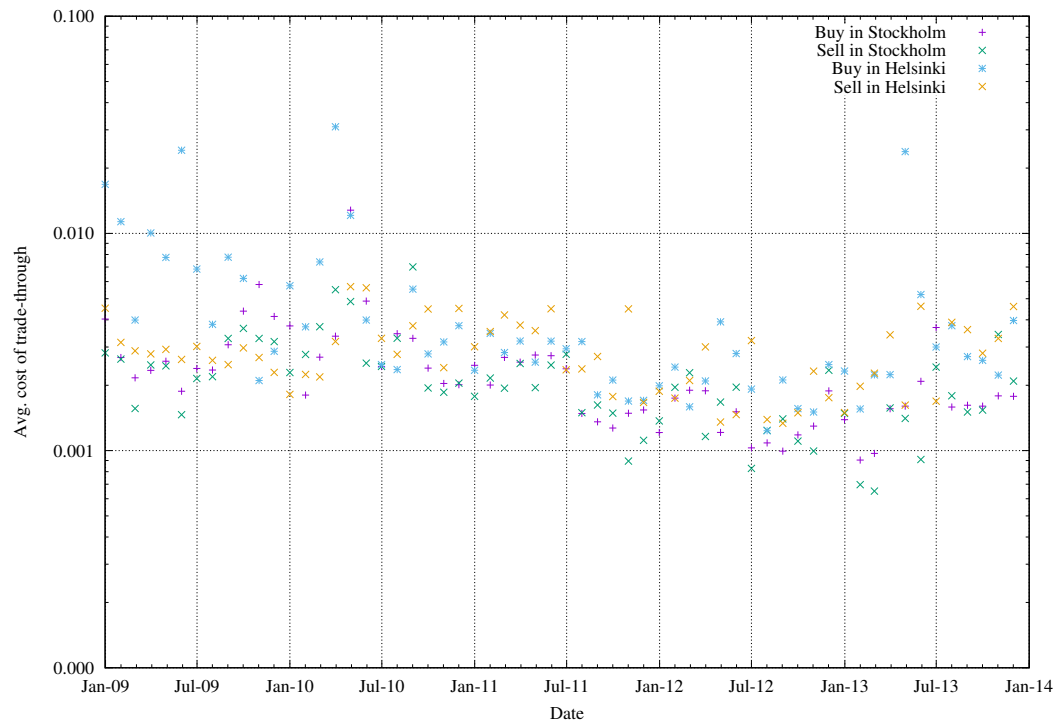


Figure 6.30: Cost of trade-throughs in Stora Enso.

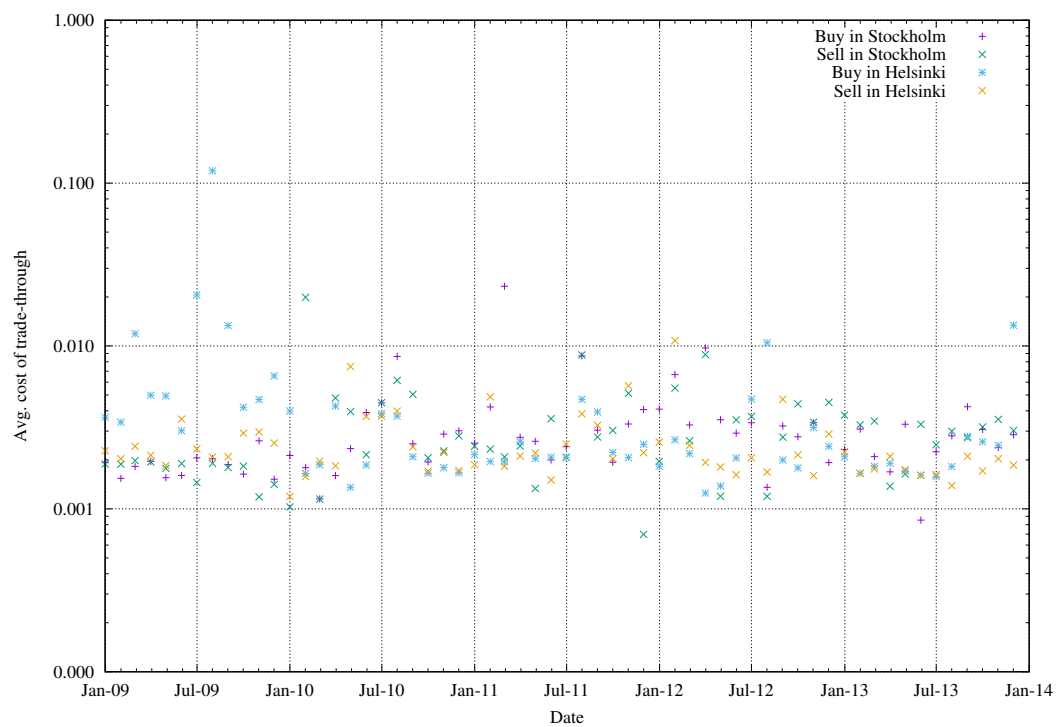


Figure 6.31: Cost of trade-throughs in Telia Sonera.

Bibliography

- [1] I. T. Frisch and H. Frank, “Computer communications: how we got where we are,” in *Proceedings of the May 19-22, 1975, national computer conference and exposition*, pp. 109–117, ACM, 1975.
- [2] N. Mills, “Nasdaq: a user-driven, real-time transaction system,” in *Proceedings of the May 16-18, 1972, spring joint computer conference*, pp. 1197–1206, ACM, 1972.
- [3] M. A. Goldstein, P. Kumar, and F. C. Graves, “Computerized and high-frequency trading,” *Financial Review*, vol. 49, no. 2, pp. 177–202, 2014.
- [4] M. J. McGowan, “The rise of computerized high frequency trading: Use and controversy,” *Duke Law & Technology Review*, vol. 16, p. 1, 2010.
- [5] W. D. Nordhaus, “The progress of computing,” 2001.
- [6] J. D. Farmer and D. Foley, “The economy needs agent-based modelling,” *Nature*, vol. 460, no. 7256, pp. 685–686, 2009.

- [7] D. Cliff and L. Northrop, “The global financial markets: an ultra-large-scale systems perspective,” in *Proceedings of the 17th Monterey conference on Large-Scale Complex IT Systems: development, operation and management*, pp. 29–70, Springer-Verlag, 2012.
- [8] S. A. Levin and A. W. Lo, “Opinion: A new approach to financial regulation,” *Proceedings of the National Academy of Sciences*, vol. 112, no. 41, pp. 12543–12544, 2015.
- [9] W. B. Arthur, *Complexity and the Economy*. Oxford University Press, 2014.
- [10] N. Ehrentreich, *Agent-based modeling: The Santa Fe Institute artificial stock market model revisited*, vol. 602. Springer Science & Business Media, 2007.
- [11] V. Darley and A. V. Outkin, *NASDAQ market simulation: insights on a major market from the science of complex adaptive systems*. World Scientific Publishing Co., Inc., 2007.
- [12] L. Harris, “Decimalization: A review of the arguments and evidence,” *Unpublished working paper, University of Southern California*, 1997.
- [13] J. Geanakoplos, R. Axtell, D. J. Farmer, P. Howitt, B. Conlee, J. Goldstein, M. Hendrey, N. M. Palmer, and C.-Y. Yang, “Getting at systemic risk via an agent-based model of the housing market,” *The American Economic Review*, vol. 102, no. 3, pp. 53–58, 2012.

- [14] E. B. Budish, P. Cramton, and J. J. Shim, “The high-frequency trading arms race: Frequent batch auctions as a market design response,” *Chicago Booth Research Paper*, no. 14-03, 2015.
- [15] E. Wah and M. P. Wellman, “Latency arbitrage, market fragmentation, and efficiency: a two-market model,” in *Proceedings of the fourteenth ACM conference on Electronic commerce*, pp. 855–872, ACM, 2013.
- [16] A. J. Menkveld and B. Z. Yueshen, “The flash crash: A cautionary tale about highly fragmented markets,” *Available at SSRN 2243520*, 2015.
- [17] M. Paddrik, R. Hayes, A. Todd, S. Yang, P. Beling, and W. Scherer, “An agent based model of the e-mini s&p 500 applied to flash crash analysis,” in *2012 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*, pp. 1–8, IEEE, 2012.
- [18] T. A. Vuorenmaa and L. Wang, “An agent-based model of the flash crash of may 6, 2010, with policy implications,” *Available at SSRN 2336772*, 2014.
- [19] S. J. Leal, M. Napoletano, A. Roventini, and G. Fagiolo, “Rock around the clock: An agent-based model of low-and high-frequency trading,” *Journal of Evolutionary Economics*, vol. 26, no. 1, pp. 49–76, 2016.
- [20] J. M. Epstein and R. Axtell, *Growing artificial societies: social science from the bottom up*. Brookings Institution Press, 1996.
- [21] V. Grimm and S. F. Railsback, “Individual-based modeling and ecology:(princeton series in theoretical and computational biology),” 2005.

- [22] L. Tesfatsion and K. L. Judd, *Handbook of computational economics: agent-based computational economics*, vol. 2. Elsevier, 2006.
- [23] S. De Marchi and S. E. Page, “Agent-based models,” *Annual Review of Political Science*, vol. 17, pp. 1–20, 2014.
- [24] J. H. Miller and S. E. Page, *Complex adaptive systems: An introduction to computational models of social life*. Princeton university press, 2009.
- [25] S. F. Railsback and V. Grimm, *Agent-based and individual-based modeling: a practical introduction*. Princeton university press, 2011.
- [26] U. Wilensky and W. Rand, *An introduction to agent-based modeling: modeling natural, social, and engineered complex systems with NetLogo*. MIT Press, 2015.
- [27] A. Namatame and S.-H. Chen, *Agent Based Modelling and Network Dynamics*. Oxford University Press, 2016.
- [28] C. M. Macal and M. J. North, “Tutorial on agent-based modelling and simulation,” *Journal of simulation*, vol. 4, no. 3, pp. 151–162, 2010.
- [29] N. Gilbert and P. Terna, “How to build and use agent-based models in social science,” *Mind & Society*, vol. 1, no. 1, pp. 57–72, 2000.
- [30] S. C. Banks, “Agent-based modeling: A revolution?,” *Proceedings of the National Academy of Sciences*, vol. 99, no. suppl 3, pp. 7199–7200, 2002.

- [31] A. P. Kirman, “Whom or what does the representative individual represent?,” *The Journal of Economic Perspectives*, vol. 6, no. 2, pp. 117–136, 1992.
- [32] A. Shaikh, *Capitalism: Competition, conflict, crises*. Oxford University Press, 2016.
- [33] A. Lehtinen and J. Kuorikoski, “Computing the perfect model: Why do economists shun simulation?,” *Philosophy of Science*, vol. 74, no. 3, pp. 304–329, 2007.
- [34] K. L. Judd, “Computational economics and economic theory: Substitutes or complements?,” *Journal of Economic Dynamics and Control*, vol. 21, no. 6, pp. 907–942, 1997.
- [35] C. W. Reynolds, “Flocks, herds and schools: A distributed behavioral model,” *ACM SIGGRAPH computer graphics*, vol. 21, no. 4, pp. 25–34, 1987.
- [36] J. Goldstein, “Emergence as a construct: History and issues,” *Emergence*, vol. 1, no. 1, pp. 49–72, 1999.
- [37] T. C. Schelling, *Micromotives and macrobehavior*. WW Norton & Company, 2006.
- [38] P. Hedström and P. Ylikoski, “Causal mechanisms in the social sciences,” *Annual Review of Sociology*, vol. 36, pp. 49–67, 2010.
- [39] R. Cont, “Empirical properties of asset returns: stylized facts and statistical issues,” *Quantitative Finance*, 2001.

- [40] A. Pagan, “The econometrics of financial markets,” *Journal of Empirical Finance*, vol. 3, no. 1, pp. 15–102, 1996.
- [41] T. Lux, “Herd behaviour, bubbles and crashes,” *The Economic Journal*, pp. 881–896, 1995.
- [42] J.-P. Bouchaud and M. Potters, “More stylized facts of financial markets: leverage effect and downside correlations,” *Physica A: Statistical Mechanics and its Applications*, vol. 299, no. 1, pp. 60–70, 2001.
- [43] S.-H. Chen, C.-L. Chang, and Y.-R. Du, “Agent-based economic models and econometrics,” *The Knowledge Engineering Review*, vol. 27, no. 02, pp. 187–219, 2012.
- [44] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz, G. Huse, *et al.*, “A standard protocol for describing individual-based and agent-based models,” *Ecological modelling*, vol. 198, no. 1, pp. 115–126, 2006.
- [45] S. F. Railsback, S. L. Lytinen, and S. K. Jackson, “Agent-based simulation platforms: Review and development recommendations,” *Simulation*, vol. 82, no. 9, pp. 609–623, 2006.
- [46] C. Nikolai and G. Madey, “Tools of the trade: A survey of various agent based modeling platforms,” *Journal of Artificial Societies and Social Simulation*, vol. 12, no. 2, p. 2, 2009.

- [47] S. Rossiter, “Simulation design: Trans-paradigm best-practice from software engineering,” *Journal of Artificial Societies & Social Simulation*, 2015.
- [48] S. N. Goodman, D. Fanelli, and J. P. Ioannidis, “What does research reproducibility mean?,” *Science translational medicine*, vol. 8, no. 341, pp. 341ps12–341ps12, 2016.
- [49] A. M. Law, W. D. Kelton, and W. D. Kelton, *Simulation modeling and analysis*, vol. 2. McGraw-Hill New York, 1991.
- [50] R. G. Sargent, “Verification and validation of simulation models,” in *Proceedings of the 37th conference on Winter simulation*, pp. 130–143, winter simulation conference, 2005.
- [51] P. L’Ecuyer and R. Simard, “Testu01: Ac library for empirical testing of random number generators,” *ACM Transactions on Mathematical Software (TOMS)*, vol. 33, no. 4, p. 22, 2007.
- [52] K. W. Comer, *Who Goes First? An Examination of the Impact of Activation on Outcome Behavior in Agent-based Models*. PhD thesis, George Mason University, 2014.
- [53] W. Radax and B. Rengs, “Timing matters: lessons from the ca literature on updating,” *arXiv preprint arXiv:1008.0941*, 2010.
- [54] B. Schönfisch and A. de Roos, “Synchronous and asynchronous updating in cellular automata,” *BioSystems*, vol. 51, no. 3, pp. 123–143, 1999.

- [55] D. Cornforth, D. G. Green, and D. Newth, “Ordered asynchronous processes in multi-agent systems,” *Physica D: Nonlinear Phenomena*, vol. 204, no. 1, pp. 70–82, 2005.
- [56] X. Gabaix, P. Gopikrishnan, V. Plerou, and H. E. Stanley, “A theory of power-law distributions in financial market fluctuations,” *Nature*, vol. 423, no. 6937, pp. 267–270, 2003.
- [57] B. Toth, Y. Lemperiere, C. Deremble, J. De Lataillade, J. Kockelkoren, and J.-P. Bouchaud, “Anomalous price impact and the critical nature of liquidity in financial markets,” *Physical Review X*, vol. 1, no. 2, p. 021006, 2011.
- [58] I. Mastromatteo, B. Toth, and J.-P. Bouchaud, “Agent-based models for latent liquidity and concave price impact,” *Physical Review E*, vol. 89, no. 4, p. 042805, 2014.
- [59] A. M. Law, *Simulation Modeling and Analysis*. McGraw-Hill, 2007.
- [60] C. Alexopoulos, “A comprehensive review of methods for simulation output analysis,” in *Proceedings of the 38th conference on Winter simulation*, pp. 168–178, Winter Simulation Conference, 2006.
- [61] C. Alexopoulos and D. Goldsman, “To batch or not to batch?,” *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 14, no. 1, pp. 76–114, 2004.
- [62] P. Heidelberger and P. D. Welch, “Simulation run length control in the presence of an initial transient,” *Operations Research*, vol. 31, no. 6, pp. 1109–1144, 1983.

- [63] S. J. Russell, P. Norvig, J. F. Canny, J. M. Malik, and D. D. Edwards, *Artificial intelligence: a modern approach*, vol. 2. Prentice hall Upper Saddle River, 2003.
- [64] B. LeBaron *et al.*, “A builder’s guide to agent-based financial markets,” *Quantitative Finance*, vol. 1, no. 2, pp. 254–261, 2001.
- [65] L. Harris, *Trading and Exchanges: Market Microstructure for Practitioners*. Oxford University Press, 2002.
- [66] J. H. Holland and J. H. Miller, “Artificial adaptive agents in economic theory,” *The American Economic Review*, vol. 81, no. 2, pp. 365–370, 1991.
- [67] S. G. Winter and R. R. Nelson, “An evolutionary theory of economic change,” *University of Illinois at Urbana-Champaign’s Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*, 1982.
- [68] C. Schinckus, “Methodological comment on econophysics review i and ii: statistical econophysics and agent-based econophysics,” *Quantitative Finance*, vol. 12, no. 8, pp. 1189–1192, 2012.
- [69] C. H. Hommes, “Heterogeneous agent models in economics and finance,” *Handbook of computational economics*, vol. 2, pp. 1109–1186, 2006.
- [70] H. P. Boswijk, C. H. Hommes, and S. Manzan, “Behavioral heterogeneity in stock prices,” *Journal of Economic dynamics and control*, vol. 31, no. 6, pp. 1938–1970, 2007.

- [71] R. G. Palmer, W. B. Arthur, J. H. Holland, B. LeBaron, and P. Tayler, “Artificial economic life: a simple model of a stockmarket,” *Physica D: Nonlinear Phenomena*, vol. 75, no. 1, pp. 264–274, 1994.
- [72] M. Andrews and R. Prager, “Genetic programming for the acquisition of double auction market strategies,” *Advances in genetic programming*, vol. 1, pp. 355–368, 1994.
- [73] R. Palmer, W. B. Arthur, J. H. Holland, and B. LeBaron, “An artificial stock market,” *Artificial Life and Robotics*, vol. 3, no. 1, pp. 27–31, 1999.
- [74] S. Martinez-Jaramillo and E. P. Tsang, “An heterogeneous, endogenous and coevolutionary gp-based financial market,” *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 1, pp. 33–55, 2009.
- [75] S. Buti, B. Rindi, and I. M. Werner, “Diving into dark pools,” *Charles A. Dice Center Working Paper*, no. 2010-10, 2010.
- [76] R. H. Day and W. Huang, “Bulls, bears and market sheep,” *Journal of Economic Behavior & Organization*, vol. 14, no. 3, pp. 299–329, 1990.
- [77] J. D. Farmer, “Market force, ecology and evolution,” *Industrial and Corporate Change*, vol. 11, no. 5, pp. 895–953, 2002.
- [78] M. D. Gould, M. A. Porter, S. Williams, M. McDonald, D. J. Fenn, and S. D. Howlson, “Limit order books,” *Quantitative Finance*, vol. 13, no. 11, 2013.
- [79] O. A. Lamont and R. H. Thaler, “Anomalies: The law of one price in financial markets,” *Journal of Economic Perspectives*, vol. 17, no. 4, pp. 191–202, 2003.

- [80] P. Hedstrom and P. Ylikoski, “Causal mechanisms in the social sciences,” *Annual Review of Sociology*, vol. 36, pp. 49–67, 2012.
- [81] D. Ladley, “Zero intelligence in economics and finance,” *The Knowledge Engineering Review*, vol. 27, no. 2, pp. 273–286, 2012.
- [82] J. D. Farmer, P. Patelli, and I. I. Zovko, “The predictive power of zero intelligence in financial markets,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 6, p. 2254–2259, 2005.
- [83] B. Toth, Y. Lemperiere, C. Deremble, J. de Lataillade, J. Kockelkoren, and J.-P. Bouchaud, “Anomalous price impact and the critical nature of liquidity in financial markets,” *Physical Review X*, vol. 1, 2011.
- [84] J. Gatheral and R. Oomen, “Zero-intelligence realized variance estimation,” *Finance and Stochastics*, vol. 14, no. 2, pp. 249–283, 2010.
- [85] E. Wah and M. P. Wellman, “Latency arbitrage, market fragmentation, and efficiency: a two-market model,” in *EC ’13 Proceedings of the fourteenth ACM conference on Electronic Commerce*, (New York, New York), pp. 855–872, ACM, 2013.
- [86] Luis A. Aguilar, “U.s. equity market structure: Making our markets work better for investors,” 2015.
- [87] K. Schacht, C. Cronin, J. C. Allen, and R. Preece, “Market microstructure: The impact of fragmentation under the markets in financial instruments directive,” tech. rep., CFA Institute, 2009.

- [88] B. Bias and T. Foucault, “HFT and Market Quality,” *Bankers, Markets & Investors*, vol. 128, pp. 5–19, 2014.
- [89] M. O’Hara, “High frequency market microstructure,” *Journal of Financial Economics*, vol. 116, no. 2, pp. 257–270, 2015.
- [90] C. Cao, E. Ghysels, and F. Hatheway, “Price discovery without trading: Evidence from the nasdaq preopening,” *Journal of Finance*, vol. 55, no. 3, pp. 1339–1365, 2000.
- [91] A. V. Shkilko, B. Van Ness, and R. Van Ness, “Locked and crossed markets on NASDAQ and the NYSE,” *Journal of Financial Markets*, vol. 11, no. 3, pp. 308–337, 2008.
- [92] R. Garvey and A. Murphy, “Crossed markets: Arbitrage opportunities in nasdaq stocks,” *Journal of Alternative Investments*, vol. 9, no. 2, pp. 46–58, 2006.
- [93] S. Sunder and D. Gode, “Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality,” *The Journal of Political Economy*, vol. 101, no. 1, pp. 119–137, 1993.
- [94] S.-H. Chen, “Varieties of agents in agent-based computational economics: A historical and an interdisciplinary perspective,” *Journal of Economic Dynamics and Control*, vol. 36, pp. 1–25, 2012.
- [95] B. Edmonds and S. Moss, “From kiss to kids – an ‘anti-simplistic’ modelling approach,” in *Multi-Agent and Multi-Agent-Based Simulation* (P. Davidsson,

- B. Logan, and K. Takadama, eds.), pp. 130–144, Berlin: Springer Berlin Heidelberg, 2005.
- [96] S. Maslov, “Simple model of a limit order-driven market,” *Physica A: Statistical Mechanics and its Applications*, vol. 278, no. 3-4, pp. 571–578, 2000.
 - [97] M. G. Daniels, J. D. Farmer, L. Gillemot, G. Iori, and E. Smith, “Quantitative model of price diffusion and market friction based on trading as a mechanistic random process,” *Physical Review Letters*, vol. 90, no. 10, 2003.
 - [98] E. Smith, J. D. Farmer, L. Gillemot, and S. Krishnamurthy, “Statistical theory of the continuous double auction,” *Quantitative Finance*, vol. 3, no. 6, pp. 481–514, 2003.
 - [99] S. Mike and J. D. Farmer, “An empirical behavioral model of liquidity and volatility,” *Journal of Economic Dynamics and Control*, vol. 32, no. 1, p. 200–234, 2008.
 - [100] R. Cont, S. Stoikov, and R. Talreja, “A Stochastic Model for Order Book Dynamics,” *Operations Research*, vol. 58, no. 3, pp. 549–563, 2010.
 - [101] F. Abergel and A. Jedidi, “A mathematical approach to order book modeling,” *International Journal of Theoretical and Applied Finance*, vol. 16, no. 5, 2013.
 - [102] H. Banks, A. Broido, B. Canter, K. Gayvert, S. Hu, M. Joyner, and K. Link, “Simulation algorithms for continuous time markov chain models,” in *Proceedings of SiMCRT 2011* (F. Kojima, F. Kobayashi, and H. Nakamoto, eds.), pp. 3–18, IOS Press, Aug. 2012.

- [103] D. E. Knuth, *Art of Computer Programming, Volume 2: Seminumerical Algorithms*. Addison-Wesley Professional, 1997.
- [104] A. Shleifer and R. Vishny, “The limits of arbitrage,” *Journal of Finance*, vol. 52, no. 1, pp. 35–55, 1997.
- [105] D. Gromb and D. Vayanos, “Limits of arbitrage: The state of the theory,” Working Paper 15821, National Bureau of Economic Research, March 2010.
- [106] J. Gipson, “Electronics threaten nyse monopoly,” *Washington Post*, Nov 1980.
- [107] “Rule 611 of Regulation NMS,” 2015. Security and Exchange Commission memorandum available at <https://www.sec.gov/spotlight/emsac/memo-rule-611-regulation-nms.pdf>.
- [108] “Best Execution under MiFID,” Tech. Rep. Ref: CESR/07-320, The Committee of European Securities Regulators, 2007.
- [109] B. Ende, P. Gomber, and M. Lutat, *Smart Order Routing Technology in the New European Equity Trading Landscape*, pp. 197–209. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.
- [110] J. Hasbrouck, “One security, many markets: Determining the contributions to price discovery,” *Journal of Finance*, vol. 50, no. 4, pp. 1175–99, 1995.
- [111] C. S. Eun and S. Sabherwal, “Cross-border listings and price discovery: Evidence from u.s.-listed canadian stocks,” *Journal of Finance*, vol. 58, no. 2, pp. 549–576, 2003.

- [112] A. De Jong, L. Rosenthal, and M. A. Van Dijk, “The risk and return of arbitrage in dual-listed companies*,” *Review of Finance*, vol. 13, no. 3, pp. 495–520, 2009.
- [113] L. Gagnon and G. A. Karolyi, “Multi-market trading and arbitrage,” *Journal of Financial Economics*, vol. 97, no. 1, pp. 53 – 80, 2010.
- [114] A. J. Menkveld and M. A. Zoican, “Need for speed? exchange latency and liquidity,” 2015.
- [115] E. Pagnotta and T. Philippon, “Competing on speed,” tech. rep., National Bureau of Economic Research, 2011.
- [116] A. J. Menkveld, “The economics of high-frequency trading: Taking stock,” *Annual Review of Financial Economics*, vol. 8, no. 1, 2016.
- [117] J. Grant, “Nasdaq renews European efforts,” *Financial Times*, 2010.
- [118] NASDAQ OMX Group, “NASDAQ OMX launches new roadmap for the Nordic equity market,” *GlobalNewswire*, 2008.
- [119] “The Firm Hand of Nordic Trading,” *FTSE Global Markets*, vol. 41, May 2010.
- [120] NASDAQ OMX Group, “NASDAQ OMX and EMCF launch central counterparty clearing in the Nordics,” *GlobalNewswire*, 2009.
- [121] NASDAQ OMX Group, “NASDAQ OMX Nordic introduces harmonized tick sizes,” *GlobalNewswire*, 2009.

- [122] F. Kjell, “Tick size reductions effect on provisions of liquidity: Research on the stockholm stock exchange,” 2015.
- [123] T. W. Epps, “Comovements in Stock Prices in the Very Short Run,” *Journal of the American Statistical Associations*, vol. 74, no. 366, pp. 291–298, 1979.
- [124] E. Budish, P. Cramton, and J. Shim, “Editor’s choice the high-frequency trading arms race: Frequent batch auctions as a market design response,” *The Quarterly Journal of Economics*, vol. 130, no. 4, pp. 1547–1621, 2015.
- [125] M. Scholes and J. Williams, “Estimating betas from nonsynchronous data,” *Journal of financial economics*, vol. 5, no. 3, pp. 309–327, 1977.
- [126] A. W. Lo and A. C. MacKinlay, “An econometric analysis of nonsynchronous trading,” *Journal of Econometrics*, vol. 45, no. 1-2, pp. 181–211, 1990.
- [127] R. Reno, “A closer look at the epps effect,” *International Journal of Theoretical and Applied Finance*, vol. 06, no. 01, pp. 87–102, 2003.
- [128] B. Tóth and J. Kertész, “On the origin of the epps effect,” *Physica A: Statistical Mechanics and its Applications*, vol. 383, no. 1, pp. 54 – 58, 2007.
- [129] B. Tóth and J. Kertész, “The epps effect revisited,” *Quantitative Finance*, vol. 9, no. 7, pp. 793–802, 2009.
- [130] A. Storkenmaier and M. Wagener, “Do we need a European “National Market System”? Competition, arbitrage, and suboptimal executions,” 2011.

- [131] A. Kohler and R. von Wyss, “Fragmentation in European Equity Markets and Market Quality – Evidence from the Analysis of Trade-Throughs,” Working Papers on Finance 1210, University of St. Gallen, School of Finance, Oct. 2012.