NORGES BANK

DOCTORAL DISSERTATIONS IN ECONOMICS

NO. 5

Essays on the microstructure of stock markets: Empirical evidence from trading arrangements without dealer intermediation

Randi Næs

The series ***Doctoral Dissertations in Economics*** is issued by Norges Bank (Central Bank of Norway) and contains doctoral works produced by employees at Norges Bank, primarily as part of the Bank's research activities.

The series is available free of charge from:
Norges Bank, Subscription Service
E-mail: central.bank@norges-bank.no
Postal address:
PO Box 1179 Sentrum
N – 0107 Oslo
Norway

Previous publication of doctoral dissertations from Norges Bank:
(Prior to 2002, doctoral dissertations were published in the series *Occasional Papers*)

Ragnar Nymoen: Empirical Modelling of Wage-Price Inflation and Employment using Norwegian Quarterly Data. (*Occasional Papers* no. 18, Oslo 1991)
Bent Vale: Four essays on Asymmetric Information in Credit Markets. (*Occasional Papers* no. 20, Oslo 1992)
Birger Vikøren: Interest Rate Differential, Exchange Rate Expectations and Capital Mobility: Norwegian Evidence. (*Occasional Papers* no. 21, Oslo 1994)
Gunnvald Grønvik: Bankregulering og bankatferd 1975 – 1991. (*Occasional Papers* no. 22, Oslo 1994)
Ingunn M. Lønning: Controlling Inflation by use of the Interest Rate: The Critical Role of Fiscal Policy and Government Debt. (*Occasional Papers* no. 25, Oslo 1997)
Tom Bernhardsen: Interest Rate Differentials, Capital Mobility and Devaluation Expectations: Evidence from European Countries. (*Occasional Papers* no. 27, Oslo 1998)
Øistein Røisland: Rules and Institutional Arrangements for Monetary Policy. (*Occasional Papers* no. 29, Oslo 2000)
Quaisar Farooq Akram: State Dependent Effects in Labour and Foreign Exchange Markets. (*Doctoral Dissertations in Economics* no. 1, Oslo 2002)
Kai Leitemo: Inflation targeting and monetary policy. (*Doctoral Dissertations in Economics* No. 2, Oslo 2002)
Carl Andreas Claussen: Four essays on the political economy of economic reform (*Doctoral Dissertations in Economics* No. 3, Oslo 2003)
Johannes A. Skjeltorp: Trading in equity markets. A study of individual, institutional and corporate trading decisions (*Doctoral Dissertations in Economics* No. 4, Oslo 2005)

 The author may be contacted at:
Randi.Nas@Norges-Bank.no

This book was printed and bound at Lobo Media AS, Oslo.
The text was written in 10-point Times by the author.

Essays on the microstructure of stock markets: Empirical evidence from trading arrangements without dealer intermediation

by Randi Næs

A dissertation submitted for the degree of dr. oecon. at the Norwegian School of Economics and Business Administration

To Paula

Contents

List of Tables

List of Figures

Preface

Understanding the competitive environment for securities trading is of great importance for Norges Bank's overall responsibility for financial stability. As a manager of the Norwegian Government Petroleum Fund, the bank also has a more specific interest in the properties of different security trading arrangements.

A notable feature of today's stock markets is the rapid development of electronic, order-driven trading arrangements. In this thesis, Randi Næs provides empirical evidence on the properties of two such arrangements. Firstly, trading through crossing networks is analyzed based on transactions data from the Petroleum Fund. Crossing networks are shown to be cost effective alternatives to regular exchanges, however, crossing may also induce adverse selection costs not accounted for in traditional cost measures. Secondly, limit order trading is analyzed based on transactions data from the Oslo Stock Exchange. A main result is that the order book contains relevant information about contemporaneous trading volume and price volatility in the market.

This thesis is part of the author's dr. oecon. exam at the Norwegian School of Economics and Business Administration, Department of Finance and Management Science. The thesis was defended on 2 September 2004, and Norges Bank is pleased to make this dissertation available to a wider audience by publishing it as Doctoral Dissertation in Economics No. 5.

Oslo, 21 July 2005

Research Department
Øyvind Eitrheim
Director

Acknowledgements

I have written this thesis while being at the Research Department of Norges Bank. I am very grateful for this opportunity, and would like to thank Norges Bank for its generous support of my studies. I would also like to thank the Department of Finance and Management Science at the Norwegian School of Economics and Business Administration for its great hospitality on my numerous stays in Bergen over the years.

I would like to express my warmest gratitude to Steinar Ekern, who has been my supervisor and a true source of inspiration ever since I left Bergen in 1991, and whose constant encouragement is the reason why I ever started on a PhD. I also gratefully acknowledge Bernt Arne Ødegaard for arousing my interest in market microstructure through his lectures and later through our joint work, and for always being prepared to send me some more data and give constructive comments. A special thank to Johannes A. Skjeltorp for being a great friend and an inspiring co-author, and for making the working on this thesis a rewarding social activity. Over the years I have become indebted to many persons at Norges Bank for their helpful comments, suggestions and encouraging remarks. Especially, I owe many thanks to Bent Vale, Eilev S Jansen, Øyvind Eitrheim, Kjersti-Gro Lindquist, and Dagfinn Rime at the Research Department. I am very grateful to James Angel at Georgetown University for his valuable support in the process of data preparation, and for many inspiring discussions and suggestions. I would also like to thank Jonas Andersson at the Norwegian School of Economics and Business Administration for valuable comments and suggestions.

Finally, I would like to thank Knut Kjær and Yngve Slyngstad at the Investment Management Department of Norges Bank and Sverre Lilleng and Thomas Borchgrevink at the Oslo Stock Exchange for providing the data samples used in the thesis, and for being very helpful during the stage of data preparation.

Oslo, April 2004

Randi Næs

Chapter 1

Introduction and overview

The securities business is an emotional business with a high degree of entertainment value for at least some of the participants.
George J.W. Goodman, The Money Game, 1968.

1 Introduction

According to the efficient market hypothesis, a well-functioning financial market will always make sure that security prices reflect all available information about security values. Since its conception in the 1960s, this hypothesis has been a most important proposition in financial economics.[1] During the last two decades, however, the earlier strong empirical success of the hypothesis as well as its theoretical foundations have been challenged.

Market microstructure constitutes one bulk of this literature.[2] Market microstructure studies how prices may diverge from (or converge to) long-term equilibrium values due to strategic behavior among the participants in the trading process.[3] The standard asset pricing theory abstracts from such mechanisms: rational investors with symmetric information about the expected risk and return characteristics of the market make prices adjust immediately as a result of new information. The trading process itself is just a "black box" with no effects on price discovery. While this may be a plausible assumption for some kind of news, other types of news are likely to be dispersed and not immediately available to all investors in aggregated form.[4] The market microstructure theory shows that, in the case of dispersed information, one can no longer as-

[1] For a review of the efficient market hypothesis, see Fama (1970) and Fama (1991).

[2] Behavioral finance constitutes another strand of this literature, see for example Schleifer (2000). Behavioral finance challenges the assumptions of investor rationality and unlimited arbitrage. For a defense of these concepts, see Rubinstein (2001).

[3] Market participants include investors, borrowers, hedgers, gamblers, brokers, dealers, and market makers.

[4] Evidence of the existence of dispersed news is given in French and Roll (1986) who document empirically that asset prices are much more volatile during exchange trading hours than during non-trading hours. This phenomenon cannot be reconciled with a standard asset pricing model, unless there is a systematic tendency for price relevant information to arrive during normal business hours only.

sume that the trading process is a black box with no effects on price determination and trading volume. Madhavan (2000) categorizes the topics covered by the market microstructure literature into four groups.[5]

1. *Studies of price formation and price discovery (looking inside the black box).* Work within this group includes theoretical and empirical studies of the determinants of transactions costs, and dynamic models of the dissemination of information into prices. Overall, our understanding of the role of dealers in price formation has been considerably enhanced by this research. Empirically, the key factors that cause price movements are identified as inventory and asymmetric information.

2. *Market structure and design issues (how do different rules affect the black box).* Madhavan (2000) separates market structure dimensions into degree of continuity, reliance on market makers, degree of automation, price discovery, order forms permitted, protocols, and transparency. There is a large degree of hetereogeneity along these dimensions in existing trading systems. Different trading arrangements emerge to serve the needs of different groups of traders, and there is no single "best" market structure for everyone. The resulting market fragmentation may reduce competition within each market center, but can enhance competition across trading arrangements.

3. *Information and disclosure issues (how does revealing the workings of the black box affect trader behaviour).* Transparency is found to have significant effects on the price discovery process. Theoretical literature emphasizes the benefits of transparency, but empirical and experimental results are ambiguous. Some disclosure is found to be better than no disclosure, but more transparency is not always better because traders may be unwilling to reveal their trading intentions. Changes in transparency will in general benefit one group of traders at the expense of others.

4. *The interface of market microstructure with other areas of finance (do models of the black box "matter").* This very important field of study is still evolving. Studies have been conducted in different areas including; asset pricing (e.g., liquidity as a priced factor in expected returns), corporate finance (e.g., pricing of IPOs and explanations for stock splits), and international finance (e.g., explaining cross border order flow and exchange rate movements).

Following O'Hara (2003), the two main functions of a market are to provide liquidity and price discovery. Liquidity is a measure of traders' possibility to trade and includes several dimensions known as width, depth, immediacy and resiliency.[6] Price discovery is the process by which new

[5]Some recent survey articles on market microstructure literature are Keim and Madhavan (1998), Madhavan (2000) and Biais et al. (2002). A detailed survey of theoretical market microstructure is provided in the book by O'Hara (1995), and a comprehensive and practical oriented overview of trading and the organization of markets is provided in the book by Harris (2003).

[6]The four dimensions of liquidity is suggested by Harris (1990). Width measures the cost per share of liquidity,

information is incorporated into asset prices. Liquidity provision and price discovery are both closely related to the execution system of a market, that is the procedures adopted for the matching of buyers and sellers.[7] Execution systems can be quote-driven, order-driven, brokered, or some combination of the three. In a pure quote-driven market, traders trade indirectly with each other through one ore more dealers.[8] Dealers quote prices and negotiate all trades. Bond and currency markets are typically quote-driven. In a pure order driven market, buyers and sellers trade with each other without the intermediation of dealers. Instead, trades are arranged using rules for order precedence and pricing. Oral auctions, single price auctions, continuous electronic auctions, and crossing networks are all examples of order-driven markets. In a brokered market, brokers actively search to match buyers and sellers.[9] The markets for large blocks of stocks and bonds are examples of brokered markets. The main US equity exchanges, NYSE and NASDAQ, are mixtures of both quote-driven, order-driven, and brokered markets. The NYSE is essentially order-driven but requires its dealers to offer liquidity if no one else will do so. The NASDAQ requires its dealers to display and sometimes execute public limit orders. In both markets large brokers sometimes arrange block trades.

Even though order-driven trading systems are quite common, the tasks of providing liquidity and securing continuity in the market place have traditionally been carried out by dealers.[10] For this reason, the overall bulk of the market microstructure research is done assuming some form of dealer intermediation. In their survey of the market microstructure of stock markets, Biais et al. (2002) divide the existing literature into two generations depending on the assumed nature of competition between dealers/market makers.[11] The first generation assumes a fully competitive environment, while the second relaxes this assumption and discusses cases where liquidity is provided by strategic agents who exploit some form of market power.[12]

The main rationale for the strong role of dealers in the securities business has been that it is efficient to delegate market monitoring to a small subset of the participating agents. While this argument is still valid, technological advances in electronic communications in recent years, which allow buyers and sellers to provide liquidity themselves at low costs, are posing a threat to the vital role of dealer intermediation.[13]

This thesis contributes to the literature by providing empirical evidence on the properties of

depth displays the ability of the market to absorb a series of trades, immediacy describes how fast a trade for a given number of shares can be executed, and resiliency expresses how fast the price reverts to its "true" value after order flow imbalances caused by non-informed trading has moved prices temporarily away from the "true" level.

[7]The discussion in the following three paragraphs is based on Harris (2003), pages 92-96 and 112.

[8]Harris (2003) defines dealers as "profit-motivated traders who allow other traders to trade when they want to trade". In some equity markets, dealers may be known as market makers or specialists.

[9]This structure is suitable in illiquid markets where dealers do not want to quote prices.

[10]Dealers often trade in order-driven markets, and in some markets they actually provide most of the liquidity. However, as long as the dealers cannot choose their clients and must arrange all trades according to the market's trading rules, the market is still known as order-driven.

[11]Cohen et al. (1981) develop a model of the bid-ask spread in a market with many competing limit order traders.

[12]Research within the second generation points to the benefits of allowing investors to compete to supply liquidity themselves.

[13]Madhavan (2000) points out that the new development is partly driven by a practical need for automation to handle the increasingly high volumes of trading.

two order-driven trading arrangements for equities which do not rely on dealer intermediation; an internal crossing network and an electronic limit order-driven stock market.

The thesis consists of four essays. The two first essays are based on transactions data from a large institutional investor's acquisition of a US stock portfolio, which involved extensive use of *crossing*. The results from these studies are of interest for two strands of the literature. Firstly, they add to the literature on measuring the costs of trading equities (group 1 above), in particular for institutional investors. We show how to utilize the information from an exact investment strategy to detect cost components which are generally extremely hard to measure. Secondly, the results add to the literature on the structure of securities markets by showing some evidence of the properties of crossing networks (group 3 above). In general, there is little available evidence on the use of crossing networks. Moreover, our data is from a particular type of network which is a "blacker box" than the few which are actually studied elsewhere.

The two last essays are based on a comprehensive sample of transactions data from an *electronic limit order-driven* stock market. There is steady growth in the availability of detailed transactions data for stock markets, however, the fact that we are able to rebuild the whole limit order book at any point in time is still quite remarkable. In electronic limit order markets, most orders are first submitted to the market as limit orders. Thus, in contrast to dealer markets, the buying and selling interest for a security can be deducted from the order-book. Our main contribution in the third essay is to document several relationships between the shape of the order book and the volume-volatility relation found in most financial markets. The order book shape is measured as the average of the elasticities of the supply and demand schedules in the book. An interesting interpretation of our findings is that this measure proxies for dispersion of beliefs about asset values among the liquidity providers. If so, our results support models where investor heterogeneity intensifies the volume-volatility relation. This result adds to the literature on price formation and price discovery (group 1 above).

In the last essay, liquidity provision in a limit order market is related to data on ownership structure. Hence, in this essay we focus on the link between market microstructure and corporate finance (group 4 above). The main contribution from the essay comes from studying this link based on considerably more comprehensive data on ownership, and under a different trading arrangement, than previous studies. In addition, we look at the Granger causality between ownership variables and liquidity measures, which is an important and little addressed topic in the existing literature.

Some distinguishing characteristics of crossing networks and electronic limit order markets are discussed in section 2. The essays on trading through crossing networks are summarized in section 3 and section 4, respectively. The essay on the volume-volatility relation is reviewed in section 5, and the essay on liquidity and ownership structure is shortly presented in section 6.

2 Two trading methods without dealer intermediation

2.1 Crossing networks

Crossing networks have existed for quite a while, however, their current high popularity is a relatively recent phenomenon often attributed to the growth of institutional investors, and the needs of these investors to obtain large quantities of equities, for example for index tracking purposes.[14]

In contrast to the other order-driven markets, a crossing network is not an auction market.[15] Traders submit *unpriced* orders to buy or sell given quantities. Quantities are then matched (or crossed), according to some given algorithm. There is an ex ante agreement that the price in the cross will be some observable price determined outside the crossing network, such as the closing price at the NYSE or NASDAQ at the day of the cross.[16] The crossing price is *not* observable at the time of order submission. When an order is submitted to a crossing network, there is therefore uncertainty both as to whether the order will be filled and at what price. Hence two distinctive characteristics of a crossing network are: (i) price discovery is taking place elsewhere, and (ii) the network cannot guarantee immediate execution, or execution at all. A third characteristic of crossing networks is low transparency. The networks are typically completely confidential and anonymous systems.

Crossing can be performed in different ways. In *external crossing systems*, orders are matched electronically at pre-specified times or time intervals. In the US, there are three external crossing systems: POSIT, Instinet Crossing, and NYSE crossing sessions.[17] But crossing is also performed regularly in more exclusive arenas, by members of specific parties. Every business day, several large fund managers and custodians run their in-house computers to match buy and sell orders from portfolios under their management.[18] This is called *internal crossing*. This type of crossing is largely a "black box", even to the institutional investors doing the trading.

Two important market microstructure issues with respect to crossing networks are (i) cost effectiveness and (ii) the nature of competition between this trading venue and the traditional trading systems.

Cost effectiveness Crossing networks are designed to be cost effective. The absence of dealer intermediation implies no spread costs, and the absence of price discovery implies no direct price impact costs. Moreover, Keim and Madhavan (1998) report that crossing commissions are

[14]One of the major crossing networks in the US, Instinet Corporation, was founded back in 1969.

[15]"In an auction market, the trading rules formalize the process by which buyers seek the lowest available prices and sellers seek the highest available prices" (Harris (2003), page 94.)

[16]This property is sometimes referred to as the "derivative pricing rule".

[17]POSIT is the largest crossing network. POSIT performs seven daily matches at the prevailing bid-ask midpoint in the stock's primary market. Instinet Crossing matches orders once a day in the afternoon. Listed stocks are matched at the NYSE closing price and the NASDAQ stocks are matched at the closing inside quote midpoint. NYSE performs an after hours crossing service each day at the NYSE closing price.

[18]Ruyter (1999)

substantially lower than commissions charged by brokers on exchanges.[19] On the other hand, there may be opportunity costs related to non-execution. Moreover, the low transparency attracts informed investors, which may induce adverse selection costs. Another potential problem is price manipulation.

Based on a large sample of institutional trades, Conrad et al. (2003) show evidence that crossing networks have a distinct cost advantage to exchange based trading. However, the data sample is not suited to study potential costs related to non-execution and adverse selection.

Competition The recent success of low cost crossing networks has raised the issue of harmful and unfair competition. The primary markets complain that crossing networks "cream skim" the orders originating from uninformed investors, and that the resulting reduction in primary market liquidity will harm all market participants in the end. Competition is claimed to be unfair because the crossing networks do not compensate the primary markets for the use of their price discovery process.

There is a growing academic literature on intermarket competition. In the case of competitive liquidity supply, Chowdhry and Nanda (1991) indicate a tendency for the "winner takes most" type of outcome. Both informed traders and liquidity traders will flock to the largest exchange, informed traders because it is easier to "hide" the bigger the liquidity order flow, and liquidity traders because the more *other* liquidity traders that are present, the lower their costs. In the case where strategic liquidity suppliers have some form of market power, they may find it optimal to provide liquidity outside the primary market.[20] The arguments used by the representatives of the primary markets are supported by models emphasizing asymmetric information, such as Easley et al. (1996). Here, off-market trading is explained as driven by "cream skimming" of orders originating from uninformed traders, and is most likely for small orders in liquid securities. In contrast, reputation models such as Seppi (1990) explain the benefits of trading outside exchanges in terms of the ability to screen out informed investors and permit mutually advantageous trades off-market. If so, trading outside exchanges will be largely complementary to exchange trading, and the off-market trading will be more likely in large orders, especially in less liquid stocks. Hendershott and Mendelson (2000) develop a theoretical model where different types of heterogeneous liquidity traders and informed traders choose between a competitive dealer market and a crossing network. The model is quite complex and provides few unambiguous implications. The effects on dealers' spread from the introduction of a crossing network are shown to depend on the types of traders in the market. With no informed trading, the negative "cream-skimming" effect dominates the positive effect of attracting new order flow. With short-lived information, the low order-submission costs ensure that the introduction of a crossing network will always raise the dealers' spread. Under most circumstances, the crossing network will also increase dealer spreads when information is long-lived. However, this can be offset if the crossing network manages to attract sufficient new liquidity traders.

[19] Crossing commissions are usually below 2 cents a share

[20] See Biais et al. (2002), page 35.

Fong et al. (1999) use data from the Australian stock exchange (ASX) to study the competition between exchanges and alternative trading mechanisms including upstairs markets, after-hours trading and electronic crossing networks, and find support for an asymmetric information explanation.[21] Conrad et al. (2003) find similar results for both the NYSE and the NASDAQ: alternative trading systems do provide significant competition for order flows from institutional investors.

2.2 Electronic limit order markets

Electronic limit order markets have emerged during the last 15 years.[22] Today, it is the typical way to arrange stock trading outside the US.[23] Electronic limit order books are closely akin to the electronic communication networks (ECNs) which are becoming increasingly popular in the US markets.[24] In contrast to the typical market structure in countries with limit order driven stock exchanges, ECNs coexist with trading in the same stock on other trading venues.[25]

In electronic limit order market traders only negotiate with each other by submitting and cancelling orders. Buyers and sellers are matched according to “order precedence rules”, and the resulting trades are priced according to “trade pricing rules”. Typically, the markets are arranged with a initial single price call auction and subsequent continuous electronic auctions. Single price auctions use the “uniform pricing rule”, while continuous auctions use the “discriminatory price rule”. The uniform pricing rule matches all trades at the same market clearing price.[26] Under the discriminatory pricing rule, prices are determined by the limit prices of the standing orders.[27] In making their order placement decisions, limit order traders must balance the trade off between a high probability of execution and the “goodness” of the price in the case of execution. The uniform pricing rule gives traders an incentive to issue more aggressively priced orders than the discriminatory pricing rule.

As was first suggested by Copeland and Galai (1983), limit order traders writes free out of the money options to the market.[28] Hence limit order traders are exposed to the risk of being “picked off” when the market valuation is changing.[29] Biais et al. (2002) note that this adverse selection component is somewhat different than the adverse selection component typically ana-

[21] In Harris (2003), the upstairs market is defined as a market serving “large traders who cannot convey credible information about their trading motives and intentions to traders in the regular market.”

[22] For a review of a pioneering automated limit order trading system in the US, see the description of Automated Trading Desk (ATD) in Whitcomb (2003).

[23] Limit order driven stock exchanges are found in, for example, Toronto, Tel Aviv, Paris, Frankfurt, Stockholm, and Oslo.

[24] According to Bloomfield et al. (2003), ECN’s “such as Island, Instinet, and Archipelago use an electronic order book structure to trade as much as 45 % of the volume on Nasdaq”.

[25] ECNs provide a higher degree of anonymity and speed of execution than regular exchange trading. For a recent examination of the competition among ECNs and NASDAQ market makers, see Barclay et al. (2003).

[26] Hence, single price auctions maximize the trading volume (sets supply equal demand).

[27] The rule is called discriminatory because large traders can discriminate among smaller traders based on their willingness to trade.

[28] A limit buy order (sell order) is equal to a free out of the money put option (call option) with a strike price equal to the limit bid price (limit ask price).

[29] The adverse selection cost related to limit orders is also discussed in Whitcomb (2003).

lyzed in theoretical models of dealer-based markets.[30]

Parallel to its growing popularity, there has also been a growing academic interest in electronic trading methods. Researchers want to know more about traders' order placement strategies and their contribution to liquidity and price formation. Moreover, transactions data from electronic markets are generally more detailed and comprehensive than the transactions data available from dealer or hybrid markets, enabling more detailed studies. Finally, Biais et al. (1995) argue that the special properties of electronic limit order markets make them a particularly appropriate testing ground: they rely solely on order placements, they strictly enforce time and price priority, they generate data which fully capture the order flow and execution process, and they provide traders with a high degree of transparency.

The literature on limit order trading is still evolving, but existing work is yet substantial. The existing literature deal mainly with (i) the competition between limit order markets and dealer markets, and (ii) the relative performance of limit orders versus market orders, including the probability of limit order execution. This literature is summarized, when relevant, throughout the thesis.

3 Essay 1: Equity trading I: To cross or not to cross

Written with Bernt Arne Ødegaard

Conrad et al. (2003) show evidence that alternative trading systems have a distinct cost advantage to exchange based trading. In fact, the magnitude of the cost advantage leads Conrad et al. (2003) to question whether these are sustainable equilibria. Large costs due to non-trading or adverse selection could explain the price differences. The prime contribution of our results is to actually shed some light on this issue, with particular application to a passive crossing system.

We analyze a particular order placement strategy followed by a large institutional investor, the Norwegian Government Petroleum Fund. The data sample includes all orders from the establishment of a US equity portfolio worth USD1.76 billion over a 6-month period from January 1998 to June 1998. The portfolio was tracking the US part of the FTSE All World index, which consists of about 500 stocks, and has a very high correlation with the S&P 500 index. The order placement strategy was to first send all orders to crossing networks, and then to place the orders that could not be crossed in the primary market.

We find that the probability of execution in the internal crossing network is related to the subsequent performance of the desired stocks. This result is documented by means of an event study comparing orders that were crossed with those that had to be filled in the market. The cumulative abnormal returns (CARs) on the stocks that were purchased in the market are found to be significantly higher in the month after the trade than the CARs for the stocks that were crossed. We interpret this result as evidence of informed trading. We also show that the costs of

[30] Traditional dealer-models focus on asymmetric information about the value of the asset, which will exist even if the dealer and the informed trader move simultaneously.

trading through an internal crossing network is somewhat lower than similar costs reported for external networks.

The availability of detailed transactions data have increased in recent years, however, as noticed by Madhavan (2000), a serious problem with many tests of microstructure theories is that the data sets do not allow the researchers to ask “what if” questions. A suggested solution to this problem is to conduct laboratory or experimental studies. We know the exact investment strategy of the Fund ex ante. Our data set is therefore something like a “controlled real world experiment”. In other studies of execution costs, the investment strategy of a trader has to be estimated from the sequence of trades. This induces a selection bias in the data. It might be that the trader has decided to send the most difficult orders to brokers and the least difficult orders to crossing networks. We do not have this selection bias problem in our data set. The data set has a weakness in being from one institution only. However, we show evidence that the Fund’s trades are quite representative for large institutional investors in the US markets. Nevertheless, we should be cautious to draw general conclusions from our study about the market structure effects from crossing networks.

4 Essay 2: Equity trading II: Evidence on order submission strategies

Written with Johannes A. Skjeltorp

This essay is a natural follow up of the first essay where we (i) try to investigate the evidence of adverse selection more closely, and (ii) investigate the costs of following alternative submission strategies.

Could it be that stocks that are not supplied in crossing networks are less liquid in general? If so, these stocks might need a higher return to induce investors to hold them, and the abnormal performance of the non-crossed stocks found in the first essay might be explained (or partly explained) by a liquidity premium. On the other hand, a wider spread may also capture a higher adverse selection component, and thus a difference in liquidity between the two groups of stocks may capture the same adverse selection effect, only measured by a different proxy. We investigate these questions by calculating a whole range of liquidity and activity measures in the primary market across the groups of stocks that were supplied/not supplied in the crossing network. Our results indicate that the difference in abnormal return between the two groups of stocks may be explained by both liquidity differences and private information.

In the second part of the paper, we simulate alternative order submission strategies to assess if the Fund could have used a better strategy. Our simulated strategies are based on real historical price/volume paths of the stocks traded. This is possible to do because we know the dates when the decision to trade was made, in addition to the desired quantities. The simulated strategies are the set of equilibrium strategies in the Hendershott and Mendelson (2000) model. Our results indicate that the actual crossing strategy was inexpensive. Even though the crossed stocks were among the most liquid stocks on the NYSE, it would have been very hard to achieve lower execution costs by submitting limit orders for the same stocks on the same dates that they were first tried to be crossed.

5 Essay 3: Order book characteristics and the volume-volatility relation

Written with Johannes A. Skjeltorp

A variety of studies document that there is a positive correlation between price volatility and trading volume for most types of financial contracts. The main theoretical explanation for this phenomena is that new information about asset values acts as the driving force (or mixing variable) for both market prices and trading volume (the mixture of distribution hypothesis). Market microstructure based theoretical models predict that dispersion of beliefs about asset values will intensify the volume-volatility relation, by increasing both trading volume and volatility. The typical finding in empirical studies is that the relation can be explained by a mixture of distribution hypothesis, where the arrival rate of information is proxied by the daily number of transactions.

The objective of this essay is to broaden our knowledge about the volume-volatility relation in an electronic limit order market. Since the demand and supply schedules in a limit order book represent the reservation prices of the liquidity suppliers in the market, it is interesting to study whether the order book contains additional information about the volume-volatility relation.

A systematic negative relation between the average slope of the order book and the price volatility is documented. Similarly, we find a significant and robust negative relationship between our slope measure and the daily number of trades. We also show that the slope of the book provides different information depending on what fraction of the book we use in the calculation. An interesting interpretation of our findings is that the shape of the book is related to the dispersion of beliefs among liquidity suppliers, i.e., steep slopes indicate that there is a high degree of agreement among investors about security values, while gentle slopes indicate greater disagreement. If so, our finding that there is increased trading activity when slopes are more gentle, supports models where heterogeneity among investors contributes to the volume-volatility relation.

6 Essay 4: Ownership structure and market liquidity

This essay is an empirical study of the relationship between ownership structure and market liquidity in the Norwegian stock market. Our data sample on ownership structure is considerably more detailed than the data sets used in comparable studies. Moreover, we are not aware of anyone who has been able to analyze this issue in using a panel regression approach.

Theoretical models relate the efficiency of a particular ownership structure to its ability to cope with the conflicts of interest raised by the separation of ownership and control. A central variable for this ability is informational advantage. Large and direct owners have an informational advantage relative to small and indirect owners which make them better suited to monitor firm managers. Similarly, the informational advantage of corporate insiders reduces the need for monitoring. The market microstructure theory predicts that informational asymmetries will be reflected in market liquidity through higher implicit costs of trading. Thus, the positive performance effect from monitoring is predicted to be mitigated by costs related to reduced market liquidity.

In line with the theoretical predictions, owner concentration and insider holdings are found to increase spread costs. Owner concentration also increases an estimate of adverse selection costs. Institutional ownership is not found to have significant effects on market liquidity, while foreign ownership is concentrated in stocks with low spreads and high depth. In general, we do not find a one-way Granger causality from ownership to liquidity.

A interesting extension of the problem studied in this essay relates to the research issue of whether illiquidity is priced in firms' cost of capital.[31] Some very preliminary calculations indicate that there is a positive relationship between transaction costs and return in the Norwegian market. This essay is the first part of a larger project aimed at linking liquidity effects from ownership variables to the firms' performance and capital costs.

[31] Different strands in the literature address this issue. Holmstrøm and Tirole (2001) develop an asset pricing model where asset prices are driven by a corporate demand for liquidity. The basic idea is that risk-neutral firms are willing to pay a premium on assets that provide liquidity in states of liquidity shortage. Several papers study whether liquidity is a common factor in stock returns, see for example Amihud (2002) and Hasbrouck and Seppi (2001). O'Hara (2003) examines the implications for asset prices of transactions costs of liquidity and the risks of price discovery. The Bodie et al. (2002) textbook generalizes the CAPM expected return - beta relationship to include a liquidity effect.

Bibliography

Amihud, Y. (2002). Illiquidity and stock returns: cross section and time series effects. Working paper, New York University.

Barclay, M. J., Hendershott, T., and McCormick, D. T. (2003). Competition among trading venues: Information and trading on electronic communications networks. *Journal of Finance*, 58:2637–2665.

Biais, B., Glosten, L., and Spatt, C. (2002). The microstructure of stock markets. CEPR Discussion Paper No. 3288.

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

Bloomfield, R., O'Hara, M., and Saar, G. (2003). The "make or take" decision in an electronic market: Evidence on the evolution of liquidity. Working Paper, Cornell University. Forthcoming in the Journal of Financial Economics.

Bodie, Z., Kane, A., and Marcus, A. (2002). *Investments 5th ed.* McGraw-Hill.

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

Cohen, K. J., Maier, S. F., Schwartz, R. A., and Whitcomb, D. K. (1981). Transaction costs, order placement strategy, and existence of the bid-ask spread. *Journal of Political Economy*, 89:287–305.

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

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

Easley, D., Kiefer, N. M., and O'Hara, M. (1996). Cream-skimming or profit sharing? The curious role of purchased order flow. *Journal of Finance*, 51:811–833.

Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25:383–417.

Fama, E. (1991). Efficient capital markets: II. *Journal of Finance*, 46:1575–1617.

Fong, K., Madhavan, A., and Swan, P. (1999). Why do securities markets fragment? Working Paper, University of Southern California.

French, K. and Roll, R. (1986). Stock return variances. the arrival of information and the reaction of traders. *Journal of Financial Economics*, 17:5–26.

Harris, L. (1990). Liquidity, trading rules, and electronic trading systems. *New York University Salomon Center Monograph Series in Finance and Economics*, 4.

Harris, L. (2003). *Trading and Exchanges. Market Microstructure for Practitioners*. Oxford University Press.

Hasbrouck, J. and Seppi, D. (2001). Common factors in prices, order flows, and liquidity. *Journal of Financial Economics*, 59:383–411.

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

Holmstrøm, B. and Tirole, J. (2001). LAPM: A Liquidity-Based Asset Pricing Model. *Journal of Finance*, 56:1837–1867.

Keim, D. B. and Madhavan, A. (1998). The cost of institutional equity trades. *Financial Analysts Journal*, 54(4):50–69.

Madhavan, A. (2000). Market microstructure: a survey. *Journal of Financial Markets*, 3:205–258.

O'Hara, M. (1995). *Market microstructure theory*. Blackwell Publishers Inc.

O'Hara, M. (2003). Presidental address: liquidity and price discovery. *Journal of Finance*, 58:1335–1354.

Rubinstein, M. (2001). Rational markets: Yes or No? The Affirmative Case. *Financial Analysts Journal*, 57(3):15–29.

Ruyter, T. (1999). The dinosaur, the rock, and the hard place. Plansponsor Magazine. March.

Schleifer, A. (2000). *Inefficient markets. An introduction to behavioral finance*. Oxford University Press.

Seppi, D. (1990). Equilibrium block trading and asymmetric information. *Journal of Finance*, 45:73–94.

Whitcomb, D. K. (2003). Applied market microstructure. *Journal of Applied Finance*, 13:77–80.

Chapter 2

Equity trading by institutional investors: To cross or not to cross

Written with Bernt Arne Ødegaard

Abstract

This paper looks at an increasingly popular market structure for trading of US equities, the crossing network. A crossing network is a satellite trading place: it uses prices from the primary market, and merely matches on quantity. Since trades are passively matched among participants, there is no guarantee of execution in a crossing network. Our prime result shows that the probability of execution in crossing networks is related to stock performance. Using data from a large institutional investor, we show that over the month following an attempt at crossing, there is a two percent difference in abnormal return between stocks that the investor managed to cross and stocks that had to be bought in the market. We interpret our results as evidence of informed trading in the most liquid stocks in the world, the constituents of the S&P500. We may, however, still be looking at an equilibrium outcome. Our result is consistent with the implications of a theoretical model of intermarket competition, in which the adverse selection in a crossing network is offset by the lower cost of the trades that one manages to cross.

1 Introduction

Traders desiring to trade US equities are faced with a bewildering menu of choices for executing their trades. The traditional bastions of trading, the NYSE and NASDAQ, are challenged by scores of alternative trading systems, the emergence of which hinges upon the possibilities provided by universal and costless electronic communications.

In this paper, we show some evidence of the properties of a particular alternative trading system, the *crossing network*. In a crossing network, participants submit unpriced orders to buy and sell stocks. Quantities are then matched at an agreed price, which typically is based on prices in a primary market, such as the NYSE closing price. The crossing price is not observable at the time of order submission. When an order is submitted to a crossing network, there is therefore uncertainty both as to whether the order will be filled and, in the event, at what price. In external crossing systems, orders are matched electronically at pre-specified times or time intervals. [1] But crossing is also performed regularly in more exclusive arenas, by members of specific parties. Every business day, several large fund managers and custodians run their in-house computers to match buy and sell orders from portfolios under their management.[2] This is called internal crossing. This type of crossing is largely a "black box," even to the institutional investors doing the trading. The evidence from these marketplaces is scant, due to the obvious lack of interest in making it public from the participants' point of view. Most of our transactions data are from an internal crossing network. Our paper is to our knowledge the first to show some evidence of trading in one of these "blacker boxes."

A stated purpose of crossing networks has been to reduce transaction costs for large traders. Current research in Conrad et al. (2003), comparing exchange based trading with alternative trading systems, suggests that they have indeed been successful in this respect: a distinct cost advantage to alternative trading system is documented. In fact, the magnitude of the differences lead Conrad et al. (2003) to question whether these are sustainable as equilibria, or whether there are issues not accounted for in their analysis. With respect to crossing networks, they point to the risk of not being able to fill orders. A large cost of non-trading can potentially explain the price differences between measured costs of crossing and exchange based trading. Keim and Madhavan (1998) and Hendershott and Mendelson (2000) suggest that the explicit cost advantage of crossing networks might be mitigated by a presence of informed traders in the networks.

The cost associated with non-trading is notoriously hard to measure. Even though the concept has been acknowledged as a potential factor ever since its introduction into the literature by Treynor (1981) and Perold (1988), little empirical work has been done. In their survey ar-

[1] In the US, there are three external crossing systems: POSIT, Instinet Crossing, and NYSE crossing sessions. POSIT is the largest crossing network, and performs seven daily matches at the prevailing bid-ask midpoint in the stock's primary market. Instinet Crossing matches orders once a day in the afternoon. Listed stocks are matched at the NYSE closing price and the NASDAQ stocks are matched at the closing inside quote midpoint. NYSE performs an after hours crossing service each day at the NYSE closing price.

[2] Ruyter (1999)

ticle on institutional trading costs, Keim and Madhavan (1998) explain this as primarily due to data limitations. The prime contribution of our results is to actually shed some light on this issue, with particular application to a passive crossing system. We show that in our case the cost of non-trading is substantial and likely to be information-based, supporting the hypothesis of Conrad et al. (2003).

The reason for our ability to make these statements is peculiarities of our data. Our data set includes all orders from the establishment of a US equity portfolio worth USD 1.76 billion over a 6-month period from January 1998 to June 1998. The portfolio was tracking the US part of the FTSE All World index, which consists of about 500 stocks, and has a very high correlation with the S&P 500 index. The data set has two unique features. First, it contains information on the investors' exact motivation for trading, the order submission strategy, the timing of the orders, as well as when the orders were actually filled and at what prices. Hence, it is close to a "controlled experiment".[3] The second feature is related to the actual implementation of the order submission strategy: stocks were first attempted crossed, and then, if they could not be crossed, bought in the market. This feature is what allows us to get at the issue of non-trading

Consider the case of an uninformed equity buyer submitting an order to a crossing network. This buyer may be facing two types of informed traders. In one case the informed trader knows the current price is "too low" (the stock is undervalued). The informed will then be buying. Because there is no price mechanism in the crossing network, matching can be assumed to take place on a random basis. The informed buyers will add to the crowd of uninformed buyers, reducing the probability of execution for an uninformed buyer. Alternatively, the informed trader may know the price is "too high" (the stock is overvalued). The informed trader will then be selling, increasing the probability that any buy order is filled. Hence, in the presence of informed traders, execution probability will be affected by expected stock performance. Since crossing networks have no price mechanism, execution probability is all that *can* be affected. Thus, the presence of informed traders in the crossing network will result in (ex post) performance differences in the stocks an uninformed trader acquires or does not acquire in the crossing network. The stocks not obtained in the crossing network will tend to perform better than the stocks purchased through the crossing network, which is what we see in the data.

This result is documented by means of an event study comparing orders that were crossed with those that had to be filled in the market. The cumulative abnormal returns (CARs) on the stocks that were purchased in the market are found to be significantly higher in the month after the trade than the CARs for the stocks that were crossed. Thus, the stocks that could not be purchased through the crossing network tended to be ones which did "better" than market expectations.

Based on the standard cost measure used in the current empirical literature on transaction costs, we also show that the investor in our study ended up with lower total execution costs

[3] In most other studies, the order submission strategy of a trader has to be inferred from the sequence of orders. This induces a selection bias in the data: it might be that the trader has decided to send the most difficult orders to brokers and the least difficult orders to crossing networks.

than are suggested for alternative order types in other studies (Keim and Madhavan (1998) and Conrad et al. (2003)). Our cost estimates for internal crossing are somewhat lower than similar costs reported for external crossing.

Section 2 looks at cost measurement and determinants of transaction costs. Section 3 describes the data sample and discusses the Fund's trading strategy. In section 4, we provide estimates of conditional cost components, and investigate the presence of adverse selection costs. In section 5, we investigate the robustness of our cost determinants by means of a regression analysis. In section 6 we discuss whether our data sample is representative. Section 7 concludes the paper.

2 Transaction costs in equity markets

The task of measuring total transaction costs associated with different trading venues is challenging and requires detailed information on the entire order submission process. In this section, we discuss empirical measures of transaction costs in general and measures of transaction costs in a crossing network in particular. Finally, we discuss the effects from the intensified competition from crossing networks on the overall cost situation and market structure in the equity markets.

2.1 The components of transaction costs

Using the classification in Keim and Madhavan (1998), total trading costs can be split into explicit and implicit costs of trading,

$$\text{Total cost} = \underbrace{\text{Broker commissions}}_{\text{Explicit Cost}} + \underbrace{\text{Spread} + \text{Price impact} + \text{Opportunity cost}}_{\text{Implicit Cost}} \tag{2.1}$$

The explicit costs are the actual out-of-pocket costs of trading, such as broker fees.[4] The implicit costs consist of spread costs, price impact costs and opportunity costs. The implicit cost components are much harder to quantify than the explicit costs, however, according to the survey article by Keim and Madhavan (1998), there is little doubt that they are economically significant. The *spread* is set to cover the specialists' costs. However, sometimes prices have to move away from the bid-ask spread to enable an order to be executed. The resulting *price impact cost* may be decomposed into a temporary component reflecting the liquidity cost of the trade, and a permanent component reflecting possible new information. Both the spread cost and the information cost are related to the adverse selection problem studied in most of the theoretical market

[4] In a study by Keim and Madhavan (1997), the reported average commission is 0.2% of trade value. Jones (2000) provides evidence of a significant lowering of commissions in the US equity markets over the last decade.

microstructure literature. There is always a risk that a given order is informed, and this risk is presumably larger for large orders. In theory, the total price impact of a trade can easily be computed if one knows what the price of the stock would have been if the trade had not occurred. In practice, this so-called "unperturbed" price is not observable. A common empirical measure of the price impact is the deviation between the transaction price and a proxy for the unperturbed price. This measure captures one-half of the bid-ask spread plus the price impact.[5] *Opportunity costs* are related to the investor not being able to accurately achieve his or her desired portfolio. First, some orders may be delayed, during which time the market price may move in an undesirable direction. Such costs are important for informed investors who need timely execution to capture the value of their information. Passive traders who break up their orders to reduce price impact may also incur such "timing costs".[6] Second, some orders are only partially filled or not executed at all.[7]

Treynor (1981) has proposed a theoretical measure of the total cost of trading which incorporates all the mentioned cost components, including the opportunity cost of not trading. This measure, which Perold (1988) called the "implementation shortfall," is defined as the difference in performance between the portfolio of actual trades and a matching "paper" portfolio where the stock returns are computed assuming that the trades were executed at the prices prevailing on the dates of the decision to trade. In addition to capturing the relevant cost components, the implementation shortfall overcomes the problem of measuring costs on an individual trade basis when the order consists of a package of sub-trades. In recent empirical studies on transaction costs, the implementation shortfall is estimated as follows,

$$\text{Implementation shortfall} = \underbrace{\frac{\text{commission per share}}{P_d}}_{\text{Explicit cost}} + \underbrace{\frac{P^a}{P_d} - 1}_{\text{Implicit cost}} \tag{2.2}$$

where P^a is the average price of all the executed trades in the order and P_d is an estimate of an unperturbed price, typically the closing price of the stock on the day before the decision to trade.[8]

[5] Keim and Madhavan (1996) show that the choice of a benchmark price makes a large difference to the estimated price impact. Using data on block trades for one institutional investor, they find that the average price impact for a seller-initiated transaction varies from -4.3% to -10.2% when the unperturbed price is defined as the previous day's close and the price three weeks before the trade, respectively. This result strongly suggests that the unperturbed price for block trades should be defined as the date on which the decision to trade was made.

[6] Wagner and Edwards (1993) define timing costs as the costs of seeking liquidity (the price movements between the initial submission to the trade desk and the exposure of that order to the broker.)

[7] Wagner and Edwards (1993) give two reasons for such non-execution. The trader cannot locate the shares or the price has moved out of the range he or she is willing to pay. Keim and Madhavan (1998) points out that there is mixed evidence on the importance of opportunity costs. The high rates of order completion found in Keim and Madhavan (1995), suggest that the opportunity costs of failing to execute are low. Wagner and Edwards (1993), however, find significant opportunity costs for a sample of institutional managers.

[8] See Keim and Madhavan (1998) and Conrad et al. (2003).

2.2 Transaction costs in a crossing network

The recent success of crossing networks is often attributed to reduced transaction costs. There are three distinctive characteristics of a crossing network which make the cost situation for crossed orders different from the cost situation for orders traded in the regular market. First, passive matching of orders makes traditional broker/dealer services unnecessary. Second, the price discovery is taking place elsewhere, i.e. there is some primary market from which prices are derived. Third, the network do not guarantee execution.

Passive order matching without the presence of brokers yields low explicit costs. Keim and Madhavan (1998) report that crossing commissions are usually below 2 cents a share. This is substantially lower than commissions charged by brokers on exchanges. The crossing participants are not using dealers to provide liquidity. Hence there are no spread costs. Furthermore, because the crossing price is set independently of the characteristics or the crossed orders, there are no direct price impact costs. There may however be an "implicit" price impact if the existence of a large crossing order is known to participants in the primary market. On the other hand, the risk of non-execution suggests that both timing costs and other forms of opportunity costs from failure to execute may be significant for crossed orders. Moreover, the anonymity provided by most networks makes crossing attractive to informed traders. Uninformed liquidity traders who use crossing networks to reduce explicit and implicit trading costs might therefore incur costs related to adverse selection. Note that, while adverse selection costs for exchange traded orders will be included in the implicit cost component of the implementation shortfall cost, this is not so for crossed orders. As will be discussed further below, the lack of a price mechanism in the network imply that the presence of informed traders can only affect the probability of getting an order executed.

The total cost for an order sent to a crossing network is conditional on whether the crossing attempt is successful or not. Assuming that the investor goes to the regular market if he or she is not able to cross, the ex ante cost situation for a crossed order can be summarized as follows,

$$E[\text{Crossing cost}] = p(\text{cross})E\left[\text{Total cost}|\text{cross}\right] + (1 - p(\text{cross}))\,E\left[\text{Total cost}|\text{market}\right]. \quad (2.3)$$

With a given probability, $p(\text{cross})$, the order is executed on the crossing network. If crossed, it will have a cost,

$$\text{Total cost}|\text{cross} = \underbrace{\text{Crossing commission}}_{\text{Explicit cost}} + \underbrace{\text{Implicit price impact}}_{\text{Implicit cost}} \quad (2.4)$$

If the order is not crossed, it is sent to the market and will have a cost as determined by equation 2.1.

Let us now turn to the problem of detecting adverse selection costs in the crossing network. We argue that the presence of informed trading in the network, will affect the subsequent performance of the stocks offered in the crossing network. To see this, suppose first that there are only liquidity traders in the network. The probability of getting an order crossed should then be a result of the random idiosyncratic preferences of these liquidity traders, and therefore completely unrelated to the subsequent performance of the desired stocks. Next, suppose that there are also informed traders in the crossing network. Consider the submission of a buy order and look at the situation where an order is not crossed. Non-execution may have one of two reasons:

1. There did not happen to be any liquidity traders who wanted to sell these stocks on these particular dates, nor did anyone have private information on these stocks indicating that they should be sold.

2. The trader was "crowded out" by other traders who wanted to buy the stocks. The other traders could have been either liquidity traders or informed traders.

Let us also assume that informed traders with (relatively) long lived information will use an opportunistic strategy of trying crossing first and then trading in the main market if the crossing attempt fails.[9] Our trader then knows that the informed traders are most likely on the same side of the market: they are also buying. If the informed traders are selling, case 1 is ruled out, and the only reason that the trader did not get a cross is that he was "crowded out" by other liquidity traders also wanting to buy. The presence of informed traders on his side will make case 1 less likely and tend to "crowd out" his trades from any liquidity traders wanting to sell in case 2. The opposite argument applies to stocks where our liquidity trader *did* manage to buy in the cross. Then it is more likely that any informed traders were on the opposite side of the trade. More specifically, suppose that the trader is facing informed traders who know that the current price is "too high," the stock is overvalued. These traders will try to sell, increasing the probability of a cross. On the other hand, suppose he is facing informed traders who know the price is "too low." These traders will try to buy, decreasing the probability of a cross. In symbols, let subscript "l" denote pure liquidity trading and subscript "l,i" denote the presence of both liquidity traders and informed traders, and let superscripts "$+$" indicate positive information (undervalued stock) and superscript "$-$" indicate negative information (overvalued stock). Thus, for an uninformed investor, we will have that

$$p(\text{cross})_{l,i}^{-} > p(\text{cross})_{l} > p(\text{cross})_{l,i}^{+} \tag{2.5}$$

The above discussion is rather informal, but gives the gist of the argument.[10] Applying these arguments about execution probability, if we look at the *difference* in ex post performance between the stocks a liquidity trader managed to buy in a crossing network and those that did not get crossed, the non-crossed stocks are likely to perform better than the crossed stocks.

[9] This is shown to be a feature of the informed's strategy in Hendershott and Mendelson (2000).
[10] The same result is shown more formally on page 2085 of Hendershott and Mendelson (2000).

If there are traders present who have information about whether a stock is overvalued (undervalued), the expected conditional costs are higher (lower) than they would otherwise be, both in the cross and the market. However, since informed traders also affect execution probability ($p(\text{cross})$ in equation (2.3)), there is no unambiguous prediction about how the presence of informed traders will affect expected total execution costs ($E[\text{Crossing cost}]$ in equation (2.3)). Thus, in the case of informed trading in the crossing network, the optimal choice of order submission strategy for an uninformed trader will depend on the cost advantage of the crossing network relative to the main market, and the magnitude of adverse selection costs in the crossing network relative to the main market.

2.3 Competition from crossing networks

The introduction of alternative trading systems has intensified the competition for order flow faced by the traditional markets. A positive effect has been an overall reduction in transaction costs over the recent past. On the other hand, because crossing networks do not contribute to price discovery, concerns have been raised that competition from this trading method eventually will have negative effects in the form of reduced liquidity in primary markets.

There is a large body of literature in the market microstructure literature comparing different exchanges,[11] however, this work is mainly concerned with settings with multiple opportunities for price discovery. There is little work concerned with venues for trading where there is no price discovery, as discussed here. Some support for the notion that competition from crossing networks may have negative effects can be found in Mendelson (1987). Mendelson shows that market fragmentation has both costs (in the form of low liquidity and high volatility) and benefits (in the form of better price signals). Because crossing networks do not contribute to price discovery, the potential benefits from better price signals are lost and only the potential costs from low liquidity and high volatility are left.

Fong et al. (1999) use detailed data from the Australian stock exchange (ASX) to study the competition between exchanges and alternative trading mechanisms such as upstairs markets, after-hours trading and electronic crossing markets. Two different explanations for why traders may decide to trade outside exchanges are summarized in their paper. Models emphasizing asymmetric information, such as Easley et al. (1996), explain off-market trading as driven by "cream skimming" of orders originating from uninformed traders. If so, trading outside exchanges will be competing directly with the primary market, and will be more likely for small orders in liquid securities. In contrast, reputation models (Seppi (1990)) explain the benefits of trading outside exchanges in terms of the ability to screen out informed investors and permit mutually advantageous trades off-market. If so, trading outside exchanges will be largely complementary to exchange trading, and will be more likely for large orders, especially in less liquid stocks. Fong et al. (1999) find support for an asymmetric information explanation.

[11] Theoretical examples include Madhavan (1992) and Glosten (1994). Empirical examples include the many comparisons across NYSE and NASDAQ: Lee (1993),Huang and Stoll (1996), LaPlante and Muscarella (1997), Keim and Madhavan (1996), Keim and Madhavan (1997) and Chan and Lakonishok (1997).

In an interesting theoretical piece, Hendershott and Mendelson (2000) look at the coexistence of exchanges and crossing networks. They show that there are subtle interactions between the two markets, and that the presence of a crossing network may have negative effects on the underlying market, in particular if the market is used as a "dealer of last resort."[12]

3 The data

The data set was provided to us by the Norwegian Government Petroleum Fund (hereafter the Fund). The Fund, which was established in 1990, is a vehicle for investing the Government's income from petroleum-related activities in international capital markets. Initially the Fund was invested in foreign government securities only. However, new criteria, applying from January 1998, stated that between 30 and 50 percent of the Fund was to be invested in equities. The composition of the Fund portfolio was changed to include equities in the first half of 1998. The investment universe includes 20 countries in Europe, North America, and Asia. The benchmark portfolio was the FT/S&P's Actuaries World Index. By the end of June 1998, the market value of the total Petroleum Fund portfolio was USD 17.7 billion, and the market value of the equity portfolio was USD 7.2 billion. Our data set is from this buildup period, when the Fund was a large buyer of equities.

We use data for US equity markets only.[13] US stocks represent 28.5 percent of the Fund's benchmark for the total equity portfolio. The US part of the FT/S&P's Actuaries World Index currently consists of slightly more than 500 different stocks. The constituents of this index are the largest companies on the exchange. The index is thus slightly broader-based than, but has a very high correlation with, the S&P 500 index.

In addition to the actual trades by the Fund we use market data. The NYSE Trades and Quotes (TAQ) data base provides trading volumes and prices. Datastream is the source of data on (longer-term) stock returns and market capitalization. We exclude stocks that split around the Fund's trades. In some cases we are not able to match the trades with market data, but this only applies to a small number of cases.[14] Table 2.1 shows some descriptive statistics for the Fund's trades.

The mean order is for 6851 shares, worth about USD 377000. Hence the Fund's trades are relatively large. This is further confirmed by two other statistics listed in the table. The average order is for 0.03 percent of a company's value, and 1.4 percent of the total NYSE volume that day. These orders are larger than in most other studies empirically investigating institutional trading costs. We will return to these issues in the robustness section.

[12] Crossing networks are also discussed in Dönges and Heinemann (2001) and the survey article of Keim and Madhavan (1998) as well as the book by Harris (2003).

[13] Focusing on the US part of the portfolio allows us to compare with other studies, most of which use US data.

[14] The problems stem from the different identifiers used for equities in the various databases, as well as problems due to Datastream removing delisted companies. Except for one date, the match percentage is in the range 82-94 percent.

Table 2.1: Descriptive statistics for the Fund's trades

The table provides descriptive statistics for the Fund's trades in the period January to June 1998. We provide means, standard deviations and medians split on crosses, market orders and all trades. *Number of Shares* is the number of shares in an order. *Trade value* is the dollar value of each trade in thousands of US dollars. *% of company* is the size of the order in percent of the total value of the company. *% of primary market volume* is the size of the order in percent of the total volume traded in the primary market (NYSE/NASDAQ) the same day. *Company Market value* is the company market value in billions of US dollars, and *Price* is the price of the stock. The last row lists the total number of observations in each category.

	Crosses			Market			All trades		
	mean	(std)	med	mean	(std)	med	mean	(std)	med
No of shares	7021	(9636)	3900	5920	(9108)	3400	6851	(9564)	3800
Trade value	390	(686)	177	304	(641)	149	377	(680)	173
% of company (%)	0.03	(0.02)	0.03	0.03	(0.02)	0.03	0.03	(0.02)	0.03
% of primary market volume	1.3	(2.7)	0.8	2.0	(3.4)	1.1	1.4	(2.8)	0.9
Company market value	18	(29)	8	13	(26)	6	17	(29)	8
Price	53	(30)	49	51	(32)	46	53	(30)	48
No of observations	2929			537			3466		

3.1 The trading strategy

The Fund did not not take any active positions during the transition period. It tracked the FT/S&P's Actuaries World Index subject to a maximum tracking error. At the beginning of each month, the benchmark index was changed to allow for a prescribed increase in the equity part of the total portfolio at the end of the month. This meant that the Fund had to buy most of the stocks in the index each month in order to comply with the restriction on the maximum tracking error.

Four index managers were hired to establish the portfolio. One of the index managers was chosen as "transition manager". The transition manager should first try to cross the Fund's desired orders internally with its own customers (internal cross). If it was not possible to find a seller in this network, the manager should try to find a seller in the customer bases of the three other index managers or send the order to an external crossing network (external cross). If it was not possible to cross the order at all, the stock should be bought in the primary market.[15] According to (Ruyter, 1999), this strategy is typical for relatively patient customers of large index managers.

What makes the Fund's strategy particularly interesting for our purpose is the following: Firstly, in contrast to other empirical studies on this subject, we do not have to *estimate* the trading strategy from the transactions data. We know that this strategy was decided by the Fund *ex ante*. This is important for a proper measurement of execution costs. It also means that we do not have to deal with the selectivity bias problem found in most other studies.[16] Secondly, because the Fund used the market as a "dealer of last resort", we know the identity of stocks that

[15]The submission strategy was explained to us in meetings with the people responsible for the transition in the Central Bank. The process of selecting managers is discussed in some detail in the Fund's 1998 annual report. The fund used Barclays Global Investors (BGI) as transition manager. Gartmore Investment Management, Bankers Trust Company and State Street Global Advisors UK were the other three managers. According to Harris (2003), BGI's internal crossing network is "probably the largest in the world". Hence, both the manager and the private network, where most of the actual crossing was performed, should be quite representative for the US market.

[16]Orders may be routed to different trading venues based on an initial conception of how hard it will be to fill them, see for example Conrad et al. (2003).

were initially tried crossed and subsequently acquired in the open market. This particular type of information enables us to look for signs of adverse selection in the crossing network. Such costs are extremely hard to detect empirically.[17] Finally, the fact that we have access to data from an internal crossing network is quite extraordinary. Evidence from these networks is scant, due to the obvious lack of interest in making it public from the participants' point of view.

Table 2.2 summarizes the magnitude and actual sequence of trades by the Fund. The portfolio was established in the period from January to June 1998. The total portfolio investment was USD 1.75 billion, of which USD 1.50 billion or nearly 86 percent, was crossed. The majority of the crossed orders, USD 1.36 billion, was executed internally. The Fund crossed externally at two occasions, while market trades to complete the desired portfolio were needed on three trading dates.[18] Transaction costs should be measured against the date when the *decision* to trade was made. Hence, the timing of the Fund's trades is important. All orders were executed within two days after the decision to trade. 18 percent of the investment amount was executed on the second day after the decision date, whereof 10 percent were market trades.[19]

The highest trading volume on one date amounted to USD 300 million, or 17.1 percent of the total portfolio investment. Note that for the period we are considering the Fund was only buying, not selling securities. For the first two months, crossing prices were set as the primary market (NYSE/NASDAQ) closing prices that day. For the remainder of the period, prices were set as the volume weighted average price (VWAP) of trades in the primary market during the day.

4 Measuring transaction costs

In this section, we first provide estimates of the conditional cost components of the Fund's crossing strategy. We then investigate the presence of adverse selection costs in the network by means of an event study.

4.1 The conditional implementation shortfall cost

We estimate the execution costs of the Fund following the empirical version of the implementation shortfall described in section 2. Use of this method also facilitates a comparison of the explicit and implicit cost components of the Fund's trades with the cost estimates reported in Keim and Madhavan (1998) and Conrad et al. (2003), which use data for the NYSE/NASDAQ and ECN's/External Crossing networks, respectively.

[17] Keim and Madhavan (1998) suggest studying the correlation between estimates of opportunity costs and future performance in order to assess the traders' information.

[18] We do not know how aggressively the market orders were traded in the market nor the actual intra-day timing of the executions.

[19] The table shows that market trades sometimes happened on the same date as internal crosses. According to the order submission strategy, these orders should be sent to external crossing before they were sent to the market. We do not know how this was done. Hence, it might be that the submission strategy was not strictly followed on this occasion.

Table 2.2: Establishing the US stock portfolio.
The table shows the characteristics of the actual orders for the Fund during the six month period January to June 1998. For reasons of anonymity we do not show the actual dates, but the date numbers are in chronological order. Some of the dates are close in time, day 3 is the day after day 2, day 6 is the day after day 5 and day 16 is the day after day 15. For each date we list the total dollar value of the trades (in millions of dollars) and the number of distinct securities traded (*n*), split on the three types of execution: Internal cross, external cross and market. Internal crosses are crosses performed with the transition manager's other customers. External crosses are trades performed with the other three institution's customers. Market trades are performed in the market, which is the NYSE or NASDAQ. The final column lists what percentage of the total trades by the fund was performed on the given date.

Date	Crosses				Market		All	
	Internal		External					
	mkt val	*n*	mkt val	*n*	mkt val	*n*	mkt val	*n*
1	174	454					174	454
2	184	453					184	453
3			115	342			115	342
4	58	153			3	29	61	182
5	19	65					19	65
6			30	105			30	105
7	163	130			73	303	236	433
8	300	634					300	634
9	14	88					14	88
10	14	49					14	49
11	231	465					231	465
12	70	386					70	386
13	8	45					8	45
14	23	48					23	48
15	97	350					97	350
16					174	418	174	418
	1356	3320	145	447	250	750	1751	4517
Percent	77		8		14			

Two problematic features of our cost calculation should be mentioned. *First*, the price difference between P_a and P_d is sensitive to general market movements between the two observation times. Essentially the measure of implicit costs assumes that the main source of this price difference is our order. When we look at averages for trades on many different dates this is not a major problem. Since the expected market movement is close to zero, the market movement is washed out in the average. However, this sensitivity becomes problematic for the external crosses and market trades in our sample because they were concentrated on only two and three days respectively. *Second*, our cost estimates are conditioned on the particular opportunistic crossing strategy chosen by the Fund. If we want to compare cost estimates across different trading venues, we have the problem that the orders were not submitted to different trading venues at the same time. For the orders that were executed on the day following the initial attempt at internal crossing, we therefore decompose the total cost into one component associated with the delay of the order in the internal crossing network, and one component associated with the final execution in an external crossing network or in the primary market.

The results of our cost estimation are summarized in table 2.3. Adding the average implicit

Table 2.3: Estimated trading costs for the Fund's transactions

Average trading costs for the Norwegian Government Petroleum Fund's transactions, January to June 1998. Costs are measured as in Keim and Madhavan (1998): Implicit trading costs are defined as $P^a/P_d - 1$, where P^a is the average price of all the executed trades in the order and P_d is an estimate of the "unperturbed" price price for the stock. Explicit trading costs are defined as (commission per stock$/P_d$). Costs are reported as percentages. Standard deviations are in parentheses. "vw avg" are value-weighted averages. The costs for the different types of orders (internal crosses, external crosses and market trades) are decomposed into (i) the average total cost *excluding* the costs associated with the delay of orders, (ii) the average delay cost, and (iii) the average implicit cost *including* the delay cost, i.e. the average implicit implementation shortfall cost. These correspond to different definitions of P_d, the unperturbed price. The difference between these cost definitions occurs when an order is first unsuccessfully attempted crossed, and then some time passes before the trade is actually completed in the market. For some of the trades this delay was one day. In these cases, for (i) one considers the date before the order submission *to the crossing network* the unperturbed price, and for (iii) one considers the date before of order submission *to the market* the unperturbed price. The two last columns show respectively the number of trading days and the number of stocks traded for each type of order.

	Total Cost			Implicit Cost			Explicit Cost			Days	Stocks
	mean	(stdev)	vw	mean	(stdev)	vw	mean	(stdev)	vw		
All orders											
Implementation shortfall	0.00	2.13	0.18	-0.03	2.13	0.17	0.03	0.14	0.01	16	3837
delay costs	-0.12	1.01	-0.10	-0.12	1.01	-0.10					
ex delay	0.12	1.95	0.28	0.09	1.95	0.27					
Internal Crosses											
Implementation shortfall	0.15	1.97	0.37	0.12	1.98	0.36	0.03	0.16	0.01	11	2861
External Crosses											
Implementation shortfall	0.07	2.99	-0.01	0.04	2.99	-0.02	0.03	0.03	0.01	2	390
delay costs	0.46	2.01	0.59	0.46	2.01	0.59					
ex delay	-0.39	2.19	-0.60	-0.42	2.18	-0.61					
Market Trades											
Implementation shortfall	-0.69	2.01	-0.87	-0.74	2.01	-0.91	0.05	0.04	0.04	3	586
delay costs	-0.99	1.63	-1.21	-0.99	1.63	-1.20					
ex delay	0.30	1.60	0.34	0.25	1.60	0.30					

and explicit cost for all orders gives an average implementation shortfall for of 0.00. Hence, the Fund's strategy seems very successful. Note, however, that the negative average implicit cost is driven by an average implicit cost for the market orders of -0.74 percent. The Fund incurred delay costs for market orders on one occasion. The fact that the drop in market values on this day was large enough to have a significant effect on the total implementation shortfall cost suggests that the true conditional implicit costs component of opportunistic crossing is underestimated.[20] Keim and Madhavan (1997) argue that one should *not* try to adjust for market movements because they are a part of the timing cost for the order submission strategy. However, they also show that, measured over some time, the daily market movements are small compared to the price impact costs. Hence, for large samples, adjusting for daily market returns does not make much difference. In our case, the delay of orders in the private internal crossing network resulted in a negative timing cost of -0.12 percent.[21]

Since there are many more dates on which internal crosses were done, our conditional cost estimates for internal crosses are not as blurred by particular market movements. The total trans-

[20] On one of the three dates when the Fund traded in the market, the market went markedly down, leaving the Fund with an implicit delay cost for the non-crossed orders of -1.79 percent. Because the non-crossed orders had to be bought in the market on the following day, an average additional cost of 0.48 percent was incurred, giving a total implementation shortfall cost of -1.31 percent. Due to the non-synchronous nature of the Fund's market trades relative to the close-to-close returns on the SP 500 index, a correct adjustment for the market movement would involve the actual timing of the trades during the day as well as the intraday SP 500 returns. None of which are easily obtainable.

[21] In the later empirical investigations we circumvent this problem by considering cross-sectional price differences at the same date.

action cost for an order that was internally crossed was on average 0.15 percent, of which the implicit cost component constituted 0.12 percent. It is not obvious how a cross can have a price impact, but one has to consider the possibility of an implicit price impact through the effects of removal of orders from the primary market. If we want to compare cost estimates across trading venues, we should exclude the costs related to delays in the crossing network. If crossing failed, and ignoring the delay component, the average cost for the subsequent market trade was 0.30 percent.[22] Hence, our results are in line with the findings of Conrad et al. (2003), that the costs for orders sent to alternative trading systems are lower than the costs for broker-filled orders.[23] The average cost of 0.15 percent for the Fund's internal crosses compares favorably with the results in Conrad et al. (2003). Externally crossed orders in their sample had an average total cost of 0.19 percent, of which explicit costs represented 0.07 percent. Hence, our results indicate that the explicit costs are lower in internal crossing networks than in external crossing networks. Our cost estimates also compare favorably with the results of Keim and Madhavan (1998), who found that the average cost for the smallest trade size quartile was 0.31 percent for exchange listed stocks, and that costs for higher quartiles were even higher. The median size of the trades in our sample is comparable to the Keim and Madhavan (1998) sample, but the Fund's trades are in stocks with larger market capitalization. This can explain part of the difference in estimated cost, since the costs in Keim and Madhavan (1998) were lowest in the stocks with highest market capitalization.

Table 2.3 also shows that the value-weighted average implicit cost for the internal crosses is considerably higher than the equally weighted average, indicating a size effect: larger orders have higher costs.[24] One would generally believe that the larger the order, the less likely it is that one will be able to fill the order in the crossing network, since the participants of the network are competing for a fixed number of offered shares. However, a size effect in costs is not self-evident, since crossing prices are set independently of order size. One cause of this size effect may be that the removal of order flows from the primary market by the crossing network has an implicit price impact.

4.2 Adverse selection costs in crossing networks

The estimates of the implementation shortfall cost do not account for potential adverse selection costs related to the Fund's order submission strategy, since they are estimated conditional on the execution. In this section, we use our knowledge of the Fund's ex ante submission strategy to investigate the existence of adverse selection by means of an event study.

We concentrate on the unambiguous prediction that the presence of informed traders in a crossing network implies that stocks that are not available in the crossing network are likely to

[22] As noted, it would be more correct to compare the crossing costs with costs of market orders submitted at the same time.

[23] However, as noted, the cost estimate of our market trades should be interpreted with caution due to the few dates on which they were traded in our sample.

[24] We have split the sample on the basis of both trade size and market capitalization, and the results show that costs increase with trade size and decline with market cap, which is similar to results in Keim and Madhavan (1998).

perform better than stocks that are easily crossed.

For a liquidity trader following an opportunistic crossing strategy, the presence of informed traders should lead to an ex post difference in performance between the stocks that were successfully crossed and the stocks that had to be purchased in the market. We investigate the presence of private information by performing an event study where we use the difference in cumulative abnormal return (CAR) for the two groups of stocks as a proxy for performance differences. We calculate different CARs based on several *event* periods to measure the information revelation. Note that by looking at the *difference* in the CAR of the two groups, crossed and market orders, done at the same date, we avoid the problem of sensitivity to market movements discussed earlier.

In order to perform the event study, for each stock i, we compare the daily actual return (R_{it}) with an estimate of its "expected return" $(\widehat{E[R_{it}]})$. The abnormal return is the difference:

$$\widehat{AR}_{it} = R_{it} - \widehat{E[R_{it}]}$$

We employ the market model as an estimate of expected return.[25] The CAR is found by aggregating these abnormal returns over time.

Some care has to be used in this event study. We want to consider cases where the Fund first tried to cross and then immediately went to the market to buy the stocks they did not get in the crossing network. There are only three such cases with significant market orders by the Fund. Of the relevant cases, one is in the beginning of the period (date 4 in table 2.2), one is in the middle of the period (date 7), and a final one is near the end of the period (dates 15 and 16). Table 2.4 summarizes these investigations.

In table 2.4 we have tried to show the robustness of the event study by performing it for a number of different cumulating periods, starting both before and at the date of the crossing attempt, and ending 10, 15, 20 and 25 trading days after the attempt to cross. For each combination of dates before and after the "event date" we list the difference in average CAR for the two groups, market orders minus crosses. A positive difference should be interpreted as the market orders having higher CAR on average. We also list the p-value for a test of whether the difference is zero.

Consider the last row in each panel, where the event study starts on the day of crossing. Looking more than 10 days out in the February study, and for all cases in the June study the difference in CAR is significantly positive. Those stocks that could not be had in the crossing network outperform the crossed ones with 2.8% over a 25 trading day period (Slightly above one calendar month). Figure 2.1 illustrates the June case. The June case is particularly interesting for two reasons. First, it is the case with the largest number of market trades. Second, since mid-June was the ending date for the establishment of the total Fund portfolio, we know for sure that

[25] According to (Campbell et al., 1997, Ch 4) the market model does not produce very different conclusions from alternative methods, and is much simpler to perform. The market model produces the estimate of expected return as $\widehat{E[R_{it}]} = \alpha_i + \beta_i R_{mt}$ where R_{mt} is the observed market return for date t. We use daily stock returns for two years preceding the "event" to estimate α_i and β_i for each stock i. The S&P 500 is used as the market index in the estimation.

Table 2.4: Summarizing event studies

The table summarizes the various event studies performed by showing the differences in average CARs between the two categories of orders, crosses and market orders, as a function of when the event study starts and ends. We calculate excess returns by using the market model to estimate the expected return $\widehat{E[R_{it}]} = \alpha_i + \beta_i R_{mt}$ where R_{mt} is the observed market return for date t. Daily returns for the two years preceding the "event" are used to estimate α_i and β_i for each stock i. The S&P 500 is used as the market index in the estimation. Excess return is the the difference between the actual return and this expected return, $AR_{it} = \widehat{E[R_{it}]} - R_{it}$. CAR is estimated by cumulating these excess returns. On the horizontal axis we consider different ending dates for the cumulation, on the vertical axis we consider different starting dates. For example, if the event study starts on date -20 and ends on date 10, we have used 20 trading days before the event and 10 trading days after the event in calculating the cumulative excess returns. The number listed in the table is the difference between the average CARs for market orders and for crosses. A positive value means that the average CAR for market orders is higher than the average CAR for crosses. In parenthesis we show the probability value of the difference being nonzero. The probability value is calculated by a slight adjustment of the J_1 statistic discussed in chapter 4.4 of Campbell et al. (1997). We exclude stocks which were partially filled in the crossing network before the remainder being filled in the market. The February study has 127 crosses and 23 market orders. The March study has 240 market orders and 91 crosses. The June study has 104 crosses and 160 market orders.

Panel A: February

	Ending			
	10	15	20	25
Starting: −20	−1.61(0.00)	−0.80(0.00)	0.09(0.62)	1.45(0.00)
−15	−3.51(0.00)	−2.70(0.00)	−1.81(0.00)	−0.45(0.01)
−10	−1.82(0.00)	−1.01(0.00)	−0.12(0.53)	1.24(0.00)
−5	−0.31(0.25)	0.51(0.03)	1.39(0.00)	2.75(0.00)
0	−0.27(0.41)	0.54(0.05)	1.42(0.00)	2.78(0.00)

Panel B: March

	Ending			
	10	15	20	25
Starting: −20	0.44(0.00)	0.46(0.00)	0.62(0.00)	−0.52(0.00)
−15	0.20(0.13)	0.22(0.07)	0.38(0.00)	−0.75(0.00)
−10	0.11(0.44)	0.13(0.32)	0.29(0.02)	−0.84(0.00)
−5	0.20(0.24)	0.22(0.15)	0.38(0.00)	−0.76(0.00)
0	−0.17(0.39)	−0.15(0.37)	0.01(0.95)	−1.13(0.00)

Panel C: June

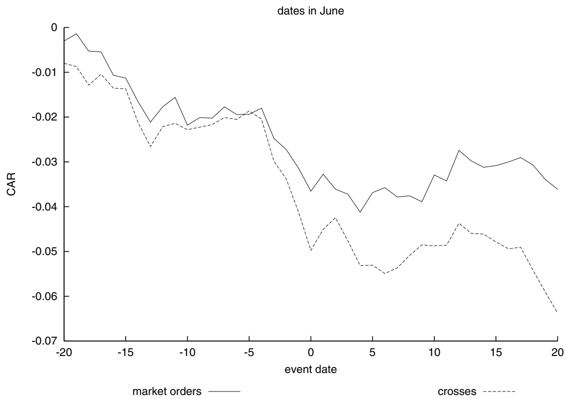
	Ending			
	10	15	20	25
Starting: −20	1.58(0.00)	1.70(0.00)	2.76(0.00)	3.81(0.00)
−15	1.29(0.00)	1.41(0.00)	2.47(0.00)	3.53(0.00)
−10	1.00(0.00)	1.12(0.00)	2.18(0.00)	3.24(0.00)
−5	1.47(0.00)	1.59(0.00)	2.65(0.00)	3.71(0.00)
0	0.60(0.02)	0.71(0.00)	1.78(0.00)	2.83(0.00)

orders not successfully crossed were sent directly to the market. The March study has no clear conclusion. This could be related to the number of stocks in the two groups. In the March case most of the stocks were traded in the market, only 91 out of a total of 331 observations are crosses. In the other cases relatively more of the trades are done as crosses.[26]

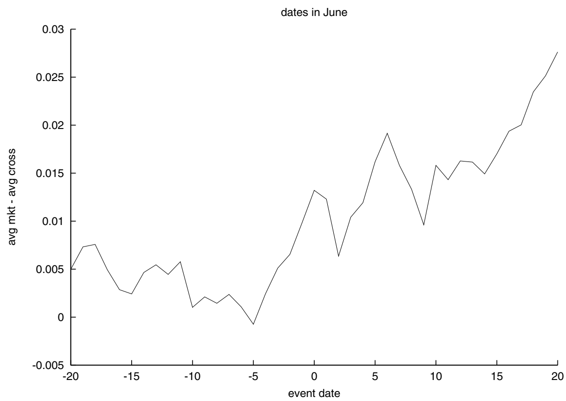
Figure 2.1: Event study

The plots illustrate the results of the June event study. An event is an occasion on which the Fund tried to cross a set of shares, and then went to the market to purchase the shares they were not able not cross. We calculate excess returns by using the market model to estimate the expected return: $\widehat{E[R_{it}]} = \alpha_i + \beta_i R_{mt}$, where R_{mt} is the observed market return for date t. Daily returns for the two years preceding the "event" are used to estimate α_i and β_i for each stock i. The S&P 500 is used as the market index in the estimation. Excess return is the the difference between the actual return and this expected return, $AR_{it} = \widehat{E[R_{it}]} - R_{it}$. CAR is estimated by cumulating these excess returns. The figures in panel A compare the cumulative abnormal returns for orders that were crossed and market orders for two subsequent dates in June (dates 15 and 16 in table 2.2). We remove equities that were partially filled in both markets. The study uses 104 crosses and 160 market orders. Panel B plots the difference between average CAR's for market orders and crosses. (The difference between the lines in the upper panel.)

Panel A: Average CARs



Panel B: Difference in CARs



[26]One possibility is that relatively few stocks in the cross for the March date are related to how aggressively the attempt to cross has been, and that the likelihood of informed trading is highest in those cases where just a few stocks could not be crossed.

We posit that the difference in CARs is evidence of the presence of informed traders. We investigate the plausibility of alternative explanations below, but we have problems in seeing how any differences in liquidity and other measures of trade difficulty can induce a 2.8% return differential over a month.

It may at this point be helpful for intuition to think about how one could construct a trading strategy to exploit these results. What we find is that the fact that one is not able to buy a stock in a crossing network is a positive signal. To exploit this, one could send a small buy order to a crossing network, and then buy in the market those stocks one did not get in the cross in the market, possibly combined with short positions in the stocks one *was* able to cross.[27]

5 Regression analysis

In section 4 we showed evidence that the implementation shortfall cost of the opportunistic crossing strategy followed by the Fund was inexpensive compared with estimates of transactions costs in other studies. However, we also found indications of adverse selection costs mitigating this cost advantage. In this section, we investigate the robustness of our cost determinants more thoroughly through several regression models. We estimate three different types of models:

- Firstly, we estimate a regression model on the conditional cost estimates of the Fund's trading strategy based on the determinants of costs found in other studies.
- Secondly, we estimate a regression model on the CARs found in the event study to check whether other variables than the crossing event might explain the return differential.
- Finally, we estimate a probit model on the probability of execution in the network where we include the CARs as well as other variables that might explain the probability of getting an order crossed.

5.1 A regression model on the conditional costs of crossing

We first estimate a standard regression equation on total trading cost based on the regression approach used in Keim and Madhavan (1997) and Jones and Lipson (1999).[28] From the determinants of trading costs in these studies we include as explanatory variables: a variable for order size, reflecting the fact that large orders are more expensive than small orders, a variable for liquidity, reflecting a negative relationship between execution costs and stock liquidity, a

[27] In fact this strategy was described to us at a conference after we presented our paper, how one could "ping" a crossing network by sending a small order, and use the lack of success in the cross as a buy signal. (The term "ping" is presumably from the UNIX program used to test communication lines, and is actually a very apt name for such a strategy. The program `ping` sends a short message to another machine on a communication network and asks for a response. The receipt of the response is a confirmation that there is a way to get to the other machine through the communication network.)

[28] Keim and Madhavan (1997) examine the magnitude and determinants of transaction costs for a sample of institutional traders with different investment styles. Jones and Lipson (1999) compare transaction costs across NYSE, NASDAQ and AMEX using a sample of institutional equity orders in firms that switch exchanges.

variable for total market activity, reflecting the potential greater ease of trading when market activity is high, a variable for "adverse momentum," reflecting the greater difficulty of executing a buy order when prices are rising, a variable for intraday volatility, reflecting the fact that it is more difficult to trade when markets are volatile, and the inverse of the stock price, reflecting the effects of general price movements on proportional costs.[29]

We estimate two regression equations. In the first we use the total trading costs of all orders as the dependent variable, and include dummies for the externally crossed orders and the market orders. Since our cost estimates for the internally crossed orders are the most robust, we run a second regression where we only consider these orders.

Table 2.5: Regression analysis of total trading costs for all orders

The table reports the results from estimation of the regression model

$$\begin{aligned} TC_i &= \beta_0 + \beta_1 IP_i + \beta_2 \ln(MCap_i) + \beta_3 \ln(Ord_i) + \beta_4 \ln(MktVlm_{i,t-1}) \\ &\quad + \beta_5 R_{i,t-3,t-1} + \beta_6 HL_{i,t-1} + \beta_7 D_i^{EC} + \beta_8 D_i^M + \beta_9 D_i^P + \varepsilon_i \end{aligned}$$

where, for trade i, IP_i is the inverse of the price per share of the stock traded, $MCap_i$ is the market capitalization of the stock traded, Ord_i is the order size measured as number of stocks, $MktVlm_{i,t-1}$ is the number of stocks traded on the NYSE on the day before the transaction, $R_{t-3,t-1}$ is total returns over the two days preceding the transaction, $HL_{i,t-1}$ is the difference between the highest and lowest mid quote on the day before the transaction. D_i^{EC} is a dummy variable equal to one for stocks that were externally crossed, D_i^M is a dummy variable equal to one for stocks that were bought in the market and D_i^P is a dummy variable equal to one if the order is partially filled. ln() is the natural logarithm. The model is estimated using all orders in the data sample for which we can extract returns data and data on market capitalization from Datastream. In the table the first column lists the variable, the second the coefficient estimate, the third the estimated standard deviation, the fourth the probability value of the coefficient being nonzero, and the last column the mean of the explanatory variable.

Dependent variable: TC	coeff	(stdev)	pvalue	mean
Constant	-0.0010	(0.0039)	0.79	
$IP_{i,t}$	-0.1693	(0.0259)	0.00	0.0248
$\ln(MCap_i)$	-0.0035	(0.0006)	0.00	9.0354
$\ln(Ord_i)$	0.0028	(0.0004)	0.00	8.2910
$\ln(Vol_{i,t-1})$	0.0012	(0.0004)	0.00	13.1239
$R_{i,t-3,t-1}$	0.0003	(0.0001)	0.04	0.2684
$HL_{i,t-1}$	0.0000	(0.0002)	0.96	1.2350
D^{EC}	-0.0061	(0.0011)	0.00	0.0987
D^M	0.0014	(0.0009)	0.12	0.1704
n	3516			
R^2	0.03			
Average (TC)	0.0012			

Table 2.5 presents the estimated coefficients of the regression model on all orders. The first thing to note is that the part of the total variation in trading costs explained by the independent variables is very small. This is natural since most of the explanatory variables measures trade difficulties in the primary market which by construction should be less important in a crossing network. The coefficients of stock liquidity and order size are both significant and have the

[29] The lower the stock price, the higher the fixed proportional costs.

expected signs. A positive coefficient of trading volume the day before the transaction is also significant, indicating higher costs the more popular the stock has been lately in the primary market. The size of the dummies are supposed to capture any "order form" effects on trading costs that are unrelated to the explanatory variables.[30] The dummy variables for externally crossed orders is negative and significant, indicating lower costs for these orders after controlling for the trade difficulty variables. As mentioned before, however, the cost estimates of both externally crossed orders and market trades should be interpreted with great caution due to the few dates on which they were traded in our sample.

Table 2.6: Regression analysis of total trading costs for internally crossed orders only

The table reports results from estimation of the regression

$$\begin{aligned} TC_i &= \beta_0 + \beta_1 IP_i + \beta_2 \ln(MCap_i) + \beta_3 \ln(Ord_i) + \beta_4 \ln(MktVlm_{i,t-1}) \\ &\quad + \beta_5 R_{i,t-3,t-1} + \beta_6 HL_{i,t-1} + \varepsilon_i, \end{aligned}$$

where, for trade i, IP_i is the inverse of the price per share of the stock traded, $MCap_i$ is the market capitalization of the stock, Ord_i is the order size measured as number of stocks, $MktVlm_{i,t-1}$ is the number of stocks traded on the NYSE/NASDAQ on the day before the transaction, $R_{i,t-3,t-1}$ is total returns over the two days preceding the transaction, and $HL_{i,t-1}$ is the difference between the highest and lowest mid quote on the day before the transaction. ln() is the natural logarithm. The model is estimated using all internally crossed orders in the data sample for which we can extract returns data and data on market capitalization from Datastream. In the table the first column lists the variable, the second the coefficient estimate, the third the estimated standard deviation, the fourth the probability value of the coefficient being nonzero, and the last column the mean of the explanatory variable.

Dependent variable: TC	coeff	(stdev)	pvalue	mean
Constant	0.0012	(0.0044)	0.79	
$IP_{i,t}$	-0.2125	(0.0293)	0.00	0.0246
$\ln(MCap_i)$	-0.0042	(0.0006)	0.00	9.0890
$\ln(Ord_i)$	0.0035	(0.0005)	0.00	8.3099
$\ln(Vol_{i,t-1})$	0.0010	(0.0004)	0.02	13.2048
$R_{i,t-3,t-1}$	0.0003	(0.0001)	0.01	0.4685
$HL_{i,t-1}$	-0.0000	(0.0002)	0.86	1.2513
n	2917			
R^2	0.03			
Average (TC)	0.0009			

Table 2.6 presents the estimated coefficients of the regression model on the total costs of internal crosses. The only difference from the results of the estimation of the regression model on all orders is that the positive coefficient for adverse momentum becomes significant, indicating that stocks that did well recently are more costly to buy in the crossing network.

One problem with the regression analysis above is that the result whether a stock was crossed or not is given as an exogenous variable. In section 5.3 below we estimate a limited dependent variable model where we treat the cross dummy as the dependent variable. This is more correct in that we assume that the probability of success in the crossing network is endogenous.

[30] In the cost estimation in section 4.1, the comparison of costs across trading venues does not control for differences in trade difficulty between venues. If e.g. the market orders were all large orders, and large orders are more difficult to fill, this could cause the results.

5.2 Determinants of the CAR differentials

To gauge the robustness of our conclusion that the difference in the CARs is due to adverse selection, we perform a regression where we explain the CARs with other factors than the cross/no cross outcome which may affect CAR differences.

Specifically, we estimate a regression model explaining the cumulative abnormal return starting at the date of trade t and accumulating for 20 trading days, $CAR_{i,t,t+20}$.

The first explanatory variable we consider is the company market value. This is included to control for the possibility that investors might demand a liquidity premium to hold smaller companies that is not fully reflected in our estimate of beta. If so, one would expect that the smaller the company, the higher the ex post CAR should be. However, as all the stocks in the sample are index constituents, there may be little difference in liquidity due to size, since they are all large firms.

Because our conditional cost estimates indicate a size effect, we also control for order size in the regression.[31] The recent return on the stock is included to control for short term return reversals or continuations.[32]

Two measures of stock variability and a measure of relative trading activity are included to proxy for the presence of private information. Variability may be related to informed trading in the sense that the value of being informed is higher for more volatile stocks. We use two measures of variability, high minus low and return volatility, because they may measure different properties.[33] We also construct a variable for relative trading activity of the stock in the primary market by dividing the trading volume on a given date by the average daily trading volume for the same stock during the relevant month. To reduce the noise of the measure we calculate it for the two trading days prior to order submission by the Fund. This number gives an indication of whether the day had a higher or lower than usual trading volume. A high value signifies a stock which is currently actively traded in the primary market. We believe that a high value for this variable is a proxy for informed trading in the primary market because informed traders add to the usual crowd of traders.

We finally include a dummy variable for whether the order was crossed. If a cross outcome is merely a proxy for the other variables, the cross dummy should not be significant; the other variables should explain all the CAR differences. We have earlier argued that we may have problems due to the fact that this investigation is only performed on a few dates. We therefore add dummy variables that measure fixed date effects.

The results of the regression analysis are presented in table 2.7. The variables that show a

[31] We have also considered the order size in dollars, which yields similar conclusions.

[32] The literature on profitability of short term momentum type strategies is extensive, and there is little consensus about its existence and direction. A recent paper on the issue, Lewellen (2002) and particularly the discussion of this paper by Chen and Hong (2002) give some pointers to relevant papers. We choose to be agnostic about the type of strategy; the sign of the coefficient will determine whether it is reversal or momentum.

[33] "High minus Low" is measured using the previous date's data, and may therefore be better at capturing short term changes in variability. Volatility is estimated using return data for the previous 3 months (60 trading days), and may be better at capturing longer term differences in variability. The correlation between these two measures of variability is as low as 0.08, indicating that they may indeed be capturing different stock properties.

Table 2.7: Determinants of CAR

The table shows the results of a regression

$$\begin{aligned} CAR_{i,t,t+20} &= \beta_0 + \beta_1 \ln(MCap_i) + \beta_2 \ln(Ord_i) + \beta_3 R_{i,t-6,t-1} + \beta_4 HL_{i,t-1} + \beta_5 \sigma(R_{i,t-60,t-1}) \\ &\quad + \beta_6 RelVlm_{i,t-2,t-1} + \beta_7 D_i^{Cross} + \sum_{j=2}^{4} \gamma_j D^j + \varepsilon_i \end{aligned}$$

where $CARi,t,t+20$, the cumulative abnormal return over the 20 trading days following the attempt to cross, is the dependent variable. Explanatory variables are $MCap_i$, the company market value, Ord_i, the size of the order in number of shares, $R_{i,t-6,t-1}$, the return of stock i over the week before the attempt to cross, $HL_{i,t-1}$, high minus low, the difference between the highest and lowest price for stock i on the date before the attempt to cross, $\sigma(R_{i,t-60,t-1})$, stock return volatility for stock i in the two months before the attempt to cross and $RelVlm_{i,t-2,t-1}$, the average relative volume for the two days before the attempt to cross, where relative volume is defined as the trading volume in the primary market divided by the average daily volume in the primary market for the month. $\ln()$ is the natural logarithm. D_i^{Cross} is a dummy variable equal to one if the trade was executed in the crossing network. We finally add dummy variables D^j for dates to adjust for fixed date effects. The regression uses data for all the dates where there are both crosses and market orders. In the table, the first column lists the variable, the second the coefficient estimate, the third the estimated standard deviation, and the last column the probability value of the coefficient being nonzero, using a normal distribution.

	coeff	(stdev)	pvalue
Constant	-0.081	(0.025)	0.00
$\ln(MktVal_i)$	0.006	(0.003)	0.06
$\ln(NoShares_i)$	0.007	(0.003)	0.04
$R_{i,t-6,t-1}$	-0.000	(0.001)	0.67
$HL_{i,t-1}$	0.013	(0.004)	0.00
$\sigma(R_{i,t-60,t-1})$	-0.182	(0.362)	0.62
$RelVol_{i,t-2,t-1}$	-0.010	(0.006)	0.10
$D_{i,Cross}$	-0.027	(0.010)	0.01
Date 2	-0.032	(0.011)	0.00
Date 3	-0.031	(0.010)	0.00
Date 4	-0.050	(0.013)	0.00
n	1100		
$\bar{R}^2$	0.05		

significant effect on the CARs are the constant, the market value of the company, the order size and two of the three proxies for private information. The constant measures effects common to all securities, such as market movement. The market value of the company has a significantly positive effect on the CARs. This is not in accordance with a liquidity premium explanation. However, as mentioned all the stocks in our sample are quite liquid. The two measures of volatility have positive coefficients, while the relative volume has a negative coefficient.

Note that the cross dummy has a significant negative coefficient of -0.027. When we control for the other determinants, the return differential between crossed stocks and stocks that had to be bought in the market is thus 2.7% over the month after the cross. This is very similar to the 2.8% we found for two of the event studies. Note that this simple dummy assumes that the effect of a cross is constant, which is clearly unrealistic. It is still an important control for the event study, in which one does not control for other properties of a given order. Our results thus lend support to the prediction that informed investors will try to exploit their information in a crossing network first.[34]

5.3 A probit model of crossing success

In this section we study what factors other than private information may affect the probability of success in the crossing network. To this end we use a limited dependent variable analysis where the probability of success in the crossing network is treated as an endogenous variable

Since we have data for one institution's trades only, we cannot test for general market structure effects of crossing networks. Nevertheless, it might be interesting to see whether factors related to trade difficulty in the primary market are related to the crossing success of the Fund. To this end, we consider a choice theoretic (probit) model where the dependent variable is the probability of a stock being crossed. Note that our probit estimation is simplified by the fact that our data only contain buy orders, we do not have to adjust for the direction of trade.

The literature on why traders may decide to trade outside exchanges tells us that a suitable model for the right-hand side of the equation should include variables for order size and stock liquidity. To measure whether crossing success is related to the activity of the underlying stock, we include the relative trading volume in the stocks the day of the order. We also include the relative trading volume two days before submitting the order. We also include short-term return on the stock to control for price reversals and continuations, and two measures of stock variability for the same reasons as in the previous investigation.

Table 2.8 presents the estimation results. These estimates show the effects of changes in one of the explanatory variables on the probability of an order being crossed. The order size effect is negative and the market capitalization effect is positive. Hence, the data support the cream skimming hypothesis in Easley et al. (1996). However, one should keep in mind that the stocks in our data set are all very liquid. Moreover, we would expect a negative effect due to order size simply because the participants in the network are competing for a fixed number of offered

[34] This is a prediction of the Hendershott and Mendelson (2000) model.

Table 2.8: Probit model estimating determinants of probability of a cross

We estimate a probit model of the probability that a given order is crossed. The probability of observing a cross is assumed to be given by the model

$$Prob(Y = Cross) = F(\beta'\mathbf{x})$$

where $\mathbf{x}$ is the vector of explanatory variables, β the vector of coefficients, and $F(\cdot)$ a cumulative distribution function. Success in the probit is defined as a cross. Explanatory variables are $MCap_i$, the company market value, Ord_i, the size of the order in number of shares, $R_{i,t-6,t-1}$, the return of stock i over the week before the attempt to cross, $HL_{i,t-1}$, high minus low, the difference between the highest and lowest price for stock i on the date before the attempt to cross, $\sigma(R_{i,t-60,t-1})$, stock return volatility for stock i in the two months before the attempt to cross and finally $RelVlm_{i,t-2,t-1}$ and $RelVlm_{i,t}$, the relative volume for the two days before the attempt to cross and the date of the cross, respectively. The relative volume is defined as the total trading volume in the primary market divided by the average daily volume in the primary market for the month. ln( ) is the natural logarithm. The model is estimated for the same set of trades as those used in the event studies. The total data set contains 950 transactions, of which 372 were crosses. In the table, the first column list the variable, the second the coefficient estimate, the third the estimated standard deviation and the fourth the estimated probability value for the coefficient being nonzero, using a normal distribution.

	coeff	(std)	pvalue
constant	-2.083	(0.43)	0.00
ln($MktVal_i$)	0.335	(0.06)	0.00
ln($NoShares_i$)	-0.150	(0.05)	0.00
$R_{i,t-6,t-1}$	-0.046	(0.02)	0.01
$HL_{i,t-1}$	-0.032	(0.06)	0.57
$\sigma(R_{i,t-60,t-1})$	6.473	(6.56)	0.32
$RelVol_{i,t-2,t-1}$	0.242	(0.11)	0.03
$RelVol_{i,t}$	-0.327	(0.13)	0.01
n	950		
pseudo R^2	0.05		

shares, i.e. the larger the order, the harder it will be to fill it in the network. Additionally, given that a crossing network exhibits positive network externalities (in the probability of a cross) it is possible that market cap is a measure of trading activity reflecting this.[35]

Interestingly, the coefficient on the last two days' return is significantly negative; equities which have had a large positive return the last couple of days are harder to find in the crossing network. The use of short-term continuation strategies by market participants may thus explain some of the CAR differential.[36]

Regarding the last two variables, which are proxying for the trading activity being above or below the normal in the main market, note that they are both significant, but have different signs. Stocks that one was able to cross had a lower than usual main market volume the day of the cross, indicated by the negative coefficient of the $RelVol_{i,t}$ variable. A possible explanation is that we see signs of reduced activity in the main market in periods with higher crossing activity. Also, the crossed stocks showed higher trading activity the two days *before* the cross. While this is consistent with anticipatory trading, participants in the crossing network who know that there is a large uninformed buy order coming may want to position themselves to exploit it, this is not the only possible explanation. Trading activity may simply be related to information, which we

[35]This point was suggested by an anonymous referee.

[36]However, this variable is sensitive to market movements. Since we are aggregating data for only two dates, differences in the market movements for these two days may be the cause of this result. When we look at the estimates split over the two separate days, the sign of this coefficient is different for the two days, and the coefficient is not significant on either day.

have seen affects cross probability.

6 Robustness

The empirical results of the previous section has shown a number of interesting results regarding crossing networks. An obvious question is the robustness of these conclusions. The prime issue is the representativeness of the data, since we are only using the trades of one institution. The applicability of our results therefore depends critically on this investor being representative. We will argue for representativeness by several means. The sheer size of the investor in terms of dollars speaks for the economic significance of the trades. The four managers performing the actual crossing are among the largest in the US market. Finally, we compare the Fund's trades to the trades used in similar studies, both in terms of the typical order, and the companies invested in, and show that the Fund's trades, if anything, are in larger companies, in more liquid stocks, and for higher amounts than in the other studies.

As mentioned in the data section, the transactions used are data on buying USD1.7 billion worth of US equities. While many US institutional investors have funds under management which are larger than this, not many of them have built up such a large position over a 6 month period. There is thus real money behind the transactions, which speaks for the economic significance of the data.

Another important point is that the managers actually performing the crossing, in particular the transition manager, BGI, are among the largest institutions of this type in the US market. In fact (Harris, 2003, page 133) claims that the BGI network is "probably the largest in the world" in its class. This again speaks for the representativeness of the trades.

We can also argue for the representativeness of the trades based on the size of both the orders and the companies traded. The trades are exclusively trades in listed stocks, and only stocks contained in the US part of the FT/S&P world index. These are thus the largest companies on the exchange. We have compared the Fund's trades with similar, recent studies of institutional investors' transaction costs in the US equity market. The most relevant samples of institutional investors with which to compare the Fund's trades are those used in Keim and Madhavan (1995, 1997) and Conrad et al. (2003). Keim and Madhavan (1995, 1997) use data on all equity transactions by 21 institutional investors from January 1991 through March 1993. This data set contains a total of 62,333 orders. The institutions vary in size. For fundamental value managers, the mean dollar value of assets under management was USD 4.8 billion, ranging from a low of USD 0.7 billion to a high of USD 12.9 billion. For index managers and technical traders, the mean dollar values of assets under management were USD 3.2 billion and USD 5.3 billion respectively. Conrad et al. (2003) have a larger data set from a more recent time period. Their sample consists of almost 800 thousand orders submitted by 59 institutions from the first quarter of 1996 to the first quarter of 1998.

If we look first at order size, the Fund's median order is for USD174 thousand. This is slightly larger than the median buy order of USD138 thousand used in Keim and Madhavan (1998), and much larger than the case for the crossed and ECN filled orders in Conrad et al. (2003). The average dollar value of the Fund's orders of USD386 thousand is also higher than the average dollar value of the orders sent to external crossing and ECN's, but considerably lower than the average dollar value of the orders filled by brokers and multiple order mechanisms in the Conrad et al. paper.

Comparing the market capitalizations of the stocks in the sample with these other papers, we find that the most liquid stocks in Conrad et al.'s study are the ones underlying the orders sent to external crossing systems. These securities have an average market capitalization of USD 12.7 billion. The average market capitalization for the stocks purchased by the Fund was USD 16.9 billion. The Fund is thus trading in larger companies.

To shed further light on the Fund's trades, we can compare them with some descriptive statistics for the trading activity in the same stocks for the same period on the NYSE. The average Fund order of 6851 stocks is much larger than the average NYSE trade of 1622 stocks. Also, over the days they traded, the Fund's average trade was 1.4 percent of the total NYSE volume that day.

We can summarize the characteristics of the Fund's trades as being large orders for the most liquid stocks on the NYSE. The orders are at least of a comparable size to other, similar studies done on US equity markets.

7 Conclusion

The US equity markets are currently in a state of flux; a significant fraction of order flow is directed to electronic marketplaces away from the traditional marketplaces, the NYSE and the NASDAQ. This paper has looked at one particular type of market structure which has experienced increased popularity in the recent past, the *crossing network*. We use the US equity trades of a large institutional investor to show some evidence of the characteristics of crossing networks.

Our data set has two unique features. First, we have detailed information on the execution costs in an internal crossing network. This type of crossing has largely been a "black box", even to the institutional investors doing the trading. Second, we have detailed information on the investor's motivation prior to the order submission process as well as on the instructions given to the Fund manager responsible for carrying out this strategy. This allows us to examine execution costs on the basis of the exact investment strategy of the investor, rather than having to infer the investment strategy from the sequence of trades as was done in e.g. Conrad et al. (2003).

Our results can be grouped into two. First, by measuring trading cost, we show estimates of average costs for internal crosses which are lower than similar cost estimates for external crossing network reported in the literature. In doing the cost estimation we also highlight a potential problem of the now standard method of estimating the cost of trading, as used in e.g. Keim and Madhavan (1995, 1997) and Conrad et al. (2003). The method, which compares the price at which a trade was performed with an "unperturbed" price, assumes that the only factor causing the price movement was the trade. In particular, market movement between the time when measuring the "unperturbed" price and the actual trade is assumed neutral. This assumption is reasonable when the trades are relatively evenly spread out in time, but it is problematic when trading is concentrated on a few dates, as in our study. In one case we actually estimate a *negative* cost of trading, but on investigation this turns out to be due to market falling, in which case it was much cheaper to buy one day later.

We also show evidence that the probability of execution in crossing networks is related to the subsequent performance of the stock. The investor in our study used the primary market as a "market of last resort". This means that the investor tried to execute all the orders in the crossing networks first, and then took the trades on to the traditional marketplaces if there were no counterparties in the network. Our evidence of informed trading is found by conducting an event study comparing the ex post performance of the two groups of stocks.

We claim that our results indicate the presence of informed traders in crossing networks. However, we wish to point out that this result is not necessarily tantamount to saying that trading in crossing networks should be avoided. As shown in Hendershott and Mendelson (2000), such an outcome can still be an equilibrium outcome, as the adverse selection can be offset by the lower cost of the trades that one *did* manage to cross.

A potential problem with our results is due to the fact they are based on the trades of only one institution. However, we have argued in the robustness section that there is no reason to believe this institution to be any way special. If anything, it is larger than the institutions used in similar studies. The trades are large, both per trade and in terms of the total amount traded. The investment advisers are among the largest US investment advisors.

Bibliography

Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton University Press.

Chan, L. K. C. and Lakonishok, J. (1997). Institutional equity trading costs: NYSE versus Nasdaq. *Journal of Finance*, 52(2):713–735.

Chen, J. and Hong, H. (2002). Discussion of Momentum and autocorrelation in stock returns. *Review of Financial Studies*, 15:565–573.

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

Dönges, J. and Heinemann, F. (2001). Competition for order flow as a coordination game. Working paper, Goethe-Universität Frankfurt.

Easley, D., Kiefer, N. M., and O'Hara, M. (1996). Cream-skimming or profit sharing? The curious role of purchased order flow. *Journal of Finance*, 51:811–833.

Fong, K., Madhavan, A., and Swan, P. (1999). Why do securities markets fragment? Working Paper, University of Southern California.

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

Harris, L. (2003). *Trading and Exchanges. Market Microstructure for Practitioners*. Oxford University Press.

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

Huang, R. and Stoll, H. (1996). Dealer versus auction markets: A paired comparison of execution costs on Nasdaq and NYSE. *Journal of Financial Economics*, 41:313–357.

Jones, C. M. (2000). A century of stock market liquidity and trading costs. Working Paper, Columbia University.

Jones, C. M. and Lipson, M. L. (1999). Execution costs of institutional equity orders. *Journal of Financial Intermediation*, 8:123–140.

Keim, D. B. and Madhavan, A. (1995). Anatomy of the trading process: Empirical evidence for the motivation for and execution of instituional equity trades. *Journal of Financial Economics*, 37:371–398.

Keim, D. B. and Madhavan, A. (1996). The upstairs market for large–block trades: Analysis and measurement of price effects. *Review of Financial Studies*, 9:1–36.

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

Keim, D. B. and Madhavan, A. (1998). The cost of institutional equity trades. *Financial Analysts Journal*, 54(4):50–69.

LaPlante, M. and Muscarella, C. J. (1997). Do institutions receive comparable execution in the NYSE and Nasdaq markets? A transaction study of block trades. *Journal of Financial Economics*, 45:97–134.

Lee, C. M. (1993). Market integration and price execution for NYSE–listed securities. *Journal of Finance*, 48:1009–10038.

Lewellen, J. (2002). Momentum and autocorrelation in stock returns. *Review of Financial Studies*, 15(2):533–563.

Madhavan, A. (1992). Trading mechanisms in securities markets. *Journal of Finance*, 48:607–641.

Mendelson, H. (1987). Consolidation, fragmentation, and market performance. *Journal of Financial and Quantitative Analysis*, 22:189–207.

Perold, A. (1988). The implementation shortfall: Paper versus reality. *Journal of Portfolio Management*, 14(3):4–9.

Ruyter, T. (1999). The dinosaur, the rock, and the hard place. Plansponsor Magazine. March.

Seppi, D. (1990). Equilibrium block trading and asymmetric information. *Journal of Finance*, 45:73–94.

Treynor, J. (1981). What does is take to win the trading game? *Financial Analysts Journal*, 47(1).

Wagner, W. H. and Edwards, M. (1993). Best execution. *Financial Analysts Journal*, 49:65–71.

Chapter 3

Equity trading by institutional investors: Evidence on order submission strategies

Written with Johannes A. Skjeltorp[1]

Abstract

The trading volume channeled through off-market crossing networks is growing. Passive matching of orders outside the primary market lowers several components of execution costs compared to regular trading. On the other hand, the risk of non-execution imposes opportunity costs, and the inherent "free riding" on the price discovery process raises concerns that this eventually will lead to lower liquidity in the primary market. Using a detailed data set from a large investor in the US equity markets, we find evidence that competition from crossing networks is concentrated in the most liquid stocks in a sample of the largest companies in the US. Simulations of alternative trading strategies indicate that the investor's strategy of initially trying to cross all stocks was cost effective: in spite of their high liquidity, the crossed stocks would have been unlikely to achieve at lower execution costs in the open market.

[1]The article is published in Journal of Banking & Finance 27 (2003) 1779-1817. A couple of minor editorial changes have been made in the current version

1 Introduction

There is currently a plethora of venues for trading US equities. Some fit the needs of small retail investors while others are more suited for the needs of large institutional investors and portfolio managers.[2] Using a detailed data set from a large institutional investor, we investigate the nature of competition between a principal exchange and one particular type of alternative trading system, the crossing network. A crossing network is a satellite trading place: it uses prices from the primary market and merely matches quantities. Passive matching of orders implies that several components of execution costs are low compared to regular market trading: commissions are lower and there are no spread costs or direct price impact costs. On the other hand, traders are not guaranteed execution in the network, and this may lead to significant opportunity costs. In addition, the execution probability may or may not be associated with adverse selection costs, depending on the type of traders in the crossing network.[3] Finally, because crossing networks compete for order flow, crossing participants may eventually incur implicit price impact costs as a result of reduced primary market liquidity. The largest crossing markets in the US include POSIT (ITG), NYSE crossing session I and II, and Instinet Global Crossing. In addition, there are less public internal crossing networks, many of which are the exclusive domains of institutional investors.[4]

Investigating execution costs associated with different trading methods is of obvious interest to investors seeking cost effective ways to trade. However, the functioning of alternative trading systems should also be of interest to academics, regulators and policy makers responsible for the design of securities markets. The recent success of electronic trading venues has intensified the competition for order flow faced by the traditional markets. In general, the increase in competition is positive because it lowers execution costs. Several empirical studies find that transaction costs decreased over the recent past.[5] However, increased competition for order flow

[2]The trading venues can be broadly classified into four groups: (i) the principal exchanges, (ii) the "over the counter" (OTC) markets, (iii) other exchanges and (iv) alternative trading systems (ATS). The *principal exchanges* include the NYSE and the NASDAQ/NNM. The *OTC markets* includes the OTC bulletin board and the "pink sheet" market. The OTC bulletin board is for companies too small to list on the NNM, and the "pink sheet" market is an internet quotation service for *very* small companies operated by Pink Sheets LLC. *Other exchanges* include the AMEX, the regional exchanges in Boston, Philadelphia, Pacific, and Chicago(Midwest), and the Cincinnati Stock Exchange. Finally the *ATSs* include Electronic Communication Networks (ECNs), the Arizona Stock Exchange, and external and internal crossing networks.

[3]A stated goal of many crossing networks is to keep the identities and trades of their participants anonymous, both before and after the trades. The following example is taken from the Instinet homepage: "With Instinet Global Crossing, the process is anonymous. Pre-trade or post-trade, neither your trading partner nor other market participants will know your identity, strategy, order size, or residual size."

[4]POSIT is by far the largest crossing market and facilitated the crossing of 7.8 billion shares in 2000 and 9.3 billion shares in 2001. POSIT performs eight daily matches at the price equal to the bid-ask midpoint of the stock's primary market at fixed times which are randomized within 5 minutes to avoid manipulation. The NYSE after hours crossing session I allows participants to submit orders until 5pm when the orders are matched using the NYSE closing price for each stock. The NYSE crossing session II is designed to facilitate trading of baskets of at least 15 NYSE securities valued at USD 1 million or more. Instinet Global Crossing began in 1986 as the first electronic crossing service in the US. Currently, its operations facilitate "end-of-day crossing" and "VWAP crossing". The "end-of-day crossing" crosses orders at the closing price in the primary market, while the "VWAP crossing" is settled before the opening of the primary market and the participants are guaranteed the VWAP price during the day.

[5]For a survey on research on transaction costs, see Keim and Madhavan (1998).

has also raised some concerns related to potentially negative effects from market fragmentation. Mendelson (1987) shows that market fragmentation has both costs and benefits. The costs are related to reduced liquidity and increased volatility in each "sub-market", while the benefits are related to increased quality of the market price signals. Because crossing networks do not contribute to price discovery, the potential benefits from better price signals are lost and only the potential costs from low liquidity and high volatility are left. These costs might also eventually harm participants in crossing networks through their reliance on primary market prices. A better understanding of the nature of the competition between crossing networks and primary markets is clearly called for, including under what circumstances and for which types of assets crossing networks will coexist with other markets.

Three recent empirical papers on alternative trading systems are Fong et al. (1999), Næs and Ødegaard (2000), and Conrad et al. (2001b). Fong et al. (1999) use detailed data from the Australian stock exchange (ASX) to study the competition between exchanges and different off-market trading mechanisms, including crossing networks (POSIT Australia). Off-market trading is found to be concentrated in the most liquid stocks. The cross-sectional differences in off-market trading seem to be driven by institutional trading interest (trading volume, index inclusion), primary market liquidity (spreads, market depth, introduction of closing auction market), and the existence of a derivative market. Conrad et al. (2001b) study explicit and implicit execution costs on externally crossed orders, orders sent to ECNs, and broker-filled orders based on a large data set from the US equity market provided by the Plexus Group.[6] Conrad et al. (2001b) also find that the most liquid stocks are the ones underlying the orders sent to external crossing systems. Moreover, the average total trade cost is found to be substantially lower for orders sent to external crossing systems and ECNs than for orders filled by traditional brokers.

Both papers suggest that crossing networks provide significant competition for order flow, especially in highly liquid stocks, and considerably lower execution costs than other trading methods. On the other hand, as hypothesized in Keim and Madhavan (1998) and Hendershott and Mendelson (2000), informed traders may be present in crossing networks, offsetting their explicit cost advantage. The existence of adverse selection costs is hard to detect based on the cost measures used in the empirical literature and the data typically available to researchers, such as the data from the Plexus Group used in Conrad et al. (2001b). Using a special data set, the relation between execution probability and adverse selection is studied in Næs and Ødegaard (2000). They find that, over the month following an attempt at crossing, there is a one percent difference in risk adjusted returns between stocks that were successfully crossed and stocks that had to be purchased in the market. This finding is interpreted as evidence that the benefits of lower costs in crossing networks are mitigated by costs related to adverse selection.

In this paper, we extend the analysis of Næs and Ødegaard using the same data set. The data set includes all orders from the establishment of a US equity portfolio worth USD 1.76

[6] The sample consist of 797,068 orders submitted by 59 institutions between the first quarter of 1996 and the first quarter of 1998.

billion over a 6-month period from January 1998 to June 1998. The portfolio was tracking the US part of the FTSE All World index[7], which consists of about 500 stocks, and has a very high correlation with the S&P 500 index. The data set is unique in that it contains information on the investors' complete order submission strategy, including the ex ante trading strategy, the dates on which the decision to trade was made, and the resulting fill rates of each order for different trading venues. Hence, the data set is close to a "controlled experiment" which is quite rare.[8] The weakness of the data set is that it is from one trader's buy orders only and covers a limited period of time. Both Fong et al. (1999) and Conrad et al. (2001b) have access to huge data sets on orders and trades and their results are therefore more robust than ours. However, we show that the investor in our study is quite representative for large institutional investors in the US markets.

First, we try to investigate the evidence of adverse selection more closely. On the one hand, the available empirical evidence suggest that crossing networks are competing in the most liquid stocks. If stocks that are not supplied in crossing networks are less liquid in general, then these stocks might need a higher return to induce investors to hold them, and the abnormal performance of the non-crossed stocks found in Næs and Ødegaard (2000) might be explained (or partly explained) by a liquidity premium.[9] On the other hand, a liquidity and an information story need not be mutually exclusive. First, in addition to being a proxy for differences in liquidity, a wider spread may also capture a higher adverse selection component. Furthermore, other measures of liquidity, such as depth, may also capture the effect that uninformed investors withdraw from the market if they are worried about being picked off by better informed investors. Thus, a difference in liquidity between the two groups of stocks may capture the same effect as found in Næs and Ødegaard (2000), but by using different proxies for adverse selection. An interesting question in this respect, is whether the liquidity characteristics are temporary or more systematic over time. Because information asymmetries are expected to vanish relatively quickly, it would be harder to interpret a systematic liquidity difference as a sign of adverse selection, especially for the largest companies in the US market. On the other hand liquidity differences may be more permanent in nature.

We investigate these questions by calculating a whole range of liquidity and activity measures in the primary market across the groups of stocks that were supplied/not supplied in the crossing network.[10] Our results indicate that the difference in abnormal return between the two groups of stocks may be explained by both liquidity differences and private information. On the one hand, we find support for the earlier finding that crossing networks are competing in the very liquid segment of listed US equities. Stocks that are successfully crossed are significantly

[7] The FTSE All-World index includes 49 different countries and about 2300 stocks. The aim of the index is to capture up to 90% of the investible market capitalization of each country.

[8] In many other studies, the exact investment strategy of a trader has to be estimated from the sequence of trades. This induces a selection bias in the data. It might be that the trader has decided to send the most difficult orders to brokers and the least difficult orders to crossing networks. We are not facing a selection bias problem in our data set.

[9] Amihud and Mendelson (1986) show that risk-adjusted returns for stocks and bonds are increasing in their illiquidity, measured by the spread.

[10] We use the crossing success of the Fund as a proxy for supply in the crossing network.

more liquid and more actively traded in the primary market than stocks that are not crossed. Moreover, we also show that the differences in liquidity and activity between the two groups of stocks are not date specific, but rather systematic throughout the entire period examined. On the other hand, the difference in spread between the groups of stocks is sometimes significant even though the measures of activity are equal. Following Easley et al. (1996b), this is evidence of informed trading in the stocks that could not be crossed. In addition, we show that the stocks in our sample have a very high correlation with the S&P 500 index. It is hard to believe that liquidity differences between the 500 largest and most liquid companies in the US can explain a difference in abnormal performance between the two groups of stocks of 1 percent over 20 days.

Second, we want to investigate the costs of following alternative submission strategies. This is done by simulating the set of equilibrium order submission strategies for liquidity traders in the Hendershott and Mendelson (2000) model. Our simulated strategies are based on real historical price/volume paths of the stocks traded. This is possible to do because we know the dates when the decision to trade was made in addition to the desired quantities. The simulations confirm the result that crossed and non-crossed stocks have different liquidity characteristics. The stocks that are not obtained through crossing are also the most difficult and expensive stocks to acquire in the market. More interestingly, we find that the actual crossing strategy was inexpensive. Even though the crossed stocks were among the most liquid stocks on the NYSE, it would have been very hard to achieve lower execution costs by submitting limit orders for the same stocks on the same dates that they were first tried to be crossed.

The paper is organized as follows. In section 2, we describe our data set. We first give a short description of the investor and the crossing strategy. Then we provide some descriptive statistics establishing that the investor is indeed representative for the group of large institutional traders in the US equity market. In section 3, we discuss the relationship between execution probability and several measures of primary market liquidity. Section 4 contains a description of the methodology and results from the simulation approach. Section 5 provides our conclusions.

2 The data

Our data set contains transactions data from an actual submission strategy carried out in the US equities market by a large institutional investor, the Government Petroleum Fund in Norway (hereafter "the Fund"). To construct liquidity measures and simulate other submission strategies, we use additional transaction data from the NYSE Trades and Quotes database (TAQ), which contains all the trades and quotes for stocks listed on the NYSE, American Stock Exchange (AMEX) and NASDAQ's National Market System. In this section, we first give a short description of the Fund and explain the opportunistic crossing strategy in some more detail. We then provide some descriptive statistics to establish that the Fund is representative for the group of large institutional traders in the US equity market.

2.1 The trading strategy

The Fund is a vehicle for investing the Norwegian Government's income from petroleum-related activities in international capital markets. Initially, the Fund was invested in foreign government securities only. However, new criteria, applying from January 1998, stated that between 30 and 50 percent of the Fund portfolio was to be invested in equities. The composition of the Fund portfolio was changed to include equities during the first half of 1998. We use transaction data for the part of the portfolio that was invested in US equities during this "buildup"/transition period.

The investment universe for the equity portfolio includes at present 28 countries in Europe, America, and Asia. US stocks represent around 29 percent of the total stock portfolio. Benchmark portfolios consist of the companies in the FTSE All-World index for these countries.[11] The US part of the index currently consists of about 480 different securities. The constituents of this index are the largest companies in the US market, and the index has a very high correlation with the S&P 500 index.

The equity portion of the total benchmark portfolio was set to 8 percent at the end of January 1998, and was then increased by another 8 percentage points at the end of each subsequent month until it reached the benchmark weight of 40 percent in June. The maximum tracking error restriction implied that the Fund was pre-committed to buy most of the stocks in the index every month.

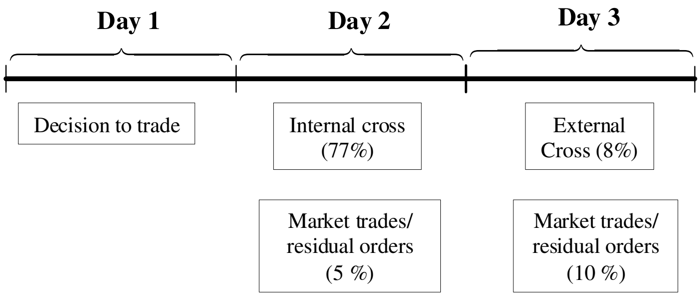
The Fund employed four index managers to establish the portfolio. One of the index managers was chosen as "transition manager". The order submission strategy was as follows: First, try to find sellers among the customers of the transition manager (internal crossing). If this is not possible, search for counterparties among the customers of the other three index managers or send the order to an electronic crossing network (external crossing). Finally, purchase residual orders that cannot be crossed (if any) in the primary market. According to the discussion in Ruyter (1999), this is the typical order submission strategy large index managers follow for their customers. Figure 3.1 illustrates the actual implementation of the Fund's order submission strategy.

The total portfolio investment was USD 1751 million. The Fund went to the primary market with USD 250 million, or 14 percent, of this investment. We do not know what part of the externally crossed orders that were sent to an electronic crossing network rather than being crossed with one of the Fund's index managers. The majority of the crossed orders, USD 1356 million of USD 1501 million, was executed internally. Market trades to complete the desired portfolio were needed on three of a total of sixteen trading dates. Looking at the transactions data, it turns out that at some occasions market trades happened on the same date as internal crosses. According to the order submission strategy, these orders should be sent to external crossing before they were sent to the market. We do not know if this was done. Hence, it might be that the submission strategy was not strictly followed with respect to the stage with external

[11] These indices used to be called the FT/S&P's Actuaries World Index.

Figure 3.1: Implementation of the Fund's order submission strategy

The Fund's order submission strategy was as follows: First, try to find sellers among the customers of the transition manager (internal crossing). If this is not possible, search for counterparties among the customers of the other three index managers or send the order to an electronic crossing network (external crossing). Finally, purchase residual orders that cannot be crossed (if any) in the primary market. The figure illustrates the actual implementation of the order submission strategy followed by the Fund. The overall part of the orders were crossed internally. All orders were executed within two days after the decision to trade. At some occasions market trades happened on the same date as internal crosses. We do not know if these orders were sent to external crossing before they were sent to the market. Hence, it might be that the submission strategy was not strictly followed with respect to the stage with external crossing. The numbers in parentheses are percent of total portfolio investment (USD 1 751 billion).



crossing. The highest trading volume on one date amounted to USD 300 million, or 17.1% of the total portfolio investment. Note that for the period we are considering the Fund was only buying, not selling securities. For the first two months, crossing prices were set as the primary market (NYSE/NASDAQ) closing prices that day. For the remainder of the period, prices were set as the volume weighted average price (VWAP) of trades in the primary market during the day.

2.2 Robustness

Our study is based on the trades of only one institution. It is therefore of crucial importance that the investor is representative for the group of institutional investors used in other studies dealing with similar issues.

The Fund used Barclay Global Investor (BGI) as a transition manager. According to Harris (2003), BGI's internal crossing network is "probably the largest in the world". Hence, both the manager and the private network, where most of the actual crossing was performed, are representative for the US market.

Most recent empirical studies of institutional investors' in the US equity market use data provided by the Plexus Group. These studies include Keim and Madhavan (1995, 1997), Jones and Lipson (1999a,b) and Conrad et al. (2001a,b). The Plexus Group is a consulting firm that monitors the costs of institutional trading. The data sets used in Jones and Lipson (1999a,b) are limited to trades executed in some specific firms. The most relevant samples of institutional investors with which to compare the Fund's trades are therefore the ones used in Keim and Madhavan (1995, 1997) and Conrad et al. (2001b).

Keim and Madhavan (1995, 1997) use data on all equity transactions of 21 institutional investors from January 1991 through March 1993. This data set contains a total of 62,333 orders. The institutions vary in size. For fundamental value managers, the mean dollar value of assets under management was USD 4.8 billion, ranging from a low of USD 0.7 billion to a high of USD 12.9 billion. For index managers and technical traders, the mean dollar value of assets under management was USD 3.2 billion and USD 5.3 billion respectively.[12] In the period we are examining, the Fund was an index tracker, and, at the end of June 1998, the US equity portfolio was worth USD 1.7 billion. Conrad et al. (2001b) have a larger data set from a more recent time period. Their sample consists of 797,068 orders submitted by 59 institutions between the first quarter of 1996 and the first quarter of 1998.

If we first look at order size, our median order is for USD 174,000. As table 3.1 shows, this is slightly larger than the median buy order of USD 138,000 in Keim and Madhavan (1995, 1997), and much larger than the crossed and ECN filled orders in Conrad et al. (2001b). One of the reasons for this may be that the orders routed through ECNs are generally much smaller than orders routed through crossing networks. The average dollar value of the Fund's orders of USD 386,000 is also higher than the average dollar value of the orders sent to external crossing and ECNs, but considerably lower than the average dollar value of the orders filled by brokers and multiple order mechanisms.

Since the Fund was tracking the US stocks included in the FTSE All-World index, the stocks in the sample are obviously the more liquid stocks in the market. The most liquid stocks in Conrad et al.'s study are the ones underlying the orders sent to external crossing systems. These securities have an average market cap of USD 12.7 billion, while the average market cap for the stocks purchased by the Fund was USD 16.9 billion. Hence, the Fund was clearly trading in the larger companies.

One more characteristic with our data set is worth noting. Unlike most other studies, there is no selection bias in our data set. The Fund did not select what orders to send to the crossing network and what orders to send to the market based on a perception of trade difficulty.

3 Execution probability and primary market liquidity

In this section, we analyze in detail the relation between the probability of getting a stock crossed and the liquidity and trading activity in the primary market. This is possible because we know that the Fund initially tried to cross all the stocks. The data set therefore reveals the date and identity of stocks that could not be crossed. We find that there are indeed significant differences in liquidity and activity between the two groups of stocks based on a wide range of liquidity and activity measures. Moreover, most of the liquidity and activity measures we calculate are

[12]Fundamental value managers are defined as managers whose investment strategies are based on assessment of long-term fundamental values, technical managers are defined as managers whose strategies are based on capturing short-term price movements, and index managers are defined as managers who seek to mimic the returns of particular stock indexes (Keim and Madhavan (1997)).

Table 3.1: Descriptive statistics for traded securities

In this table, we make a comparison of the data used in this study and in Keim and Madhavan (1995, 1997) and Conrad et al. (2001b). In our study and in Keim and Madhavan (1995,1997), the numbers are for buyer-initiated trades only. "Multiple mechanism orders" in the Conrad et al. (2001b) paper are orders in which more than one of the three trading mechanisms (brokers, ECNs or external crossing systems) are used to fill the order. Market cap values are in USD billion. "Listed %" is the percentage of total orders that is in listed stocks. "*n*" is the total number of orders.

	Order size				Liquidity			
	Dollar value		No. of shares		Market cap		Listed	n
	mean	med.	mean	med.	mean	med.	%	
Our study								
All orders	386	174	6 898	3 800	16.9	7.5	100	4 200
- Cross	396	177	7 013	3 800	17.6	7.8	100	3 494
- Market order	339	157	6 329	3 550	13.6	6.1	100	706
KM [1995,1997]								
All orders		138		4 800		1.1	82.6	36 590
Conrad et al. (2001b)								
All orders								723 998
- External cross	187	45			12.8		>90.0	112 159
- ECN's	194	53			3.0			51 127
- Broker filled	1474	137			11.1			560 712

significantly different across the groups of stocks, both on the days when they were first crossed and for the month prior to and after the actual trading dates. These results are confirmed in a probit model. After a proper orthogonalization of the independent variables, the probability of a successful cross is shown to be higher the lower the effective spread, the higher the liquidity ratio, and the higher the dollar trading volume in the primary market.

3.1 Liquidity measures

Market liquidity is a comprehensive concept that covers several transactional properties of the marketplace. Harris (1990) defines four interrelated dimensions of the concept: width, depth, immediacy and resiliency. *Width* is defined as the bid-ask spread for a given number of shares, and measures the cost per share of liquidity. *Depth* is defined as the number of shares at the bid-ask quotes, *immediacy* describes how fast a trade for a given number of shares can be executed, and *resiliency* describes how fast the price reverts to its "true" value after order flow imbalances caused by liquidity trading has moved prices temporarily away from the "true" level. We try to capture the width, depth and resiliency dimensions by calculating several spread, volume, and volatility measures.[13]

[13] A discussion of data issues and the formulas for calculating the different liquidity and activity measures are provided in appendix A.

Spread measures We consider three measures of the spread to capture the width of the market. The most commonly used spread measure is the *quoted dollar spread*. It measures the average difference between the inside quoted ask and bid for a stock over the trading day and can be thought of as the absolute "round trip" cost of trading a small amount of shares at the inner quotes. The *quoted percentage spread* is calculated as the quoted spread relative to the spread midpoint, or the "true" value, at each trade time. The *effective spread* takes into account the fact that trades are often executed inside (price improvement) or outside the spread ("walking the book"), and is often considered a more appropriate measure of trading costs than are quoted spreads, especially for large trades.[14] The effective spread is calculated as the average absolute dollar difference between the execution price and the bid ask midpoint multiplied by two. The spread measures the handling of a single trade, and does not capture the ability of a market structure to absorb a series of trades without perturbing prices excessively. We therefore need to supplement the spread estimates with measures of depth and volatility.

Depth and resiliency To capture market depth and resiliency, we calculate the average quoted number of shares at the inner quotes and the daily and intraday Amivest liquidity ratio.[15] The daily liquidity ratio reflects the average trading volume that would be needed to move the price by one percent during a trading day, while the average intraday liquidity ratio measures the same ability over 15 minute intervals. A high liquidity ratio indicates ability of the market to absorb large trades without affecting the price.[16] To get a broader picture of the volume and trading activity in the primary market across the groups of stocks, we also calculate total shares traded, the dollar value of shares traded, and the average trade size.

Volatility As an additional liquidity measure we calculate two measures of volatility. Volatility captures a dimension of liquidity in the sense that high depth at the inner quotes makes the trade prices less volatile since there is more depth to absorb the liquidity demand. The first volatility measure we calculate is the standard deviation of daily returns over the 10 days prior to the date when the Fund was trying to cross the stock. The other measure tries to capture the intraday volatility (15 minute return standard deviation) in each stock. When interpreting short term volatility, it is important to keep in mind that the sources of volatility may vary. From the viewpoint of a trader, high volatility can increase the probability of filling a limit order. This could attract liquidity suppliers to volatile stocks. However, high volatility may also be associated with news and informed trading so that the risk of an adverse price movement after a fill is higher ("pick off risk"). Furthermore, informed trading would also induce the specialist to increase his spread which would make the trading costs higher. From a liquidity perspective, high volatility may also be a sign of low liquidity in the sense that the market is unable to absorb large trades without excessive price movements.

[14] See for example Angel (1997) and Bacidore et al. (1999).

[15] Amivest Capital Management introduced this measure of liquidity.

[16] This ratio is applied in several studies (see e.g. Khan and Baker (1993), Amihud et al. (1997)).

3.2 Results

In order to investigate whether stocks that are easy/hard to cross have different liquidity and activity characteristics, we split the orders into three categories on each sample date: (i) *Crossed stocks*: orders in this group were fully crossed, (ii) *Cross/Market*: orders in this group could not be fully crossed, and the residual order was purchased in the open market the next day, and (iii) *Market stocks*: orders in this group could not be crossed at all, and the whole order was therefore purchased in the open market the next day. A market trade means that the Fund was either "crowded out" by other traders who wanted to buy the stock or (the rather unlikely case) that the supply of the stock in the network was less than the size of our order.

Table 3.2 shows the different liquidity measures for the three order categories on two of the three dates when the Fund was not able to obtain all the required stocks in the crossing network.[17] In table 3.3 we have averaged the liquidity measures in table 3.2 according to the number of stocks traded by the Fund on each date. To examine whether our sample of stocks differs from the stocks in the S&P 500 index, we calculate the average liquidity measures for the S&P 500 index over the same dates as well as for the entire period when the Fund was trading (first half of 1998). For each liquidity measure, we perform tests for differences in means between the S&P 500 index stocks and the stocks purchased by the Fund. Except for the quoted percentage spread and the volatility measures, none of the liquidity measures are significantly different at the 1% level. Hence, the two samples have quite similar liquidity and activity characteristics. We also find that the S&P 500 stocks average for the entire half year is not significantly different from the S&P 500 stocks average on the particular dates when the Fund was trading.

The numbers in both tables strongly indicate that stocks that were easy to cross had lower spread costs than stocks that were hard to cross. The average spread difference is 22%, which is both economically and statistically significant. Interpreting spreads as a proxy for liquidity, this means that stocks that could not be crossed were less liquid than the stocks supplied in the crossing network. The group of non-crossed stocks was also less liquid measured by the intraday and daily liquidity ratios. Moreover, measured by the number of trades, the trading volume, and the number of shares traded, the trading activity was lower in the non-crossed stocks over the entire sample.[18] Stocks that were hard to cross were also more volatile than stocks that were easy to cross. As we would expect, the liquidity of the stocks underlying the group of orders that were partly crossed and partly filled in the market lies in between the two other groups.

Using the result in Easley et al. (1996b) that higher spread for stocks with similar trading volume is an indication of informed trading, our results give some support to the evidence of informed trading in the crossing network found in Næs and Ødegaard (2000). On the other hand, if there are systematic differences in liquidity between the two groups of stocks also on other dates, this would be less supportive to an informed trading story. To check this, we calculate the

[17] We do not report the liquidity measures separately for one of the three days because the number of orders purchased in the market on this day was too small to perform reliable statistical tests of the differences between the two groups.

[18] This difference was insignificant for one of the trading dates, however.

Table 3.2: Liquidity in the primary market on the trading dates

The table shows different measures of liquidity and activity in the primary market on the dates when the Fund did not fill all orders in the crossing network. “Crossed stocks” means that the whole order of a stock was crossed. “Crossed/Market” means that part of the order was crossed and part of the order had to be purchased in the open market. “Market stocks” means that the stock could not be crossed at all. The calculation and explanation of the different measures are found in Appendix A. The t-stat and p-value are the test statistics of a two-sided t-test, where the null is that the mean for the ”Crossed stocks” and ”Market stocks” are equal. Similarly, for the “Crossed/Market” group, the null is that the mean for the “Crossed stocks” and ”Crossed/Market” stocks are equal. The test depends on whether the population variances of the two groups are equal or not. If the variances are equal, then the t-stat is calculated as $t = (\bar{x}_c - \bar{x}_m)/\sqrt{s^2(1/n_c + 1/n_m)}$ where $\bar{x}_c$ and $\bar{x}_m$ are the means for the two groups respectively, n_c and n_m are the number of stocks in each group while s^2 is the pooled standard deviation calculated as $s^2 = [(n_c - 1)s_c^2 + (n_m - 1)s_m^2]/[n_c + n_m - 2]$, where s_c^2 and s_m^2 are the standard deviation of measure for the cross and market stocks respectively. We use the SAS package to perform all tests. If the variances are significantly different, the standard approximation supplied in SAS is used. For the *Daily volatility* measure, we use an F-test to test for differences in variance between the two groups, where the null is that the ratio of the two sample variances is equal to 1.

DATE 1	S&P 500 stocks	Fund stocks	Crossed stocks	Market stocks	Diff. test p-value	Crossed/ Market	Diff. test p-value
Spread measures							
Effective spread	0.1112	0.1063	0.0931	0.1118**	0.0103	0.0893	0.6514
Quoted USD spread	0.1322	0.1315	0.1135	0.1395**	0.0069	0.0910	0.0969
Quoted % spread (midp.)	0.3270	0.2566[a]	0.1916	0.2852**	<.0001	0.1200*	0.0111
Volume measures							
Trades	807	861	1317	575**	0.0002	4985**	<.0001
Shares traded (1000)	1180	1274	2039	868**	0.0001	5487**	0.0007
Volume (USD mill.)	61	67	116	39**	<.0001	434**	<.0001
Trade size (USD 1000)	79	85	88	67**	0.0011	103	0.6159
Liquidity ratios and depth							
Daily LR (USD mill.)	117	148	293	76**	<.0001	785	0.1221
Intraday LR (USD mill.)	13	16	28	8**	<.0001	103	0.0576
Depth at quotes (shares)	1198	1841	2126	1692	0.0965	3351	0.1942
Volatility and return							
Daily volatility	0.0275	0.0256[a]	0.0225	0.0265**	<.0001	0.0342**	<.0001
Intraday volatility (%)	0.2601	0.2367[a]	0.2573	0.2296*	0.0307	0.2044*	0.0361
N stocks	454	368	100	261		7	

DATE 2	S&P 500 stocks	Fund stocks	Crossed stocks	Market stocks	Diff. test p-value	Crossed/ Market	Diff. test p-value
Spread measures							
Effective spread	0.1174	0.1139	0.1027	0.1327**	0.0082	0.1039	0.7636
Quoted USD spread	0.1396	0.1420	0.1299	0.1605*	0.0445	0.1326	0.5257
Quoted % spread (midp.)	0.3903	0.3375[a]	0.3255	0.3724*	0.0380	0.3136	0.5391
Volume measures							
Trades	737	678	515	692	0.1830	763*	0.0447
Shares traded (1000)	1015	929	847	875	0.8390	1025	0.2390
Volume (USD mill.)	53	48	40	46	0.5469	54	0.1914
Trade size (USD 1000)	68	67	78	66**	0.0065	67	0.0917
Liquidity ratios and depth							
Daily LR (USD mill.)	100	101	81	92	0.5182	120	0.0598
Intraday LR (USD mill.)	10	9	8	8	0.9397	11	0.2481
Depth at quotes (shares)	1572	1506	1524	1464	0.7546	1532	0.9676
Volatility and return							
Daily volatility	0.0220	0.0263[a]	0.0250	0.0271**	0.0025	0.0263*	0.0449
Intraday volatility (%)	0.3494	0.3217[a]	0.3298	0.3220	0.6551	0.3167	0.4233
N stocks	454	478	114	171		193	

a Equality of the measure between the S&P 500 and Fund stocks is rejected at the 5% level.
* Equality of the measure between the Crossed and Non-crossed stocks (Market stocks) is rejected at the 5% level.
** Equality of the measure between the Crossed and Non-crossed stocks (Market stocks) is rejected at the 1% level.

Table 3.3: Average liquidity over all stocks

The first and second columns in the table show the cross-sectional average measures for the S&P 500 constituent stocks during the first half of 1998 (S&P 500 June-July) and on the same dates as for the Fund trades (S&P 500). The third column (Fund) is the cross-sectional average for all the stocks that the Fund traded on the two dates. The t-stat and p-value are the test statistics of a two-sided t-test, where the null is that the mean for the "Crossed stocks" and "Market stocks" are equal. Similarly, for the "Crossed/Market" group, the null is that the mean for the "Crossed stocks" and "Crossed/Market" stocks are equal. The test depends on whether the population variances of the two groups are equal or not. If the variances are equal, then the t-stat is calculated as, $t = (\bar{x}_c - \bar{x}_m)/\sqrt{s^2(1/n_c + 1/n_m)}$ where $\bar{x}_c$ and $\bar{x}_m$ are the means for the two groups respectively, n_c and n_m are the number of stocks in each group while s^2 is the pooled standard deviation calculated as $s^2 = [(n_c - 1)s_c^2 + (n_m - 1)s_m^2]/[n_c + n_m - 2]$, where s_c^2 and s_m^2 are the standard deviation of measure for the cross and market stocks respectively. We use the SAS package to perform all tests. If the variances are significantly different, the standard approximation supplied in SAS is used. For the *Daily volatility* measure, we use an F-test to test for differences in variance between the two groups, where the null is that the ratio of the two sample variances is equal to 1.

ALL TRADES	S&P 500 Jan-July	S&P 500 stocks	Fund stocks	Crossed stocks	Market stocks	Diff. test p-value	Crossed/ Market	Diff. test p-value
Spread measures								
Effective spread	0.1151	0.1143	0.1106	0.0982	0.1201**	0.0006	0.1034	0.1157
Quoted $ spread	0.1360	0.1359	0.1374	0.1222	0.1478**	0.0021	0.1312*	0.0204
Quoted % spread (midp.)	0.3414	0.3586	0.3023[a]	0.2629	0.3197**	<.0001	0.3068**	0.0037
Volume measures								
Trades	758	772	758	890	621*	0.0255	911	0.8897
Shares traded (1000)	1073	1097	1079	1404	871**	0.0015	1181	0.2454
Volume (mill. USD)	58	57	56	76	42**	<.0001	68	0.5154
Trade size (1000 USD)	77	74	74	83	67**	0.0004	68**	<.0001
Liquidity ratios and depth								
Daily LR (mill.USD)	116.62	108.23	121.45	179.70	82.28**	<.0001	143.74	0.2189
Intraday LR (mill.USD)	12.69	11.81	12.02	17.68	8.34**	<.0001	13.92	0.1771
Depth at quotes (shares)	1851	1385	1652	1805	1601	0.2018	1596	0.2369
Volatility and return								
Daily volatility	0.0238	0.0250	0.0261[a]	0.0241	0.0269**	<.0001	0.0266**	<.0001
Intraday volatility (%)	0.2785	0.3048	0.2847[a]	0.2960	0.2662**	0.0078	0.3128	0.2163
Average stocks per date	452	454	423	107	216		100	

a Equality of the measure between the S&P 500 and Fund stocks is rejected at the 5% level.
* Equality of the measure between the Crossed and Non-crossed stocks (Market stocks) is rejected at the 5% level.
** Equality of the measure between the Crossed and Non-crossed stocks (Market stocks) is rejected at the 1% level.

liquidity measures on each date across a window stretching from 20 business days before to 20 business days after the actual trading date. The results are shown in figure 3.2 with the values and tests in table 3.4. As can be seen from the figure and table there are systematic differences in most of the liquidity and activity measures. A notable exception is the intraday volatility measure which is quite similar between the two groups, except on the actual trade date when it is significantly higher for the crossed stocks. If a market cannot absorb trades without large price movements, the intraday volatility increases. If this is the reason for the change in intraday volatility on the trade dates, the stocks that were supplied in the crossing network did experience a decline in primary market liquidity. Note also that the quoted depth is significantly higher for the crossed stocks than for the non-crossed stocks during the days prior to the crossing date, but not significantly different on the actual crossing dates. These findings lend some support to the hypothesis that crossing networks removes order flow from the primary market.

To investigate the relationship between primary market liquidity and the outcome of the attempt at crossing the stocks more formally, we estimate a probit model of the probability of getting a stock crossed as a function of various liquidity measures. More specifically, we assume that the probability of observing a cross is given by the model

$$y = Pr(\text{cross}) = F(\beta_0 + \beta_1 \text{eff\_spread}_i + \beta_2 \text{depth}_i + \beta_3 \text{LR}_i + \beta_4 \text{volume}_i + \beta_5 \text{vola}_i + \varepsilon_i) \quad (3.1)$$

where $F(\cdot)$ is the cumulative normal distribution function, and the β's are coefficients of the

Table 3.4: Time series of liquidity and activity measures over all sample stocks

The time series of liquidity and activity measures of 40 trading days surrounding the actual trade dates for the crossed stocks (C) and non-crossed stocks (M) that were bought in the market in the portfolio. A t-test for differences in means between the two groups is performed for each date for each measure. Numbers in bold indicate a significant difference at the 5% level. The test applied is a two-sided t-test, where the null is that the mean for the "Crossed stocks" and "Market stocks" are equal. The test depends on whether the population variances of the the two groups are equal or not. If the variances are equal, then the t-stat is calculated as $t = (\bar{x}_c - \bar{x}_m)/\sqrt{s^2(1/n_c + 1/n_m)}$ where $\bar{x}_c$ and $\bar{x}_m$ are the means for the two groups respectively, n_c and n_m are the number of stocks in each group while s^2 is the pooled standard deviation calculated as $s^2 = [(n_c - 1)s_c^2 + (n_m - 1)s_m^2]/[n_c + n_m - 2]$, where s_c^2 and s_m^2 are the standard deviation of measure for the cross and market stocks respectively. We use the SAS package to perform all tests. If the variances are significantly different, the standard approximation supplied in SAS is used.

	Effective spread		Quoted spread (USD)		Quoted % spread		Shares traded		Volume USD		Trades		Trade size (Shares)		LR (daily)		LR Intraday		Depth (shares)		Intraday Volatility	
Date	C	M	C	M	C	M	C	M	C	M	C	M	C	M	C	M	C	M	C	M	C	M
-20	**0.094**	**0.114**	**0.121**	**0.142**	**0.239**	**0.309**	**1388**	**832**	**76**	**43**	**879**	**560**	**91**	**77**	**168**	**86**	**17**	**9**	1864	1724	0.250	0.233
-19	**0.097**	**0.114**	**0.122**	**0.140**	**0.241**	**0.310**	**1346**	**798**	**71**	**41**	**916**	**610**	**79**	**67**	**161**	**85**	**16**	**8**	1820	1660	0.278	0.272
-18	**0.093**	**0.115**	**0.121**	**0.143**	**0.241**	**0.312**	**1122**	**771**	**59**	**40**	781	559	78	73	**161**	**86**	**15**	**9**	1931	1634	0.233	0.238
-17	**0.095**	**0.115**	**0.119**	**0.141**	**0.239**	**0.309**	**1308**	**804**	**69**	**42**	**905**	**570**	**77**	**74**	**168**	**83**	**17**	**9**	1962	1745	0.261	0.243
-16	**0.096**	**0.116**	**0.120**	**0.143**	**0.237**	**0.309**	**1235**	**762**	**68**	**40**	**841**	**561**	**80**	**73**	**167**	**82**	**16**	**8**	1974	1654	0.244	0.246
-15	**0.095**	**0.110**	**0.119**	**0.136**	**0.236**	**0.299**	**1295**	**763**	**71**	**40**	**899**	**583**	**75**	**70**	**171**	**80**	**17**	**8**	2154	1860	0.241	0.231
-14	**0.095**	**0.116**	**0.119**	**0.142**	**0.238**	**0.309**	**1309**	**822**	**72**	**43**	853	626	**81**	**69**	**175**	**81**	**17**	**9**	**1939**	**1620**	0.244	0.244
-13	**0.095**	**0.117**	**0.120**	**0.145**	**0.242**	**0.321**	**1356**	**835**	**73**	**43**	**955**	**669**	**79**	**65**	**176**	**81**	**18**	**9**	1799	1562	0.292	0.282
-12	**0.097**	**0.117**	**0.123**	**0.145**	**0.244**	**0.320**	**1400**	**727**	**75**	**37**	**950**	**556**	**82**	**67**	**180**	**82**	**17**	**8**	1773	1567	0.270	0.264
-11	**0.096**	**0.110**	**0.120**	**0.137**	**0.237**	**0.305**	**1375**	**759**	**72**	**38**	**910**	**521**	**82**	**73**	**176**	**88**	**17**	**8**	**2094**	**1572**	0.257	0.242
-10	**0.093**	**0.114**	**0.117**	**0.145**	**0.237**	**0.317**	**1380**	**861**	**70**	**43**	886	631	**80**	**69**	**184**	**86**	**17**	**8**	**1836**	**1541**	0.283	0.262
-9	**0.096**	**0.120**	**0.121**	**0.149**	**0.242**	**0.319**	**1441**	**855**	**74**	**46**	**1002**	**641**	**78**	**71**	**189**	**86**	**17**	**9**	1771	1539	**0.275**	**0.249**
-8	**0.093**	**0.118**	**0.118**	**0.147**	**0.239**	**0.319**	**1285**	**876**	**67**	**46**	880	649	**81**	**70**	**188**	**87**	**16**	**10**	**1941**	**1545**	0.259	0.254
-7	**0.093**	**0.118**	**0.119**	**0.145**	**0.237**	**0.316**	**1324**	**862**	**68**	**43**	844	632	**83**	**68**	**191**	**84**	**17**	**8**	1950	1684	0.256	0.249
-6	**0.093**	**0.118**	**0.117**	**0.148**	**0.233**	**0.315**	**1375**	**819**	**70**	**42**	**849**	**582**	**83**	**74**	**184**	**85**	**17**	**9**	**2267**	**1756**	0.246	0.236
-5	**0.094**	**0.129**	**0.119**	**0.154**	**0.228**	**0.318**	1374	965	75	50	944	732	**76**	**68**	**204**	**98**	**20**	**10**	1879	1636	0.241	0.245
-4	**0.095**	**0.117**	**0.121**	**0.148**	**0.236**	**0.316**	**1291**	**846**	**67**	**41**	860	617	**76**	**65**	**185**	**89**	**16**	**8**	**1953**	**1618**	0.248	0.243
-3	**0.095**	**0.116**	**0.118**	**0.144**	**0.234**	**0.307**	**1239**	**874**	**66**	**42**	819	627	**81**	**68**	**182**	**90**	**15**	**8**	1858	1594	0.272	0.255
-2	**0.096**	**0.115**	**0.120**	**0.142**	**0.237**	**0.309**	**1499**	**806**	**76**	**39**	**948**	**562**	**81**	**69**	**176**	**88**	**17**	**7**	**2074**	**1668**	0.283	0.267
-1	**0.096**	**0.122**	**0.120**	**0.145**	**0.237**	**0.315**	**1231**	**779**	**66**	**38**	**854**	**570**	**77**	**67**	**173**	**86**	**16**	**8**	**1955**	**1629**	0.281	0.272
0	**0.098**	**0.120**	**0.122**	**0.148**	**0.263**	**0.320**	**1404**	**871**	**76**	**42**	**890**	**621**	**83**	**67**	**180**	**82**	**18**	**8**	1805	1601	**0.296**	**0.266**
1	**0.095**	**0.117**	**0.122**	**0.146**	**0.244**	**0.316**	**1405**	**824**	**76**	**41**	**877**	**602**	**84**	**68**	**177**	**84**	**19**	**8**	1866	1563	0.276	0.255
2	**0.091**	**0.113**	**0.114**	**0.139**	**0.227**	**0.304**	**1291**	**858**	**71**	**43**	**837**	**613**	**83**	**70**	**179**	**84**	**17**	**9**	**2115**	**1647**	0.275	0.256
3	**0.092**	**0.111**	**0.117**	**0.139**	**0.232**	**0.303**	**1443**	**917**	**83**	**44**	**828**	**606**	**93**	**74**	**183**	**84**	**20**	**9**	**2078**	**1597**	0.259	0.255
4	**0.095**	**0.113**	**0.118**	**0.139**	**0.237**	**0.304**	**1441**	**842**	**80**	**41**	**898**	**554**	**94**	**73**	**186**	**87**	**18**	**8**	**1983**	**1596**	**0.272**	**0.247**
5	**0.090**	**0.112**	**0.116**	**0.139**	**0.232**	**0.304**	**1258**	**795**	**69**	**39**	**871**	**596**	**78**	**65**	**188**	**88**	**18**	**8**	1883	1582	0.237	0.245
6	**0.094**	**0.112**	**0.119**	**0.140**	**0.236**	**0.302**	**1511**	**946**	**84**	**48**	**971**	**671**	**87**	**72**	**189**	**87**	**21**	**10**	1799	1545	0.265	0.266
7	**0.095**	**0.113**	**0.120**	**0.141**	**0.236**	**0.307**	**1310**	**874**	**72**	**45**	842	626	**83**	**70**	**190**	**88**	**17**	**9**	**1938**	**1541**	0.273	0.260
8	**0.097**	**0.117**	**0.122**	**0.146**	**0.240**	**0.313**	**1311**	**869**	**75**	**44**	**848**	**629**	**86**	**69**	**189**	**85**	**17**	**9**	1741	1576	0.286	0.280
9	**0.094**	**0.115**	**0.121**	**0.148**	**0.238**	**0.320**	**1047**	**641**	**58**	**32**	**745**	**455**	**76**	**69**	**188**	**88**	**16**	**7**	1747	1514	0.246	0.235
10	**0.094**	**0.118**	**0.115**	**0.142**	**0.229**	**0.303**	**1299**	**804**	**74**	**40**	**853**	**559**	**82**	**72**	**191**	**89**	**18**	**8**	**1987**	**1557**	0.266	0.264
11	**0.097**	**0.118**	**0.121**	**0.145**	**0.239**	**0.308**	**1547**	**920**	**84**	**46**	**905**	**587**	**94**	**77**	**185**	**89**	**19**	**9**	1839	1588	0.278	0.273
12	**0.095**	**0.116**	**0.118**	**0.144**	**0.235**	**0.305**	**1521**	**889**	**84**	**44**	**987**	**636**	**85**	**71**	**183**	**92**	**20**	**8**	**1975**	**1592**	0.262	0.268
13	**0.092**	**0.116**	**0.119**	**0.142**	**0.242**	**0.309**	**1416**	**730**	**80**	**38**	**978**	**532**	**80**	**72**	**182**	**103**	**20**	**9**	**1718**	**1479**	0.253	0.263
14	**0.095**	**0.117**	**0.119**	**0.147**	**0.236**	**0.311**	**1395**	**852**	**80**	**44**	**968**	**625**	**80**	**70**	**172**	**92**	**18**	**10**	**1932**	**1523**	0.253	0.263
15	**0.096**	**0.119**	**0.120**	**0.146**	**0.241**	**0.308**	**1665**	**974**	**89**	**49**	**1084**	**696**	**81**	**71**	**166**	**95**	**20**	**10**	**1877**	**1550**	0.298	0.276
16	**0.095**	**0.117**	**0.120**	**0.146**	**0.243**	**0.314**	**1429**	**920**	**79**	**48**	**963**	**666**	**79**	**71**	**171**	**96**	**18**	**9**	**1870**	**1541**	0.297	0.279
17	**0.093**	**0.120**	**0.120**	**0.150**	**0.243**	**0.321**	**1232**	**806**	**68**	**43**	**844**	**592**	**79**	**72**	**171**	**94**	**16**	**10**	1853	1549	0.232	0.238
18	**0.095**	**0.122**	**0.119**	**0.152**	**0.242**	**0.318**	**1252**	**790**	**72**	**44**	**896**	**652**	**76**	**67**	**168**	**93**	**17**	**10**	1967	1656	0.275	0.266
19	**0.092**	**0.118**	**0.117**	**0.145**	**0.237**	**0.308**	**1387**	**925**	**81**	**53**	920	723	**87**	**73**	**166**	**93**	**20**	**12**	1987	1623	0.276	0.275
20	**0.095**	**0.120**	**0.120**	**0.150**	**0.242**	**0.314**	1466	1130	**84**	**59**	953	839	**86**	**71**	**178**	**100**	**19**	**11**	**1880**	**1528**	0.301	0.305
All dates	**0.095**	**0.116**	**0.119**	**0.144**	**0.238**	**0.311**	**1356**	**844**	**74**	**43**	**896**	**613**	**82**	**70**	**179**	**88**	**18**	**9**	**1918**	**1604**	0.266	0.258

Figure 3.2: Time series average of liquidity and activity measures

The figures show average time series plots of the different liquidity and activity measures. The actual trading days are aligned at t=0. From the figures there seem to be a systematic difference in both liquidity and activity over time between the group of stocks that were fully crossed and those that were not crossed at all. Similar plots of the measures around the separate dates show the same systematic patterns.

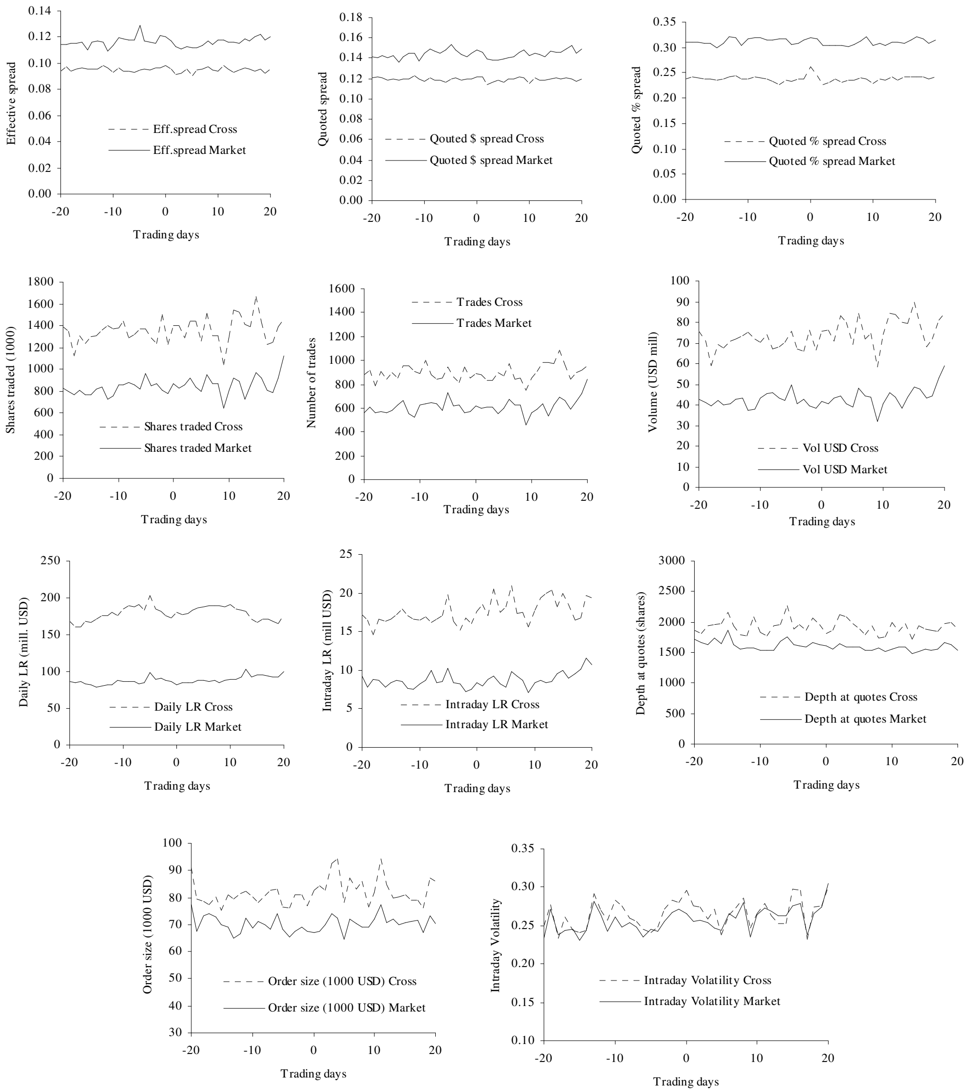


Table 3.5: Probit model estimating determinants of probability of a cross

We estimate a probit model of the probability that a given order is successfully crossed. The probability of observing a cross is assumed to be given by the model

$$y = Pr(\text{cross}) = F\left(\beta_0 + \beta_1 \text{eff\_spread}_i + \beta_2 \text{depth}_i + \beta_3 \text{LR}_i + \beta_4 \text{volume}_i + \beta_5 \text{vola}_i + \varepsilon_i\right)$$

where $F(\cdot)$ is the cumulative normal distribution function, and the β's are coefficients of the explanatory variables. Explanatory variables include the effective spread ("eff_spread"), the average depth at the inner quotes ("depth"), the intraday liquidity ratio ("LR"), the trading volume measured in USD ("volume"), and the standard deviation of daily returns measured over the last 10 days ("' vola"). The total data set contains 646 transactions, of which 214 were crosses. The intraday liquidity variable is highly correlated with the dollar volume of trading. We therefore use orthogonal versions of these two variables in the regression model. $\frac{dy}{dx}$ is the slope estimates (marginal effects) at the means of the regressors. These estimates predict the effects of changes in one of the explanatory variables on the probability of belonging to a certain trade category.

	coefficient	std deviation	pvalue	dy/dx
β_0: constant	0.0888	0.1887	0.6380	-
β_1: eff_spread	-4.8483	1.4834	0.0010	-1.7173
β_2: depth	-0.0002	0.0314	0.9940	-0.0001
β_3: LR	0.1926	0.0528	0.0000	0.0682
β_4: volume	0.2424	0.5630	0.0000	0.0858
β_5: vola	-1.4638	3.3163	0.6590	-0.5185
n	646.0000			
Wald $\chi^2(5)$	27.9400			
Prob > χ^2	0.0000			
Log likelihood	-389.0788			
pseudo R^2	0.0516			
Observed P	0.3313			
Predicted P (at means)	0.3129			

explanatory variables. Explanatory variables include the effective spread ("eff_spread"), the average depth at the inner quotes ("depth"), the intraday liquidity ratio ("LR"), the trading volume measured in USD ("volume"), and the standard deviation of daily returns measured over the last 10 days ("' vola"). The total data set contains 646 transactions, of which 214 were crosses.[19]

The model is estimated on all orders that were either fully crossed or fully filled in the primary market. The explanatory variables capture many dimensions of primary market liquidity and trading activity. The effective spread is considered the most appropriate measure of trading costs or market width. Average depth at the inner quotes is a frequently used depth measure, see for example Chordia et al. (2001). The intraday liquidity ratio captures part of the market resiliency dimension, and dollar trading volume and return volatility capture different aspects of the trading activity.[20] The estimation results are presented in table 3.5.

When interpreting the model, we calculate slope estimates (marginal effects) at the means of the regressors ($\frac{dy}{dx}$ in table 3.5).[21] These estimates predict the effects of changes in one of the explanatory variables on the probability of belonging to a certain trade category. Note also

[19] We use STATA 7 to estimate the model. The intraday liquidity variable is highly correlated with the dollar volume of trading. We therefore use orthogonal versions of these two variables in the regression model.

[20] We also estimate a multinomial logistic regression model using the same set of explanatory variables, but with an additional category consisting of the partly crossed orders. Because the results from this model do not provide any additional insight, we only report the results from the probit model.

[21] For non-linear probability models such as the probit and the logit model, we have that the effects of changes in one of the explanatory variables will vary with the value of the regressors.

that our estimation is simplified by the fact that our data only contains buy orders; we need not adjust for the direction of trade.

The estimated probit model in table 3.5 confirms the result in 3.3 that the probability of finding a counterparty in the crossing network is positively related to the liquidity of the stock in the primary market. The probability of a cross is higher the lower the effective spread, the higher the intraday liquidity ratio, and the higher the dollar trading volume in the primary market. This implies that stocks that are easy to cross are also highly traded in the market and have low costs measured by the effective spread.[22]

To sum up, our results indicate that the most liquid and actively traded stocks in the primary market also have the highest probability of being crossed. Both liquidity differences and private information may explain the difference in ex post abnormal return between the crossed and non-crossed stocks found in Næs and Ødegaard (2000). A significant difference in liquidity between the two groups of stocks, also on other dates than the trading dates, may indicate that investors need a higher return to hold the non-crossed stocks. On the other hand, it is hard to believe that liquidity differences between the 500 largest and most liquid companies in the US can explain a difference in abnormal performance between the two group of stocks of 1 percent over 20 days as found in Næs and Ødegaard (2000).

4 Limit order simulation

To judge whether trading in the primary market is more expensive than crossing, we need additional information on the costs of obtaining the stocks directly in the market. Since the Fund was trading in the 500 largest and most liquid companies in the US market, it could well be that a strategy of buying them directly in the market would have been less expensive than the crossing strategy followed by the Fund.

In this section, we examine the cost of the opportunistic crossing strategy relative to alternative submission strategies. In addition to a cost comparison, the simulations allow us to obtain a measure of immediacy. This is an important dimension of liquidity which is crucial for transaction costs, and which is not directly captured by the measures used in the previous section.

4.1 Literature

Crossing networks There are two theoretical papers on crossing networks; Hendershott and Mendelson (2000) and Dönges and Heinemann (2001). There is also closely related literature on the ability of multiple competing trading venues to coexist, see for example Chowdhry and Nanda (1991), Easley et al. (1996a) and Seppi (1990).[23]

[22] Market depth and return volatility do not have significant effects on the probability of getting a stock in the crossing network.

[23] There is an extensive literature on related subjects such as (i) the costs of using electronic communication networks (ECNs) (see Barclay et al. (2002), Barclay and Hendershott (2002), Coppejeans and Domowitz (1999), Domowitz and

Hendershott and Mendelson (2000) develop a complex model where different types of heterogenous liquidity traders and informed traders choose between a competitive dealer market and a crossing network. There are two types of informed traders: one type with short-lived information and one type with long-lived information. Short-lived information cannot be exploited in the crossing network, but traders with long-lived information can first try trading in the crossing network and then go to the dealer market if they are not able to cross. Trader strategies are modeled as Nash strategies: each trader chooses his or her best response given her expectation of all other traders' strategies.[24] The model solution consists of multiple equilibria. All equilibria are characterized by three cutoff values that segment liquidity traders into the following four (some possible empty) sets of strategies:

- do not trade,
- trade exclusively on the crossing network,
- trade opportunistically on the crossing network, i.e. attempt to trade on the crossing network, and then go to the dealer market if you cannot get an execution in the crossing network, and
- trade only in the dealer market.

The implications on dealers' spread from the introduction of a crossing network are shown to depend on the types of traders in the market. With *no informed trading*, the negative "cream-skimming" effect dominates the positive effect of attracting new order flow. This is because the crossing network has a negative impact on the dealers' inventory and fixed costs, and because orders going first to the crossing network impose higher costs on the dealer market than those going directly to the dealer market.[25] With *short-lived* information, the low order-submission costs ensure that the introduction of a crossing network will always raise the dealers' spread. This is because the crossing network reduces the order flow from liquidity traders without affecting the order flow from informed traders. Under most circumstances, the crossing network will also increase dealer spreads when information is *long-lived*. However, this can be offset if the crossing network manages to attract sufficient new liquidity traders.

The Dönges and Heinemann (2001) model is considerably simpler than the Hendershott and Mendelson (2000) model. Competition for order flow is modeled as a coordination game. The central variable is the value of trading, or, equivalently, the disutility from non-executed orders in the crossing network. Three different settings are analyzed. In the first setting all traders face an identical and certain cost of not getting an order executed in the crossing network. In this case, there are multiple equilibria as in the Hendershott and Mendelson model. In the second

Steil (1998)), and Hasbrouck and Saar (2001) and (ii) why some traders may want to trade outside the primary market (see Easley et al. (1996a) and Seppi (1990)).

[24] Trading decisions are based on the trader's reservation value, the spread cost, a crossing commission, the probability of getting a cross executed, and an impatience factor.

[25] Order flow sent to the crossing network leaves the dealers with fewer orders to cover the inventory and fixed costs, leading to higher average costs per order.

setting, all traders face an identical, but unknown cost of non-execution. By introducing private signals on the value of this cost, a unique equilibrium with market consolidation is shown to exist. According to Dønges and Heinemann, assets with low price volatility and large turnovers will be traded at a crossing network, while assets with high volatility or small volumes will be traded at dealer markets. In the third setting, the cost of non-execution is no longer assumed to be common among the traders. In this case, and provided that the disutility from non-execution differs sufficiently, there exists a unique equilibrium with market fragmentation. The two models provide few unambiguous implications. Rather, they form a framework for discussing important questions.

Limit order simulations The probability of non-execution is a central variable for both limit orders and orders submitted to a crossing network, especially for investors who are precommited to trade. Much cited papers on the modeling of execution probability and execution time of limit orders are Angel (1994), Lo et al. (2002), and Hollifield et al. (1999).[26] Angel (1994) derives closed form solutions for the probability of limit order execution when orders arrive according to a Poisson process and prices are discrete. Lo et al. (2002) develop an econometric model of limit order execution times using survival analysis and estimate it using actual limit order data. Hollifield et al. (1999) also develop, estimate, and test an econometric model of a pure limit order market. Their model describes the tradeoff between the limit order price and the probability of execution.

There are also several interesting empirical papers on the use of limit orders. Cho and Nelling (2000) investigate the probability of limit order executions for a selection of stocks at the NYSE. They find that the probability of execution is higher for sell orders than for buy orders, lower when the limit price is farther away from the prevailing quote, lower for larger trades, higher when spreads are wide and higher in periods of higher volatility. In addition, they find that the longer a limit order is outstanding, the less likely it is to execute, and that limit orders tends to be submitted at the bid-ask midpoint. Examining order flow and limit order submission strategies in a pure limit order market (the Paris Bourse), Biais et al. (1995) find that traders' limit order strategies depend on the market conditions: traders submit more market orders when spreads are narrow and submit more limit orders when spreads are wide, as shown by Angel (1994). Harris and Hasbrouck (1996) compare the performance of limit orders relative to market orders using the TORQ database. They find that limit orders placed at the quotes or further into the market outperform market orders when the spread is larger than the tick size. They therefore argue that limit orders in some cases can reduce execution costs compared to market orders. Handa and Schwartz (1996) approach the problem from a different angle by examining the performance of limit orders versus market orders by "submitting" hypothetical limit orders on the actual price paths of the thirty Dow Jones Industrial firms traded on the

[26] There is also an extensive theoretical literature on the effect of limit orders on the price discovery process as well as the relative profitability of limit orders compared to market orders. Important contributions include Foucault (1999), Glosten (1994), Easley and O'Hara (1992), Parlour (1996), Chakrevarty and Holden (1995), Seppi (1997).

NYSE. Since they are using simulations, they can also evaluate the cost of non-executed limit orders. Their main finding is that non-execution costs are positive, but not always significant.

4.2 Simulation design

We base our simulations on the strategies followed by the liquidity traders in the Hendershott and Mendelson (2000) model, ignoring the ”no trade” category. The first strategy, *opportunistic crossing*, is the actual strategy followed by the Fund. The second strategy, *pure cross*, is the case where the trader only submit orders to the crossing network. In this case, the trader has a low demand for immediacy/liquidity.

The third strategy is the case where the orders are only submitted to the market. Orders submitted to the market can be market orders or limit orders. An uninformed investor such as the Fund would generally prefer the lower costs and lower execution probability associated with limit orders to the immediacy provided by market orders. On the other hand, orders that are worked into the market may help reducing transactions costs. Domowitz (2001) shows that when the trader is ”monitoring the book”, and thus strategically searching for liquidity and favorable execution possibilities, a market order strategy (working the order) may reduce transaction costs considerably and reduce the price impact cost for large orders. Angel (1997) shows that about 30 percent of the market orders submitted through the SuperDot system experienced a price improvement of about USD 0.04 per share.

The best way to simulate a market order strategy would probably be to set up and estimate a dynamic model that minimizes transaction costs given the stock and market characteristics at the time of submission, such as the order flow, the depth of the limit order book, the volatility etc. The realism of such an ex post optimized strategy would be very hard to judge, however. Moreover, an “in sample” optimized strategy based on data from a limited period of time have restricted interest “out of sample”. Due to the obvious difficulties in constructing a market order simulation taking into account the plethora of strategic decisions involved, we restrict our analysis to simulating different limit order strategies. In this way, we get an interesting additional liquidity statistic and a realistic “lower bound” on the implicit execution costs of alternative submission strategies in the primary market.[27]

The closest proxy to a market order strategy in our simulations is a marketable limit order strategy (MLO). A MLO strategy is a limit order strategy that is more aggressive (”in to the market”) than an ”at the quote” (ATQ) limit order strategy. The main difference between an ATQ and MLO strategy is that the limit price is set at the bid and ask prices respectively. The higher limit price of the MLO strategy increases the execution probability and speed relative to an ATQ strategy. However, this increased immediacy may come at a cost.[28]

[27] As noted by Lo et al. (2002), there will be a general bias in favor of early execution of simulated limit orders compared to actual limit orders. Moreover, the simulation does not track where in the limit order queue our order is at any point in time, only the price priority. This probably affects the fill rate and execution time of the orders in favor of the simulated orders compared to actual limit order execution.

[28] The cost differential between the two types of strategies may vary over time depending on market conditions.

Note that both limit orders and crossing orders have a potentially costly adverse selection component. From the buyer's perspective, a limit order is filled when there is adverse price movement and not filled when the stock value increases. Both cases may or may not be due to new information. Similarly, the probability of being a successful buyer in a crossing network increases with the number of investors on the selling side of the market. As for limit orders, if there are informed investors (with long-lived information) in the crossing network, the execution probability of a buy order decreases if the information is positive.[29]

Limit order simulations All limit order submissions are simulated using the same stocks and dates that applied when the Fund first tried to cross the orders. The first limit order simulation (LO1) is identical to the simple simulation strategy in Handa and Schwartz (1996), i.e. we do not take into account the actual order sizes traded by the Fund. In other words, we assume that only one share is traded in each stock. At the beginning of each crossing date, a limit order is submitted with a limit price equal to the opening bid-quote (”at the quote” limit order strategy) for each stock that the Fund tried to cross. If a trade with a price lower than the limit order price is observed during the day, the order is assumed to be filled. If an order is not filled, we assume that it is executed at the opening price the next day. Thus, we implicitly assume an investor who is pre-committed to trade the stocks. During the transition period, the Fund was tracking an index with a limit on the relative volatility between the transition portfolio and the benchmark. Thus, even though the trades probably could have been worked more carefully into the market the next day, the penalty for unexecuted orders which follow from our assumptions is not completely unrealistic. Because we are ignoring order size, the first limit order simulation constitutes a lower bound on transaction costs.

In the second simulation (LO2), we split the actual order size into suborders. The number and size of the suborders are determined by the average order size traded in the stock at $t-1$. In addition, we have one residual suborder of a smaller size (if necessary). All the suborders are assumed to be submitted sequentially. Thus, at the beginning of the trading day, the first suborder is submitted as an ”at the quote” limit order. A suborder is assumed filled if the observed execution price is less than the limit-price without taking into account the size of the suborder. When a suborder is filled, the next suborder is submitted at the bid quote following the fill (”chasing the market”). Unfilled orders are assumed to be executed at the opening price the next day.[30]

The third limit order simulation (LO3) is the most realistic because here we also take into account the size of the suborders. The strategy is similar to LO2 except that we also examine whether the size of the suborder is less than or equal to the size of the actual order executed in the market. A suborder is only assumed filled if the observed execution price is less than the

Obviously, the execution probability of a marketable limit order is lower in a bear market relative to a bull market.

[29] Næs and Ødegaard (2000) find evidence that the Fund was “crowded out” by informed investors on the same side of the market.

[30] The unexecuted orders are assumed submitted to the pre-trade auction without affecting the opening price.

limit price *and* the size is equal to or larger than the size of our order. Due to price priority, our hypothetical order would under most circumstances execute before the observed trade since our order would be the last in the queue at our limit price.

A problem with this type of simulation is that the hypothetical orders most likely would have changed the structure of the market in the stocks if they had actually been submitted. Furthermore, Lo et al. (2002) note than the results from simulations with actual limit-order data underestimate the execution times in a real world trading situation. The execution time for a real limit order is a function of the order size, the limit price and the current market conditions, and a trader would generally vary the order submission strategy based on current and expected market conditions. Such factors are obviously very hard to capture in a simulation approach like ours. On the other hand, we do know the order sizes of the actual strategy and we do take these into account in the LO2 and LO3 simulations, which probably reduces the bias.

Pure crossing simulation A pure crossing strategy is defined as a strategy where the trader only trades in the crossing network. According to Hendershott and Mendelson (2000), the low liquidity preference traders who would follow this type of strategy are most likely to benefit from the existence of a crossing network. To simulate this strategy we use the actual price data for the stocks that the Fund was able to cross. For the stocks that the Fund was not able to cross, we assume crossing over the next 10 days. Hence, the opportunity costs are simulated, but the identity of stocks that could not be crossed are not. The choice of a 10-day trading window for calculating the opportunity costs is based on the statistics on order fills in Conrad et al. (2001b): the 95th percent confidence interval for getting an order filled in an external crossing system is reported to be 10 days. Thus, on each crossing date we take the stocks that did not cross and assume that they were crossed over the next 10-day period to the equally weighted close price over the 10-day period.

4.3 Measuring trading costs

In order to compare the performance of different submission strategies we must apply a measure of transaction costs. Current empirical academic literature on transaction costs are to a large degree based on versions of a theoretical measure which was first proposed by Treynor (1981) and which Perold (1988) later called the "implementation shortfall". The implementation shortfall is defined as the difference in performance between the portfolio of actual trades and a matching "paper" portfolio in which the stock returns are computed assuming that the trades were executed at the prices prevailing on the date of the decision to trade. In this way, both explicit cost components such as broker's fees, and implicit components such as spread costs, price impact costs, and costs related to delayed or uncompleted trading (opportunity costs) are captured. The approach also overcomes the problem of measuring costs on an individual trade basis when the order consists of a package of sub-trades [31]. Keim and Madhavan (1998) and Conrad et al. (2001b) suggest an empirical version of the implementation shortfall approach:

$$\begin{aligned} \text{total cost} &= \text{explicit cost} + \text{implicit cost} \\ &= \left\{ \frac{\text{commission per share}}{P_d} \right\} + \left\{ \left[\alpha \frac{P_a}{P_d} + (1-\alpha) \frac{P_{d+x}}{P_d} \right] - 1 \right\} \end{aligned} \tag{3.2}$$

where P_d is the closing price for the stock on the day before the decision to trade, P_a is the average price for all the executed trades in the order, α is the fill rate, and P_{d+x} is the closing price x number of days after the decision date, i.e. the unfilled portion of an order is assumed settled x days after the decision date.

We use the same measure as in Conrad et al. (2001b), except that we assume that the non-crossed orders in the pure crossing strategy are settled at the average of the closing prices over the x days after the decision date. In addition, since we cannot easily get good estimates for the explicit costs related to the trades that we simulate, the cost comparison is made on the basis of implicit costs only. Thus, our cost comparison is *not* based on total execution costs. A more serious problem is related to the limited number of trading days in our data set. The implicit cost estimate is intended to account for the price impact of orders. However, the price difference between P_a and P_d will also be affected by general market movements between the two observation times. Essentially, the measure assumes that the main source of price impact is our order. When we look at averages for trades on many different dates, this is not a big problem, because the market movement will tend to wash out in the average[32]. However, if we look at trades concentrated on a few dates, the general market movements at these dates will affect the measured costs. As we shall see, this is a particular problem for the market orders in our data set because they are concentrated on only three days.

Empirical studies document that the magnitude of different cost components vary with factors such as order size, intraday timing of the trade, stock liquidity, market design and investment style. Hence, to measure costs properly, detailed data on the entire order submission process is required. For the actual submission strategy followed by the Fund, we have access to such data. For the simulated strategies, however, the results will necessarily be driven to some extent by our own assumptions.

4.4 Results

For the orders that were executed on the day following the initial attempt at internal crossing, the total cost should be decomposed into one component associated with the delay of the order in the internal crossing network, and one component associated with the final execution in an

[31] Much of the relevant research on the measurement of transaction costs is summarized in Keim and Madhavan (1998)

[32] Keim and Madhavan (1997) show that the average daily return on stocks is small compared to the price impact from a trade.

external crossing network or in the primary market. Table 3.6 decompose the implicit costs for the Fund's order submission strategy into these two components.

Including the delay costs, the average implicit cost for all crossed orders was 0.11 percent, and the average implicit cost for all market orders was -0.74 percent. This implies an average implicit cost for all orders of -0.03. Some care should be taken when interpreting the negative implicit costs for the market orders. Because the orders purchased in the primary market are concentrated on three trading days only, the cost estimates are quite sensitive to the market movements on these days. Ignoring the delay component, the average implicit cost for all market orders was about 0.25 percent. The Fund incurred delay costs for market orders on one occasion. The market went markedly down on this day, leaving the Fund with an implicit delay cost for the non-crossed orders of -1.79 percent. Because the non-crossed orders had to be bought in the market on the following day, an average additional cost of 0.48 percent was incurred, giving a total implementation shortfall cost of -1.31 percent.

Measured over some time, the daily market movements are small compared to the price impact costs, as shown in Keim and Madhavan (1997). Hence, for large samples, adjusting for daily market returns does not make much difference. However, in our case, the cost measure is likely to be largely driven by the market movement. Keim and Madhavan (1997) argue that one should *not* try to adjust for market movements because they are a part of the timing cost for the order submission strategy. If so, the average implicit cost associated with the delay of orders in the private internal crossing network of -0.121 percent should be interpreted as a negative timing cost. On the other hand, the fact that the drop in market values on one of the trading days was large enough to have a significant effect on the total implementation shortfall cost of the actual strategy, suggests that the true costs of opportunistic crossing may be underestimated.

What the discussion above highlights most of all is that cost measures based on the implementation shortfall over a few days should be interpreted with great caution. Due to the non-synchronous nature of the Fund's market trades relative to the close-to-close returns on the SP 500 index, a correct adjustment for the market movement would involve the actual timing of the trades during the day as well as the intraday SP 500 returns. None of which are easily obtainable. What we want is to set up a horse race between the opportunistic crossing strategy and certain alternative order submission strategies. If the alternative strategies cannot beat the strategy when the negative delay costs are excluded, they surely cannot beat the strategy when these costs are included. In Table 3.7, we have therefore compared the estimated execution costs for the simulated strategies with the actual average execution costs *excluding* the delay costs.[33] That is, all cost estimates in the table are in percent of the closing price on the day before the trade.[34]

[33] What we ignore, however, is that the high volatility in the market at this particular day may have affected the outcome with respect to what stocks we were able to achieve in the crossing networks, as suggested in Domowitz (2001).

[34] Næs and Ødegaard (2000) also estimate the explicit costs for the Fund's strategy. The equally weighted average explicit costs for all orders were 3 percent. For the crossed orders and the non-crossed orders, the explicit costs were 3 percent and 5 percent respectively.

Table 3.6: Decomposition of the implicit costs for the opportunistic crossing strategy

Estimates of the average implicit costs for the opportunistic crossing strategy are decomposed into (i) the average implicit cost *excluding* the costs associated with the delay of orders, (ii) the average delay cost, and (iii) the average implicit cost *including* the delay cost, i.e. the average implicit implementation shortfall cost. The two last columns show respectively the number of trading days and the number of stocks traded for each type of orders.

Average implicit costs	Costs ex delay	Delay costs	Impl. shortfall	Days	Stocks
All orders	0.088	-0.121	-0.033	16	4 517
Crossed orders	0.055	0.056	0.111	15	3 767
Non-crossed orders	0.254	-0.998	-0.744	3	750
Delayed orders:					
All delayed orders	0.018	-0.620	-0.603	3	865
Delayed crossed orders	-0.415	0.465	0.049	2	447
Delayed non-crossed orders	0.483	-1.787	-1.304	1	418

Table 3.7: Estimates of implicit costs for different trading strategies - pre-trade benchmark

The table shows the execution cost estimates for four alternative submission strategies in addition to the original strategy (*Opportunistic Cross*). The estimates are based on the implementation shortfall methodology. The second strategy in the table, *Pure cross*, is the result of a hypothetical strategy where we assume that the entire residual order would have been crossed in equal amounts over the 10 days after the decision to trade. We split the non-crossed part of the order into 10 equal orders, each one of which is assumed crossed at the closing price each of the 10 days. The three last strategies in the table show the implicit cost estimates for the three submission strategies in the primary market. The first limit order strategy (LO1) is the most passive strategy which assumes that limit orders are submitted at the opening bid ("At-the-quote" limit order strategy), ignoring order sizes (no sub orders) as in Handa and Schwartz (1996). Whenever we observe a trade at our limit price or better, we assume the entire order is filled at that price. The second limit order strategy (LO2) assumes that limit orders are submitted sequentially at the prevailing bid following the filling of a suborder ("chasing the market"). However, in this case we ignore the size of each suborder. The third limit order strategy (LO3) is the most realistic strategy where all limit orders (also suborders) are submitted sequentially at the prevailing bid following the filling of a suborder as for LO2, but this simulation also takes the size of each suborder into account when evaluating the fill. If we observe a trade that is larger or equal in size to our order, we assume that our order would have been filled at that price. If there is a fill, the next suborder is submitted at the following bid. For all strategies, we assume that the remaining/unfilled part of an order is bought at the opening price the next day. Numbers in bold are estimates that are significantly different from zero at the 1% level. For each strategy and original group of stocks, tests of difference in means between the original submission strategy and the respective strategies are performed where ** indicates a significant difference in implicit costs at the 1% level.

Implicit costs	Opport. Cross	Pure Cross	LO 1	LO 2	LO 3
EW					
All orders	**0.0879**	**0.1443**	**0.0626**	**0.1303**	**0.2435****
Crossed orders	0.0553	0.0553	-0.0147**	**0.0520**	**0.1729****
Non-crossed orders	**0.2536**	**0.5867**	**0.4317****	**0.5048****	**0.6143****
VW					
All orders	0.2028	0.2534	0.0836	0.2849	0.3885
Crossed orders	0.1837	0.1837	0.0141	0.2007	0.3025
Non-crossed orders	0.3101	0.5867	0.4298	0.6615	0.7892

Examining the execution costs for the simulated strategies in table 3.7, we find that neither the pure crossing strategy nor the two first limit order strategies (LO1 and LO2) have significantly different execution costs from the opportunistic crossing strategy. Thus, not even the most simplistic and unrealistic limit order simulation (LO1), which constitute our ”lower bound” on primary market execution costs, is able to significantly beat the opportunistic crossing strategy. The most realistic limit order strategy (LO3) is significantly more expensive than the opportunistic crossing strategy, with costs of about 0.24 percent. In addition, we have not taken into account that the explicit costs in crossing networks are lower than in the primary market. Hence, the total execution costs would overwhelmingly favor the opportunistic crossing strategy, or potentially the pure crossing strategy.

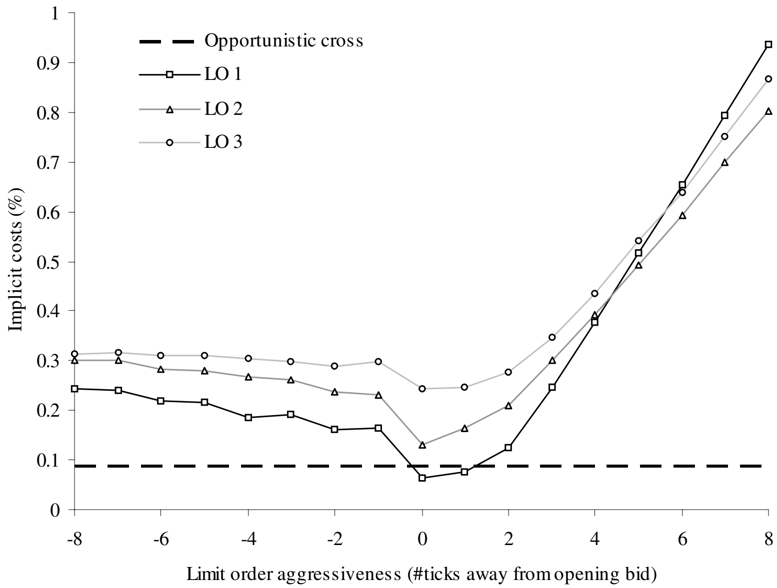
An additional choice variable for an investor is the aggressiveness of the limit order. In figure 3.3, we have plotted the implicit costs for the three limit order strategies LO1, LO2 and LO3 assuming more or less aggressive limit prices. In addition, the figure includes the implicit cost (ex delay costs) of the opportunistic crossing strategy (straight line across all aggressiveness levels). The ATQ limit order strategy is at 0 on the x-axis (indicating that the limit price is 0 ticks away from the opening bid). The MLO strategy is located between 1 to 3 ticks away from the bid, depending on the spread and tick sizes of the different stocks at the time of submission. An interesting observation in figure 3.3 is that the LO1 line forms a lower bound on execution costs. In addition, we see that the implicit costs across all strategies and aggressiveness levels reaches a minimum around 0 and 1 ticks away from the opening bid. This is in line with the results in Harris and Hasbrouck (1996), that limit orders generally are cheaper than market orders. More specifically, they find that when the spread is larger than one tick, limit orders placed in the market (improving the best bid or ask) perform better with respect to costs. Furthermore, Cho and Nelling (2000) show that the majority of limit orders are in fact submitted at the bid-ask midpoint.

By looking more carefully at the crossed/non-crossed groups, we find that the non-crossed stocks have the highest execution costs regardless of submission strategy. In the previous section, we found that stocks that are not supplied in the crossing network are less liquid than stocks that are easily crossed. The higher execution costs for these stocks support this finding: these stocks are also the most difficult to fill in the primary market. Note also that the opportunity costs constitute a large part of the execution costs for orders in these less liquid stocks. Since unfilled limit orders generally are for stocks that rise in value, these orders are penalized by the execution at the opening price the next day. This result, together with the high costs found for the pure crossing strategy, supports the finding in Næs and Ødegaard (2000) that the stocks bought in the market had a high ex post return.

Overall, our results strongly favor the opportunistic crossing strategy as a cost-effective submission strategy, especially when the difference in explicit costs between the crossing network and primary market is taken into account. Furthermore, it is important to recognize that the orders examined here are for the most liquid and largest companies in the US. Thus, even

Figure 3.3: Limit order simulation for varying aggressiveness levels.

The figure shows the implicit costs of the three types of limit order simulations we perform (LO1, LO2 and LO3) for varying aggressiveness levels, where aggressiveness is measured in ticks relative to the "at the quote" limit order strategy. A limit order aggressiveness of 0 indicates that the limit order price is set at the opening bid price. An aggressiveness larger (lower) than 0 means that the limit order price is set x number of ticks higher (lower) than the opening bid price. The horizontal line shows the implicit cost of the opportunistic crossing strategy excluding delay costs.



the stocks with the potentially lowest execution costs in the primary market would have been cheaper to obtain in the crossing network.

In table 3.8, we have calculated the fill rates for all orders in panel (a), and the fill rates across groups of orders in panel (b). The execution times (in minutes since open) for the simulated strategies are shown in panel (c). As expected, the fill rate decreases and the execution time increases as we impose more restrictions on the limit order strategy. It is interesting to note that the fill rates across groups of stocks in panel (b) are higher for the non-crossed orders than for the crossed orders. Thus, even though the fill rate is higher for the non-crossed stocks, the execution costs are higher. This indicates that the stocks in the non-crossed group that were not filled in the limit order simulation had a very high opportunity cost. This result provides further support to the information hypothesis in Næs and Ødegaard (2000).

Table 3.8: Fill rates and order execution time for different trading strategies

Panel (a) shows the fill rates for the different strategies with respect to the total number of shares and the number of orders filled. Panel (b) shows the fill rates across the groups of crossed/non-crossed stocks. Panel (c) shows the average execution time (in minutes) for the entire strategy with respect to the opening time of the market (minutes since open). The numbers in parenthesis are the average execution time of the orders (minutes since submission). For the opportunistic and pure crossing strategies these numbers are ignored since they are over several days. For LO1, the measure of "minutes since open" and "minutes since submission" is equal because only one order is submitted for each stock.

(a) Fill rates for submission strategies

	Opport. Cross	Pure Cross	LO1	LO2	LO3
Orders					
Filled (%)	83.2%	100.0%	85.6%	71.9%	65.1%
Not filled (%)	16.8%	0.0%	14.4%	28.1%	34.9%
Submitted orders	3909	3909	3909	11864	11289
Filled orders	3316	3909	3346	8528	7347
Unfilled orders	594	0	563	3336	3942
Shares					
Filled (%)	84.8%	100.0%	88.5%	49.7%	42.5%
Not filled (%)	15.2%	0.0%	11.5%	50.3%	57.5%
Shares in submitted orders	26776710	26776710	26776710	26776710	26776710
Shares in filled orders	22714683	26776710	23693158	13303893	11372729
Shares in unfilled orders	4070060	0	3083552	13472817	15403981

(b) Fill rates across groups

	Opport. Cross	Pure Cross	LO1	LO2	LO3
Orders					
Cross group:					
Filled (%)	83.2%	100.0%	84.8%	70.9%	64.0%
Not filled (%)	16.8%	0.0%	15.2%	29.1%	36.0%
Non-crossed group:					
Filled (%)	100%	-	89.7%	76.3%	70.2%
Not filled (%)	0	-	10.3%	23.7%	29.9%
Shares					
Cross group:					
Filled (%)	84.8%	100.0%	88.1%	48.9%	41.5%
Not filled (%)	15.2%	0.0%	11.9%	51.1%	58.5%
Non-crossed group:					
Filled (%)	100%	-	90.5%	53.6%	47.7%
Not-filled (%)	0	-	9.6%	46.4%	52.3%

(c) Execution time (minutes)

	Opport. Cross	Pure Cross	LO1	LO2	LO3
Mean	-	-	30 (30)	42 (22)	71 (38)
Median	-	-	7 (7)	9 (5)	24 (10)
Minimum	-	-	0 (0)	0 (0)	0 (0)
Maximum	-	-	389 (389)	390 (390)	390 (390)
First quartile	-	-	3 (3)	4 (1)	7 (1)
Third quartile	-	-	19 (19)	31 (14)	80 (34)
Standard deviation	-	-	67 (67)	80 (56)	102 (73)

5 Conclusion

In this paper, we use data from an actual order submission strategy using crossing networks to investigate execution costs and primary market liquidity. The data includes all orders from the establishment of a US equity portfolio worth USD 1.76 billion in the period from January 1998 to June 1998. The investor in our study was following an "opportunistic" crossing strategy, meaning that an attempt was made to cross all stock orders initially, and residual orders were purchased in the open market. Because we know the identity of stocks and timing of stock orders that failed to be executed in the crossed network, we can investigate whether stocks that are supplied in crossing networks and stocks that can only be traded in the market have systematically different characteristics. In addition, the costs of alternative, more traditional, submission strategies can be estimated and compared.

By calculating several measures of liquidity for the different groups of stocks in the data set, we show that there are significant differences in liquidity between stocks that are crossed and stocks that have to be bought in the market. For one date, spreads are significantly different even though the trading volume in the two groups of stocks was similar. According to the market microstructure literature, this might be an indication of informed trading in the stocks that could not be executed in the crossing network, a result which is also suggested in Næs and Ødegaard (2000). We also find, however, that there are systematic differences in liquidity between the two groups of stocks on other dates than the trading dates of the actual crossing strategy. This result suggests that there are systematic differences in the characteristics of the two groups of stocks that are not related to private information.

To evaluate the performance of the actual crossing strategy against other submission strategies, we perform limit order simulations on transactions data from the NYSE. The simulations can also be viewed as an additional measure of trading difficulty. The non-crossed orders turn out to be significantly more expensive than the crossed orders across all simulations. Hence, the stocks that the Fund could not get in the crossing network would also have been the most difficult to buy in the market. We also show that it would have been very hard to beat the actual opportunistic crossing strategy. The only simulation which gives us a lower implicit cost estimate is when we completely ignore the size of our orders. However, the explicit cost differential between the crossing network and regular market would probably even this difference out. Finally, it should be stressed that the significant differences found in crossing probability, liquidity and primary market execution costs are for the 500 largest and most liquid stocks in the US market.

Appendix A

Calculation of liquidity and activity measures

To calculate the liquidity statistics in the primary market for all securities traded by the fund we use the TAQ database (NYSE Trades and Quotes database). However, before we perform the calculations, the data has to be filtered to remove erroneous records both in the quotes file and the trades file.

1 Data issues and filtering

Quotes data

All the spread measures are calculated with respect to the inside quotes (best bid and ask) reported in the TAQ database between 9:30 and 16:00. There are several filters applied to ”clean” the data. We mainly use the quote conditions (MODES) in the TAQ data[1] to do this. An observation is removed if one of the following conditions applies;

- **Closing quote** The last quote from a participant during the trading day (Mode = 3)
- **News dissemination** A regulatory halt when price sensitive news arrives (Mode = 4)
- **Fast trading** Indicating that there is extreme activity (quotes are entered on a ”best efforts” basis) making the time stamps unreliable (Mode = 5)
- **Order imbalance** A non-regulatory trading halt due to large order imbalances (Mode = 7)
- **Non-firm quote** A regulatory halt when the Exchange is unable to collect, process and disseminate quotes that accurately reflect market conditions (Mode = 9)

[1] A more detailed description can be found in the TAQ User Guide which can be downloaded from the NYSE homepage at http://www.nyse.com/marketinfo/marketinfo.html

- **News pending** A regulatory trading halt or delayed opening due to an expected news announcement (Mode = 11)
- **Trading halt due to related security** A non-regulatory halt used when there is news related to one security which will affect the trading and price in another security (Mode = 13)

In addition we remove quotes where the bid price is larger than the ask price, quotes are negative, or the average quoted spread is zero over the trading day. Also quotes with a price higher than USD 10,000 are removed both due to possible errors as well as to remove securities with extreme prices which could affect our statistics. Lastly, when quotes from several different exchanges are reported at the same time (down to the second), we use the lowest ask or highest bid among these as a proxy for the NBBO (National Best Bid and Offer).

Trades data

The trades reported in TAQ may contain corrections and errors. If so, the record has a *Correction Indicator* (Corr) attached to it. The requirement is that a trade must have a correction value less than 2 (Corr < 2). If Corr=0, then the trade record is a regular trade that was not corrected, changed, cancelled or was erroneous. If Corr=1, then the observation was later corrected, but the record contains the original time and the corrected data for the trade. If Corr > 2, then the record is either out of time sequence, cancelled due to error or cancelled due to wrong timestamp. Thus, we remove all records with Corr ≥ 2.

There are also *Sale Conditions* (Corr) connected to each trade. We apply a filter removing records with conditions that make the timing and reliability of the records questionable. A record is removed if one of the following conditions applies;

- **Bunched sold** A bunched trade not reported within 90 seconds of execution (Cond=G)
- **Sold last** A trade reported later than the actual transaction time (Cond=L)
- **Opened last** An opening trade with delayed reporting (Cond=O)
- **Sellers option** Delivery date is between 2 and 60 days after the trade (Cond=R)
- **Pre- and Post-Market Close Trades** A trade that occurred within the current trading day, but is executed outside of the current market hours (Cond=T)
- **Sold sale** A transaction that is reported to the tape at a time later than it occurred and when other trades occurred between the time of the transaction and its report time (Cond=Z)
- **Crossing session** NYSE Crossing Session matches (Cond=8 and 9)

After the filtering is performed, we use the remaining quotes and trades to calculate the following liquidity and activity measures.

2 Spread measures

Effective spread

The effective spread takes into account the transaction prices (and accounts for the fact that many trades are executed within the quoted spread due to price improvement). The number of trades in the security, i, on date, t, is denoted by $N_{i,t}$. The index τ defines the time of the day when a trade is observed, $P_{i,\tau}$ is the trade price, and $bid_{i,\tau}$ and $ask_{i,\tau}$ is the bid and ask, respectively, at the time of the trade. The first valid trade is normalized to $\tau = 1$. Then, for security i on date t, the average effective spread is calculated as,

$$ES_{i,t} = \frac{1}{N_{i,t}} \sum_{\tau=1}^{N_{i,t}} \left\{ 2 \mid P_{i,\tau} - \frac{ask_{i,\tau} + bid_{i,\tau}}{2} \mid \right\}$$

The effective spread takes into account the relationship between execution price and quoted spread, and is often considered a more appropriate measure of trading costs than quoted spreads, especially for large trades.

Quoted dollar spread

The average quoted dollar spread is defined as the average difference between the inside quoted ask and bid for a firm over the trading day. The quoted dollar spread is calculated with respect to each trade observed at time τ. The inner ask and bid is defined as $ask_{i,\tau}$ and $bid_{i,\tau}$ respectively, and $N_{i,t}$ is the total number of trades in security i during the trading day t. Thus, the quoted dollar spread is calculated as,

$$QS_{i,t} = \frac{1}{N_{i,t}} \sum_{\tau=1}^{N_{i,t}} (ask_{i,\tau} - bid_{i,\tau})$$

Quoted percentage spread

The quoted percentage spread calculates the absolute spread relative to the spread midpoint at each valid trade record τ. Thus,

$$RS_{i,t} = \frac{1}{N_{i,t}} \sum_{\tau}^{N_{i,t}} \left\{ \frac{ask_{i,\tau} - bid_{i,\tau}}{(ask_{i,\tau} + bid_{i,\tau})/2} \right\}$$

3 Volume and depth measures

It is widely argued that spreads should not be examined in isolation when using it as a liquidity measure. This is because liquidity shocks both widen spreads as well as reduce depths. Furthermore, spreads may also widen as a response to adverse selection without liquidity necessarily decreasing. Therefore, we also look at volume and depth measures.

Trading Volume (Shares)
The total number of shares traded in security i during day t.

$$VOL\_shares_{i,t} = \sum_{\tau=1}^{N_{i,t}} Q_{i,\tau}$$

Trading volume (USD)
The total dollar value of trades during day t in security i.

$$VOL\_USD_{i,t} = \sum_{\tau=1}^{N_{i,t}} Q_{i,\tau} \cdot P_{i,\tau}$$

Trades
The total number of trades during day t in security i.

$$Trades_{i,t} = \sum_{\tau=1}^{N_{i,t}} \tau_i$$

Trade size
The average trade size in USD 1000 on day t in security i.

$$Trade\_size_{i,t} = \frac{VOL\_USD_{i,t}}{Trades_{i,t} \cdot 1000}$$

Quoted depth
The quoted depth is calculated as average of the quoted bid and ask depths during the day t in security i,

$$QD_{i,t} = (\bar{q}_{i,t}^{bid} + \bar{q}_{i,t}^{ask})/2$$

where $\bar{q}_{i,t}^{bid}$ and $\bar{q}_{i,t}^{ask}$ is the average depth on the bid- and the ask-side respectively in security i on day t.

4 Liquidity ratios and volatility measures

Daily Liquidity Ratio

The Amivest Liquidity Ratio is one type of liquidity measurement which represents the dollar value of trading associated with a one percent change in the share price. Amivest is the "creator" of this liquidity measurement. The liquidity ratio measures the average trading volume necessary to move the price by one percent during a trading day. We calculate the average daily liquidity ratio over the 10-day period prior to the Fund's trading date, t_0. The daily liquidity ratio for security i on date t is thus calculated as,

$$LR(D)_{i,t} = \frac{1}{10} \sum_{t=t_0-11}^{t_0-1} \frac{VOL\_USD_{i,t}}{|\%r_{i,t}|} / 1000$$

where $|\%r_{i,t}|$ is the absolute "midpoint return" over day t calculated using the bid-ask midpoints at opening and closing to avoid biases with respect to the bid-ask bounce. $VOL\_USD_{i,t}$ is the total USD trading volume in security i on date t.

Intraday Liquidity Ratio

To measure liquidity on one date, we also calculate the liquidity ratio using intraday data. To do this, we first discretisize the data to get a common time frame. Consistent with several other studies we use 15-minute windows, starting from 9:30am until 16:00pm. Thus, we have 26 15-minute intervals during each trading day. During each interval, denoted by ω, we calculate the midpoint returns using the bid-ask midpoint price at the beginning (or closest to the beginning) of each window. Thus, $\omega \in [1, 26]$, and the average ratio for security i on date t is calculated as,

$$LR(I)_{i,t} = \frac{1}{26} \sum_{\omega=1}^{26} \frac{VOL\_USD_{i,\omega}}{|\%r_{i,\omega}|} / 1000$$

where $VOL\_USD_{i,\omega}$ is the total USD volume traded in security i in window ω, and $|\%r_{i,\omega}|$ is the 15-minute absolute midpoint return over window ω. Generally, the liquidity ratio measure assumes that there is a linear relationship between the trade size and price change which is not necessarily the case. In addition, the ratio is positively correlated with the general price trend in the market and negatively correlated with volatility.

Average 10-day volatility

Calculates the 10-day average volatility prior to the actual trading date (t_0) as,

$$V(D)_{i,t} = \sqrt{\frac{1}{10} \sum_{t=t_0-11}^{t_0-1} (r_{i,t} - \bar{r}_i)^2}$$

where $r_{i,t}$ is the return on day t and $\bar{r}_i$ is the average return over the 10-day period prior to the actual crossing date.

Intraday volatility

When calculating intraday volatility, we use the same discretization as for the intraday liquidity ratio calculations described above. Thus, we calculate the volatility of 15-minute returns over the trading day, using the bid-ask midpoint price at the beginning of each window, such that,

$$V(ID)_{i,t} = \sqrt{\frac{1}{26}\sum_{\omega=1}^{26}(r_{i,\omega} - \bar{r}_{i,t})^2}$$

where $r_{i,\omega}$ is the midpoint return over 15-minute window ω, and $\bar{r}_{i,t}$ is the average return over all windows during trading day t in security i.

Bibliography

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

Amihud, Y., Mendelson, H., and Lauterbach, B. (1997). Market microstructure and securities values: Evidence from the tel aviv stock exchange. *Journal of Financial Economics*, 45:365–390.

Angel, J. (1994). Limit versus market orders. Unpublished Working Paper, School of Business Administration, Georgetown University.

Angel, J. (1997). Who gets price improvement on the NYSE? Working Paper FINC-1377-12-694, Georgetown University.

Bacidore, J. M., Ross, K., and Sofianos, G. (1999). Quantifying best execution at the New York Stock Exchange: Market orders. NYSE Working Paper 1999-05.

Barclay, M. J. and Hendershott, T. (2002). Price discovery and trading costs after hours. Working Paper, University of Rochester, Forthcoming in Review of Financial Studies.

Barclay, M. J., Hendershott, T., and McCormick, D. T. (2002). Electronic communications networks and market quality. Working Paper, University of Rochester.

Biais, B., Hillion, P., and Spatt, C. (1995). An empirical analysis of the limit order book and the order flow in the Paris Bourse. *Journal of Finance*, 50:213–231.

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

Cho, J. W. and Nelling, E. (2000). The probability of limit order execution. *Financial Analysts Journal*, 56:28–33.

Chordia, T., Roll, R., and Subrahmanyam, A. (2001). Market liquidity and trading activity. *Journal of Finance*, 56:501–30.

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

Conrad, J., Johnson, K. M., and Wahal, S. (2001a). Institutional trading and soft dollars. *Journal of Finance*, 56:397–416.

Conrad, J., Johnson, K. M., and Wahal, S. (2001b). Institutional trading costs and alternative trading systems. Working Paper, University of North Carolina at Chapel Hill, Forthcoming in Journal of Financial Economics.

Coppejeans, M. and Domowitz, I. (1999). Pricing behavior in an off-hours computerized market. Working Paper, Pennsylvania State University.

Domowitz, I. (2001). Liquidity, transaction cost and reintermediation in electronic markets. eBusiness Research Center Working Paper 4-2001, Penn State University.

Domowitz, I. and Steil, B. (1998). Automation, trading costs and the structure of the trading services industry. Working Paper, Penn State University.

Dönges, J. and Heinemann, F. (2001). Competition for order flow as a coordination game. Working paper No. 64, Goethe-Universität Frankfurt.

Easley, D., Kiefer, N. M., and O'Hara, M. (1996a). Cream-skimming or profit sharing? The curious role of purchased order flow. *Journal of Finance*, 51:811–833.

Easley, D., Kiefer, N. M., O'Hara, M., and Paperman, J. B. (1996b). Liquidity, information, and infrequently traded stocks. *Journal of Finance*, 51:1405–1436.

Easley, D. and O'Hara, M. (1992). Time and the process of security price adjustment. *Journal of Finance*, 47:577–606.

Fong, K., Madhavan, A., and Swan, P. (1999). Why do securities markets fragment? Working Paper, University of Southern California.

Foucault, T. (1999). Order flow composition and trading costs in a dynamic limit order market. *Journal of Financial Management*, 1(2):99–134.

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

Handa, P. and Schwartz, R. A. (1996). Limit order trading. *Journal of Finance*, 51:1835–1861.

Harris, L. (1990). Liquidity, trading rules, and electronic trading systems. *New York University Salomon Center Monograph Series in Finance and Economics*, 4.

Harris, L. (2003). *Trading and Exchanges. Market Microstructure for Practitioners*. Oxford University Press.

Harris, L. and Hasbrouck, J. (1996). Market vs. limit orders: The superDOT evidence on order submission strategy. *Journal of Financial and Quantitative Analysis*, 31:231–231.

Hasbrouck, J. and Saar, G. (2001). Limit orders and volatility in a hybrid market: The island ECN. Stern Working Paper, New York University.

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

Hollifield, B., Miller, R., and Sandas, P. (1999). An empirical analysis of a pure limit order market. Working Paper 029-99, University of Pennsylvania.

Jones, C. M. and Lipson, M. L. (1999a). Execution costs of institutional equity orders. *Journal of Financial Intermediation*, 8:123–140.

Jones, C. M. and Lipson, M. L. (1999b). Sixteenths: Direct evidence on institutional execution costs. Working Paper, Columbia Business School.

Keim, D. B. and Madhavan, A. (1995). Anatomy of the trading process: Empirical evidence for the motivation for and execution of instituional equity trades. *Journal of Financial Economics*, 37:371–398.

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

Keim, D. B. and Madhavan, A. (1998). The cost of institutional equity trades. *Financial Analysts Journal*, 54:50–69.

Khan, W. A. and Baker, H. K. (1993). Unlisted trading privileges, liwuidity, and stock returns. *Journal of Financial Research*, 16:221–236.

Lo, A. W., MacKinlay, A. C., and Zhang, J. (2002). Econometric models of limit-order executions. *Journal of Financial Economics*, 65:31–71.

Mendelson, H. (1987). Consolidation, fragmentation, and market performance. *Journal of Financial and Quantitative Analysis*, 22:189–207.

Næs, R. and Ødegaard, B. A. (2000). Equity trading by institutional investors. To cross or not to cross? Discussion Paper 3/2000, Norwegian School of Management BI.

Parlour, C. (1996). Price dynamics in limit order markets. Mimeo, Carnegie-Mellon University.

Perold, A. (1988). The implementation shortfall: Paper versus reality. *Journal of Portfolio Management*, 14(3):4–9.

Ruyter, T. (1999). The dinosaur, the rock, and the hard place. Plansponsor Magazine. March.

Seppi, D. (1990). Equilibrium block trading and asymmetric information. *Journal of Finance*, 45:73–94.

Seppi, D. (1997). Liquidity provision with limit orders and a strategic specialist. *Review of Financial Studies*, 10:103–150.

Treynor, J. L. (1981). What does is take to win the trading game? *Financial Analysts Journal*, 47(1).

Chapter 4

Order book characteristics and the volume-volatility relation: Empirical evidence from a limit order market

Written with Johannes A. Skjeltorp

Abstract

We examine empirically the relationship between the demand and supply schedules in a limit order book and the volume volatility relation. Several empirical studies find support for the hypothesis that the volume-volatility relation is driven by the arrival rate of new information, proxied by the number of transactions. Our results show that the number of trades and the price volatility are also negatively related to the slope of the order book. One possible interpretation for this finding is that the slope of the book is proxying for dispersed beliefs about asset values. If so, this is in line with models where investor heterogeneity increases both trading volume and volatility.

1 Introduction

In this paper, we examine empirically the relationship between the demand and supply schedules in a limit order book and the volume volatility relation.

A variety of studies document that there is a positive correlation between price volatility and trading volume for most types of financial contracts including stocks, Treasury bills, currencies and various futures contracts. The main theoretical explanation for this phenomenon is that new information about asset values acts as the driving force (or mixing variable) for both market prices and trading volume. Harris (1986) links this "mixture of distributions hypothesis" to asset pricing theory, and suggests that the mixing variable is the process that directs the rate of flow of information from systematic risk factors into prices and trading volume. However, for many types of financial contracts, movements in prices seem much "too large" to be attributed to movements in the fundamental values of the underlying securities.[1] A suggested explanation for this puzzle is that prices do not change merely because of changes in systematic risk factors and asset payoffs but also because investors have dispersed beliefs about asset values. This dispersion may be due to asymmetric information or to differences of opinion about symmetric information. In any case, theoretical models by Shalen (1993) (asymmetric information) and Harris and Raviv (1993) (symmetric information) show that dispersion of beliefs will intensify the volume-volatility relation, by increasing both trading volume and volatility.

The theoretical explanations for the volume-volatility relation are hard to test. The essence of the mixture of distributions hypothesis is that prices adjust to new equilibria over time as new information is being reflected through trades. Since the arrival rate of information is unobservable, it is difficult to set up an alternative hypothesis. Several empirical studies find support for the explanation under the assumption that the arrival rate of information can be proxied by the daily number of transactions.[2] Since the daily number of transactions may be driven by factors other than new information, these studies do not rule out the other explanations for the volume-volatility relation. Specifying data implications from the models with dispersed beliefs is also very challenging. Daigler and Wiley (1999) perform an indirect test of Shalen (1993)'s model and find evidence that uninformed traders contribute to price volatility. Ghysels and Juergens (2001) measure dispersion of beliefs directly by dispersion of analysts' earnings forecasts. They also find that dispersion is positively related to volatility.

The objective of this paper is to broaden our knowledge about the volume-volatility relation in an electronic limit order market. Since the demand and supply schedules in a limit order book represent the reservation prices of the liquidity suppliers in the market, it is interesting to study whether the book contains additional information about the volume-volatility relation. We have access to exceptionally rich transactions data from the Norwegian equity market in the period from February 1999 through June 2001. The market operates as a fully automated limit order-driven trading system, and our data sample enables us to rebuild the full order book at any

[1] A standard reference for the stock market is Shiller (1981).

[2] See Harris (1987) and Jones et al. (1994)

point in time. We are not aware of anyone who has investigated this issue with a data set as rich as ours.

Several papers investigating order book data are relevant for our work. Biais et al. (1995) analyze in detail the interaction between the order book and order flow on the Paris Bourse. One relevant finding is that the status of the order book is important for order flows and trading volume. Biais et al. (1995) only have data on the cumulative trading interest near the inner quotes. We show that the whole order book contains additional, interesting information. Goldstein and Kavajecz (2000) provide evidence of a negative relation between the shape of the order book and volatility during a case of an extreme market movement. However, they do not attempt to investigate this relation over a longer time period with varying trading conditions. Our data set spans a relatively long period which included the boom and burst of the internet bubble. Kalay et al. (2003) estimate the demand and supply elasticities for stocks on the Tel Aviv Stock Exchange using data for order placements at the opening of the market. Their main findings are that the order book is more elastic at the beginning of the day, and that the demand side is more elastic than the bid side.[3] Kalay et al. (2003) only have data for order placements at the opening of the market. Our estimates of supply and demand schedules are also based on the continuous trading session.

We first establish that the standard volume-volatility relation exists in a limit order market, and investigate in detail the composition of the order book at the intra-day level. This exercise documents that the trading structure on the Norwegian Stock Exchange exhibits the same features as are found in empirical studies of other countries' stock markets. The features suggest that: informational asymmetries are more pronounced at the beginning of the trading day, there is competition among informed traders, and uninformed traders require a compensation for the higher pick-off risk at the beginning of the day.[4] These results are systematic across sub-periods, firm sizes, and tick-sizes.

The main contribution of our study is that we are able to document several relationships between the volume-volatility relation and the shape of the order book. We measure the order book shape by the average elasticity of the supply and demand schedules in the book. The lower the elasticity (steeper the slope), the less dispersed are the bid and ask prices in the order book.[5] To examine the effects of the order book slope on volume and volatility, we first include the slope measure as an independent variable in a cross sectional time series version of the standard regression model used to document the volume-volatility relation. To investigate the relationship between the slope of the book and the trading activity, we estimate a cross-sectional time series regression with the number of trades as the dependent variable. A systematic negative relation

[3] The first result is interpreted as supportive to sequential trading models with asymmetric information which predict higher adverse selection at the opening (Glosten and Milgrom (1985)). The second result is interpreted as supportive to the empirical finding that buy orders have larger price impacts than sell orders.

[4] Our results are in accordance with the results in Kalay et al. (2003) as well as with the results in several studies of time-of-day effects in spreads and price impacts, for example French and Roll (1986), Harris (1986), and Niemeyer and Sandas (1995).

[5] This is in the case of direct demand and supply curves (prices on the x-axis and accumulated volume on the y-axis). In the case of inverted demand and supply curves, the relationship would be opposite.

between the average slope of the order book and the price volatility is documented in a daily time series cross-sectional analysis. These results are also shown to be robust to the choice of time period. Similarly, we find a significant and robust negative relationship between our slope measure and the daily number of trades.

If the slope of the book is essentially a liquidity measure, most of the information contained in the slope should be reflected by the volume close to the inner quotes. To check this, we calculate the slope measure based on different fractions of the order book and re-estimate all the regression models. When we investigate the relation between different slope measures and trading activity, an interesting pattern emerges. In line with the findings in Biais et al. (1995) that thick books at the inner quotes result in trades, we find a significant *positive* relationship between the slope of the book and the number of trades when the slope is calculated based on the volume at the inner quotes. This result is the opposite of what we get when we use a slope measure based on the full order book. Thus, the slope of the book provides different information depending on what fraction of the book we use in the calculation.

A possible interesting interpretation of the full order book slope is related to the dispersion of beliefs hypothesis. Harris (1987) notes that, if trades are self generating, the number of daily transactions will be the true mixing variable rather than a proxy for the arrival rate of new information. It could be that the slope of the limit order book capture dispersion of beliefs about asset values, i.e. steep slopes of the supply and demand schedules indicate that there is a high degree of agreement among investors about the fair value of the security, while gentle slopes indicate that there is greater disagreement among investors about the value of the security. If so, our finding that there is increased trading activity when slopes are more gentle (greater disagreement about valuation) could reflect a situation of self generating trades, i.e. that the volume-volatility relation is not merely driven by new information. This interpretation support models where heterogeneity among investors contributes to the volume-volatility relation.

On the other hand, our results can also be explained within a Glosten (1994) type of model where all liquidity suppliers are homogeneous: for a given level of liquidity motivated trading and a given probability of informed trading, the slopes will be more gentle the more volatile assets are, while a positive relation between the slope at the inner quotes and trading activity could be explained by price sensitive liquidity traders.[6] However, the results from the test of the Glosten model in Sandås (2001) do not provide empirical support for a model with homogeneous liquidity suppliers.

The paper is organized in the following way. Section 2 surveys the relevant literature. Section 3 describes our data sample. Section 4 examines in detail the order flow and order book on an intra-daily basis. Section 5 provides the results from our analysis of the volume-volatility relation in the Norwegian equity market. Section 6 concludes the paper.

[6]We are grateful to an anonymous referee for pointing this out to us.

2 Literature

The mixture of distributions hypothesis The early research into the volume-volatility relation is reviewed in Karpoff (1987). The main theoretical explanation from this period is known as the "mixture of distributions hypothesis" (hereafter the MDH). According to the MDH, there is a positive correlation between daily price changes and trading volume because both variables are mixtures of independent normals with the same mixing variable. Originally, the MDH was suggested by Clark (1973) as an alternative explanation for the observed leptokurtosis in the distribution of log price changes.[7] The basic idea underlying the hypothesis is that prices and trading volume are driven by a time-varying arrival rate of information.[8] Let $\Delta p_{i,t}$ and $v_{i,t}$ be respectively the intraday price change and volume of trade resulting from information event number i on date t, and let n_t be the total number of information events during day t. Assume that (i) the number of events each day, n_t, varies across days, and that (ii) the intraday price changes, Δp, and trading volumes, v, are jointly independently and identically distributed with finite variances. Our explanation of the MDH is largely based on Harris (1987). The daily price change and trading volume are equal to the sum of respectively the intraday price changes and trading volumes, i.e.

$$\Delta P_t = \sum_{i=1}^{n_t} \Delta p_{i,t} \quad \text{and} \quad V_t = \sum_{i=1}^{n_t} v_{i,t} \tag{4.1}$$

where ΔP_t is the daily price change and V_t is the daily trading volume. Given equation (4.1), and provided that n_t is large, the joint distribution of the daily price change and volume of trade will be approximately bivariate normal conditional on n_t.[9] The volume-volatility relation arises because both price changes and trading volume are likely to be large when the number of information events is large and small when the number of information events is small.[10]

Harris (1986) finds empirical support for the MDH based on cross-sectional tests of common stocks continuously traded on the NYSE or one of the regional exchanges in the period 1976-1977. The critical assumption behind the tests is that the distribution of the mixing variable is not identical for all securities. Assuming that transactions take place at a uniform rate in event time, Harris (1987) finds both theoretical motivation and empirical support for the use of the daily

[7] Mandelbrot (1963) and Fama (1963) showed that the return distributions of commodity and stock prices were leptokurtic, and well approximated by symmetric stable distributions with characteristic exponents between 1 and 2 (the normal distribution has a characteristic exponent equal to 2). An examination of the stable distributions hypothesis for the Norwegian market is provided in Skjeltorp (2000) who shows that a characteristic exponent between 1.6 and 1.7 best characterizes the Norwegian data.

[8] Copeland (1976, 1977)'s "sequential arrival of information" model which is later extended by Jennings et al. (1981) and Jennings and Barry (1983) also predicts a positive relationship between volume and absolute price changes. The main feature of the model is that information is disseminated to only one trader at a time, and the main criticism of the models is that traders cannot learn from the market prices as other traders become informed.

[9] See Harris (1987), page 129.

[10] The variation in the daily number of information events implies that the expectation of the unconditional distribution is a weighted average (or "a mixture") of the conditional distributions.

number of transactions as a proxy for the time-varying unobserved information evolution rate.[11] However, since the arrival rate of new information is unobservable, we do not know whether a part of the volume-volatility relation may be a result of the actions of heterogeneous traders. As suggested by Harris (1987), if trading is self-generating, the daily number of transactions would be the *true* mixing variable rather than a proxy for the unobserved information evolution rate.

Using a simple regression approach for daily data on Nasdaq-NMS securities over the 1986-1991 period, Jones et al. (1994) find that both volatility and trading volume are positively correlated with the number of daily transactions. However, the average size of trades contains no additional information about volatility beyond that contained in the number of transactions. If the number of transactions is a good proxy for the mixing variable, this result is supportive of a pure MDH; "..volatility and volume are positively correlated only because both are positively related to the number of daily information arrivals (the mixing variable)." The problem caused by a lack of ability to interpret the mixing variable can be further illustrated by this study. If informed traders camouflage their information, for example by splitting their orders into medium sized trades as suggested by the "stealth trading hypothesis" of Barclay and Warner (1993), the number of daily transactions would be the true mixing variable and the results in Jones et al. (1994) would also support explanations of the volume-volatility relation based on heterogeneous traders.[12]

Dispersion of belief The MDH simply states that price changes and trading volume are directed by the flow of new information. It does not say anything about what type of information or how this information is revealed to investors. Hence, an important limitation of the hypothesis is that it does not address the role of economic agents or market structure for prices and trading volume. Later theoretical work on the volume-volatility relation centers around these issues. Harris (1986) links the MDH to asset pricing theory by suggesting that the mixing variable directs the rate of flow of information from systematic risk factors. A problem with this interpretation is that the movements in prices for many types of financial contracts seem much "too large" to be attributed to movements in the fundamental values of the underlying securities only. This fact suggests that prices are not merely driven by changes in systematic risk factors and asset payoffs, but also by changes in the expectations of heterogeneous agents. Figure 4.1 illustrates the information structure in a standard asset market for the two main types of such models. Panel (a) in the figure describes a "differences in opinion" model, while panel (b) describes a market microstructure model with asymmetric information.

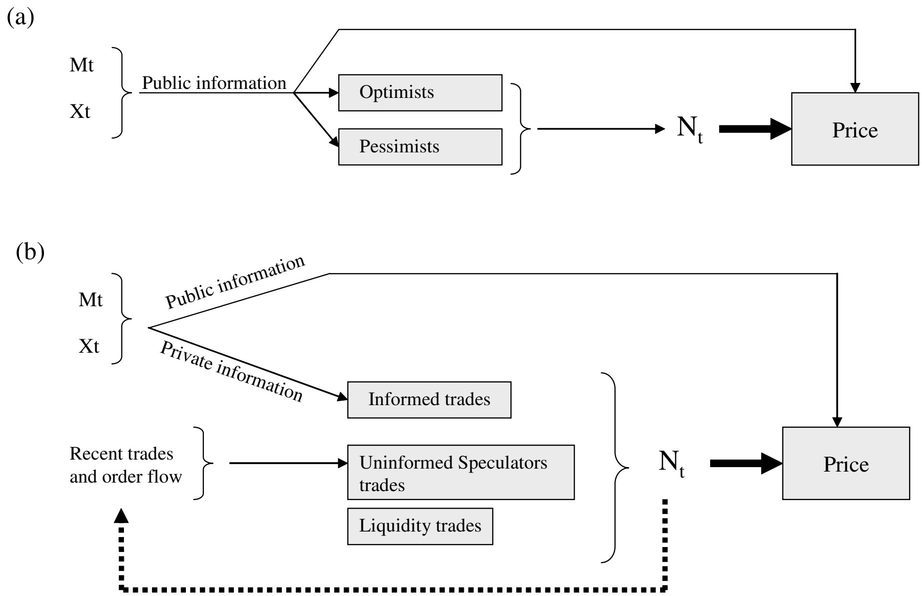
Figure 4.1 (a) illustrates a "differences of opinion" model. In this model, investors are assumed to act differently on the same news, i.e. trading is induced by differences of opinion

[11] Harris (1987) derives and tests several implications of the MDH for transactions data on a sample of 50 NYSE stocks that traded between December 1, 1981 and January 31, 1983. The results from the tests are supportive of the MDH.

[12] In addition, in order-driven markets, a large order is often automatically executed against many smaller orders by the automatic matching system. Thus, even though the original order is large, it may show up as many small trades as it is matched against several smaller orders rendering the average daily trade size unimportant in explaining volatility.

Figure 4.1: The information structure

The figure illustrates the assumed information structure in a "differences in opinion" model (panel a) and a market microstructure model (panel b). From the fundamental asset pricing equation, $P_{i,t} = E_t[\sum_{j=0}^{\infty} M_{t+j} X_{i,t+j}]$, we know that relevant information about the price, P, of an asset, i, may come from either news about the stochastic discount factor, M_i, or news about the payoff, X_i. In the "differences in opinion" model in panel a, all news arrives as public information. Some types of information are immediately incorporated into the asset price. For other types of information, traders disagree on the effects on the valuation of the underlying assets. Trading occurs whenever the cumulative information for a particular type of trader switches from favorable to unfavorable. In the market microstructure model in panel b, new information arrives as either public or private information. Public information is immediately incorporated into the asset price. Informed traders trade on the basis of private information. Uninformed investors are either liquidity traders or speculators. The uninformed investors are trying to infer the private information from the trades, N_t. However, they are not able to separate informed and uninformed trades.



about publicly available information. Beliefs are updated using Bayes rule. All traders are rational, but they view others as having irrational models. Harris and Raviv (1993) explain the volume-volatility relation by a model of this kind. Two groups of risk-neutral speculators receive the same information but disagree on the extent to which it is important (but agree to disagree). As long as one of the groups remains more optimistic than the other, there is no trading. Trading occurs only and whenever the cumulative information for one of the trader groups switches from favorable to unfavorable, or vice versa.[13]

In the standard asset pricing models, the trading process itself does not convey information which is relevant to price determination. Prices adjust immediately as a result of new information. This is a plausible assumption for some kind of news. Other types of news are likely to be dispersed and not immediately available to all investors in aggregated form.[14] Modelling

[13]Prices change every period whether or not trading occurs. The volume-volatility relation arises because the price changes are larger on average when trading occurs.

[14]Evidence of the existence of dispersed news is given in French and Roll (1986) who document empirically that asset

dispersed information is the essential feature in the market microstructure models illustrated in figure 4.1 (b). In these models, there is a group of investors who trade on the basis of private information. The market maker and the uninformed investors can only infer this information from trades and order flows. The room for strategic behavior among agents differs in different models.[15] Shalen (1993) uses a market microstructure model to study the volume-volatility relation. In her model, both trading volume and price volatility increase with the dispersion of traders' expectations about fundamental values. This is called the "dispersion of beliefs hypothesis"(hereafter the DBH). In this version, dispersion of beliefs about the value of a security is assumed to be wider the larger the share of the traders in the security that consists of uninformed investors. Uninformed traders cannot distinguish informed trades from liquidity trades. Instead they react as if all trades were informative, and thus they increases both volatility and volume relative to equilibrium values.

Daigler and Wiley (1999) perform an indirect test of the DBH. Facilitating the possibility of distinguishing traders with different types of information in the futures markets, they test whether uninformed traders contribute to volatility. The results of their study support Shalen (1993): "..uninformed traders who cannot differentiate liquidity demand from fundamental value increase volatility." In a similar study, Bessembinder and Seguin (1993) examine the relation between the volume-volatility relation and market depth, proxied by open interest, in eight physical and financial futures markets. Unexpected volume is found to have a larger effect on volatility than expected volume, and large open interest is found to mitigate volatility.

Limit order markets In this paper, we investigate the information about trading volume and price volatility contained in the slope of a limit order book. In an electronic limit order market, liquidity is not supplied by designated specialists or market makers, but rather by the traders themselves. The majority of trades are first submitted to the market as limit orders, which accumulate into the limit order book. Hence, at any point in time, the limit order book reflects an aggregate of buying and selling interests at various prices. Each ask (bid) price reflects the lowest (highest) price at which different investors are willing to sell (buy) the security.[16]

Theoretical models of limit order markets differ in their assumptions about investor heterogeneity. In Glosten (1994), privately informed investors are assumed to submit market orders while homogeneous uninformed investors provide the limit order book. Hence, the shape of

---

prices are much more volatile during exchange trading hours than during non-trading hours. This phenomenon cannot be reconciled with a standard asset pricing model unless there is a systematic tendency for price-relevant information to arrive during normal business hours only.

[15] In Kyle (1985), informed investors attempt to camouflage their trades by spreading them over time. Kyle's model implies that larger volumes support more informed traders. In Admati and Pfleiderer (1988), a certain amount of the uninformed investors are allowed to act strategically by having the discretion to time their trading. This is shown to imply that within-day trading becomes concentrated. Hence, price changes and transactions are bunched in time, and the effect of volume on price movements will depend on recent volume levels.

[16] Biais et al. (1995) note that the shape of the order book may reflect the competition among buyers/sellers as well as the correlation in their valuations. If the supply and demand curves are inelastic and volume is concentrated around the inner quotes, this may reflect that the valuations among various investors are correlated on each side of the market relative to the case where the valuations are more dispersed and the order book is more elastic.

the limit order book reflects the information characteristics of the incoming market flow. In the dynamic model proposed by Parlour (1998), all traders are assumed to have different valuations for the traded asset. Parlour shows that, when the choice between a limit order or a market order depends both on the past (through the state of the order book) and the future (through expected subsequent order flow), then systematic patterns in order placement strategies will be generated even in the absence of asymmetric information. Moreover, both sides of the book will matter for optimal order placement strategies. Foucault et al. (2003) model a limit order market where liquidity suppliers have asymmetric information on the risk of being picked off by traders with superior information. This feature is shown to affect the shape of the order book. When the book is thin, uninformed liquidity traders are reluctant to add depth because it may be an indication of high pick-off risk. The informed liquidity traders exploit this by bidding less aggressively than in the case where the liquidity traders have symmetric information. Sandås (2001) tests a version of Glosten (1994) empirically.[17] The results do not lend support to the model. Relative to the theoretical predictions, the empirical price schedules of the limit order book offer insufficient depth.

3 The data

3.1 The Norwegian stock market

Our data set is from the the Oslo Stock Exchange (OSE) in Norway.[18] Norway is a member of the European Economic Area, and its equity market is among the 30 largest world equity markets by market capitalization.[19] Table 4.1 reports some general statistics for all the companies listed on the OSE. At the end of 2001, 212 firms were listed on the exchange with a total market value of about NOK 657 bill. The OSE is the only regulated marketplace for securities trading in Norway. Since January 1999, it has operated as a fully computerized centralized limit order book system similar to the public limit order book systems in e.g. Paris, Toronto, Stockholm and Hong Kong.

The OSE allows the use of limit orders, market orders, and various customary order specifications. Participants can also submit hidden orders. When an order is submitted as a hidden order, only a specified fraction of the underlying order is visible to the market. As is normal in most electronic order-driven markets, the order handling rule follows a strict price-time priority.[20] All orders are submitted at prices constrained by the minimum tick size for the respective stocks which is determined by the price level of the stock. For prices lower than NOK 9.99

[17] The tests are based on updating restrictions that link the market order flow to the order book dynamics and break-even conditions for the marginal bid and offer prices that define the price schedule.

[18] We obtained the data directly from the exchange's surveillance system. The SMARTS© system is the core of the exchange's surveillance operations. Through access to the SMARTS© database, we obtained all the information on orders and trades in the market

[19] Source is FIBV (International Federation of Stock exchanges). Notable Norwegian listings include Norsk Hydro, Telenor, and Statoil.

[20] In the case of hidden orders, when the visible part of the order is executed, it loses time priority.

Table 4.1: The Oslo Stock Exchange (OSE) - General statistics

Descriptive statistics for the Oslo Stock Exchange for the period 1999 to 2001. All numbers in the table are official statistics obtained from the OSE annual reports (available at www.ose.no).

	1999	2000	2001
Number of listed firms	215	214	212
Market capitalization (bill. NOK)	582.94	637.86	677.03
NOK/USD exchange rate[a]	7.81	8.81	8.99
Turnover velocity[b]	88.6	96.7	86.4
Market development			
Market index level (TOTX)	1153.74	1366.05	933.22
OSE benchmark index	189.76	195.79	167.18
OSE benchmark index return (%)	48.45	3.18	-14.61

[a]Average midpoint rate for the respective year. [b]Turnover velocity: Average of annualized turnover per month divided by market value at the end of each month.

(Norwegian kroner) the tick size is NOK 0.01, between NOK 10 and NOK 49.9 the tick size is NOK 0.1, between NOK 50 and NOK 999.5 the tick size is NOK 0.5 and for prices above NOK 1000 the tick size is NOK 1.

The trading day of the OSE comprises two sessions: the “pre-trade” session starting at 9:30 and ending with an opening auction at 10:00, and the “continuous trading” session from 10:00 until the trading closes at 16:00. During the pre-trade session, brokers can register trades that were executed after the close on the previous day as well as new orders. At the opening auction at the end of the pre-trade session, all orders registered in the order book are automatically matched if the prices are crossing or equal. The quoted opening price is thus the price that clears the market. During the continuous trading session, electronic matching of orders with crossing or equal price generates transactions. Orders without a limit price (market orders) have automatic price priority and are immediately executed at the best available prices. At the OSE, market orders are allowed to “walk the book” until they are fully executed. Any remaining part left of the market order is removed from the order book. This is different from the treatment of market orders on e.g. the Paris Bourse, where any remaining part of an unfilled order is automatically converted to a limit order at the current quote. The difference implies that market orders on OSE are more aggressive than market orders at the Paris Bourse. On the Paris Bourse, market orders are essentially marketable limit orders.

3.2 The data sample

The dataset consists of every order and trade that occurred on the OSE in the period from February 1999 through June 2001.

The trade data contain, quantity transacted, a time stamp, brokerage house ID on each side, and an ID for the house that initiated the trade as well as whether the house was the buyer or the seller in the transaction. Every trade is linked to the underlying orders through an order ID. Thus, if a large order is executed against many smaller orders resulting in several smaller trades, we can trace each executed part back to the initial order. There are also additional flags attached to each trade that identify special features of the trade such as whether it was an odd-lot trade, an off-exchange trade, a cross (within the same or different brokerage houses), and whether a trade results from a market order or a limit order. The order data contain all order entries as well as all deletions and amendments of orders already in the order book.

In table 4.2 we provide some descriptive statistics of the trade data throughout our sample period. A large part of the listed firms are traded quite infrequently. Since we examine intraday data, including infrequently traded firms would introduce a large amount of noise into our analysis. We therefore filter the firms based on their trading activity through the sample period. The first filtering criterion is that the firm must have been traded in at least 400 out of 597 days, or about 70 percent of the trading days, and the second criterion is that the firm must have an average of 5 trades per day to be included in our sample. Once the first criterion is applied, the second criterion only removes a few companies from our sample. After the filtering we are left with 108 firms, which constitute our sample throughout the paper. Note that there were 215, 214 and 212 listed firms at the end of 1999, 2000 and 2001 respectively. Table 4.2 shows that there has been increasing trading activity during the sample period with the total number of trades having tripled and the volume in Norwegian kroner (NOK) having doubled. Further, the average number of daily trades across firms has more than doubled from 32 in the first half of 1999 to 79 in the first half of 2001.[21] The increase in activity has also been accompanied by a decrease in the average percentage spread. As found in most markets, the average effective spreads are lower than the average quoted spreads. To give a better picture of the diversity of the sample, we divide the sample into four portfolios based on their market capitalization value. The firms are assigned to a market capitalization group based on their market capitalization value at the beginning of each year. The general picture is that the number of trades, the trading volume (both in shares and NOK), the prices and the quoted spread increase across firm size portfolios, while the average daily volatility, the average trade size and the quoted percentage spread decrease.

We also report the average correlations between the trading volume, the trade size and the number of transactions. The correlation structure in our sample is quite similar to the one documented for the US market in Jones et al. (1994). The correlation between the average trade size and the number of trades is low, and both the average trade size and the number of trades have high positive correlations with share volume. Hence, the two components of share volume seem to contain different information about volume. The same structure is evident when we calculate correlations over sub-periods of half a year.

[21] At the same time, the average trade size has gone down from 3429 shares to 2648 shares. This decline is most likely related to the introduction and growth of online trading in the sample period, since these traders generate a lot of trades of small sizes. During our period, the fraction of total trades coming from pure online brokerage houses has increased from 0% to almost 10%.

Table 4.2: Descriptive statistics of trades

The table provides some descriptive statistics of trades for the whole sample, the five half-year sub-periods, and the four market capitalization groups. Group 1 consists of the 25% smallest firms while group 4 consists of the 25% largest firms. Some firms have experienced large changes in capitalization value during the sample period. To take account of this, we re-sort the market capitalization groups at the beginning of each year. At the bottom of the table we report the Pearson correlation coefficients between the trading activity variables. The number of trades (N) is the average number of daily trades across all firms. The share volume (V) is the average daily share volume (in 1000 shares) across all firms. The average trade size (AV) is the average number of shares in each trade averaged across all firms for the sample period. The quoted spread is calculated as a percent of the spread midpoint. The effective spread is calculated as the difference between the execution price and the spread midpoint (in percent of the spread midpoint) multiplied by two.

		Sub-periods (half years)					Market Capitalization groups			
	Whole sample	1999.1	1999.2	2000.1	2000.2	2001.1	1	2	3	4
Aggregate statistics:										
Number of firms	108	107	108	108	108	104	27	27	27	27
Trades (in thousands)	3724	328	545	946	953	953	390	522	504	2309
Shares traded (mill.)	9585	1339	2300	2027	2072	1847	1707	1922	919	5037
NOK volume (bill.NOK)	648	67	131	152	153	146	21	44	68	516
Cross-sectional averages:										
Market cap (mill.NOK)	5259	4120	4714	5507	6127	5836	354	938	2339	13978
Price	88.4	71.8	82.7	102.7	102.3	81.9	23.34	62.43	105.66	150.73
Daily volatility (%)	2.71 %	2.64 %	2.89 %	2.98 %	2.48 %	2.57 %	3.49 %	2.98 %	2.30 %	2.29 %
Shares traded (thousands)	151	130	167	155	151	153	116	171	78	288
Trades	58	32	40	72	69	79	28	41	41	148
Trade size (AV) in shares	2890	3429	3365	2453	2551	2648	4859	2684	1549	1912
Quoted spread (NOK)	1.65	1.55	1.62	1.79	1.78	1.50	0.94	1.63	2.11	1.57
Effective spread (NOK)	1.22	1.12	1.14	1.34	1.36	1.13	0.68	1.20	1.59	1.16
Quoted % spread (midpt.)	3.04 %	3.66 %	3.49 %	2.62 %	2.55 %	2.89 %	4.74 %	2.77 %	2.40 %	1.34 %
Effective % spread (midpt.)	2.22 %	2.67 %	2.48 %	1.92 %	1.89 %	2.15 %	3.38 %	2.03 %	1.85 %	0.99 %
Correlations:										
Corr(AV,N)	-0.061	0.045	0.051	-0.091	-0.085	-0.084	-0.116	0.280	0.172	-0.066
Corr(V,N)	0.525	0.660	0.591	0.724	0.568	0.442	0.690	0.480	0.847	0.365
Corr(V,AV)	0.330	0.358	0.438	0.290	0.288	0.201	0.393	0.932	0.504	0.759

3.3 Composition of orders

Our order data are quite rich. For each order, we have a time stamp, a unique order ID, the disclosed/hidden orders as well as flags indicating whether the order was a buy or sell order, whether the order is a new order, a deletion of an order or an amendment to an existing order (price change and/or volume change). In addition, a unique brokerage house ID is attached to each order. Moreover, compared to the Paris Bourse data in Biais et al. (1995), our data are not restricted to include placements, amendments and deletions of orders within the 5 best quotes. We have access to all orders, which makes it possible to reconstruct the full order book at any point of time. The descriptive statistics discussed in this section are based on 6 hourly spaced snapshots of the entire order book for each company at each trading day in the sample period. The order book is rebuilt at 10:30, 11:30, 12:30, 13:30, 14:30 and 15:30 each trading day for each firm. We exclude order volume above/below 100 ticks away from the inner quotes. For a stock trading at NOK 100 with a minimum tick size of NOK 0.5 this would mean that orders above NOK 150 and below NOK 50 are excluded from our calculations. The limit on 100 +/- ticks means that we disregard less than 5 percent of our sample.

To get a general view of the composition of the order-flow, we group the orders into four types based on their trading aggressiveness. “Market orders” are orders with no limit price. “Aggressive limit orders” are orders that are placed at the opposite quote (marketable limit order) or at a price further away from the best quote on the opposite side. “Quote improving orders” are orders that are placed in between the inner quotes, and “Passive orders” are orders that are placed at the best (same side) quote or further away from the market. Panel A in table 4.3 shows the composition of orders and the order book activity for our data sample. The numbers in the table are daily cross-sectional time series averages of order volumes (in shares), and the number of orders submitted. The numbers are averaged over each of the three years in the sample as well as over market capitalization quartiles. Each firm is assigned to a market capitalization quartile at the beginning of each year.

The table shows the distribution of order placements in the market. The use of market orders is modest. However, market orders and aggressive limit orders together constitute around 40 percent of the average daily number of submitted orders. Measured in number of shares, there is considerable variation in the size of the submitted orders across order groups. A part of this variation can probably be explained by differences in the price level of the stocks, both over time and over firm size. Quote improving orders are the largest order group, while market orders are the smallest order group. This holds for the entire sample as well as for each market capitalization group, and is also a systematic pattern across sub-periods (not shown in the table). Measured over the whole sample, on average 94 orders are submitted during a trading day for one firm. The activity is considerably higher for the largest firms than for firms in the other three groups. The average daily number of orders submitted for the largest firms was 224, while the similar average for the three other groups ranged from 45 to 53. For comparison, Biais et al. (1995) report an average of 160 orders for the Paris Bourse in 1995.

Table 4.3: Descriptive statistics of the order book

Panel A shows the daily average number of submitted orders and the daily average order size for different types of orders. The numbers are averaged over companies and time. We also report averages over the four market capitalization groups. Group 1 consists of the 25% smallest firms while group 4 consists of the 25% largest firms. Some firms have experienced large changes in capitalization value during the sample period. To take account of this, we re-sort the market capitalization groups at the beginning of each year. Limit orders are classified into three different types based on their aggressiveness. Passive orders are orders that are submitted at the best (same side) quote or further away from the market. Quote improving orders are orders that are submitted in between the inner quotes prevailing at order submission, and aggressive orders (Aggr.) are orders that are submitted at the opposite quote (marketable limit order) or at a price further away from the market on the opposite side. Market orders (MO) constitute a separate group. The numbers in parentheses are each order class' fraction of total orders. Panel B provides descriptive statistics on the distribution of order book volume. The numbers are daily average fractions of accumulated volume, and are reported for all firms, for the bid and ask side separately, for minimum tick sizes, and for the four market capitalization groups.

PANEL A: Order types and order sizes

		Submitted orders					Order sizes			
	Firms	Total orders	Passive	Quote impr.	Aggr.	MO	Passive	Quote impr.	Aggr.	MO
All firms	108	94	42 (0.44)	15 (0.16)	34 (0.36)	4 (0.04)	6428	7063	5882	1715
Market capitalization quartiles										
1 (small)	27	45	22 (0.45)	10 (0.21)	14 (0.31)	3 (0.06)	10708	11501	9824	4341
2	27	52	23 (0.43)	10 (0.19)	18 (0.34)	3 (0.05)	6244	7460	5634	1382
3	27	53	22 (0.41)	10 (0.19)	19 (0.37)	3 (0.05)	3437	3900	3038	531
4 (large)	27	224	100 (0.45)	31 (0.14)	87 (0.39)	7 (0.03)	5324	5392	5032	605

PANEL B: Order book volume distribution (normalized)

Minimum tick size	ATQ	+/- 1 tick	+/- 5 tick	+/- 10 tick	+/- 20 tick	+/-50 tick	+/-100 tick
All firms	20.9 %	34.7 %	56.8 %	69.4 %	78.4 %	88.6 %	100.0 %
Bid side	23.0 %	40.0 %	62.9 %	73.8 %	81.4 %	89.7 %	100.0 %
Ask side	20.8 %	29.3 %	50.8 %	64.9 %	75.5 %	87.4 %	100.0 %
Minimum tick size							
0.01	20.2 %	30.8 %	37.8 %	49.0 %	60.1 %	82.2 %	100.0 %
0.1	22.2 %	34.2 %	53.2 %	67.4 %	79.4 %	91.7 %	100.0 %
0.5	22.3 %	39.1 %	65.8 %	78.4 %	88.1 %	95.5 %	100.0 %
1	7.0 %	10.7 %	17.6 %	25.1 %	38.8 %	70.0 %	100.0 %
Market capitalization quartiles							
1 (small)	19.1 %	29.7 %	45.2 %	56.6 %	68.2 %	84.0 %	100.0 %
2	21.6 %	34.9 %	56.3 %	69.6 %	79.9 %	91.1 %	100.0 %
3	23.6 %	38.3 %	62.7 %	75.5 %	83.8 %	92.6 %	100.0 %
4 (large)	19.3 %	34.6 %	62.9 %	75.9 %	84.3 %	91.0 %	100.0 %

In Panel B in table 4.3, we show the distribution of volume in the order book averaged across all firms and dates. At each tick level, the fraction of total shares in the order book is averaged over the 6 order book snapshots.[22] The table shows the order book distribution across minimum tick sizes and market capitalization quartiles.[23] Around 35 percent of the order book depth is concentrated at the quotes or plus/minus one tick from the quotes. This is quite stable both across tick sizes and across market cap quartiles. However, when we separate the bid and ask sides, we find that the volume on the bid side is more concentrated at the inner quotes than the volume on the ask side. This is in line with the findings in several other empirical papers, and is consistent with the interpretation that the price impact is larger for buy orders than for sell orders.[24] Note that the depth within +/- 5 ticks, which is what Biais et al. (1995) investigate, only includes 56 percent of the total order book depth in our sample. There does not seem to be large differences in order depth across market capitalization quartiles. The largest tick size category is special in that it only contains one, highly volatile and very actively traded, company.[25] One interesting thing to note about this firm is that as much as 30 percent of the order book depth lies between 50 and 100 ticks away from the quotes, even though it has been one of the most heavily traded companies at the exchange during our sample period.

4 Intraday analysis of the order book

In this section, we discuss how to measure the shape of the order book, and present statistics on the limit order book at an intraday level.

4.1 The shape of the order book

Figure 4.2 shows the average order books for two companies listed on the OSE. The order books are averaged over the five last trading days in May 2001, and are normalized in the sense that they show the percentage of shares in all orders within an increasing/decreasing number of ticks away from the quotes (zero in the figure is the best quote on each side of the market). The upper graph shows the average order book for Norsk Hydro (NHY) while the lower graph shows the average order book for Opticom (OPC). Both companies are among the most liquid on the exchange.[26] Norsk Hydro is a leading energy, aluminium and fertilizer company, based

[22]For instance, on the ask side of the book for one company/snapshot, we divide the aggregate number of shares at each tick by the total number of shares supplied (offered) at that time/snapshot. We do this for each snapshot, and average across all snapshots on the particular date to obtain the average fraction supplied on each tick for the security. Since we limit the order book to orders within +/- 100 ticks from the bid/ask midpoint, the fraction of aggregate volume at +/- 100 ticks is 100%.

[23]If a firm trades across two minimum tick size regimes on the same day, we remove that company for that day from the sample. The results do not change if we include these observations.

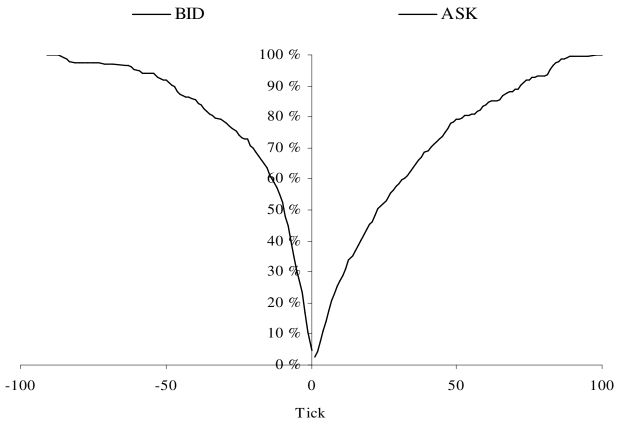
[24]See Chan and Lakonishok (1993), Chan and Lakonishok (1995), and Kalay et al. (2003)

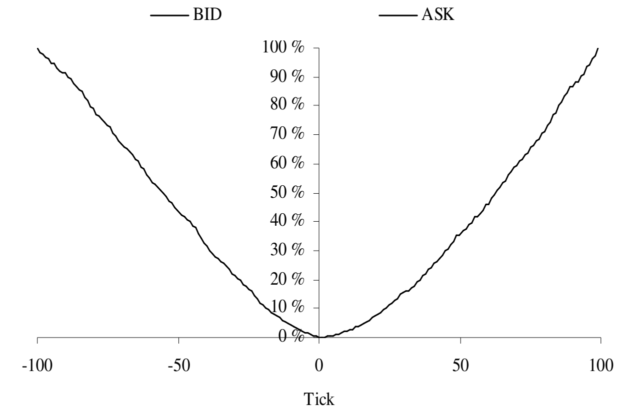
[25]The company is Opticom (OPC).

[26]During the period illustrated in the figure, both companies traded in prices around NOK 400-500 and had a tick size of NOK 0.5. For Norsk Hydro the calculated average order book is based on around 2000 orders with a share volume of around 400 000 shares. For Opticom the similar calculations are based on around 4000 orders with a share volume of around 200 000 shares.

Figure 4.2: Average order books for Norsk Hydro and Opticom

The figure illustrates the order books for two different companies listed on the OSE. The upper graph shows the average normalized (with respect to the total number of orders in the order book) order book for Norsk Hydro (NHY), a large Norwegian blue chip company, and the lower picture shows the average normalized order book for Opticom, a Norwegian IT company. The order books are averaged over the last five days of May 2001. (For each day the average order book is calculated from hourly snapshots of the book.) The picture shows the percentage of shares in all orders within varying ticks away from the quotes. Zero represents the best quote on each side of the market.





in Norway. It has 50,000 employees in 60 countries worldwide. The company's operations are well known and there is a large amount of available information about the company, including experts' analyzes. Opticom, on the other hand, is a relatively new IT company which currently has under 100 employees. The company describes its business concept as pioneering research and development in new technology in electronics. The company has no cash flow and very uncertain future income possibilities. Discussions in the popular press have been largely focused on how difficult it is to value the company, and there have been large differences in analysts' valuations. The picture shows that the order book of the two companies are quite different: while on average about 50 percent of the orders for Norsk Hydro has limit prices which lie within 5 ticks from the quoted spread, the similar percentage for Opticom is only about 10 percent.

This difference in the average shape of the order book results from the fact that traders systematically submit orders further away from the midpoint in Opticom than in Norsk Hydro. One possible reason for this is that investors are more uncertain about the true value of Opticom than Norsk Hydro, and that this higher valuation uncertainty in Opticom is reflected in orders being submitted across a wider range of prices than in Norsk Hydro. The difference in the order book shapes may also come from pick-off risk, i.e. the reservation prices reflect a compensation for the risk of being picked off by better informed traders. Probably both effects contribute to explaining the pictures we see in figure 4.2. However, while it is obvious that there are huge differences in valuation uncertainty between the two companies, it is not so obvious that there should be such a big difference in pick-off risk. More importantly, pick-off risk should mainly concern the orders submitted close to the midpoint price. Thus, pick-off risk should affect the spread and volumes at the inner ticks, not the distribution of orders across the the entire order book. The figure also illustrates the difference in order book liquidity between the bid and the ask side, which we documented for the whole sample in panel B in table 4.3. Although it is more pronounced for Norsk Hydro, both pictures indicate that the ask side of the book is more elastic than the bid side.

Measuring the order book slope To capture the shape of the order book, we use the average elasticity/slopes of the supply and demand schedules in the order book. The more gentle (steeper) the slope, the more widely distributed (concentrated) are the bid and ask prices in the order book. Note that we use the inverse of the elasticity, with prices on the x-axis and accumulated volumes on the y-axis, as in Biais et al. (1995).

To obtain an average slope of the order book, we divide the trading day into hourly spaced intervals. At the end of each interval, we take a snapshot of the order book. These snapshots occur at 10:30, 11:30, 12:30, 13:30, 14:30 and 15:30 each trading day for each firm. Note that the first snapshot is half an hour after the regular trading session starts. Alternatively, we could end the last snapshot at 16:00, but then the order book would be affected by the large amount of order cancellations at the end of the trading day. To rebuild the order book we start at the beginning of the trading day with the orders still remaining after the opening auction has been executed at 10:00. Then we track all types of orders being submitted throughout the day, and update the order book accordingly. Thus, all deletions and/or amendments of earlier orders as well as new orders are accounted for when we update the order book.[27] After having obtained the full order book for each snapshot we calculate our slope/elasticity estimate for each company of the order book in the following steps:

1. Firstly, for each side of the order book, and each snapshot, we accumulate the aggregate number of shares supplied/demanded at each price level, such that at each price level we get the total volume supplied (demanded) at that price or lower (higher).

2. To account for large differences in liquidity between firms, we normalize the accumulated shares at each tick level (on the ask and bid side separately) relative to the total number of shares supplied/demanded at the relevant snapshot. Thus, the percentage of the shares in the order book supplied (demanded) at the highest (lowest) ask (bid) price/tick is 100 percent.

3. Next, we calculate the "local" elasticity at each price level (illustrated in equation A.2 and equation A.3 in the appendix).

4. Then, we average across all price levels (local slopes) to obtain an average elasticity/slope for the bid and ask side for that snapshot.

5. Finally, we take the average of the bid and ask slope to get one slope measure for the snapshot and average across all the snapshots during the trading day to obtain the average slope for each company on that day.

We normalize the order book because we want to take into account that there is a close relationship between our slope measure and the liquidity of the underlying stock. Less liquid firms generally have a higher volatility since the order book does not contain enough volume to absorb large trades without moving prices too much. In addition, less liquid stocks generally have a higher spread since investors require a discount when buying and a premium when selling the stock. Thus, a positive relationship between order book elasticity and volatility is expected a priori. By normalizing the order book, we get the fraction of total shares supplied/demanded at each price level regardless of the total volume in the order book. This makes the order books more comparable across firms and time.

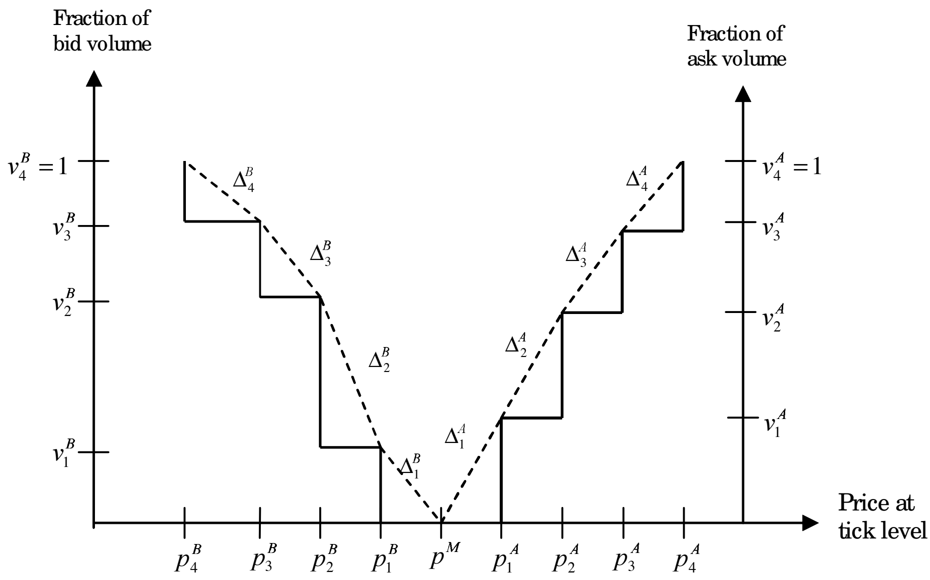
In addition to the equally weighted slope of the order book (across tick levels), we calculate a slope measure where we weight each local slope by its distance (in ticks) from the inner quote. The tick-weighting implies that local slopes further out in the book have a lower impact on the average slope than local slopes closer to the midpoint price. The main reason for doing this is to reduce the effects that "stale" orders may have on the "tails" of the order book.

Figure 4.3 illustrates how the local elasticities, Δ_τ^A and Δ_τ^B, are calculated. For illustrative purposes, the order book in the figure stretches only across 4 price levels on each side. In the

[27] The original dataset from the Oslo Stock Exchange includes order book data for the best 5 quotes on each side whenever a new order is submitted or there is a deletion or amendment of an existing order. We use this information to check that our order book is correct for these 5 levels of the book.

Figure 4.3: Calculation of the demand and supply elasticities

The figure illustrates how the local slopes/elasticities on the bid and ask side of the order book are calculated for one "snapshot" time on one date for one company. There are only 4 price levels on both sides of the book. The left y-axis shows the fraction of aggregate share volume on the demand (bid) side of the order book at each tick level. Similarly, the right y-axis shows the fraction of aggregate share volume on the supply (ask) side of the order book at each tick level. The solid step-line is the supply (right) and demand (left) curves over the various price levels. On the x-axis, we have the various price levels. p^M is the bid/ask midpoint. Prices greater than p^M are ask prices and prices below p^M are bid prices. The difference between p_1^B (best bid) and p_1^A (best ask) is the quoted spread. The dotted line-segments connecting each level of the order book have local slopes denoted by Δs. These are the normalized local elasticities of the demand and supply curves calculated in equation A.2 and equation A.3 in the appendix.



figure, p_1^A is the best available ask price (inner ask quote) with volume fraction of v_1^A supplied at that ask price. The volume fraction at the next tick level (v_2^A) is thus the accumulated volume supplied at price p_1^A and p_2^A relative to the total volume in the order book on each side. The local elasticity of the supply curve at p_1^A would thus be the slope Δ_2^A in the figure. A more specific explanation of the calculation is provided in the appendix.

When we normalize the order book, the slope measures the average percentage change in normalized volume when the price level changes by one percent. For example, suppose that the current bid price is 49 (ask price is 50), the normalized depth is 10 percent and the slope is 10. If the bid-price decreases by 1 percent to 48.5 (or the ask price increases by 1 percent to 50.5), the normalized depth will increase by 10 percent to 11 percent.

4.2 Intraday Statistics

Table 4.4 shows intraday statistics for our slope measure (calculated at the end of each time interval), the price volatility (measured as the absolute hourly return between midpoint prices closest to the end of each time interval), the quoted and the effective spread, the number of trades

Table 4.4: Intraday statistics

The table provides intraday statistics for the data sample, including the slope measure, the volatility (the absolute hourly return between trade prices closest to the end of each interval), the quoted spread, the effective spread, the number of trades executed during the time interval, the trade size (in shares), the number of orders submitted during the time interval, and the order size (in shares). All numbers are daily averages across all firms in the sample. Note that the first and last time windows are half an hour while the rest of the time windows are hourly. The slope is calculated at the end of each interval.

	Time window						
	10:00 to 10:30	10:30 to 11:30	11:30 to 12:30	12:30 to 13:30	13:30 to 14:30	14:30 to 15:30	15:30 to 16:00
Slope (end of time-window)	30.51	34.37	35.78	36.34	36.80	36.97	-
Volatility (absolute return)	-	1.34 %	0.81 %	0.72 %	0.74 %	0.88 %	0.86 %
Quoted spread	2.36	1.73	1.47	1.37	1.33	1.31	1.39
Effective spread	1.79	1.27	1.05	1.00	0.95	0.95	1.05
Trades	10.38	11.81	9.38	9.05	9.52	10.81	10.40
Trade size (shares)	2314	2653	2759	2774	2834	3027	3123
Orders	15.45	18.16	13.10	12.36	12.47	14.02	11.66
Order size (shares)	6858	6385	5723	5818	5795	6383	6706

executed during the time interval, the trade and order sizes measured in shares, and the number of orders submitted during the time interval. All numbers are daily averages across all firms in the sample, and the time intervals correspond to those used for rebuilding the order book.

Notable characteristics of the intraday statistics in table 4.4 are;

- The average slope increases at a decreasing rate throughout the day.
- The quoted and the effective spread both have a U-shape, with the highest spread at the beginning of the day.
- The average trade size is smallest at the beginning of the day, and increasing throughout the trading day.
- The average number of orders and trades both follow a U-shape, with fewer orders being placed and trades being executed in the middle of the day, and most orders being placed and trades being executed at the beginning of the day.

These regularities are also systematic across sub-periods.[28] Similar systematic intraday regularities have been found in other markets (e.g. US, France, Hong Kong, Sweden).[29] Following the sequential trading model in Glosten and Milgrom (1985), these data features can be explained by higher uncertainty about other traders' valuations at the beginning of the trading day than during the day. If this explanation is correct, a patient liquidity trader who fears being picked off by informed investors at the beginning of the day has two main options. If she believes that the probability of trading with informed traders will diminish during the day, she can act strategically and delay her trading. Alternatively, she can submit her orders at the beginning

[28] We also calculate the statistics across sub-periods of years, half-years and quarters and find that the results are both qualitatively and quantitatively similar.

[29] For the US, see French and Roll (1986) and Harris (1986). For Sweden, see Niemeyer and Sandas (1995).

of the day and take account of the increased probability of incurring a loss by placing them at prices including a discount (buys) or a premium (sells). This can explain the higher spread at the beginning of the trading day. The increase in spreads towards the end of the day may be due to higher liquidity demand and possibly more cancellation of orders just before the close. Assuming that the informed traders are trying not to reveal their information too quickly, we would also expect to see a higher number of small trades at the beginning of the trading day (stealth trading).

To obtain a measure of order aggressiveness during the trading day, we calculate a separate index similar to Harris and Hasbrouck (1996), where the aggressiveness of an order is measured by the average number of ticks the order is placed away from the best quote (on the same side). Thus an index number of zero means that the average order is placed at the quote, a positive index number means that the order is placed above (below) the bid (ask), and a negative number means that the average order is placed below (above) the bid (ask).[30] Formally, for an order of type k, the aggressiveness of a buy order with a limit price p^B is calculated as,

$$\lambda_k^{buy} = (p^B - bid)/ticksize \tag{4.2}$$

Similarly, a sell order with a limit price p^S is calculated as,

$$\lambda_k^{sell} = (ask - p^S)/ticksize \tag{4.3}$$

where *bid* and *ask* are the best bid quote and best ask quote, respectively, when the order is submitted.

Table 4.5 shows the intraday pattern in order aggressiveness, average number of orders, fraction of order types, and order sizes. Figure 4.4 illustrates graphically the intraday patterns in order aggressiveness, order size, order book slope, quoted and effective spread, and fraction of order types.

If uninformed investors believe that there is more asymmetric information at the beginning of the trading day, we would expect to see that they place orders at limit prices further away from the midpoint price at the beginning of the trading day, and then, closer to the midpoint prices later in the day, as the market price adjusts to reflect the private information. This is consistent with a Glosten and Milgrom (1985) type of model where trading is sequential and uncertainty is greatest at the opening of the trading session. Moreover, we would expect that the orders placed by better informed investors were most aggressive at the beginning of the day, especially if informed investors are competing to extract profits from the same information. This is exactly what is indicated in our data sample. Table 4.4 and figure 4.4 show that there are systematic differences in the aggressiveness of different types of orders in the course of the trading day.

"Away from market" orders, which make up a large part of the order book, are placed further away from the inner quotes at the beginning than at the end of the day. If this type of order is

[30] We cannot calculate the aggressiveness for market orders since these orders do not have a price limit.

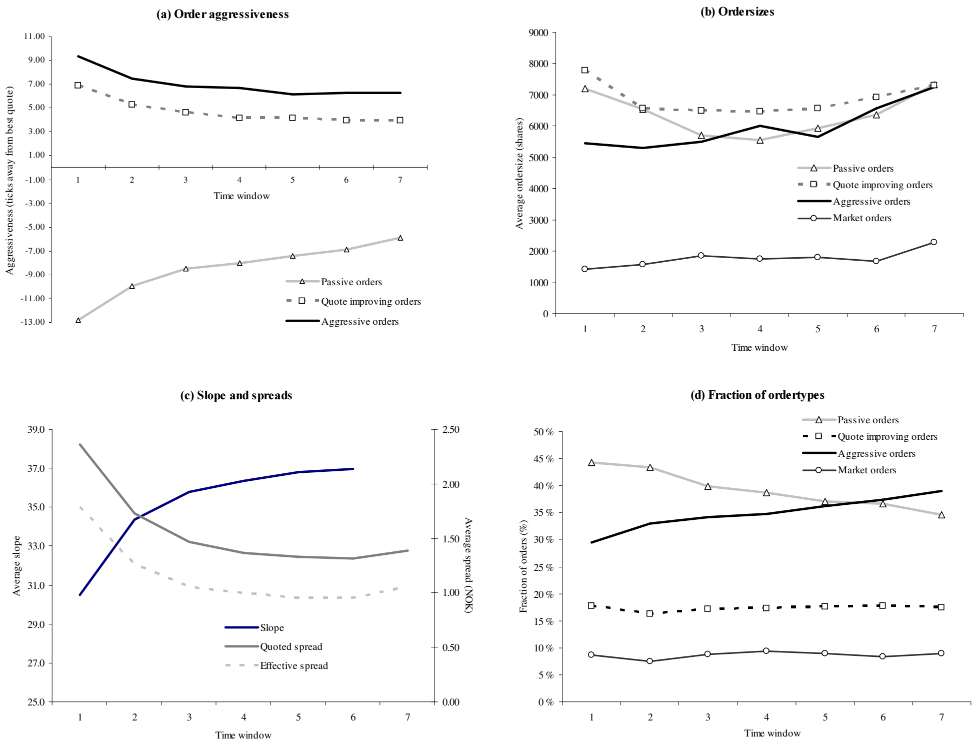
Table 4.5: Order aggressiveness

In the table, all orders within each time interval are decomposed into four groups based on their aggressiveness. The least aggressive orders, "away from market", are orders placed at or away from the quote on the same side of the book. This would be e.g. a buy order with a price (bid) equal to or lower than the current best bid, or a sell order with a price (ask) equal to or higher than the best ask price. The second type of orders, "quote-improving orders", are orders that improve the best quotes. This would be e.g. a buy order with a price higher than the current best bid, but lower than the best ask quote. The third type of orders, "aggressive orders", are orders placed at the opposite quote or higher(buys)/lower(sells). The table reports the average number of orders of each type within each time window, the percentage of all orders of each type, and the average order size in shares and NOK. For each type of order we also calculate an aggressiveness index equal to the average number of ticks away from the best quote (on the same side) that an order is submitted. Thus an index number of zero means that the average order is placed at the quote, a positive index number means that the order is placed above/below the bid/ask, and a negative index number means that the average order is placed below/above the bid/ask. We do not calculate the aggressiveness for market orders since these by definition do not have any limit price. Note that the first and last time windows are half an hour while the rest of the time windows are hourly.

	Time window						
Order type	10:00 to 10:30	10:30 to 11:30	11:30 to 12:30	12:30 to 13:30	13:30 to 14:30	14:30 to 15:30	15:30 to 16:00
				Aggressiveness (avg. ticks away from best quote)			
Passive orders	-12.81	-9.96	-8.45	-8.02	-7.44	-6.90	-5.87
Quote improving orders	6.90	5.30	4.65	4.17	4.16	3.94	3.96
Aggressive orders	9.36	7.46	6.82	6.68	6.17	6.29	6.29
Average aggressiveness (weighted)	-1.69	-1.00	-0.25	-0.06	0.20	0.52	1.11
				Average number of orders			
Passive orders	8.2	9.1	6.4	5.9	5.7	6.2	5.0
Quote improving orders	3.3	3.4	2.7	2.6	2.7	3.0	2.5
Aggressive orders	5.4	6.9	5.5	5.3	5.5	6.3	5.6
Market orders	1.6	1.6	1.4	1.4	1.4	1.4	1.3
				% fraction of orders of type			
Passive orders	44.2 %	43.4 %	39.9 %	38.7 %	37.2 %	36.7 %	34.6 %
Quote improving orders	17.7 %	16.3 %	17.1 %	17.3 %	17.6 %	17.7 %	17.4 %
Aggressive orders	29.4 %	32.9 %	34.2 %	34.7 %	36.2 %	37.3 %	39.0 %
Market orders	8.6 %	7.4 %	8.8 %	9.3 %	9.0 %	8.3 %	8.9 %
				Order size (shares)			
Passive orders	7202	6548	5716	5557	5938	6370	7317
Quote improving orders	7793	6568	6486	6470	6561	6915	7294
Aggressive orders	5461	5301	5498	6008	5649	6569	7239
Market orders	1412	1576	1855	1751	1795	1678	2281
				Order size (1000 NOK)			
Passive orders	275	235	222	221	242	267	346
Quote improving orders	274	253	258	265	274	290	328
Aggressive orders	188	204	204	214	227	376	307
Market orders	36	39	46	42	40	43	69

Figure 4.4: Intraday characteristics of the order book

The figures shows cross-sectional averages across 7 intraday windows for various measures. The windows and numbers correspond to those in tables 4.4 and 4.5. Note that windows 1 and 7 are half-hour intervals from 10:00 to 10:30 and 15:30 to 16:00 respectively, while windows 2 to 6 are hourly intervals starting every half hour. Figure (a) shows the average aggressiveness of different order types. The first type of orders, ”passive orders”, are placed at or away from the quote on the same side of the book. This would be e.g. a buy order with a price (bid) equal to or lower than the current best bid, or a sell order with a price (ask) equal to or higher than the best ask price. The second type of orders, ”quote-improving orders”, are orders that improve the best quote (on the same side). This would be e.g. a buy order with a price higher than the current best bid, but lower than the best ask quote. The third type of orders, ”aggressive orders”, are orders placed at the opposite quote or higher(buys)/lower(sells). Figure (b) shows the average order size within each limit order group and the average order size of market orders. Figure (c) show the average slope on the left axis and the average quoted and effective spreads on the right axis. Note that the slope is calculated from the order book snapshot taken at the end of each window. Figure (d) shows the fraction of each order category which is placed within each window.



mainly submitted by uninformed traders, it indicates that they require a higher compensation for trading early in the day relative to later in the day. Another interpretation is that uninformed traders have not yet processed all publicly available information (e.g. newspapers, new analyzes, gossip etc.), and are more passive when submitting their orders before they have been able to read and interpret this information. Orders that are more aggressive, and likely to stem from better informed investors or pre-committed liquidity traders, are relatively more aggressive at the beginning of the day than later in the day. Thus, a pre-committed trader or informed trader, demanding liquidity, needs to be relatively more aggressive at the beginning of the day to get his order executed since the liquidity suppliers submit their orders relatively much further away from the midpoint. At the end of the trading day all types of orders are submitted closer to the inner quotes, indicating that the adverse selection cost is reduced. Assuming that all other cost components of the spread, except the adverse selection component, are fixed through the day, the decrease in spreads may also reflect that the adverse selection cost is the largest at the beginning of the day and smaller at the end of the day.

The average number of passive orders ("away from market") and market orders decreases throughout the day, while the average number of quote-improving orders and aggressive orders has a U-shape. The intraday pattern in the relative fraction of each order type indicates that more orders are submitted closer to the midpoint at the end of the day. "Away from the market" orders are the largest at the open and close, while the most aggressive limit orders and market orders are the smallest and increase in size throughout the day. If informed investors mainly use aggressive limit orders and market orders, this may indicate that they submit smaller orders when their information is the most valuable (stealth trading).

The evidence that there is more asymmetric information at the beginning of the trading day is also captured by the intraday pattern of our slope estimate. The slope increases (at a diminishing rate) across the day, with a minimum at the beginning of the day and a maximum at the end of the day, which indicates that the order book is more dispersed in the morning relative to later in the day. Note that the average slope is calculated from the normalized order book, i.e. the slope does not merely reflect that there are fewer orders in the order book early in the day, but rather that orders are submitted across a wider price range.[31] Over time windows, the average slope increases at a diminishing rate as the order book becomes more concentrated and inelastic at the end of the day.

5 The volume-volatility relation

In this section, we first document that there exist a volume-volatility relation in the Norwegian equity market as has been found for the US by e.g. Jones et al. (1994) and in the UK by Huang and Masulis (2003). When we decompose volume into trades and order size, and interpret the

[31] A lower average slope reflects that the order book is more elastic, which implies that a lower fraction of the order volume is close to the inner quotes relative to further out in the book.

number of trades as a proxy for the mixing variable, we find support for the MDH. We then investigate the relationship between volume, volatility and the slope of the order book.

5.1 The volume-volatility relation in a limit order market

To investigate if there is a volume-volatility relation in our data sample, we follow the regression approach in Jones et al. (1994). First, we measure the daily return volatility using the standard procedure in similar empirical studies,[32] by running the following regression for each firm i,

$$R_{i,t} = \sum_{k=1}^{5} \alpha_{i,k} D_{k,t} + \sum_{j=1}^{12} \beta_{i,j} R_{i,t-j} + \hat{\varepsilon}_{i,t} \tag{4.4}$$

where $R_{i,t}$ is the return of security i on day t, and $D_{k,t}$ is a day-of-the-week dummy for day k. To avoid measurement errors due to the bid-ask bounce, we calculate returns from the average of bid-ask prices at the close. The 12 lagged return regressors estimate short-term movements in conditional expected returns. The residual, $\hat{\varepsilon}_{i,t}$, is our estimate of the unexpected return of security i on date t. The absolute value of this measure constitute our measure of volatility. Next, we estimate the regression equations suggested in Jones et al. (1994) to determine the relative effects of number of trades (N) and trade-size (AV) for volatility,

$$\textbf{Model I: } |\hat{\varepsilon}_{t,i}| = \alpha_i + \alpha_{i,m} M_t + \beta_i AV_{i,t} + \sum_{j=1}^{12} \rho_{i,j} |\hat{\varepsilon}_{i,t-j}| + \eta_{i,t} \tag{4.5}$$

$$\textbf{Model II: } |\hat{\varepsilon}_{t,i}| = \alpha_i + \alpha_{i,m} M_t + \gamma_i N_{i,t} + \sum_{j=1}^{12} \rho_{i,j} |\hat{\varepsilon}_{i,t-j}| + \eta_{i,t} \tag{4.6}$$

$$\textbf{Model III: } |\hat{\varepsilon}_{t,i}| = \alpha_i + \alpha_{i,m} M_t + \beta_i AV_{i,t} + \gamma_i Ni,t + \sum_{j=1}^{12} \rho_{i,j} |\hat{\varepsilon}_{i,t-j}| + \eta_{i,t} \tag{4.7}$$

The $\rho_{i,j}$'s measure the persistence in volatility across 12 lags. M_t is a dummy variable that is equal to 1 for Mondays and 0 otherwise, $AV_{i,t}$ is the average trade size (total number of shares traded divided by the number of transactions for stock i on date t), and $N_{i,t}$ is the number of transactions in security i on date t. The regressions are run for each firm and then the parameter estimates are averaged across firms.

The first part of table 4.6 provides the results from the estimation of regression equations 4.5-4.7 using daily returns for all companies in our filtered sample. Overall, our results are very much in line with the results in Jones et al. (1994). The explanatory power of model 2 (with respect to the adjusted R-squared), where volume is measured by the average number of daily trades, is almost the double of the explanatory power of model 1, where volume is measured by the average trade size. Moreover, the average trade size has little marginal explanatory power when volatility is conditioned on the number of transactions in model 3. These results are further

[32] See Schwert (1990), Bessembinder and Seguin (1993), Jones et al. (1994), and Daigler and Wiley (1999).

supported by the characteristics of the sampling distributions of individual-firm coefficients and t-statistics of the two variables. In model 3, 95.4 percent of the coefficients for the average number of trades are statistically significant, and 99.1 percent of the average number of trades coefficients were greater than zero. Similar numbers for the average trade size are respectively 24.1 percent and 57.4 percent.

As a robustness check we also estimate the equations for sub-periods of half-years. Although not reported in a table, the results from the whole sample regression are confirmed in the sub-sample regressions. Most notably, the $\hat{\gamma}$ estimates of the effect of trades (N), as well as their distributional properties, are very stable across sub-periods. The $\hat{\beta}$ estimates, however, vary considerably across sub-periods and are less significant than $\hat{\gamma}$ for model 1 relative to model 3.

Jones et al. (1994) find that trade size has some information content for some of the smaller Nasdaq-NMS firms. This finding is interpreted as supportive of the notion that private information based trading is important only for the smallest firms on the stock market. To check for similar features in our data sample, we re-estimate the three regression models on the four size portfolios. The results from these estimations are presented in the second part of table 4.6. In general, the results from estimating separate regression models for each size portfolio are similar to the results from running one regression for the whole sample. However, we find the opposite result from Jones et al. (1994) that the explanatory power of trade size is the strongest for the largest firms. On the other hand, only about half of the parameter estimates for trade size in the single firm regressions are greater than zero, indicating that the effect may not be very systematic across firms.

5.2 Volume, volatility and the limit order book

We now turn to the question whether the slope of the order book affect volatility and trading activity. The reported results are based on the equally weighted slope calculated from the normalized order book. As discussed in section 4 and appendix A, we also calculate a tick-weighted slope measure. The two slope measures are highly correlated (0.98), and the results from using the weighted slope measure are quite similar to those obtained using the equally weighted measure.[33] The correlations between the equally weighted slope and the other variables used in our analysis are reported in table 4.7. Table 4.8 provides some descriptive statistics on the distribution of the daily slope estimate over the whole sample, for the separate years, and for the four market capitalization groups.

Table 4.7 shows that the slope measure has the expected close relationship to measures of liquidity such as market capitalization (positive correlation of 0.44) and quoted percentage spread (negative correlation of -0.32). Thus, larger firms are generally more liquid, with a smaller spread and a steeper slope. One reason for this may be that larger firms generally are easier to value, making the dispersion of prices in the order book more concentrated around the midpoint price. In addition, we see that there is a positive correlation of 0.13 between the

[33] Estimation results for when we use the weighted slope version are reported in appendix 3.

Table 4.6: A volume-volatility regression model

The table reports the results from the estimation of three regression models of the volume/trade size -volatility relation. The models are estimated on the whole data sample and separately for each market capitalization group. The models are based on Jones et al. (1994):

$$\textbf{Model I: } |\hat{\varepsilon}_{t,i}| = \alpha_i + \alpha_{i,m}M_t + \beta_i AV_{i,t} + \sum_{j=1}^{12} \rho_{i,j}|\hat{\varepsilon}_{i,t-j}| + \eta_{i,t}$$

$$\textbf{Model II: } |\hat{\varepsilon}_{t,i}| = \alpha_i + \alpha_{i,m}M_t + \gamma_i N_{i,t} + \sum_{j=1}^{12} \rho_{i,j}|\hat{\varepsilon}_{i,t-j}| + \eta_{i,t}$$

$$\textbf{Model III: } |\hat{\varepsilon}_{t,i}| = \alpha_i + \alpha_{i,m}M_t + \beta_i AV_{i,t} + \gamma_i Ni,t + \sum_{j=1}^{12} \rho_{i,j}|\hat{\varepsilon}_{i,t-j}| + \eta_{i,t}$$

Using the Jones et al. (1994) notation we have that $|\varepsilon_{t,i}|$ is the absolute value of the return of security i in period t, conditional on its own 12 lags and day-of-week dummies, M_t is a dummy variable that is equal to 1 for Mondays and 0 otherwise, $AV_{i,t}$ is the average trade size, $N_{i,t}$ is the number of transactions for security i on day t, and the coefficients $\rho_{i,t}$ measure the persistence in volatility. Column 3-5 show parameter estimates averaged across all individual firm regression equations, while columns 6-9 show the parameter distribution across firms. $\hat{\beta}$ is the average parameter estimate for the average trade size variable (AV), $\hat{\gamma}$ is the average parameter estimate for the number of trades variable (N). In the distribution of estimates column we report, respectively, the percentage of $\hat{\beta}$ and $\hat{\gamma}$ estimates over all single firm regression equations that are significant. In the last two columns we report the percentage of parameter estimates that are greater than zero. The first part of the table shows the results from running the regression equations over the whole sample. The second part of the table shows the similar results when we split the sample into four size portfolios.

		Parameter estimates			Distribution of estimates			
Model	Firms	$\hat{\beta}$ (AV)	$\hat{\gamma}$ (N)	adj. R^2	% t($\hat{\beta}$)>2	% t($\hat{\gamma}$)>2	% $\hat{\beta}$ >0	% $\hat{\gamma}$ >0
Model I: Trade size (AV)	108	0.145	-	0.057	26.9 %	-	81.5 %	-
Model II: Trades (N)	108	-	0.031	0.145	-	95.4 %	-	100.0 %
Model III: Both (AV,N)	108	0.053	0.031	0.149	22.2 %	94.4 %	58.3 %	100.0 %
Model I: Trade size (AV)								
1 (small)	27	0.145	-	0.080	16.2 %	-	78.4 %	-
2	27	0.219	-	0.055	18.2 %	-	77.3 %	-
3	27	0.274	-	0.048	19.0 %	-	64.3 %	-
4 (large)	27	1.021	-	0.038	30.8 %	-	79.5 %	-
Model II: Trades (N)								
1 (small)	27	-	0.052	0.174	-	89.2 %	-	97.3 %
2	27	-	0.028	0.147	-	75.0 %	-	95.5 %
3	27	-	0.036	0.136	-	81.0 %	-	95.2 %
4 (large)	27	-	0.014	0.174	-	79.5 %	-	92.3 %
Model III: Both (AV,N)								
1 (small)	27	0.079	0.053	0.175	10.8 %	86.5 %	64.9 %	97.3 %
2	27	0.076	0.030	0.148	4.5 %	75.0 %	54.5 %	95.5 %
3	27	0.075	0.036	0.140	16.7 %	78.6 %	45.2 %	95.2 %
4 (large)	27	0.237	0.014	0.179	30.8 %	82.1 %	35.9 %	94.9 %

Table 4.7: Variable correlations

The table shows Pearsons' correlation coefficients between our elasticity variable (*SLOPE*) and various trading activity and liquidity variables.

	Trades (N)	Trade size shares (AV)	MCAP	SPREAD	SLOPE	Ordervol. (OV)
Trade size shares (AV)	-0.02					
MCAP	0.25	-0.04				
SPREAD	-0.20	0.16	-0.17			
SLOPE	0.13	-0.08	0.44	-0.32		
Order volume shares (OV)	0.19	0.16	0.06	-0.06	0.04	
Trade volume shares (V)	0.43	0.33	0.14	-0.13	0.08	0.45

Table 4.8: Distribution of equally weighted slope estimates

The table shows the distribution of the slope estimates where each local slope is equally weighted, and each side of the order book is normalized with respect to the total number of shares on each side. Panel A report the estimates for the entire sample and across minimum tick sizes. Panel B report the estimates across market capitalization groups and years. Each company is assigned to a market capitalization quartile at the end of every trading day. N reflects the number of firm/date observations, MCAP is the average market capitalization in NOK millions, price is the average price, P5, P10, P25, P75, P90, and P95 are the 5th, 10th 25th 75th, 90th and 95th percentiles respectively.

PANEL A

				Distribution of daily SLOPE estimates							
	N	MCAP	Price	P5	P10	P25	Median	Mean	P75	P90	P95
All firms	51015	7294	145	9.1	11.9	18.3	29.2	37.2	46.7	70.9	91.5
1999	16968	5948	110	9.4	12.6	20.3	33.2	41.4	53.0	79.2	101.3
2000	23853	7737	180	9.6	12.2	18.0	27.6	35.3	43.5	66.3	86.0
2001	10194	8498	122	7.8	10.6	16.5	27.0	34.7	43.4	65.2	85.7

PANEL B

				Distribution of daily SLOPE estimates							
	N	MCAP	Price	P5	P10	P25	Median	Mean	P75	P90	P95
MCAP Q1 (small)	12532	259	21	5.7	7.4	11.4	17.2	20.9	25.9	38.2	47.9
1999	4163	213	19	5.9	7.5	11.4	17.5	21.2	26.7	39.3	48.2
2000	5864	282	22	6.3	8.2	12.1	17.6	20.9	25.7	36.6	45.9
2001	2505	283	22	4.6	6.0	10.0	15.9	20.5	25.2	40.4	50.9
MCAP Q2	12828	1005	64	10.6	12.9	18.1	26.3	31.8	39.0	56.4	70.9
1999	4264	869	50	11.3	14.0	19.7	28.8	34.0	41.8	59.3	74.4
2000	5999	1035	69	10.7	12.8	17.7	25.1	30.2	36.9	53.1	66.8
2001	2565	1158	76	9.8	12.0	16.9	25.2	31.8	39.2	57.8	73.2
MCAP Q3	12672	2786	121	12.0	15.3	22.2	32.5	39.0	48.1	69.9	87.2
1999	4210	2289	106	15.3	19.1	26.5	38.1	45.2	55.8	78.9	98.1
2000	5934	2914	133	11.6	14.7	21.1	30.4	36.5	44.3	65.0	82.4
2001	2528	3315	121	10.6	13.1	19.5	28.9	34.5	43.0	61.3	75.8
MCAP Q4 (large)	12983	24698	369	18.0	22.1	31.6	47.1	56.5	69.4	101.6	128.8
1999	4331	20016	261	23.0	28.2	38.9	55.7	64.2	79.0	111.6	136.6
2000	6056	26320	491	16.7	21.0	29.2	44.0	53.1	64.6	94.6	120.3
2001	2596	28727	263	15.7	19.9	27.8	41.1	51.5	61.5	96.6	125.9

slope and the number of trades. Further, table 4.8 shows that larger and more liquid stocks have a higher fraction of the order book volume concentrated at or around the best quotes, while smaller firms have more elastic order books. This is also evident from panel B in table 4.3.

Figure 4.5 illustrates the relationship between the daily slope and the contemporaneous daily price changes at an aggregate level. Daily price changes are measured as the average daily absolute return over the trading day.[34] Both variables are daily equally weighted averages across all traded securities. Interestingly, even at this aggregate level the figure indicates that the price volatility is higher (lower) when the average daily slope of the order book is low (high). Another notable feature is that the average slope is steeper in the first half of the sample, with an average slope of about 41, than during the second part of the sample when the average slope drops to about 35. These two periods coincide quite well with the boom and burst of the internet bubble. It is not obvious that increased trading activity due to arrival of new information can explain the volatility pattern during this period. If the slope proxies for valuation uncertainty, the pattern in the figure reflects greater agreement among traders about asset values during the build-up of the bubble than during the subsequent market down-turn.

5.3 Daily volatility and order book shape

To examine whether our slope measure can explain the contemporaneous volatility across firms and time, we estimate modified versions of the volume-volatility regression equations in section 5.1. More specifically, we estimate 3 different versions of the following cross-sectional time-series regression model with one-way fixed effects,

$$\mid \varepsilon_{it} \mid = \sum_{k=1}^{K} \beta_k X_{itk} + \eta_{i,t} \tag{4.8}$$

where $\mid \varepsilon_{it} \mid$ is the daily volatility estimate from equation 4.4, X_{itk} is the matrix of explanatory variables (k) across time (t) for each company (i) and $\eta_{i,t} = \nu_i + \varepsilon_{i,t}$ defines the error structure with ν_i as the non-random fixed, firm-specific, effect. Since we use one-way fixed effects specification, the estimation is analogous to a least-squares dummy variable (LSDV) regression with firm-specific constants ν_i. Since not all firms are traded every day, our sample is unbalanced[35]. However, results from estimating the same models on a balanced sample are quantitatively similar.[36]

As indicated by the correlation structure in table 4.7, our slope measure may also proxy for liquidity. We therefore control for other liquidity measures in the regression model. The estimated model can be written as;

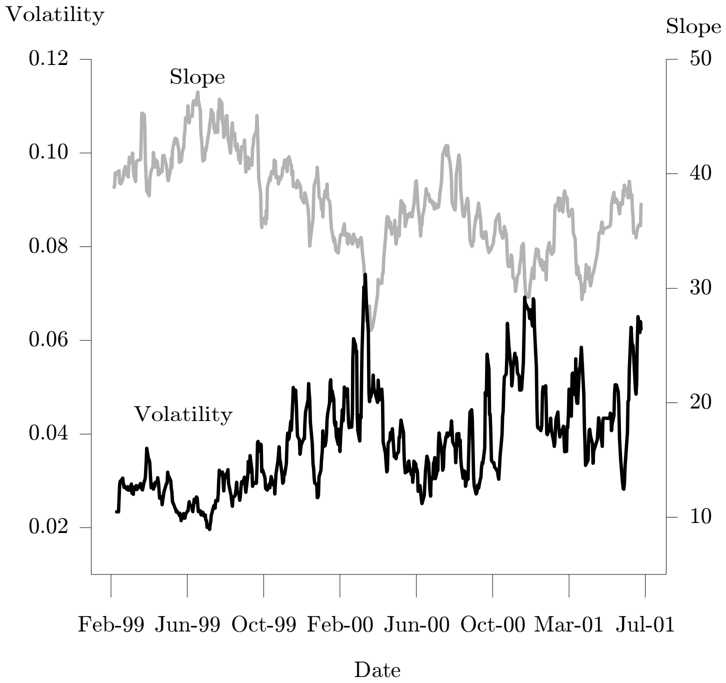
[34] cf equation 4.4.

[35] We use the TSCSREG procedure supplied with SAS v.8.2 for estimating the models. The procedure is capable of processing data with different numbers of time-series observations across different cross sections.

[36] In the unbalanced sample, all firms with 400 trading days or more throughout the sample period of 597 days are included. In the balanced sample, we filter out all firms which are not traded every day during the sample period. This filter reduces the sample to 25 firms. See appendix 2 for estimation results for the balanced sample.

Figure 4.5: Average slope and volatility

The figure illustrates the relationship between the estimates of the average daily slope of the order book and the contemporaneous daily price changes. The left axis measures the equally weighted average absolute return across firms traded on the respective date. The right axis measures the slope estimate calculated as the daily equally weighted slope, averaged over all companies that were traded during the trading day.



$$
\begin{aligned}
|\,\varepsilon_{i,t}\,| = {} & \beta_0 M_{i,t} + \beta_1 N_{i,t} + \beta_2 AV_{i,t} + \beta_3 MCAP_{i,t} + \beta_4 SPR_{i,t} + \\
& \beta_5 OV_{i,t} + \beta_6 SLOPE_{i,t} + \sum_{j=1}^{12} \rho_{i,j}\,|\,\varepsilon_{i,t-j}\,| + \eta_{i,t}.
\end{aligned}
\tag{4.9}
$$

where *SLOPE* is our slope estimate, *MCAP* is the market capitalization value (in mill. NOK), *SPR* is the quoted percentage midpoint spread, and *OV* is the average order book volume in thousand shares and $\eta_{i,t} = \nu_i + \varepsilon_{i,t}$ defines the error structure with ν_i as the non-random fixed, firm-specific, effect. Results from the estimation for the full sample period are provided in panel A of table 4.9. Model 1 is essentially the same as in the analysis in section 5.1, but with the addition of the slope variable and the additional variables accounting for stock liquidity (*SPR*, *MCAP*, and *OV*). In model 2, we estimate the model excluding the two variables which are highest correlated with the slope (*SPR* and *MCAP*), and in model 3 we exclude the trading activity (mixing) variables. We also estimate the same regression equation across 3-month sub-periods. The results from this estimation are reported in panel B of the table. Because we use lagged versions of the dependent variable, $|\varepsilon_{i,t-j}|$, as explanatory variables to adjust for autocorrelations in volatility, we choose a fixed effects model.[37]

The first thing to note in Panel A in table 4.9 is that the slope variable (*SLOPE*) is negative and highly significant across all three model specifications. Thus, volatility increases the more gentle the slope are. This may be linked to differences of opinion about public news, "noise trading" from uninformed investors[38], or pick-off risk.[39] We will discuss several interpretations of our findings at the end of the section.

Both the number of trades (N) and the trade size (AV) have a positive significant effect on volatility as we found earlier. When we remove trade size from the regression model, the reduction in R-squared is small (not shown in the table). Thus, the Jones et al. (1994) result that trade size does not include information that is not already included in the number of trades, is also evident in the panel analysis after we have controlled for additional liquidity variables. Moreover, the total volume in the order book (OV) is shown to have a significant positive effect on volatility. This result is consistent with the result in Biais et al. (1995) that more trades are executed when the order book is thick. The correlations shown in table 4.7 between order book volume and trade volume (45 percent) and between the order book volume and trades (19 percent) also suggest that the volume-volatility relation depends on the incoming order flow and the state of the order book. Finally, the estimation results show that larger firms are less volatile and that higher spreads coincide with higher volatility.

One important issue to note is that there is an indeterminacy with respect to the causality between volatility and several of the explanatory variables such as the average order book volume, number of trades, the spread and the slope measure. Although this probably is most important at the transaction level, several of our measures are averages across hourly snapshots. Thus, dynamic interactions between order submissions and the status of the order book, as examined in detail by Biais et al. (1995), is left out of our regression model. For instance, Biais et al. (1995) find that a thin book attracts new orders while a thick book increases trading activity. Another example is that a higher volatility may reduce the number of orders coming into the market, which again lowers the average slope of the book on that day. To examine this issues, we run simple Granger causality tests between our slope measure and various order types and trading activity variables, both on an hourly and a daily frequency. Overall, we are unable to determine a clear one-way causality relation between the variables, rather we find a two-way causality for most variable combinations.

The estimation results for models 2 and 3 are essentially the same as for model 1. The important thing to note is that the parameter estimate for *SLOPE* is significantly negative and

[37] For a random-effects model to be applicable, the firm-specific constants, ν_i, must be uncorrelated with the regressors. This requirement is likely to be violated by the lagged variables. Later in the paper, we test whether we should use a random-effects model more formally by running Hausman tests.

[38] A problem could be that a steeper slope implies a less pronounced bid-ask bounce, and thus a lower volatility. However, as outlined in section 5.1, we try to avoid measurement errors due to the bid-ask bounce by calculating returns using the average of bid-ask prices.

[39] If some liquidity suppliers are informed about the volatility, as in the Foucault et al. (2003) model, they may find it optimal to bid less aggressively when they know that the volatility is high.

Table 4.9: A volume-volatility regression model including the (full) order book slope

The table shows the results from estimating a panel regression model with one-way fixed effects (least squares dummy variable estimation) for the whole sample (Panel A) and for sub-periods of 3 months (Panel B). The estimated model (model 1) is,

$$| \varepsilon_{i,t} | = \beta_0 M_{i,t} + \beta_1 N_{i,t} + \beta_2 AV_{i,t} + \beta_3 MCAP_{i,t} + \beta_4 SPR_{i,t} + \beta_5 OV_{i,t} + \beta_6 SLOPE_{i,t} + \sum_{j=1}^{12} \rho_{i,j} \mid \hat{\varepsilon}_{i,t-j} \mid + \eta_{i,t}.$$

where $\eta_{i,t} = \nu_i + \varepsilon_{i,t}$ defines the error structure with ν_i as the non-random fixed, firm-specific, effects. $| \varepsilon_{i,t} |$ is the absolute return adjusted for day of week effects and autocorrelation in returns. *M* is a dummy variable for Monday, *N* is the number of transactions, *AV* is the average trade size in shares, *MCAP* is the market capitalization (in mill. NOK), *SPR* is the relative spread (quoted spread as a percent of the midpoint price), *OV* is the total number of shares in the order book (sum of all orders on bid and ask side of the order book) and *SLOPE* is the average slope of the bid and offer side of the order book. Panel A, shows the parameter estimates for 3 variations of the full model (model 1), t-values, and standard errors for the parameter estimates. In model 2, we do not control for the market capitalization (*MCAP*) and spread (*SPR*) variables, and in model 3 we exclude the trading activity (*N*) and trade size (*AV*) variables. The table shows the associated t-values as well as the R^2 for each portfolio regression. The autoregressive estimates have been excluded from the table. For the F-tests, ** denotes significance at the 1 percent level. Panel B, shows the sub-period estimates for model 1 for the *SLOPE*, *N* and *AV* variables with associated t-values. For each period, the model R-squared, F-test for fixed effects, and number of cross-sectional observations (N) and number of time series observations (T) are reported in the last four rows of the table.

PANEL A: Whole sample regression

	MODEL 1			MODEL 2			MODEL 3		
Variables	Est.	t-value	std.err	Est.	t-value	std.err	Est.	t-value	std.err
M (Monday dummy)	0.021	0.60	0.035	0.037	1.05	0.036	-0.014	-0.38	0.036
N (trades)	0.005	43.95	0.000	0.005	41.06	0.000	-	-	-
AV (avg. trade size)	0.025	6.18	0.004	0.023	5.68	0.004	-	-	-
MCAP (market cap.)	-0.013	-2.79	0.005	-	-	-	-0.001	-0.18	0.005
SPR (% quoted spread)	0.234	24.65	0.009	-	-	-	0.181	18.86	0.010
SLOPE (avg. slope)	-0.007	-11.82	0.001	-0.009	-14.43	0.001	-0.008	-12.35	0.001
OV (order-book vol.)	0.023	6.32	0.004	0.023	6.34	0.004	0.050	13.77	0.004
R^2	21.8 %			20.7 %			18.3 %		
N (cross section)	98			98			98		
T (time series)	572			572			572		
F-test no fixed effects	17.5**			15.13**			11.34**		

PANEL B: Sub-period regression

	SLOPE		*N* (trades)		*AV* (trade size)		Model			
Quarter	β_6	t-value	β_1	t-value	β_2	t-value	R^2	F test	N	T
1999.1	-0.008	-1.39	0.016	4.90	0.082	1.50	37.6 %	2.57**	61	14
1999.2	-0.005	-2.74	0.013	10.97	0.059	3.00	26.6 %	4.31**	87	59
1999.3	-0.005	-3.19	0.011	11.24	0.061	3.99	36.7 %	7.71**	96	66
1999.4	-0.006	-2.67	0.014	16.04	0.039	2.90	27.5 %	5.54**	97	64
2000.1	-0.007	-3.22	0.013	26.42	0.032	1.65	31.0 %	7.96**	98	65
2000.2	-0.006	-2.89	0.013	18.86	0.019	1.06	30.7 %	5.23**	98	58
2000.3	-0.007	-4.36	0.010	20.86	0.004	0.52	29.6 %	6.41**	98	65
2000.4	-0.009	-4.29	0.007	16.12	0.018	2.08	21.6 %	4.11**	97	63
2001.1	-0.008	-4.41	0.003	6.26	-0.005	0.05	25.6 %	5.17**	93	64
2001.2	-0.008	-4.63	0.002	8.92	0.027	2.08	25.9 %	4.67**	88	54
Average	-0.007	-3.38	0.010	14.06	0.034	1.88	29.3 %			

relatively stable across the three model specifications. The slope parameter is most negative and most significant in model 2, when we remove the spread (*SPR*) and market capitalization (*MCAP*) variables. This suggests that the slope captures liquidity effects captured by these variables. Both the F-test of no firm-specific effects (firm-specific constants) and the Hausman specification test of whether a random-effects model would be more appropriate relative to the fixed effects specification, are rejected at the 1 percent level for all three models.[40] This suggests that our firm-specific dummies are correlated with the regressor, such that a fixed-effects specification is more appropriate. The reason for this is that we have lagged versions of the dependent variable, which makes ν_i correlated with the regressors.

To examine the stability of the slope measure, we estimate model 1 for non-overlapping sub-periods of three months through the entire sample period. The results from these regressions are reported in panel B in table 4.9. We only report the parameter estimates and tests for the slope variable, number of trades, and trade size. The *SLOPE* parameter is remarkably stable across the sub-samples. In addition, it is significant at the 1 percent level within all sub-samples, except for the first. Also, the number of trades is highly significant across all sub-periods while the average trade size is significant at the 1 percent level only in half of the sub-sample regressions, suggesting that the number of trades is the important component of volume in the volume-volatility relation, as also suggested by our analysis in section 5.1. The parameter estimate for the number of trades decreases over the sample period. This is most likely due to the fact that the mean number of trades across companies increases through the sample period.

F-tests of no fixed effects within each sub-period regression is rejected at the 1 percent level.[41] Our results suggest that both the order flow and the status of the order book are significantly related to contemporaneous volatility in addition to trading volume.

A robustness check In a limit order market, most trades originate from limit orders, i.e. there must be a strong relationship between order book shape, volatility and trading volume. One interpretation of our slope measure is that it is essentially a liquidity measure, and that the inner part of the order book is capturing the main effect on volatility. A useful way to check this is to examine whether the slope calculated from different sets of the order book contain different information about volatility. One way of doing this is to calculate the slope based on truncated versions of the order book.

We re-calculate the slope measure based on two different subsets of the book. The resulting estimate distributions are shown in figure 4.6. Figure 4.6a shows the frequency distribution of slope estimates calculated from an order book which is truncated to 5 ticks away from the best quotes.[42] Figure 4.6b shows the distribution of daily slope estimates when we calculate the

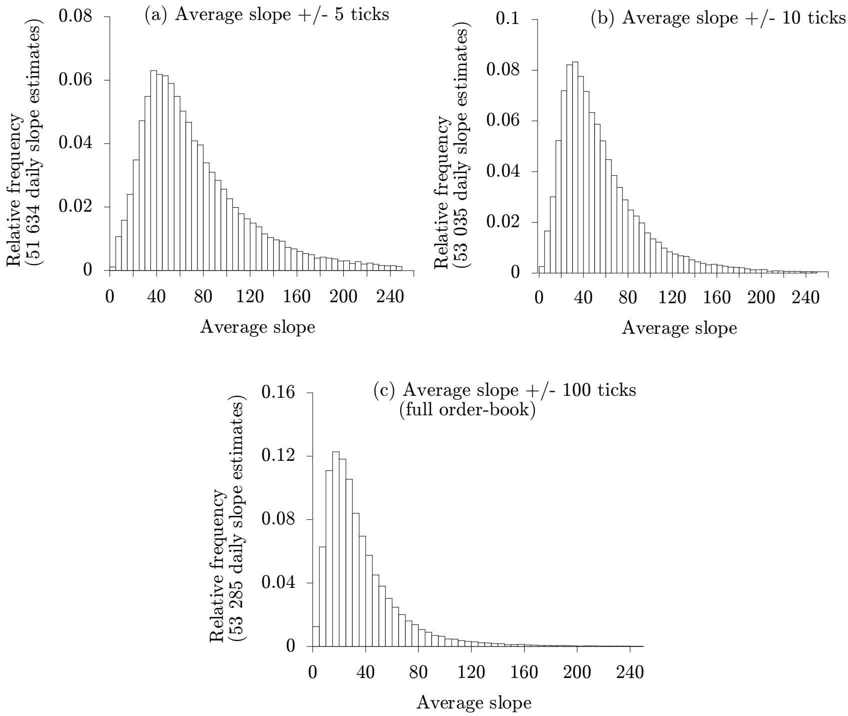
[40] The Hausman test compares an inefficient but consistent OLS estimator (the fixed effects case) to an efficient GLS estimator (the random effects case). Thus, the Hausman test is a test of H_0, that random effects would be consistent and efficient, versus H_1, that random effects would be inconsistent. Rejecting H_0 would suggest that we should use a fixed effects specification.

[41] In addition, the Hausman test rejects a random effects specification at the 1 percent level for each sub-sample model.

[42] That is, we use only the cumulative volume at the five first ticks on each side of the order book when we calculate

Figure 4.6: Frequency distribution of slope estimates

The figures show the frequency distributions for (average) daily equally weighted normalized slope estimates for all firms for the entire sample period. In figure (a) the slope calculations are calculated using only the first 5 levels of the order book, in figure (b) we use the first 10 levels of the order book and in figure (c) we use the entire order book up to 100 tick levels.



average slope based on twice as much of the order book (+/- 10 ticks). Finally, figure 4.6c shows the frequency distribution when we base our slope estimates on the entire order book (+/- 100 ticks). The slope decreases the more of the order book we use. This is expected, if the supply and demand curves in the order book are concave.[43] The mean slope when we use the full order book is about 37 (median 28), while it increases to 57 (median 46) and 76 (median 62) when we calculate it from the order book truncated to +/- 10 and +/- 5 ticks respectively.

To examine whether the inner part of the order book captures the relationship between volatility and slope, we re-estimate the regression models in table 4.9 with slope measures calculated from the two sub-sets of the order book. Panel A in table 4.10 reports the estimation results. The results when we use the slope calculated from the order book truncated to +/- 10 ticks ($SLOPE_{10}$) from the best quote on each side are reported in model 1a, and the results when we truncate the order book to +/- 5 ticks ($SLOPE_5$) are reported in model 1b. All other variables are identical to the previous analysis. Panel B of the table shows the correlation be-

the average slope.

[43]Concave when we have price on the x-axis and volume on the y-axis.

tween each slope measure and other variables. The main result from the estimation is that the slope parameter remains negative and significant. In addition, the parameter estimates become smaller compared to the case where we used the full order book. The decrease in parameter size is mainly due to the fact that the mean of the slope estimates increases (as shown in figure 4.6) while the dependent variable remains unchanged. Thus, the relationship seems to be similar when we use only the inner levels of the order book to calculate the slope. Also R-squared of the different models does not change when we change the slope variable. Overall, our results suggest that the different slope measures capture mainly the same relationship.

Panel B in table 4.10 shows the correlations between the three slope measures and the activity and liquidity variables. One interesting thing to note is that the correlation between the slope and number of trades and between the slope and the trade volume in shares increases substantially the more we truncate the order book. This may reflect that the relationship found in Biais et al. (1995), that a thicker (more concentrated) order book results in trades, is more pronounced when we evaluate the relationship closer to the inner quotes.

Table 4.11 shows the estimation results when estimating the model using the truncated slope measures across sub-periods. Similar to our findings when we estimate the model over the whole sample period, we find that the size of the parameter estimate as well as its significance declines the more we truncate the order book. For model 1b, when we truncate the order book to 5 ticks, the slope estimate is only significantly different from zero for half of the sub-samples. Thus, the significance of the slope variable is greatly reduced within sub-periods when we only use the inner part of the book. Note also that when we use the slope based on the full order book, the relationship between volatility and slope is stronger across sub-periods, as shown in panel B in table 4.9.

5.4 Number of trades and order book shape

In this sub-section, we examine the relationship between the slope and the contemporaneous trading volume.

In table 4.12 we estimate a cross-sectional time series regression with the number of trades as the dependent variable. As before, we control for liquidity variables which are expected to be important with respect to the number of trades. When we base our slope measure on the the full order book (model 1 in panel A), a significant negative relationship between the slope and the number of trades is documented. Thus, the more gentle the order book slope, the higher the trading volume represented by the number of trades. Models 2 and 3 in table 4.12 are estimated with slope measures calculated from the truncated order books. Interestingly, we find that the parameter estimate switches sign and becomes more positive the closer we get to the inner quotes. Thus, the slope at the inner quotes is positively related to the number of trades, while the average slope for the full book is negatively related to trade execution. In other words, the relationship between liquidity and trading activity becomes more evident when we restrict the analysis to the inner part of the order book.

Table 4.10: The relationship between volatility and slope for truncated order book

The table shows the results from the estimation of two panel regression models where we use slope measures calculated from different sub-sets of the full order book. In model 1a in Panel A, the average slope variable is calculated from the order book truncated to contain only the first 10 tick levels on the bid and ask side. In model 1b in Panel A, the slope is calculated from the order book truncated to quotes and volumes including the first 5 tick levels. All other variables are the same as the ones we use in the regression model described in table 4.9. The models are estimated with one-way fixed effects for the whole sample. The estimated model is,

$$| \varepsilon_{i,t} | = \beta_0 M_{i,t} + \beta_1 N_{i,t} + \beta_2 AV_{i,t} + \beta_3 MCAP_{i,t} + \beta_4 SPR_{i,t} + \beta_5 OV_{i,t} + \beta_6 SLOPE_{i,t} + \sum_{j=1}^{12} \rho_{i,j} \, | \hat{\varepsilon}_{i,t-j} | + \eta_{i,t}.$$

where $\eta_{i,t} = \nu_i + \varepsilon_{i,t}$ defines the error structure with ν_i as the non-random fixed, firm specific, effects. $| \varepsilon_{i,t} |$ is the absolute return adjusted for day of week effects and autocorrelation in returns. *M* is a dummy variable for Monday, *N* is the number of transactions, *AV* is the average trade size in shares, *MCAP* is the market capitalization (in NOK mill.), *SPR* is the relative spread (quoted spread as % of the midpoint price), *OV* is the total number of shares in the order book (sum of all orders on bid and ask side of the order book) and *SLOPE* is the average slope of the bid and offer side of the truncated order book. The autoregressive estimates have been excluded from the table. Panel B shows the correlation between various variables and the three slope measures calculated from the full order book as well as the two slope measures calculated from the restricted order books. ** indicates that the F-test from a test of no fixed effects is rejected at the 1 percent level.

PANEL A: Whole sample regression

	MODEL 1a (+/- 10 ticks)			Model 1b (+/- 5 ticks)		
Variables	Estimate	t-value	std.err	Estimate	t-value	std.err
M (Monday dummy)	0.023	0.6	0.035	0.028	0.8	0.036
N (trades)	0.005	45.2	0.000	0.005	45.5	0.000
AV (avg. trade size)	0.023	5.7	0.004	0.022	5.5	0.004
MCAP (market capitalization)	-0.010	-2.0	0.005	-0.009	-1.8	0.005
SPR (% quoted spread)	0.236	24.8	0.010	0.234	23.1	0.010
$SLOPE_{10}$ (+/- 10 ticks)	-0.005	-11.4	0.000	-	-	-
$SLOPE_5$ (+/-5 ticks)	-	-	-	-0.003	-9.9	0.000
OV (order-book volume)	0.023	6.3	0.004	0.023	6.3	0.004
R^2	21.8 %			21.8 %		
N (cross section)	98			98		
T (time series)	572			572		
F-test no fixed effects	18.33**			17.93**		

PANEL B: Variable correlations

	SLOPE (Full order-book)	$SLOPE_{10}$ (+/- 10 ticks)	$SLOPE_5$ (+/- 5 ticks)
N (trades)	0.13	0.25	0.28
Trade volume shares (V)	0.08	0.45	0.45
AV (avg. trade size)	-0.08	-0.09	-0.11
MCAP (market capitalization)	0.44	0.42	0.41
SPR (% quoted spread)	-0.32	-0.31	-0.33
OV (order-book volume)	0.04	0.04	0.03

Table 4.11: The relationship between volatility and slope across sub-periods for truncated order book

The table shows the results from estimating the two panel regression model in table 4.10 for sub-periods of three months. In model 1a, the average slope variable is calculated using an order book truncated to the first 10 tick levels on the bid and ask side. In model 1b, the average slope is calculated using an order book truncated to prices and volumes within the first 5 tick levels. All other variables are the same as the variables used in table 4.9. The models are estimated with one-way fixed effects for each sub-period. The estimated model is,

$$| \varepsilon_{i,t} | = \beta_0 M_{i,t} + \beta_1 N_{i,t} + \beta_2 AV_{i,t} + \beta_3 MCAP_{i,t} + \beta_4 SPR_{i,t} + \beta_5 OV_{i,t} + \beta_6 SLOPE_{i,t} + \sum_{j=1}^{12} \rho_{i,j} \mid \hat{\varepsilon}_{i,t-j} \mid + \eta_{i,t}.$$

where $\eta_{i,t} = \nu_i + \varepsilon_{i,t}$ defines the error structure with ν_i as the non-random fixed, firm-specific, effects. $| \varepsilon_{i,t} |$ is the absolute return adjusted for day of week effects and autocorrelation in returns. *M* is a dummy variable for Monday, *N* is the number of transactions, *AV* is the average trade size in shares, *MCAP* is the market capitalization (in NOK mill.), *SPR* is the relative spread (quoted spread as % of the midpoint price), *OV* is the total number of shares in the order book (sum of all orders on bid and ask side of the order book) and *SLOPE* is the average slope of the bid and offer side of the restricted order book. The table shows the parameter estimates for 2 variations of model 1 in table 4.9 with the associated t-value, std.error of the estimate, the model R-squared. ** indicates that the F-test for no fixed effects is rejected at the 1 percent level.

	Model 1a (+/- 10 ticks)					Model 1b (+/- 5 ticks)				
Quarter	$SLOPE_{10}$	t-value	std.err	model R^2	F-value	$SLOPE_5$	t-value	std.err	model R^2	F-value
1999.1	0.000	-0.1	0.004	39 %	3.4**	-0.003	-1.1	0.003	31 %	2.2**
1999.2	-0.004	-2.4	0.002	27 %	4.2**	-0.001	-0.9	0.001	27 %	4.5**
1999.3	-0.002	-1.6	0.001	33 %	5.2**	-0.001	-0.7	0.001	34 %	5.0**
1999.4	-0.003	-1.7	0.002	28 %	6.1**	-0.001	-1.1	0.001	28 %	6.5**
2000.1	-0.004	-2.6	0.001	31 %	8.4**	-0.002	-1.9	0.001	31 %	8.5**
2000.2	-0.003	-2.4	0.001	30 %	5.1**	-0.002	-2.0	0.001	30 %	5.1**
2000.3	-0.003	-2.8	0.001	31 %	6.9**	-0.001	-1.2	0.001	31 %	6.8**
2000.4	-0.006	-4.6	0.001	24 %	4.1**	-0.005	-4.2	0.001	24 %	4.0**
2001.1	-0.004	-3.1	0.001	26 %	5.2**	-0.003	-2.6	0.001	26 %	5.2**
2001.2	-0.006	-4.0	0.001	25 %	4.8**	-0.004	-3.3	0.001	26 %	4.8**
Average	-0.004	-2.5	0.002	29.4 %		-0.002	-1.9	0.001	28.8 %	

We also find that the number of trades is lower on Mondays, that the average trade size is unrelated to the number of trades, and that larger firms are more frequently traded. In addition, we find that there is less trading when the quoted percentage spread is large, and that there are more trades when the volume of shares in the order book is high. Again, one caveat with respect to the analysis is that we do not take into account the dynamic interactions between the order flow and status of the order book. For example, as found by Biais et al. (1995), a thinner book may attract new orders which in the next step increases the number of transactions. The most interesting result from the estimation is that a slope measure calculated on the basis of the full order book seems to provide different information compared to a slope measure calculated on the basis of the volume at the inner quotes.

In the previous sub-section, we found that the relationship between price volatility and the slope was well proxied by a slope measure based on the inner part of the book. In this section, we have documented a significant difference between slope measures based on different order book truncations.

5.5 Interpretation of the results

The relationships documented in this study are interesting in several respects. First, although most of the activity occur at the inner part of the order book, the order book data shows that the liquidity provided at the inner quotes in many cases reflects only a modest part of the total liquidity supplied in the full order book. Second, the characteristics of the order book vary systematically over the trading day as well as across firms. Third, as far as we know, no previous studies have examined in detail the relationship between the characteristics of the full order book and volume and volatility in a cross-sectional time series setting.

One question is why orders persist further out in the book? One reason may be that traders are slow in revising their orders in response to new information. Another reason suggested by Sandås (2001) is that the placement of orders deep in the book are based on strategic choices where, in a multi-period setting, the gains from obtaining price priority of the orders further out in the book are traded off the costs of monitoring them. When we examine the slopes of the order books across companies over time, we find that there are marked differences across firms in the amount of volume provided throughout the order book, and that these differences persist through time. As shown in table 4.8 some firms have generally a very large fraction of their liquidity concentrated close to the best quotes, while other firms have a relatively larger fraction of the order volume further out in the book. A second question is why such differences in slope estimates across firms appear? The systematic patterns found may indicate that the shape of the order book capture some underlying characteristics of the the trading strategies of liquidity suppliers across firms. One possible explanation is related to asymmetric information. In general, we find that smaller firms have order books with a gentler slope than larger firms, which is in line with the hypothesis that there is more private information in smaller than in larger firms.

Table 4.12: The relationship between the number of trades and the order book slope

The table shows the results when we estimate the relationship between different slope measures and the number of trades. The model is estimated as a one-way fixed effects model. The estimated model is,

$$N_{i,t} = \beta_0 M_{i,t} + \beta_1 AV_{i,t} + \beta_2 MCAP_{i,t} + \beta_3 SPR_{i,t} + \beta_4 SLOPE_{i,t} + \beta_5 OV_{i,t} + \eta_{i,t}.$$

where $\eta_{i,t} = \nu_i + \varepsilon_{i,t}$ defines the error structure with ν_i as the non-random fixed, firm specific, effects. The dependent variable, N is the number of transactions, M is a dummy variable for Monday, AV is the average trade size in shares, $MCAP$ is the market capitalization SPR is the relative spread (quoted spread in % of the midpoint price), OV is the total number of shares in the order book (sum of all orders on bid and ask side of the order book) and $SLOPE$ is the average slope of the bid and offer side from the full order book, $SLOPE10$ is the slope calculated from the order book truncated to +/- 10 ticks, $SLOPE5$ is the slope calculated from the order book truncated to +/- 5 ticks. Panel A shows the estimation results from the whole sample, while panel B shows the estimation results from sub-periods. ** indicate that the F-test for fixed effects is significant at the 1 percent level.

PANEL A: Whole sample regression

	MODEL 1			MODEL 2			MODEL 3		
Variables	Est.	t-value	std.err	Est.	t-value	std.err	Est.	t-value	std.err
M (Monday dummy)	-6.17	-3.97	1.55	-6.01	-3.85	1.56	-6.25	-3.92	1.59
AV (avg. trade size)	-0.06	-0.36	0.18	-0.17	-0.96	0.18	-0.14	-0.79	0.18
$MCAP$ (company value)	3.34	15.73	0.21	2.94	13.65	0.22	2.79	12.75	0.22
SPR (% quoted spread)	-9.13	-22.02	0.41	-8.49	-20.25	0.42	-9.03	-20.01	0.45
$SLOPE$ (full book)	-0.30	-11.11	0.03	-	-	-	-		
$SLOPE10$ (+/- 10 ticks)	-	-	-	0.14	6.96	0.02	-		
$SLOPE5$ (+/- 5 ticks)	-	-	-	-	-	-	0.20	12.92	0.02
OV (order-book volume)	5.15	32.34	0.16	5.17	32.40	0.16	5.16	32.02	0.16
R^2	39.4 %			39.3 %			39.3 %		
N cross section	95			95			95		
Time series	572			572			572		
F-test no fixed effects	235.5**			214.3**			200.3**		

PANEL B: Sub-period regression

	Avg.slope - Full order book			Avg.slope - +/- 10 ticks			Avg.slope - +/- 5 ticks		
Quarter	$SLOPE$	t-val.	R^2	$SLOPE10$	t-val.	R^2	$SLOPE5$	t-val.	R^2
1999.1	-0.049	-0.77	84 %	-0.057	-1.19	84 %	-0.058	-1.48	83 %
1999.2	-0.103	-3.72	77 %	-0.091	-3.89	77 %	-0.059	-2.95	76 %
1999.3	-0.124	-4.95	83 %	-0.088	-4.23	83 %	-0.061	-3.40	83 %
1999.4	-0.010	-0.28	71 %	0.061	2.17	71 %	0.074	3.35	71 %
2000.1	-0.372	-5.82	56 %	-0.183	-4.14	56 %	-0.076	-2.12	56 %
2000.2	-0.109	-2.16	73 %	-0.061	-1.72	73 %	-0.029	-1.00	73 %
2000.3	-0.126	-2.70	72 %	-0.081	-2.37	72 %	-0.078	-2.89	72 %
2000.4	-0.123	-2.02	70 %	-0.054	-1.26	70 %	-0.001	-0.02	70 %
2001.1	-0.024	-0.42	85 %	0.030	0.72	85 %	0.043	1.32	85 %
2001.2	-0.208	-1.70	69 %	0.075	0.81	69 %	0.154	1.98	69 %
Average	-0.125	-2.454	74 %	-0.045	-1.510	74 %	-0.009	-0.721	74 %

In summary, our main findings about the relationship between the slope of the book and the volume-volatility relation are;

- A more gentle slope (more dispersed order book) coincide with a higher volatility across firms and over time.
- The relationship between the number of trades and the slope of the book depends on which subset of the order book is used.
  - When we use the slope from the inner book (+/- 5 ticks) there is a positive relationship in which a steep slope coincide with a high number of trades.
  - When the entire order book is used (+/- 100 ticks), the relationship is reversed, i.e. a more gentle slope coincide with a higher number of trades.

One interesting interpretation of these results is that the differences in the limit order books across firms reflect valuation uncertainty and heterogeneous valuations among the liquidity suppliers. Although no models exist that offer any predictions to how the full limit order book would look like in a market with heterogeneous liquidity suppliers, models assuming strategic behavior of uninformed investors provide an interesting framework which could motivate such an interpretation. Shalen (1993) shows that the strategic behavior of liquidity traders may be an important contributor to both volume and volatility in addition to information arrivals. In her model, when uninformed investors has dispersed beliefs about asset values, they are faced with a signal extraction problem, making them react to all types of trades in the order-flow which may or may not be related to informed trading. Due to this, they increase both trading volume and price volatility above what would be expected in equilibrium. Thus, the relationship between volume and volatility is not merely due to the information arrival process (as in the mixture of distributions framework), but also due to strategic trading by uninformed traders. The higher the fraction of uninformed traders in the population, the greater the dispersion of beliefs, and the greater the excess volume and excess volatility.

Valuation uncertainty (dispersion of beliefs) may to some degree be captured by the shape of the order book, as different levels of the book reflect the reservation prices of liquidity suppliers. This provides an interesting interpretation for why the order volumes observed in the limit order book are more dispersed than predicted by theoretical models such as Glosten (1994). If the uncertainty about the value of a firm is high and liquidity traders differ in their private valuations, they may submit their orders across a wider range of prices relative to the case when there is greater agreement about the true value, cf the example in figure 4.2 where we show the difference in the average order books between two companies which obviously differ in their valuation uncertainty.

If valuation uncertainty coincide with a more gentle order book slope, our results support several predictions from Shalen (1993). First, increased dispersion of beliefs is predicted to increase (excess) trading volume. Our finding that a more gentle slope coincide with a greater number of trades is in line with this prediction. Second, a peak volume and volatility is predicted at the beginning of the trading day because dispersion of beliefs is greater when the price signal is more noisy. As shown in section 4, the slope of the order book is relatively more gentle in the beginning of the trading day than later in the day. However, this feature of the data may also be due to adjustments in liquidity demand.

There are also models that relate the status of the order book to the order submission strategies of homogeneous liquidity providers and how they provide the limit order book when there is a probability of informed trading. If one takes the one period model by Glosten (1994) as a benchmark, the slope of the supply and demand schedules in the order book results from the probability of informed trading. Sandås (2001) tests the predictions in Glosten (1994) in the Swedish market which is very similar to the Norwegian market. He finds strong evidence that there is insufficient depth in the observed order book relative to the theoretical prediction. In other words, the slope of the demand and supply schedules in the order book, at the inner quotes, is much too gentle to be explained by theory.

Our results for the inner part of the order book are consistent with models where a higher liquidity at the inner quotes increases the number of trades. However, our results for the full order book provide some additional results that are not captured by any theoretical model. Our finding that the volume and volatility in financial markets may be affected by valuation uncertainty and heterogeneous beliefs by liquidity suppliers provides a motivation for future research on this topic. From a more practical point of view, the discussions in the popular press about the value of companies, and sometimes very different buy and sell recommendations by analysts for the same stock, suggest that the differences in valuations may be an important factor driving trading activity in financial markets.

6 Conclusions

A positive correlation between price volatility and trading volume has been documented in a variety of studies. Investigating plausible explanations for this relation is important because it can enhance our understanding of how information is disseminated into market prices.

There are two, mainly complementary, hypotheses relating trading volume and volatility. The mixture of distributions hypothesis states that the volume-volatility relation is driven by a directing process that can be interpreted as the flow of information. The dispersion of beliefs hypothesis states that both trading volume and volatility should be higher the greater the dispersion of beliefs about security values among investors. One explanation behind this statement is based on asymmetric information and strategic investor behavior. Uninformed traders cannot distinguish informed trades from liquidity trades, and by reacting to trades with no information content, they increase both volume and volatility relative to equilibrium values in a situation with symmetric information. A positive relation between dispersion of beliefs and the volume-volatility relation can also be explained in a non-informational setting where investors have different opinions about the value of the same news. Thus, while the mixture of distributions hypothesis states that trading volume and price movements result from new information arrivals, the dispersion of beliefs hypothesis also relates a part of the volume-volatility relation to increased trading by uninformed traders or symmetrically informed investors who disagree on the same news.

Using a detailed data sample from the Oslo Stock Exchange (OSE), we examine whether information about the volume-volatility relation is contained in the shape of a limit order book. We first document that our data exhibit a standard volume-volatility relation. Moreover, we show that the result in Jones et al. (1994), that the average size of trades has little marginal explanatory power when volatility is conditioned on the number of daily transactions, also applies in a limit order market. A unique feature of our data sample is that we can rebuild the whole order book at any time during the trading day. This enables us to investigate whether the characteristics of the limit order book contain information about the volume-volatility relation.

Our main findings show that more gentle demand and supply schedules increase volatility and trading volume in a cross-sectional time series setting. One possible interpretation of this is that the number of trades is not a proxy for the mixing variable, but the mixing variable itself as suggested in models with heterogenous agents.

Appendix A

Calculation of slope measures and additional estimation results

1 Calculating slope measures

To explain the slope calculation more specifically, let N_A and N_B be respectively the total number of bid and ask prices (tick levels) containing orders. Let τ denote the tick level, with $\tau = 1$ representing the best quote with a positive volume. Furthermore, let p_1^B and p_1^A be respectively the best bid and ask prices, and p^M denote the bid-ask midpoint (which is the average of p_1^B and p_1^A). Let v_τ^B and v_τ^A be respectively the percentage of total share volume at each tick level on the bid and ask side of the book. E.g. $v_{\tau=1}^A$=0.1 would mean that 10% of the total number of shares supplied on the ask side of the order book is located at the best ask quote at that point in time. Finally, let ω_τ^B and ω_τ^A denote the weight of the local slope calculated at tick level τ for respectively the bid and the ask side of the book. These weights are set equal in the case when we equally weight the local slopes across all tick levels. In the case when we weight each local slope differently, we use a simple linear weighting scheme where the weight at each tick level, τ, is calculated as,

$$\omega_\tau^A = \frac{|\tau^{max}| - |\tau| + 1}{\sum_\tau (|\tau^{max}| - |\tau|) + 1} \tag{A.1}$$

for the ask side, and similarly for the bid side. τ^{max} is the maximum tick level with non-zero volume. Thus, the quotes which are the furthest out in the order book (e.g. at τ=80) get a relatively smaller weight than orders closer to the midpoint (e.g. at τ=10). The summation is done across all ticks with a non-zero volume. This ensures that the weights sum to one on each side of the book.

The average elasticity for the supply curve, SE, on day t at snapshot time $s \in [1..6]$ for company i can then be represented as,

$$SE^s_{i,t} = \left\{ \frac{v^A_1}{p^A_1/p^M - 1}\omega^A_{i,1} + \sum_{\tau=1}^{N_A} \frac{v^A_{\tau+1}/v^A_\tau - 1}{p^A_{\tau+1}/p^A_\tau - 1}\omega^A_{i,\tau} \right\} \tag{A.2}$$

Similarly, the demand curve, DE, can be represented as,

$$DE^s_{i,t} = \left\{ \frac{v^B_1}{\mid p^B_1/p^M - 1 \mid}\omega^B_{i,1} + \sum_{\tau=1}^{N_B} \frac{v^B_{\tau-1}/v^B_\tau - 1}{\mid p^B_{\tau-1}/p^B_\tau - 1 \mid}\omega^B_{i,\tau} \right\} \tag{A.3}$$

The first term of both equations expresses the slope between the bid-ask midpoint and the best bid and ask prices, while the second term of both equations expresses the sum of the local elasticities for the rest of the order book. The average elasticity in the order book at snapshot s is just the average of $SE^s_{i,t}$ and $DE^s_{i,t}$,

$$SLOPE^s_{i,t} = \frac{SE^s_{i,t} + DE^s_{i,t}}{2} \tag{A.4}$$

The order book is rebuilt at 10:30, 11:30, 12:30, 13:30, 14:30 and 15:30 each trading day for each firm. We exclude order volume above/below 100 ticks away from the inner quotes. For a stock trading at NOK 100 with a minimum tick size of NOK 0.5 this would mean that orders above NOK 150 and below NOK 50 are excluded from our calculations. Also, if we based our estimates of daily elasticities on one snapshot only (e.g. at noon), they could easily be biased due to large trades having temporarily reduced the liquidity of one side of the book or systematic time of day effects. To obtain a less noisy representation of the average daily supply and demand curves for each firm on each date, we therefore average the slopes across the 6 snapshots, i.e.

$$SLOPE_{i,t} = \frac{1}{6}\sum_{s=1}^{6} SLOPE^s_{i,t} \tag{A.5}$$

2 Balanced sample estimation

To examine the robustness of our results, we restrict our sample to firms that were traded every day through the sample period of 572 trading days. This leaves us with a balanced sample of 25 firms with 572 time series observations each. In addition, the filtering leaves us with a sample of the largest, most liquid and actively traded firms on the exchange. If the previous results are mainly due to noise or outliers introduced by small illiquid firms or the unbalanced dataset, the balancing of the sample should reveal this. In panel A of table A.1 we re-estimate model 1 in panel A of table 4.9 and model 1a and 1b in panel A of table 4.10 for the balanced sample. The estimation results are quantitatively similar to the results when we use the full sample. Most interestingly, the parameter estimate for *SLOPE* is negative and of similar size as before. In addition, the *SLOPE* estimate becomes smaller (less negative) the more we truncate the order book. As before, this is mainly due to the increase in the size of the slope estimates the more the order book is truncated. The largest difference between the models estimated for the balanced and unbalanced sample is that the R-squared of the models is much higher for the balanced sample, suggesting that there is more noise in the unbalanced sample.

In panel B of table A.1 we re-estimate the trading activity model in panel A of table 4.12 using the balanced sample. Although the parameter estimates change more in size than what was the case for the volatility models, the parameters are qualitatively similar. Most importantly, the *SLOPE* parameter estimate is negative when it is calculated using the full order book, and becomes increasingly positive the more the order book is truncated. Thus, also for the balanced sample, our results suggest that the more dispersed prices are across the order book, the more trades. Furthermore, when the slope is calculated from the truncated order book, only using the volume at the inner levels of the book, the results suggest that a thick book coincides with high trading activity.

Table A.1: Model estimations with balanced data sample

Panel A of the table shows the estimation results of a cross-sectional time series model for the relationship between different slope measures and the daily volatility when using a balanced sample. The models estimated in panel A are similar to those estimated (for the unbalanced sample) in table 4.9 (Model 1) and table 4.10. Panel B shows the results from estimating the relationship between the slope measures based on different truncations of the order book and the number of trades using a balanced sample. The models in panel B are similar to those estimated for the full (unbalanced) sample in table 4.12. The estimated model in panel A is,

$$|\varepsilon_{i,t}| = \beta_0 M_{i,t} + \beta_1 N_{i,t} + \beta_2 AV_{i,t} + \beta_3 MCAP_{i,t} + \beta_4 SPR_{i,t} + \beta_5 OV_{i,t} + \beta_6 SLOPE_{i,t} + \sum_{j=1}^{12} \rho_{i,j} |\hat{\varepsilon}_{i,t-j}| + \eta_{i,t}.$$

and the estimated model in panel B is,

$$N_{i,t} = \beta_0 M_{i,t} + \beta_1 AV_{i,t} + \beta_2 MCAP_{i,t} + \beta_3 SPR_{i,t} + \beta_4 SLOPE_{i,t} + \beta_5 OV_{i,t} + \eta_{i,t}.$$

where $\eta_{i,t} = \nu_i + \varepsilon_{i,t}$ defines the error structure with ν_i as the non-random fixed, firm-specific, effects. $|\varepsilon_{i,t}|$ is the absolute daily return, *N* is the number of transactions, *M* is a dummy variable for Monday, *AV* is the average trade size in shares, *MCAP* is the market capitalization (in NOK mill.), *SPR* is the relative spread (quoted spread as % of the midpoint price), *OV* is the total number of shares in the order book (sum of all orders on bid and ask side of the order book) and *SLOPE* is the average slope of the bid and offer side from the full order book, *SLOPE*10 is the slope calculated from the order book truncated to +/- 10 ticks, *SLOPE*5 is the slope calculated from the order book truncated to +/- 5 ticks. $|\hat{\varepsilon}_{i,t-j}|$ are lagged absolute returns to take into account autocorrelations.

PANEL A: Volatility/slope regressions for balanced sample

	Volatility/slope (full book)			Volatility/slope (+/- 10 ticks)			Volatility/slope (+/- 5 ticks)		
Variables	Est.	t-value	std.err	Est.	t-value	std.err	Est.	t-value	std.err
M (monday dummy)	-0.083	-1.5	0.057	-0.081	-1.4	0.057	-0.078	-1.4	0.057
N (trades)	0.004	34.6	0.000	0.004	35.6	0.000	0.004	35.7	0.000
AV (avg. trade size)	0.017	2.5	0.007	0.014	2.0	0.007	0.013	1.9	0.007
MCAP (market cap.)	-0.015	-2.7	0.006	-0.009	-1.6	0.006	-0.008	-1.4	0.006
SPR (% quoted spread)	0.345	16.3	0.021	0.358	16.9	0.021	0.351	16.6	0.021
SLOPE (full book)	-0.008	-10.0	0.001	-	-	-	-	-	-
*SLOPE*10 (+/- 10 ticks)	-	-	-	-0.007	-10.6	0.001	-	-	-
*SLOPE*5 (+/-5 ticks)	-	-	-	-	-	-	-0.005	-9.6	0.001
OV (order-book volume)	0.011	3.1	0.004	0.010	3.0	0.004	0.011	3.0	0.004
R^2	37.6 %			37.6 %			37.5 %		
N (cross section)	25			25			25		
T (time series)	572			572			572		
F-test (no fixed effects)	41.6**			41.2**			40.9**		

PANEL B: Volume/slope regressions for balanced sample

	Trades/slope (full book)			Trades/slope (+/- 10 ticks)			Trades/slope (+/- 5 ticks)		
	Est.	t-value	std.err	Est.	t-value	std.err	Est.	t-value	std.err
M (monday dummy)	-15.05	-3.56	4.224	-14.01	-3.31	4.227	-13.98	-3.32	4.216
AV (avg. trade size)	-0.50	-0.97	0.514	-0.49	-0.96	0.514	-0.34	-0.67	0.513
MCAP (market cap.)	8.10	19.69	0.411	7.41	17.77	0.417	6.89	16.41	0.420
SPR (% quoted spread)	-32.39	-20.99	1.543	-32.47	-21.02	1.545	-32.07	-20.83	1.540
SLOPE (full book)	-0.50	-8.07	0.062	-	-	-	-	-	-
*SLOPE*10 (+/- 10 ticks)	-	-	-	0.33	6.74	0.049	-	-	-
*SLOPE*5 (+/- 5 ticks)	-	-	-	-	-	-	0.44	11.27	0.039
OV (order-book volume)	3.33	12.75	0.262	3.43	13.13	0.262	3.45	13.22	0.261
R^2	33 %			33 %			34 %		
N cross section	25			25			25		
Time series	572			572			572		
F-test (no fixed effects)	233.2**			224.0**			211.9**		

3 An alternative slope measure and separating the bid/ask side

As a final exercise, we examine whether the weighted slope estimate, where we weight each local slope with the distance from the inner quote (tick-weighted slope)[1], changes our results. The two slope measures are highly correlated (98%) so we do not expect to see any large differences. However, in addition to examining the alternative slope estimate, we also estimate the effect of the slope of the bid and ask side of the order book separately, both in the balanced and unbalanced case as well as for the various truncations of the order book. In table A.2 we report slope estimates for the volatility regressions, and in panel B we report slope estimates for the trading activity regressions. To preserve space we only report the slope estimates for the average slope, the bid and ask slope as well as the p-value from an F-test for equality between the slope estimate for the bid and ask side.

Starting with the first column (slope calculated from the full order book), we see that the estimate for both the equally weighted slope and the tick-weighted slope are negative and significant both for the balanced and the unbalanced sample. The main difference is that the weighted slope estimate is more negative than for the equally weighted estimate. Furthermore, when examining the slope estimates for the bid and ask side separately, they are only significantly different at the 5% level for the tick-weighted slope in the unbalanced case. In the two next columns (when we use the truncated order books for calculating the slope) the results are essentially similar, in the sense that the parameter estimate for the tick-weighted slope is more negative than for the equally weighted slope. Both slopes become less negative the more the order book is truncated. With respect to differences between the ask and bid slopes, the parameter estimates are significantly different at the 5% and 10% level for both measures in the balanced case, while they are not different in the unbalanced case.

In table A.3 we perform a similar analysis for the trading activity regressions. Looking first at the first column (slope calculated from the full order book), we see that the parameter estimate is significantly negative both for the equally weighted and tick-weighted slope in the balanced and unbalanced case. Thus, in both cases a more dispersed order book coincides with high trading activity. Interestingly, for the equally weighted slope (both in the balanced and unbalanced case) the ask slope estimate is significantly positive while it is significantly negative for the bid slope. For both slope measures, the bid slope is more significant than the ask slope in explaining the number of trades. This may reflect the asymmetry in the order book. With respect to our interpretation of the slope as a proxy for dispersion of beliefs, it may reflect that the bid side of the market is more important with respect to dispersion of beliefs, while the ask side may be more related to liquidity supply. Furthermore, when we examine the parameter estimates, both for the equally weighted and tick-weighted slope estimates, for the slope measures calculated from the truncated order books (in the last two columns), we see that

[1] See appendix A for explanation.

the importance of the ask slope increases while the bid slope becomes more positive and less significant. This is in line with an interpretation, that the ask side is important in facilitating trading (and may capture increased trading activity by impatient liquidity traders) while the bid side reflects valuation uncertainty.

Table A.2: Alternative slope measures and separate effects of bid and ask slope on volatility

The table shows slope estimates from the volatility/slope regressions for the equally weighted slope measure and tick-weighted slope measure calculated from different order book truncations in the balanced and unbalanced sample.

	Volatility/slope (full book)		Volatility/slope (+/- 10 ticks)		Volatility/slope (+/- 5 ticks)	
	Estimate	t-value	Estimate	t-value	Estimate	t-value
Balanced sample:						
Equally weighted slope (*SLOPE*)	-0.008	-10.0	-0.007	-10.6	-0.005	-9.6
-Ask slope	-0.006	-4.0	-0.005	-6.2	-0.004	-6.2
-Bid slope	-0.004	-8.8	-0.003	-7.3	-0.002	-5.2
p-value (bid-ask=0)	0.22		0.03		0.01	
Tick-weighted slope	-0.016	-9.5	-0.014	-9.8	-0.010	-8.9
-Ask slope	-0.008	-1.8	-0.012	-4.8	-0.008	-4.8
-Bid slope	-0.008	-9.0	-0.006	-7.6	-0.004	-5.9
p-value (bid-ask=0)	0.97		0.03		0.06	
Full sample:						
Equally weighted slope (*SLOPE*)	-0.007	-11.8	-0.005	-11.4	-0.003	-9.9
-Ask slope	-0.002	-2.3	-0.003	-5.9	-0.002	-5.8
-Bid slope	-0.004	-11.4	-0.002	-8.1	-0.001	-5.8
p-value (bid-ask=0)	0.07		0.538		0.23	
Tick-weighted slope	-0.015	-11.5	-0.010	-10.4	-0.006	-8.8
-Ask slope	-0.001	-0.4	-0.004	-3.6	-0.003	-3.5
-Bid slope	-0.008	-11.5	-0.005	-8.8	-0.003	-6.8
p-value (bid-ask=0)	0.02		0.62		0.60	

Table A.3: Alternative slope measures and separate effects of bid and ask slope on trading activity

The table show slope estimates from the trading activity/slope regressions for the equally weighted slope measure and tick-weighted slope measure calculated from different order book truncations in the balanced and unbalanced sample.

	Trades/slope (full book)		Trades/slope (+/- 10 ticks)		Trades/slope (+/- 5 ticks)	
	Estimate	t-value	Estimate	t-value	Estimate	t-value
Balanced sample:						
Equally weighted slope (*SLOPE*)	-0.50	-8.1	0.33	6.7	0.44	11.3
-Ask slope	0.95	8.6	1.55	26.4	1.05	22.8
-Bid slope	-0.40	-11.8	-0.29	-9.8	-0.15	-5.7
p-value (bid-ask=0)	0.00		0.00		0.00	
Tick-weighted slope	-1.41	-11.0	-0.05	-0.4	0.54	6.4
-Ask slope	-0.62	-1.8	2.48	13.0	1.54	12.7
-Bid slope	-0.71	-10.6	-0.41	-6.9	-0.10	-1.9
p-value (bid-ask=0)	0.80		0.00		0.00	
Full sample:						
Equally weighted slope (*SLOPE*)	-0.30	-11.11	0.14	7.0	0.20	12.9
-Ask slope	0.08	2.1	0.46	22.5	0.30	18.6
-Bid slope	-0.19	-12.8	-0.12	-9.2	-0.02	-1.7
p-value (bid-ask=0)	0.00		0.00		0.00	
Tick-weighted slope	-0.78	-13.9	-0.07	-1.8	0.21	6.3
-Ask slope	-0.86	-7.1	0.29	5.3	0.19	5.16
-Bid slope	-0.34	-11.3	-0.13	-5.2	0.07	2.91
p-value (bid-ask=0)	0.00		0.00		0.01	

Bibliography

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

Barclay, M. and Warner, J. (1993). Stealth trading and volatility. Which trades moves prices? *Journal of Financial Economics*, 34:281–305.

Bessembinder, H. and Seguin, P. (1993). Price volatility, trading volume, and market depth: Evidence from futures markets. *Journal of Financial and Quantitative Analysis*, 28:21–39.

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

Chan, L. and Lakonishok, J. (1993). Instituional trades and intra-day stock price behavior. *Journal of Financial Economics*, 33:173–199.

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

Clark, P. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41:135–155.

Copeland, T. (1976). A model of asset trading under the assumption of sequantial information arrival. *Journal of Finance*, 31:1149–1168.

Copeland, T. (1977). A probability model of asset trading. *Journal of Financial and Quantitative Analysis*, 12:563–578.

Daigler, R. and Wiley, M. (1999). The impact of trader type on the futures volatility-volume relation. *Journal of Finance*, 54:2297–2316.

Fama, E. (1963). Mandelbrot and the stable paretian hypothesis. *Journal of Business*, 36:420–429.

Foucault, T., Moinas, S., and Theissen, E. (2003). Does anonymity matter in electronic limit order markets. Unpublished paper.

French, K. and Roll, R. (1986). Stock return variances. the arrival of information and the reaction of traders. *Journal of Financial Economics*, 17:5–26.

Ghysels, E. and Juergens, J. L. (2001). Stock market fundamentals and heterogeneity of beliefs: Tests based on a decomposing of returns and volatility. Unpublished paper, University of North Carolina.

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

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

Goldstein, M. and Kavajecz, K. (2000). Liquidity provision during circuit breakers and extreme market movements. Working Paper, The Rodney L. white Center for Financial Research.

Harris, L. (1986). Cross-security tests of the mixture of distributions hypothesis. *Journal of Financial and Quantitative Analysis*, 21:39–46.

Harris, L. (1987). Transaction data tests of the mixture of distributions hypothesis. *Journal of Financial and Quantitative Analysis*, 22:127–141.

Harris, L. and Hasbrouck, J. (1996). Market orders vs. limit orders: The superdot evidence on order submission strategy. *Journal of Financial and Quantitative Analysis*, 31:213–231.

Harris, M. and Raviv, A. (1993). Differences of opinion make a horse race. *Review of Financial Studies*, 6:473–506.

Huang, R. D. and Masulis, R. W. (2003). Trading activity and stock price volatility: Evidence from the London Stock Exchange. *Journal of Empirical Finance*, 10:249–269.

Jennings, R. and Barry, C. (1983). Information dissemination and portfolio choice. *Journal of Financial and Quantitative Analysis*, 18:1–19.

Jennings, R., Starks, L., and Fellingham, J. (1981). An equilibrium model of asset trading with sequential information arrival. *Journal of Finance*, 36:143–161.

Jones, C., Kaul, G., and Lipson, M. (1994). Transactions, volume, and volatility. *Review of Financial Studies*, 7:631–651.

Kalay, A., Sade, O., and Wohl, A. (2003). Measuring stock illiquidity. An investigation of the demand and supply schedules at the TASE. *Journal of Financial Economics*, 74(3):461–486.

Karpoff, J. (1987). The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis*, 22:109–126.

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

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36:394–413.

Niemeyer, J. and Sandas, P. (1995). An empirical analysis of the trading structure at the stocholm stock exchange. Working Paper No. 44, Stockholm School of Economics.

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

Sandås, P. (2001). Adverse selection and competitive market making: Empirical evidence from a limit order market. *Review of Financial Studies*, 14:705–734.

Schwert, G. (1990). Stock volatility and the crash of '87. *Review of Financial Studies*, 3:77–102.

Shalen, C. (1993). Volume, volatility, and the dispersion of beliefs. *Review of Financial Studies*, 6:405–434.

Shiller, R. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71:421–436.

Skjeltorp, J. (2000). Scaling in the Norwegian stock market. *Physica A*, 283:486–528.

Chapter 5

Ownership structure and market liquidity

Abstract

This paper studies the relationship between company ownership and market liquidity using a panel regression approach. The data sample contains detailed transactions data from a limit order driven stock market, and a full breakdown of company ownership into five distinct owner types as well as outside owner concentration and insider holdings. In line with theoretical predictions, owner concentration is found to be negatively related to spreads and information costs. A somewhat weaker negative relation is also found between spreads and insider holdings. No strong relationship can be documented between liquidity and institutional ownership. Ownership variables which affect spreads do not in general jointly affect depth in the predicted way, suggesting that spread and depth measure different dimensions of liquidity. Finally, a one-way Granger causality relation from ownership structure to liquidity is hard to document.

1 Introduction

This paper examines empirically the relationship between ownership structure and market liquidity in the Norwegian equity market.

The Norwegian equity market is medium sized by European standards and among the 30 largest world equity markets by market capitalization value. The stock exchange has become increasingly liquid during the last two decades, and is currently operating a fully automated computerized trading system similar to the public limit order book systems in Paris, Stockholm, and Toronto. Compared with the typical European firm, Norwegian firms have a low personal ownership and a flat power structure among the major owners. Another notable characteristic is a high aggregated holding by foreign investors.[1]

According to agency theory, the efficiency of a particular ownership structure depends on its ability to cope with the conflicts of interest raised by the separation of ownership and control. A positive relationship is predicted between performance and the ability to monitor firm managers (large owners and direct ownership), and between performance and a reduction in the need for monitoring (insider holdings). A central variable in both cases is informational asymmetry. Market microstructure theory predicts that informational advantages will be reflected in market liquidity through higher implicit costs of trading; the larger the fraction of owners with privileged access to information in a firm, the larger the adverse selection component of the bid-ask spread and the lower the quoted depth. Thus, the positive effect of monitoring is predicted to be mitigated by costs related to reduced liquidity. The relationship between ownership structure and liquidity is obviously important for traders searching cost efficient ways to trade. Empirical evidence on this subject also constitutes a valuable input to the problem of determining the net impact of ownership structure on economic performance. Finally, investigating the link between liquidity and company ownership relates to the important research issue of whether illiquidity has an impact on firms' costs of capital.[2]

Existing empirical literature from the US equity market studies the liquidity effects of insider holdings, institutional holdings and block ownership. Chiang and Venkatesh (1988) and Sarin et al. (2000) find a negative relationship between liquidity and insider holdings, while Glosten and Harris (1988) find no significant relationship. The relationship between liquidity and institutional holdings are also mixed. Heflin and Shaw (2000) find evidence of a positive relationship between liquidity and block holder ownership.

The main contribution from this study is to investigate the issue based on much more comprehensive data on ownership structure than used in previous studies. We use monthly data on ownership structure for the period from February 1999 to June 2001. In addition to owner concentration and insider holdings, we have access to a full breakdown of ownership into five owner types. Moreover, we are not aware of anyone who has been able to analyze this issue with panel

[1] At the end of December 2003, foreign investors holdings amounted to 48 percent of the market value of listed firms when fixed state holdings are excluded.

[2] cf. footnote 31 in chapter 1.

regression models, and conduct tests of the Granger causality between ownership variables and liquidity measures. A second contribution is that our study is based on transaction data from a pure limit order-driven market, while existing studies are based on liquidity measures from trading systems with some form of dealer intermediation. Limit order-driven trading systems are becoming increasingly popular, and there is a growing interest in the properties of this trading arrangement.

Owner concentration is found to matter for liquidity, both measured by the spread and by the adverse selection component of the spread measured by Kyle's lambda. There is also a similar, but weaker negative relationship between insider holdings and spreads. We are not able to detect any significant Granger causality relation between owner concentration and spread, nor between insider holdings and spread. Thus, one may suspect that there are some variables that jointly determine the two relations. Another interesting finding is that there is no general tendency for the ownership variables having significant effects on spreads also to have significant effects on market depth. This suggests that spread and depth are truly different dimensions of liquidity. The holdings of the two largest owner groups in the market, non-financial companies and foreign investors, have opposite effects on liquidity. While the aggregate holding of non-financial companies has a significant positive (negative) effect on the spread (the depth), the opposite is true for the aggregate holding of foreign owners. Our results with respect to the holdings of foreign investors are in accordance with the hypothesis that international owners invest mainly to capture gains from diversification.

The paper is organized as follows. Section 2 reviews some basic literature on the relation between ownership structure, performance, and liquidity. Section 3 describes the Norwegian equity market. Section 4 presents the data sample. Section 5 discusses the results from the analysis of the relationship between ownership structure and liquidity. Section 6 concludes the paper.

2 Literature

This section reviews the main theory and empirical evidence on the relationship between ownership structure, economic performance and market liquidity.

Both economic theory and public policy in most countries suggest that the structure of company ownership is important for economic performance. The standard theoretical predictions about the relative efficiency of different ownership structures are based on the principal-agent model.[3] According to this model, a monitoring problem arises because the owners of a firm (the principal) delegate the control over business decisions to the management of the firm (the agent). Thus, the main role of owners is monitoring. The incentives and capabilities to monitor a firm's business decisions are thought to depend on the *owner concentration* and the *owner*

[3]The theoretical arguments presented in the next three paragraphs are based on the classic ideas of the agency theory, see for example Jensen and Meckling (1976). See also the survey article on corporate governance by Shleifer and Vishny (1997).

type. A third relevant characteristic of the ownership structure is the division between outside owners and the *insiders*. Insiders are owners or others who, for some reason, have access to privileged information about the firm, and who typically also have the power to make changes inside the firm. In addition to the monitoring problem viv-a-vis the firm management, there are potentially similar conflicts of interests among sub-groups of owners. These conflicts typically go along the dimensions small versus large owners, direct versus indirect owners, and outside owners versus insiders.

Large owners are assumed to have more resources and stronger incentives to monitor the managers than small owners, while small owners have incentives to free-ride on the monitoring of large owners. Direct owners, represented by personal investors who monitor the agent directly, are predicted to perform more efficient monitoring than indirect owners. Typical examples of indirect ownership are widely held private firms, or private or public institutional investors who make investment decisions on others' behalf. On the other hand, large indirect investors may potentially be more professional and have better access to information than small direct investors. For example, the holdings of institutional investors tend to be larger than the holdings of the typical shareholders. If so, the information acquisition costs are spread over a larger investment, and this creates an incentive for the institutions to acquire information. Domestic versus international ownership is another owner type dimension. Assuming that international investors have an informational disadvantage vis-a-vis domestic owners and invest mainly to capture diversification benefits, there will be a negative effect on performance from increased foreign ownership due to reduced monitoring.

As long as company insiders have the same incentives as the outside owners to maximize the value of the firm, theory predicts that insider holdings and performance are positively related ("the convergence of interest" hypothesis). On the other hand, an insider may also have incentives to expropriate wealth from the outside owners. Typically, it is assumed that an increase in insider holdings has a positive (negative) effect at low levels (high levels) of insider holdings. Note that the role of insiders is not so much to monitor as to reduce the *need* for monitoring.

In general, agency theory can not answer whether the expected net impact on performance from a certain constellation of ownership is positive or negative. Hence, the net effects must be determined empirically. Empirical studies on this subject are surveyed in Bøhren and Ødegaard (2003a). Performance is typically measured by Tobin's Q, book return on assets, or market return on equity. Most papers analyze owner concentration, and a few analyze insider holdings. The results are inconclusive, but most studies find no link or a positive link between outside concentration and performance, and an initially increasing, but non-monotone relationship between insider holdings and performance. The studies assume that ownership structure is exogenously determined. This assumption is questioned in Cho (1998), who finds empirical evidence suggesting that corporate values affect ownership structure, and not vice versa.

Bhide (1993) and Maug (1998) deal explicitly with the relationship between liquidity and the efficiency of corporate governance mechanisms. Bhide (1993) argues that a liquid stock market is an obstacle to effective monitoring because it reduces the costs of "exit" by unhappy shareholders. Maug (1998) derives a theoretical model for investigating this negative liquidity effect against an opposite positive effect from reducing the problem of free-riding by small shareholders (better liquidity makes it less costly to hold large stakes). The model suggests that the positive effect dominates the negative, i.e., that a more liquid market makes corporate governance more effective.

A central variable behind the assumed ability to monitor firm management is informational advantages: insiders, large owners, and direct owners have an informational advantage relative to small owners and indirect owners, and domestic owners have an information advantage relative to international owners. Theoretical implications of informational asymmetries for financial market equilibrium is an essential topic in the market microstructure literature.[4] Market microstructure models derive how the fear of trading with someone with privileged access to information is reflected in the liquidity of stocks through higher implicit costs of trading.[5] Considerable effort is also expended to develop empirical techniques for measuring such costs.

Keim and Madhavan (1998) document that the implicit costs of trading, including spread costs, price impact costs, and timing costs, are economically significant.[6] Thus, detecting factors that effect market liquidity is important on its own grounds. Moreover, Amihud and Mendelson (1986) derive and find empirical support for a model where the expected return on a stock is an increasing and concave function of the spread. Brennan and Subrahmanyam (1996) find similar results using several empirical measures of the adverse selection component.[7]

Empirical studies from the US markets find mixed evidence on the hypothesis of reduced liquidity caused by informational asymmetries among company owners.[8] Using a sample of 75 NYSE stocks for 251 trading days from January through December 1973, Chiang and Venkatesh (1988) study how the market views corporate insiders and institutional holdings through their effects on the spread. Insider holdings are found to be positively related to the dealer's information costs after controlling for other holding costs and firm size, while institutional holdings are not found to have any impact on the spread. On the other hand, Glosten and Harris (1988) find an insignificant relation between spreads and insider holdings for a sample of 250 NYSE stocks in the period 1981-1983. Using a sample of 786 listed US stocks for the period from April to December 1985, Sarin et al. (2000) find that higher insider and institutional ownership are both associated with wider spreads and smaller quoted depth. Based on a sample of 260 listed US

[4] Classical articles are Glosten and Milgrom (1985) and Kyle (1985).

[5] Holmstrøm and Tirole (1993) derive a theoretical model where market liquidity and owner concentration are negatively related, without the assumption that large owners have an informational advantage. In this model, when a large owner decreases his or her ownership, liquidity increases because it opens up for an increasing number of liquidity traders in the stock. The increased liquidity makes it easier for privately informed investors to disguise their information and make money, which in turn encourages the search for information and increases the information content of the stock price.

[6] Implicit costs are significant both relative to explicit costs (commissions) and to portfolio returns.

[7] Chordia et al. (2001) find that there is a negative relationship between stock returns and the variability of dollar trading volume and share turnover, a result which does not support a hypothesis that agents care about the risk associated with fluctuations in liquidity.

[8] Another reduced liquidity hypothesis is based on a supply side argument; the more owners with large stakes in a company, the fewer number of stocks available for trading in the market.

stocks with transactions data on the 1988 ISSM database, Heflin and Shaw (2000) find that firms with greater block holder ownership have larger quoted and effective spreads, a larger adverse selection spread component, and smaller quoted depths.

3 The Norwegian equity market

This section describes the Norwegian equity market. First, some general statistics on the size and trading activity in the market, and the main characteristics of the trading system are presented. Then, some main features of the corporate governance structure in the market are summarized, and a motivation for looking further into the relationship between ownership structure and liquidity in this market is provided.

3.1 General statistics

The Oslo Stock Exchange (OSE) is the only regulated market place for trading equities in Norway. Table 5.1 reports some general statistics on market values and trading activity for the companies listed on the exchange in the period from 1994 to 2003.

Table 5.1: The Norwegian stock market - 1989-2003

The table reports som general statistics for the companies listed on the OSE in the period 1997-2003. Numbers are presented in nominal terms and in fixed 1998 prices. The nominal numbers are official statistics obtained from the web site www.ose.no, while the fixed prices are based on an official price index obtained from the web site www.ssb.no. The table shows the number of companies listed at year-end, the market capitalization values, the number of transactions, turnover by value, and turnover velocity. The market values include all capital registered with the Norwegian Central Securities Depository (VPS). Before 1995, this only included Norwegian companies. Dividend values include dividends in companies listed at year-end. Turnover velocity is defined as the average of annualized turnover per month divided by market value at the end of each month.

	No of listed	Market value, NOK mill		No of	Turnover, NOK bill		Turnover
Year	companies	nominal	1998 prices	trans.	nominal	1998 prices	velocity
1994	146	246606	268342	304622	124.4	135.4	-
1995	165	289804	307648	394052	156.7	166.4	-
1996	172	389397	408601	569806	231.7	243.1	-
1997	217	556002	568509	829794	341.1	348.8	69.3
1998	235	413673	413673	846535	322.7	322.7	63.0
1999	215	582941	569835	1330674	445.6	435.6	88.6
2000	214	637856	604603	2418219	609.1	577.4	96.7
2001	212	677032	622845	2529182	566.4	521.1	86.4
2002	203	502938	456801	2047861	444.4	403.6	74.7
2003	178	689734	611466	2348086	552.5	489.8	97.7

The market has grown substantially during the last 10 years. Measured in real terms, the total market capitalization value at the end of 2003 was more than the double of the value at the end of 1994, and the turnover value in 2003 was 3.6 times the turnover value in 1994. Another notable characteristic of the market, which is not shown in the table, is a very high concentration of values and trading activity in a few large companies. At the end of 2003, the five largest companies (by market value and including the fixed state holdings) accounted for 64 percent of the market value of all listed firms, and around 53 percent of the total turnover value.

3.2 Trading at the OSE

Since January 1999, the OSE has operated a fully automated computerized trading system similar to the public limit order book systems in Paris, Stockholm, and Toronto. The trading day comprises two sessions; the "pre-trade" session starting at 9:30 and ending with an opening auction at 10:00, and the "continuous trading" session from 10:00 until the trading closes at 16:00. During the pre-trade session, brokers can register trades that were executed after the close on the previous day as well as new orders. The opening auction at the end of the pre-trade session matches all registered orders at the price which maximizes the trading volume. During the continuous trading session, limit orders, market orders, and various customary order specifications are allowed. Automated order matching implies strict enforcement of the order handling rule. As is normal in most other electronic order driven markets, the order handling rule follows a price-time priority.[9]

3.3 The corporate governance structure

All listed firms in Norway must report every transaction of its outstanding equity to the Norwegian Central Securities Depository (VPS). The notification specifies the identity of the buyer and the seller, the exact time of the transaction, the number of securities traded, and the price per security. In addition, any change in the number of securities outstanding must be reported, such as stock splits, treasury stock issues, and issues of new shares.

Based on a large data sample from the VPS for the period 1989-1997, Bøhren and Ødegaard (2000) and Bøhren and Ødegaard (2001) provide a detailed description of the ownership structure of Norwegian firms.[10] The two largest owner groups of Norwegian firms in the sample period were foreign investors and non-financial domestic firms.[11] On average, foreign investors, institutional investors, and the state invest in larger companies than individuals and non-financial domestic firms. Compared with the typical European firm, the ownership structure of Norwegian firms exhibits two special features: a low personal ownership and a flat power structure among the major owners. The authors suggest that these findings may be partly explained by "a long social-democratic tradition and strong legal protection of stockholders".[12] The average aggregate holdings of different owner types and average percentage holdings of the mean owner,

[9] A new, similar trading system was introduced in the spring 2002. The reason for replacing the 1999 system was an agreement signed by OSE with the stock exchanges of Stockholm, Copenhagen and Iceland to establish a joint Nordic marketplace, known as NOREX. The NOREX exchanges are still independent entities, but the alliance has made it possible to create a joint Nordic marketplace with a common trading platform and harmonized regulations. For more information about trading on the OSE, see www.ose.no.

[10] A summary of this work (in Norwegian) is found in Bøhren and Ødegaard (2003b).

[11] Several large companies with a high fixed state holding have recently been listed on the OSE. As a consequence, the aggregate state holding is currently the largest owner group.

[12] Bøhren and Ødegaard (2001), page 1.

the largest owner and the five largest owners over the sample period are provided in table A.1 in the data appendix.

3.4 Motivation

Bøhren and Ødegaard (2003a) study the relationship between corporate governance structure and performance in the Norwegian market. Their findings support several predictions from agency theory. Insider ownership is value creating up to a holding fraction of 60 percent, and direct ownership implies a higher performance than indirect ownership through private or state intermediaries. On the other hand, a highly significant negative relationship is documented between outside ownership concentration and economic performance, suggesting that the negative effects from owner concentration outweights the benefits of monitoring.[13]

To the extent that these findings are caused by differences in monitoring efficiency resulting from informational asymmetries, they should be accompanied by liquidity effects in the market. The question of how performance effects and liquidity effects are interrelated is an important research issue. There is no straight forward way to compute the impact of illiquidity on performance or the cost of capital. However, the idea that liquidity is a priced factor in expected returns has theoretical as well as empirical support. In table 5.2 we show some rough calculations of the relationship between the bid-ask spread, measured as a percentage of the midpoint price, and returns for the Norwegian equity market in the period from 1980 to 2002.[14] The table shows average monthly percentage returns for five portfolios sorted on the relative bid-ask spread in the period from 1980 to 2002. Portfolios are grouped at the beginning of each year, using the average relative spread in the previous year as the criterion for grouping. The table shows an economically significant difference in returns for the portfolio with the lowest bid-ask spread and the portfolio with the largest bid-ask spread for both the equally weighted and value weighted portfolios. The equally weighted portfolios give the best picture of the average stock. For these portfolios, the difference in returns for the portfolio with the lowest bid-ask spread and the portfolio with the largest bid-ask spread was 2.20 percent. Table A.2 in the data appendix verifies that this relationship is quite robust over five years sub-periods. The numbers also indicate that the higher bid-ask spread portfolios have higher volatility. Thus, there seems to be a positive relationship between the size of the bid-ask spread and expected returns, similar to the relationship documented for the US market in Amihud and Mendelson (1986) and Brennan and Subrahmanyam (1996). This simple analysis does not prove the existence of such a relationship. However, it does provide a good motivation for making a first step and figuring out whether there are in fact liquidity effects from ownership structure in the market.

The relationship between liquidity and ownership structure in the Norwegian market is previously studied in Sjo (1998) and Tobiasson et al. (1999). Using data for 1995, Sjo (1998)

[13]Negative effects of owner concentration include majority-minority conflicts, reduced manager initiatives, reduced benefits of diversification and reduced market liquidity.

[14]I am grateful to Bernt Arne Ødegaard for providing table 5.2 and table A.2 in the data appendix.

Table 5.2: Monthly returns for liquidity based portfolios, 1980-2002

The table shows monthly returns and the number of securities for five portfolios sorted on relative spread in the period from 1980 to 2002. Portfolios are grouped at the beginning of each year, using the average relative spread in the previous year as the criterion for grouping. The sample includes all listed securities on the OSE which comply with the following three filtering criteria: (i) the stock price is above NOK 10, (ii) the total value outstanding of the company is at least NOK 1 million, and (iii) the security is traded at least 20 days during one year. The filtering criteria imply that, on average, a year contains 121 companies. Panel A(B) of the table shows the results for equally weighted(value weighted portfolios).

Panel A	Returns					No of securities		
EW Portfolios	mean	std	min	median	max	avg	min	max
1 (smallest)	0.83	6.93	-24.90	1.00	18.65	28.1	10.0	45.0
2	0.96	6.66	-22.45	1.63	18.80	27.2	10.0	44.0
3	1.38	6.44	-21.24	1.24	21.45	27.1	11.0	45.0
4	2.02	6.39	-15.51	1.51	21.87	27.0	10.0	44.0
5	3.03	7.22	-15.92	1.88	35.04	27.0	10.0	44.0

Panel B	Returns					No of securities		
VW Portfolios	mean	std	min	median	max	avg	min	max
1 (smallest)	1.42	6.89	-26.89	1.86	22.26	27.3	9.0	44.0
2	1.94	7.16	-25.53	2.49	31.11	27.0	10.0	44.0
3	2.15	7.29	-21.70	1.86	26.48	26.8	10.0	44.0
4	3.02	7.71	-15.08	2.26	45.79	26.8	10.0	44.0
5	4.10	8.50	-19.81	2.46	36.77	28.5	12.0	46.0

studies the liquidity of 61 industrial companies listed on the OSE. Liquidity, measured by the relative bid-ask spread, trade frequency, and turnover velocity, is found to be positively related to company size, low concentration of ownership, high fraction of foreign owners, high beta risk, and high market value relative to the book value of equity.[15] Moreover, relative bid-ask spread and turnover velocity are both positively related to returns. Tobiasson et al. (1999) study the relationship between liquidity and ownership structure using transaction data for two periods, 20 companies in the period from February 1 to March 20 1996, and 131 companies in the period from September 1 1997 to February 22 1998, and ownership data for year-end 1997.[16] A negative relationship is found between liquidity and the holdings of company insiders. The relationship between liquidity and the largest company owner is weak. No significant relation is found between liquidity and institutional ownership or the fraction owned by foreign investors.

Hence, existing studies show some evidence of a negative relationship between liquidity and insider holdings and some weak evidence of a negative relation between liquidity and owner concentration. A problem with both studies is that the data samples are quite limited. Neither of the papers have access to intraday data, meaning that they cannot focus on the most relevant liquidity measures (effective spread and the information component of the spread). Moreover, neither of the papers has access to time series data of ownership, and neither looks at the Granger causality issue. A final motivation is that the existing studies are based on trade data before the introduction of a fully decentralized trading system with a strict price-time priority rule.[17]

[15] Concentration of ownership is measured as the fraction of the company which is *not* owned by the three largest owners.

[16] Transactions data are from the OSE, and ownership data are from VPS, except the holdings of insiders which are prepared at the Norwegian School of Management BI.

[17] The OSE has operated an electronic trading system with continuous trading in all listed securities since 1988.

4 The data sample

Our transaction data consist of every order and trade at the OSE during the period from February 5, 1999, shortly after the implementation of the new trading system, through May 2001.[18] From the VPS we have monthly ownership data for the same period. The ownership data include a complete breakdown of firm ownership into five owner types as well as aggregated holdings of the 1-5 largest owners. We also have estimates of the aggregate holdings of primary insiders. Primary insiders include company managers and members of the Board of Directors. The holdings are estimated at the Norwegian School of Management BI based on statements given to the OSE by the insiders. We apply the following filter criteria on our data sample,

- We only look at the "continuous trading" session from 10:00 until 16:00.
- To avoid that infrequently traded firms introduce noise into our intraday liquidity measures, we filter out companies which were traded on less than two thirds of the trading days in the sample period.[19]
- Low valued stocks are problematic because they tend to have exaggerated returns. The exaggerated returns are caused by the minimum tick size and the fact that these stocks typically trade at prices close to zero. To avoid that such securities affect average returns, we exclude stocks that trade for less than NOK 10.

From the resulting data sample we remove two companies, one due to its special trading characteristics during the sample period[20], and the other one due to lack of data on ownership. This leaves us with a total of 94 securities in 88 companies.[21]

4.1 Estimation of adverse selection costs

There are many suggestions in the literature on how to estimate adverse selection costs.[22] A potential problem for this study is that the methods are designed for a different institutional setting (competitive quote-driven markets) than ours (order-driven market). We use a version of the Glosten and Harris (1988) method (hereafter the GH-method) without inventory costs and

However, the old system did not enforce priority rules. A broker could freely choose what orders he or she wanted to match, independent of price. Moreover, since there was no time priority rule, traders had no incentives to submit orders "first". While competition among brokers implied that price priority was enforced in practice, the lack of time priority presumably had a negative impact on market depth.

[18] The order data contain a ticker, the time of submission, the quantity, the order side (buy or sell), the disclosed and hidden parts of the order volume, and a flag indicating whether it's a new order, a revision of an existing order, or a cancellation of an existing order. If an order is revised, information on the previous price and volume of the order is attached to the observation. An order id enables us to track whether different parts of the order is executed against several orders. The trade data include ticker, quantity, time, the member firms on each side of the trade, and an identification of the member firm initiating the trade.

[19] More specifically, we filter out companies which were traded on less than 400 of the total 597 trading days in the sample period.

[20] The stock was extremely volatile during the sample period with prices ranging from NOK 184 to NOK 2094

[21] Six companies are represented in the sample with both A and B-shares. In contrast to A-shares, B-shares do not give the owners a right to vote.

[22] See Glosten and Harris (1988), George et al. (1991), Madhavan and Smidt (1991), and Huang and Stoll (1997).

one of the methods suggested in George et al. (1991) (hereafter the GKN-method). The main difference between the two methods is that the GH-method assumes that the adverse selection component increases with order size, while the GKN-method assumes that the adverse selection component remains a constant proportion of the spread.[23]

The GH-method assumes competitive risk-neutral market makers, but not complex dealer strategies such as those allowed in the Madhavan and Smidt (1991) method.[24] The adverse selection component is estimated as a coefficient measuring the impact on intraday price changes from signed order flow ("Kyle"s lambda),

$$\Delta P_t = \lambda q_t + \psi[D_t - D_{t-1}] + y_t \tag{5.1}$$

where ΔP_t is the intra-day change in the transaction price P_t from $t-1$ to t, q is the order flow, D is a dummy variable taking the value $+1/-1$ if the trade at t is buyer-initiated/seller-initiated, and y is an information signal. λ is the adverse selection component, and ψ is a measure of the compensation for per share execution costs. Following Brennan and Subrahmanyam (1996), we proxy the variable proportional cost of transacting as $VC = \lambda\bar{q}/P$, where $\bar{q}$ is the average transaction size in the stock.[25]

The GKN-method is based on the method of measuring effective spreads, $\hat{S_e}$, from the serial covariance of price changes, which was first suggested by Roll (1984),

$$\hat{S_e} = 2\sqrt{-cov(\Delta P_t, \Delta P_{t-1})} \tag{5.2}$$

The assumptions underlying the above estimate of the effective spread estimate are no inventory costs, no information events, and a probability of trade reversals equal to 0.5. Since the information component is not included in the estimate, we can write,

$$\hat{S_e} = \psi S_q \tag{5.3}$$

where S_q is the quoted spread. The estimated adverse selection cost is found as one minus the estimated coefficient of ψ. The GKN-extension consists in an allowance for time varying expected returns. One of the suggested ways to implement this is by exchanging the serial covariance of price changes with the serial covariance of the *difference* in trade-to-trade returns and subsequent bid-to-bid returns. The point is to get a pure measure of the bid-ask bounce by extracting the time variation in expected returns.

Note that neither of the two estimation methods we use prevent the estimates of adverse selection costs from being negative. A more detailed description of the two methods is provided

[23] If only the adverse selection component of the spread varies with trade size, then the GKN-measure will only be valid for small trades.

[24] In Madhavan and Smidt (1991), specialists use Bayesian updating to revise their quotes.

[25] Brennan and Subrahmanyam (1996) calculate a second proxy based on the firm's number of shares outstanding. This proxy overcomes the problem that very small trade sizes in very illiquid firms may yield a lower estimated variable cost for illiquid firms that for relatively liquid firms. Since our sample includes relatively liquid firms only, we do not calculate the second proxy.

in appendix B.

4.2 Descriptive statistics

Table 5.3 shows some basic statistics on the market liquidity of sample companies during the sample period. The quoted spread measures the absolute “round trip” cost of trading a small amount of shares at the inner quotes. The effective spread takes into account that trades are often executed inside (price improvement) or outside the spread (”walking the book”). The effective spread is calculated as the absolute difference between the execution price and the bid-ask midpoint, multiplied by two. This spread measure is considered the most appropriate measure of costs, especially for large trades.[26] The time weighted relative spread is measured relative to the spread midpoint. Following Sarin et al. (2000), we calculate time weights as the number of seconds a quote was outstanding divided by the total number of seconds during the trading day. Market capitalization values, prices, quoted spreads, effective spreads, and average daily trade volume are reported in Norwegian kroner (NOK). During the sample period USD 1 was equal to roughly NOK 8.5.

Table 5.3: Market liquidity

The table reports statistics on the market liquidity of the sample companies during the sample period from February 5 to June 30. The “quoted spread” is the average difference between the inside ask and bid prices for executed trades in a stock over the trading day. The “effective spread” is the average absolute difference between the execution price and the bid-ask midpoint multiplied by two. The “relative spread (time weighted)” is the time weighted quoted spread relative to the spread midpoint, where the time weights are calculated as the number of seconds the quotes were outstanding divided by the total number of seconds during the trading day. Adverse selection costs are estimated according to Glosten and Harris (1988) (GH) and George et al. (1991) (GKN). For the GH-method, the adverse selection cost is reported (in percent) as $\lambda q/P$, where λ is the estimated adverse selection component according to the Glosten and Harris (1988) method, q is the (monthly) average transaction size in the stock, P is the (monthly) average close price for the stocks. For the GKN-method, the adverse selection costs is reported as a percentage of the spread. The “quoted depth (time weighted)” is the time weighted sum of the depth at the best bid price and the best ask price divided by two, where the weights are calculated as described above for the relative spread. Market capitalization values, price, quoted spread, effective spread, and average daily trade volume are reported in Norwegian kroner (NOK). During the sample period USD 1 was equal to roughly NOK 8.5. Relative spread are in percent.

	Mean	Std	Min	Median	Max
Market cap (bill NOK)	5.95	11.82	0.12	2.29	89.13
Price	101.62	74.83	14.42	101.76	345.66
Quoted spread	1.75	1.29	0.16	1.50	7.90
Effective spread	1.30	0.98	0.11	1.13	5.75
Relative spread (time weighted)	1.91	1.12	0.23	1.78	5.47
Adverse selection component:					
- GH-method	0.02	0.08	-0.56	0.00	0.31
- GKN-method	0.53	0.90	-6.24	0.51	2.39
Quoted depth (time weighted)	10236	18597	1173	7708	124214
Avg daily trade size	1826	1315	501	1218	8330
Avg daily no of trades	57	68	7	48	366
Avg daily trade volume, in shares	134204	183693	3527	104494	1156907
- in 1000 NOK	10834.86	22193.73	478.49	8582.36	178323.19

The average firm in our sample has a value of NOK 5.95 billion, an average share price of

[26] See for example Angel (1997) and Bacidore et al. (1999).

around NOK 102, and experiences an average of 57 trades per day with an average trade size of 1826 shares. Measured by the effective spread, the average cost of trading during the sample period was NOK 1.30. As expected, this cost was lower than the average quoted spread. The average quoted depth is 10236 shares. For the GH-method, the adverse selection cost is reported as a percentage of the share price for a trade of average size, i.e., an average trade of 1826 shares in a stock with the average price of NOK 102 yields an adverse selection cost of about NOK 2.[27] For the GKN-method, the costs are reported as a percentage of the spread.[28]

Firm size varies considerably in the sample from NOK 120 million to over NOK 89 billion. Firm size is of obvious importance for market liquidity. We therefore recalculate the liquidity measures for four portfolios of firms which are grouped based on their market capitalization value at the beginning of each year. The results of these calculations are provided in table A.3 in the data appendix. The table shows that the firms in the group of the largest firms are much larger than the firms in the other three groups. The mean firm size of the largest companies varied from NOK 15.66 million in 1999 to NOK 19.62 million over the first half of 2001, while the mean market cap for the rest of the sample varied from NOK 0.34 million for the portfolio of the smallest companies in 1999 to NOK 3.47 million for the medium largest companies over the first half of 2001. Some typical features of market liquidity are also evident; spreads are reduced over time, and spreads are lowest/depths are highest for the largest firms. Average trade size does not vary a lot over the four size portfolios indicating that investors split large orders into a series of orders of smaller size.[29]

Table 5.4 shows some basic descriptive statistics of the average ownership structure over the sample period. A “state owner” represents the government (central or local) including their pension funds. “Institutional owners” consists of private Norwegian banks, insurance firms, pension funds, and mutual funds. ”Non-financial” owners are private domestic firms which are not classified as institutional owners. An “individual owner” is a personal (non-corporate) investor with Norwegian residency. Finally, a “foreign owner” is any organization not registered in Norway or a non-resident individual.[30] On average, a firm in our sample had 11 state owners, 102 institutional owners, 354 company owners, 5531 individual owners, and 306 foreign owners. The median number of individual owners is half the mean number, suggesting that the variable has a positively skewed distribution. The weighted number of individual owners is almost three times larger than the mean number suggesting that large companies have a larger number of individual owners than small companies.

Compared with the ownership structure during the 1989-1997 period, the average aggregated holdings of the five owner types have been stable. Foreign investors and non-financial domestic

[27] For comparison, Glosten and Harris (1988) report an adverse selection cost of USD 0.0133 for a 1000 share lot, that is roughly NOK 0.11 for a trade of about half the size.

[28] For comparison, George et al. (1991) find that the proportion of the spread due to adverse selection ranges between 8 and 13 percent, whereas we find that the average proportion is only 0.53 percent.

[29] In addition, in an electronic trading system, large orders will generally be partially executed against smaller orders, making the average trade size smaller the average order size.

[30] This group contains investors who are registered by their name and investors who own anonymously through a nominee account.

firms are still the largest owner types, and individual holdings are small, especially in large firms. On average, 7.8 percent of a firm is owned by primary insiders. The value weighted mean holding is half this number, suggesting that primary insiders are concentrated in the smaller firms. The five largest owners hold on average 44 percent of a firm, while the largest owner hold around 20 percent. Hence, the power structure is fairly flat. In the cases where the largest owner is an institutional investor, the largest holding is on average 10 percent. This is much lower than for the other owner groups and suggests that institutional owners hold diversified portfolios. Table A.4 provides descriptive statistics of the ownership structure for the four size portfolios over each year. The general picture is that the ownership structure has been stable over the sample period. Owner concentration is fairly similar over the size groups, while insider holdings are highest in the smallest firms. State ownership is concentrated in the largest firms, while individual investors are concentrated in the smallest firms.

Table 5.4: The corporate governance structure

The table reports some statistics on the ownership structure of the sample companies over the sample period from February 1999 to June 2001. The statistics include the number of owners and aggregate holdings of different owner types, as well as the holdings of five largest owners and the largest owner, including the cases where the largest owner belongs to a given owner type. Primary insiders is a subset of the corporate insiders and include company managers and members of the Board of Directors. The table shows the equally weighted average, the standard deviation, the value weighted average, the median observation, and the number of firms. The value weighted averages are weighted based on the value of the firm's equity. All holding numbers are in percent.

Ownership structure	EW mean	Std	Median	VW mean	n
Owner types					
No of owners:					
State	11	14	7	24	94
Institutional	102	66	89	178	94
Non-financial	354	323	243	660	94
Individual	5531	8527	2101	14728	94
Foreign	306	479	142	647	94
Aggregate holdings:					
State	6.57	12.22	2.38	17.57	94
Institutional	24.49	13.21	21.57	22.48	94
Non-financial	27.93	17.07	26.16	22.22	94
Individual	17.33	16.01	12.35	7.63	94
Foreign	23.84	18.22	21.52	30.02	94
Primary insiders	7.80	14.33	6.48	3.80 94	
Owner concentration					
Five largest owners	44.01	16.85	42.87	47.95	94
Largest owner	20.21	13.02	15.44	25.31	94
- state	25.42	16.59	19.71	.	.
- institutional	10.00	5.37	9.26	.	.
- non-financial	20.37	11.67	18.43	.	.
- individual	28.60	21.15	22.21	.	.
- foreign	18.78	14.27	13.96	.	.

To check whether our data sample is biased against certain industry groups, we also split the sample according to the FTSE global classification system. Table A.5 in the data appendix provides some descriptive statistics on liquidity and ownership structure based on this classification of the sample. The sample includes firms from all the 10 economic groups, the largest group, cyclical services, represent 22 percent of the sample. Measured by the effective spread, the costs of trading varies from an average of NOK 0.65 for resource companies to an average of NOK 1.98 for financial companies. Six companies are represented in the sample with both A and B-shares. Overall, the liquidity seems to be quite similar between these two types of shares.

Table 5.5 shows the relationship between the effective bid-ask spread, measured relative to the midpoint price, and monthly returns for our data sample. A positive relationship between costs and returns is evident, however, the relationship is not so strong as the one shown for the much longer time period in table 5.2 in section 3. Table A.6 in the data appendix shows the results of similar calculations over 5 sub-periods of six months starting from the second half of 1999. A positive relationship is evident for the two first sub-periods. When the market starts to fall in the second half of 2000, the relationship disappears.

Table 5.5: Monthly returns for portfolios sorted on effective relative spread, 1999.2-2001.2

The table shows characteristics of the return distribution of monthly returns for four equally weighted liquidity portfolios. The companies included in the data sample are all firms with price greater than NOK 10 which are traded on at least 400 days out of the 597 trading days from February 5 1999 to June 30 2001. The portfolios are grouped at the beginning of each half year, using the average relative effective spread in the previous half year as the criterion for grouping. The portfolios are assumed to be held the whole period from 1999.2 to 2001.2 and rebalanced every half year.

	Effective spread					Return				
1999.2-2001.2	mean	std	min	median	max	mean	std	min	median	max
Portfolio 1	0.55	0.17	0.14	0.56	0.82	0.01	6.03	-19.92	0.77	15.26
Portfolio 2	1.06	0.18	0.74	1.06	1.39	0.05	6.03	-25.98	0.57	16.52
Portfolio 3	1.74	0.26	1.28	1.72	2.28	2.05	5.81	-12.35	1.74	17.07
Portfolio 4	3.42	1.26	2.06	3.04	9.66	1.47	8.38	-17.34	1.00	38.62

5 Results

This section presents the results from analyzing the relationship between ownership structure and liquidity. First, a reference model is estimated where we regress liquidity measures on common control variables only. We then present some predicted relationships between liquidity and company ownership, and report the results from three regression models; one where we include owner concentration and insider holdings, and two where we also include different owner type variables. Finally, we present the results from several tests of the Granger causality between the ownership variables and the liquidity measures.

5.1 A reference model

Table 5.6 shows the results from an estimated “reference model” where we regress several liquidity measures on common control variables. More specifically, we estimate five versions of a panel regression model with one-way fixed effects of the form,

$$LM = \sum_{k=1}^{K} \beta_k X_{i,t,k} + \eta_{i,t} \tag{5.4}$$

where LM is the liquidity measure, $X_{i,t,k}$ is the matrix of explanatory variables (k) over time (t) for each company (i), and $\eta_{i,t} = \nu_i + \varepsilon_{i,t}$ defines the error structure with ν_i as the non-random fixed, firm-specific, effect. Since we use a one-way fixed effects specification, the estimation is analogous to a least-squares dummy variable (LSDV) regression with firm-specific constants ν_i.[31] The liquidity measures used are the relative time weighted spread, the relative effective spread, the time weighted depth, and the two estimates of the information component of the spread, the GH-measure and the GKN-measure, as defined in subsection 4.1. Spreads measure the costs per share of liquidity (market width), while depth measures the ability of the market to absorb a series of trades. We use a logarithm transformation of the percentage spread variables to reduce heteroscedasticity. We also take logarithms of the depth measure to reduce skewness. The control variables are firm size, stock price, return volatility, and trading frequency.[32] We use logarithms of market capitalization values to reduce skewness. Market volatility is measured as the standard deviation of daily returns (from midpoint prices), and trading frequency is the average daily number of trades.

A priori, we expect spreads to be decreasing in firm size, price, and trading activity, and increasing in volatility. If depth and spread are jointly determined (low spreads are accompanied by high depth end vice versa), we would expect depth to be positively related to firm size and trading activity, and negatively related to volatility. On the other hand, we expect a negative correlation between price and depth, which makes the relation between firm value and depth hard to predict. Two hypotheses about the relationship between the information component of the spread and the control variables are that there is more private information in small firms than in large firms, and that private information is more valuable in high risk companies. If so, the coefficient for firm size should be negative, and the coefficient for return volatility should be positive.

The results of the estimation show that all the standard properties of market liquidity apply. Spreads are lower the larger the firm size, the higher the price, and the higher the trading frequency, and higher the higher the volatility. Depth increases with trading activity and decreases with price and volatility. The positive relation between trading activity and depth is in accordance with the result in Biais et al. (1995) that thick books at the inner quotes result in trades. The GH-measure of adverse selection costs decreases with firm size and increases with volatility, supporting the ideas that there is less private information in large firms, and more valuable information in risky firms. The explanatory power of the GKN-estimates of the information component is low. The results are fairly robust when the model is estimated over three

[31] Since not all firms are traded every day, our sample is unbalanced. We use the TSCSREG procedure supplied with SAS v.8.2 for estimating the models. The procedure is capable of processing data with different numbers of time-series observations across different cross sections.

[32] These controls are typical in studies of spreads, see Chiang and Venkatesh (1988), Sarin et al. (2000), and Heflin and Shaw (2000).

Table 5.6: A reference model for market liquidity

The table reports results from estimating a panel regression model with one-way fixed effects (least squares dummy variable estimation) for five measures of liquidity as the dependent variable, and the logarithm of the market capitalization value, the average closing price, the standard deviation of daily returns, the average daily number of trades, and dummies for the fixed effect of each company as the independent variables. The dependent variables are the log(relative weighted spread), the log(relative effective spread), the log(weighted depth), and the adverse selection costs according to the GH-method (variable proportional costs), and the GKN-method. For each model, we report the estimated coefficients for the four control variables, R-squared, and the F-test for no fixed effects. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, and * defines significance at the 10 percent level. The sample includes 29 time series observations covering 93 companies.

	Dependent variables				
Independent variables	Rel weighted spread	Rel eff spread	Weighted depth	GH info comp	GKN info comp
Market cap	−0.2165***	−0.1966***	−0.0820*	−0.0002***	0.0007
Price	−0.0009***	−0.0009***	−0.0044***	0.0000	0.0000**
Return volatility	2.7521***	2.4857***	−1.1740*	0.0025***	−0.0144
Trades per day	−0.0020***	−0.0018***	0.0015***	0.0000**	0.0000
R square	0.8410	0.8471	0.7812	0.2689	0.0991
F-test no fixed effects	26.63***	27.61***	19.99***	3.48***	2.30***

sub-periods using the available observations for each year, cf table A.7 in the data appendix.[33]

5.2 Predicted relationships between ownership structure and liquidity

Market microstructure theory suggests a negative relationship between liquidity and the holdings of investors with privileged access to information. From principal agent theory we suspect that these investors include large owners, direct owners and insiders. Concentrated ownership has a negative impact on liquidity also in the absence of informational asymmetries, because there will be less available shares to trade.

Based on the predictions from the agency theory, one can discuss the likely effects of different owner types on market liquidity. What is regarded as a signal of informational asymmetry in the market is, however, ultimately an empirical question.

We have no clear prediction from theory on the net impact on liquidity from *institutional* owners. On the one hand, large institutional investors potentially have an informational advantage because they have resources to acquire and analyze information. On the other hand, institutional ownership is indirect, and the typical investment policy is to hold diversified portfolios. The latter argument suggests that the causality may go from liquidity to institutional ownership and not vice versa. The typical prediction about *foreign* owners is that they have an informational disadvantage vis-a-vis domestic owners and mainly invest to obtain gains from diversification. If so, there should be a positive relationship between foreign ownership and liquidity, and possibly a causality from liquidity to foreign ownership rather than vice versa. Theory predicts a negative effect on liquidity from large *individual* owners, while firms with many small individual investors should have high liquidity. A prediction which is gaining a lot

[33]The three sub-periods consist of 11 months in 1999, 12 months in 2000 and 6 months in 2001.

of popularity in many political environments, is that private ownership is more effective than *public* ownership. The inefficiency of public ownership is claimed to follow from factors such as a slow decision making process, too much focus on political goals, less familiarity with business management, a passive role in the board room, and conflicts between the dual role of being an owner and the governing authority. These factors are not based on asymmetric information, unless one argues that public owners for the reasons specified above have less *capacity* to acquire and analyze information. Thus a negative effect from inefficient public ownership should probably effect performance directly. However, it may also be reflected in the market liquidity. A problem with our ownership data is that the group of state owners includes the public pension funds, which probably have characteristics very similar to institutional investors.

Based on the discussion above we choose the following ownership variables:

- The holdings of the primary insiders.
- The aggregate holdings of the five largest owners. We occasionally use a "free float" variable instead, which we define to be the aggregate holding which is *not* owned by the five largest owners.
- We split the holding of the largest owner into five separate variables depending on to which owner group the largest owner belongs.
- The percentage aggregate holding and the number of owners for each owner group.
- The absolute value of the change in the number of owners. This variable should to some extent capture differences in trading activity among the owner groups.

Table 5.7 shows the correlation structure between liquidity measures and ownership variables. The spread measures are positively correlated with the holdings of primary insiders and owner concentration (except when the largest owner is public), and negatively related to the number of owners. As expected, depth and spread are negatively correlated, however, a correlation coefficient of around six percent seems modest. The GKN-measure is weakly correlated to the ownership variables as well as to the other liquidity measures.

5.3 Outside owner concentration and primary insiders

In table 5.8, we present the results from including the aggregate holding of the five largest owners and the aggregate holdings of the primary insiders in the regression model. Table A.8 shows the results for estimation of this model for each of the years 1999 to 2001.

As hypothesized from theory, there is a significant relationship between outside owner concentration and liquidity measured by the spread. This result is also quite robust over sub-periods. Assuming the mean effective spread of 1.63 percent, a coefficient estimate of 1.3796 predicts that a 1 percent increase in the holdings of the five largest owners increases the effective spread

Table 5.7: Correlation between liquidity measures and ownership structure

The table shows Pearson correlation coefficients between liquidity measures and ownership variables. "Rel spr. weig." is the average time weighted NOK spread, in percent of the midpoint price, where the weights are equal to the number of seconds the quotes were outstanding divided by the total number of seconds during the trading day. "Rel spr. eff." is the average absolute NOK difference between the execution price and the bid-ask midpoint multiplied by two, in percent of the midpoint price. "Depth weig." is the average time weighted sum of the volume at the inner bid and ask quotes divided by two, where the weights are as defined for the relative weighted spread above. "GH info. comp" and "GKN info comp." are costs of adverse selection according to the methods of, respectively, Glosten and Harris (1988) and George et al. (1991). "Prim insid" is the aggregate holdings of the primary insiders (managers and board members). "Five lar. own." is the aggregate holding of the five largest owners. "Large, state", "Large, inst.", "Large, comp.", "Large, ind.", and "Large, for." are the holding of the largest owner when the owner is, respectively, the government (central or local) including their pension funds, institutional investors (private Norwegian banks, insurance firms, pension funds, and mutual funds), private domestic firms, personal (non-corporate) Norwegian investors, and organizations not registered in Norway or non-resident individuals. Finally "No of" are the number of owners of each of the five owner groups defined above. Correlation coefficients in **bold** are significant at the 5 percent level.

	Rel. spr. weig.	Rel. spr. eff.	Depth weig.	GH info. comp.	GKN info. comp.	Prim. insid.	Five lar. own.	Large State	Large Inst.	Large comp.	Large ind.	Large for.	No of state	No of inst.	No of comp.	No of ind.	No of for.
Spr. w.	**1.0000**																
Spr. e.	**0.9491**	**1.0000**															
Depth	**-0.0634**	**-0.0571**	**1.0000**														
GH	**0.4639**	**0.4331**	-0.0431	**1.0000**													
GKN	**0.0482**	**0.0508**	0.0158	-0.0333	**1.0000**												
Insid.	**0.1313**	**0.1004**	-0.0365	**0.0654**	**-0.0614**	**1.0000**											
5 L.	**0.2787**	**0.2822**	**0.0769**	**0.1173**	0.0020	0.0323	**1.0000**										
L. sta.	**-0.0763**	**-0.0690**	**0.0670**	**-0.0791**	0.0142	**-0.1664**	**0.3009**	**1.0000**									
L. inst.	**0.0488**	0.0280	-0.0114	-0.0159	0.0082	**-0.1104**	**-0.1060**	**-0.1002**	**1.0000**								
L. com.	**0.0586**	**0.0624**	**-0.0447**	0.0050	0.0073	-0.0053	**0.3633**	**-0.2195**	**-0.1793**	**1.0000**							
L. ind.	**0.1708**	**0.1596**	-0.0156	**0.1811**	0.0360	**0.1655**	**0.1330**	**-0.0571**	**-0.0467**	**-0.1023**	**1.0000**						
L. for.	**0.0737**	**0.0835**	**0.1262**	**0.0514**	-0.0365	**0.0459**	**0.2837**	**-0.1807**	**-0.1476**	**-0.3234**	**-0.0842**	**1.0000**					
No sta.	**-0.4139**	**-0.4264**	0.0286	**-0.2037**	-0.0282	**-0.1807**	**-0.1257**	**0.1398**	**0.1243**	**-0.1289**	**-0.1270**	**-0.1118**	**1.0000**				
No inst.	**-0.5925**	**-0.6168**	0.0318	**-0.3203**	-0.0425	**-0.1417**	**-0.1214**	**0.2194**	**0.0995**	**-0.0504**	**-0.1948**	**-0.1594**	**0.7425**	**1.0000**			
No com.	**-0.4899**	**-0.4992**	0.0861	**-0.2419**	-0.0377	**-0.1624**	**-0.1613**	**0.1569**	**0.0713**	**-0.1295**	**-0.1338**	**-0.1120**	**0.7006**	**0.7798**	**1.0000**		
No ind.	**-0.3757**	**-0.7181**	**0.1079**	**-0.1740**	**-0.0477**	**-0.1053**	**-0.0927**	**0.2631**	0.0322	**-0.1537**	**-0.0934**	**-0.1032**	**0.6181**	**0.5485**	**0.8192**	**1.0000**	
No for.	**-0.3300**	**-0.3402**	**0.0450**	**-0.1443**	**-0.0535**	**0.0587**	**-0.1128**	**0.1086**	0.0077	-0.0300	**-0.0882**	**-0.0864**	**0.4400**	**0.4040**	**0.5314**	**0.7843**	**1.0000**

Table 5.8: Market liquidity, owner concentration and holdings of primary insiders

The table reports results from estimating a panel regression model with one-way fixed effects (least squares dummy variable estimation) for five measures of liquidity as the dependent variable: log(relative weighted spread), log(relative effective spread), log(weighted depth), adverse selection costs according to the GH-method (variable proportional costs), and the adverse selection component of the spread according to the GKN-method. The independent variables are the aggregate holdings of the five largest owners, the aggregate holdings of the primary insiders, the logarithm of the market capitalization value, the average closing price, the standard deviation of daily returns, the average daily number of trades, and dummies for the fixed effect of each company. For each model, we report the estimated coefficients (except the coefficients for the fixed effect dummies) , R-squared, and the F-test for no fixed effects. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, and * denotes significance at the 10 percent level. The sample includes 28 time series observations covering 93 companies. To check the residuals of the panel regression models, we estimate a standard OLS regression model for each company and test, for each company, whether the residuals are autocorrelated. We then re-estimate the panel regression models without the companies with autocorrelated residuals. At the 1 percent level (5 percent level), the number of companies with autocorrelated residuals vary from 1-2 companies (7-9 companies). Removing these companies does not seriously change our results.

	Dependent variables				
Independent variables	Rel weighted spread	Rel eff spread	Weighted depth	GH info comp.	GKN info comp
Five largest	1.1527***	1.3796***	0.2030	0.0008***	0.0089
Primary Insiders	0.2579***	0.2149***	−0.1700	−0.0001	0.0006
Market cap	−0.2387***	−0.2244***	−0.0893***	−0.0003***	0.0005
Price	−0.0009***	−0.0009***	−0.0044***	0.0000	0.0000**
Return volatility	2.4423***	2.1032***	−1.2628***	0.0022***	−0.0163
Trades per day	−0.0018***	−0.0015***	0.0016***	0.0000	0.0000
R square	0.8501	0.8592	0.7817	0.2748	0.1039
F-test no fixed effects	24.31***	26.17***	19.42***	3.22***	2.30***

by around 2.2 basis points. For the average trade size value of around NOK 200000, this means increased execution cost of NOK 44. Assuming a daily number of orders in the market of around 10000, this corresponds to more than NOK 110 millions annually. Outside owner concentration is also positively related to adverse selection cost measured by the GH-method; a 1 percent increase in concentration leads to an increase in the variable proportional costs of 8 basis points.

There is also a significant positive relationship between the spread measures and the holdings of the primary insiders. The coefficient estimate here is smaller though, and the result is not robust over sub-periods. Depth does not seem to be related to either owner concentration nor insider holdings.

5.4 Owner types

So far, we have found evidence that owner concentration matters for liquidity. Table 5.9 shows the results from an estimation where we try to determine whether the type of the largest owner also matters. In addition, we include variables for the number of owners of each owner type. Table A.9 shows the results from estimating this model over each year in the sample period.

As expected, the coefficients of the largest owners are all positive for the spread measures. The most “costly” largest owner seems to be a private, non financial company. When we re-estimate the model and evaluate the other owner types relative to this group, they are all signif-

Table 5.9: Market liquidity and owner types

The table reports results from estimating a panel regression model with one-way fixed effects (least squares dummy variable estimation) for five measures of liquidity as the dependent variable: log(relative weighted spread), log(relative effective spread), log(weighted depth), adverse selection costs according to the GH-method (variable proportional costs), and the adverse selection component of the spread according to the GKN-method. The independent variables are the aggregate holdings of the primary insiders, the holding of the largest owner split on owner type, the logarithm of the number of owners of each owner type, the logarithm of the market capitalization value, the average closing price, the standard deviation of daily returns, the average daily number of trades, and dummies for the fixed effect of each company. For each model, we report the estimated coefficients (except the coefficients for the fixed effect dummies), R-squared, and the F-test for no fixed effects. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, and * defines significance at the 10 percent level. The sample includes 28 time series observations covering 93 companies. To check the residuals of the panel regression models, we estimate a standard OLS regression model for each company and test, for each company, whether the residuals are autocorrelated. We then re-estimate the panel regression models without the companies with autocorrelated residuals. At the 1 percent level (5 percent level), the number of companies with autocorrelated residuals vary from 4-12 companies (15-22 companies). Removing these companies does not seriously change our results.

	Dependent variables				
Independent variables	Rel weighted spread	Rel eff spread	Depth	GH comp.	GKN comp
Primary Insiders	0.1907**	0.1194	−0.0829	−0.0002	0.0002
Largest owner, state	0.7328***	0.4763***	−0.5073*	0.0007*	−0.0100
Largest owner, institutional	0.1556	0.2068	−0.0849	−0.0004	−0.0009
Largest owner, non-financial	0.9647***	1.1285***	−0.7727**	−0.0002	0.0098
Largest owner, individual	0.5580**	0.5288**	0.4743	0.0030***	0.0136
Largest owner, foreign	0.2551*	0.4170***	−0.1636	0.0001	0.0078
No of state owners	−0.0149	−0.0017	0.0007	0.0000	0.0014
No of institutional owners	−0.0202	−0.0863*	−0.2653***	−0.0002**	−0.0020
No of non-financial owners	−0.2059***	−0.3072***	0.1014	0.0002**	−0.0028
No of individual owners	−0.1216***	−0.0493	−0.1833**	−0.0001	0.0035
No of foreign owners	−0.0152	0.0313	0.1369**	−0.0001	−0.0013
Market cap	−0.1574***	−0.1440***	0.0199***	−0.0001**	0.0013
Price	−0.0015***	−0.0014***	−0.0048***	0.0000	0.0000**
Return volatility	2.7464***	2.4089***	−1.6021***	0.0024***	−0.0110
Trades per day	−0.0012***	−0.0009***	0.0017	0.0000*	0.0000
R square	0.8615	0.8722	0.7770	0.2982	0.1047
F-test no fixed effects	18.86***	22.41***	15.37***	3.57***	1.91***

icantly negative. When the largest owner is an institutional investor, the effect on liquidity is not significant. Table 5.4 in section 4 showed that the average largest holding is much lower for this group. Hence, the reason for the lack of significance is most likely that these firms are more widely held. Large individual owners have a significant impact on the GH-measure of adverse selection.

A potential problem with the number of owners variables is that they are all highly correlated with firm size as well as with each other. To check the robustness of our results, we re-estimate the model without the firm size variable. Removing firm size does not seriously change our results. Potential problems with the high mutual correlations among the variables still remain though. The reason for keeping them all in the model is to include the total composition of the ownership structure. As expected, the number of owners has a negative effect on the spread. The numbers of institutional and individual owners have a negative effect on the market depth, while the depth is positively related to the number of foreign investors. These results are not very easy to interpret. The number of institutional investors in a company seems to reduce information costs, while the number of non-financial company owners has the opposite effect. This result may be explained by the tendency for institutional investors to invest less in smaller firms.

The percentage aggregate holdings of the owner types are also problematic due to their mutual correlation structure. We therefore study the effects on liquidity of the total holdings of different owner groups separately. Table 5.10 presents the results from estimating five regression models where we include the aggregate holding of each owner group one at a time. The other independent variables in the models are the free float variable defined in section 5.2, insider holdings, and the four control variables.

Institutional ownership does not significantly affect any of the liquidity measures. Hence, neither of the two opposite hypothesized effects on liquidity from this type of ownership seem to dominate the other. State ownership is negatively related to market depth, and positively related to information costs. These results are a bit surprising, given that the state generally invests in the largest companies. Non-financial company ownership has a positive effect on the spread, and a negative effect on information costs, while individual ownership has a (weakly significant) negative effect on the relative quoted spread and a significant positive effect on information costs. It is hard to come up with very good explanations for these patterns. Finally, the aggregate holding of foreign ownership is negatively related to spreads and positively related to depth in line with the theoretical prediction. Table A.10 in the data appendix shows that the relationships between owner type holdings and liquidity are not very robust over sub-periods.

Table 5.11 presents results from a model where we have included the absolute value of the change in the number of owners as explanatory variables in addition to owner concentration, proxied by the five largest owners, and primary insider holdings. The absolute change in the number of owners should capture turnover in ownership structure. As expected, this variable has a negative impact on the spread and a positive impact on the depth. The coefficients are highly significant for state ownership and institutional ownership. A possible explanation for the significant effect from state ownership is that it reflects the activity of the public pension funds.

To sum up, a large non-financial company owner has a more negative effect on liquidity then a large owner of any other type. Neither a dominating institutional owner nor a high aggregated institutional ownership has any significant effect on liquidity. The number of owners has a positive effect on liquidity measured by the spread, while the effects on market depth are mixed and hard to interpret. Ownership turnover affects both spread and depth in the expected way.

5.5 Granger causality

A potential problem with our regression approach is that it does not account for the possibility that the causality may run the other way around, i.e., liquidity may affect ownership. In fact it is hypothesized that some investors, most notably foreign investors and institutional investors, tend to purchase stocks with low spread and high depth. Another potential problem is that we do not control for the possibility that ownership structure and liquidity are simultaneously determined by the same variables.

Table 5.10: Market liquidity and aggregate holdings of owner groups

The table reports results from estimating five panel regression model with one-way fixed effects (least squares dummy variable estimation) for five measures of liquidity as the dependent variable: log(relative weighted spread), log(relative effective spread), log(weighted depth), adverse selection costs according to the GH-method (variable proportional costs), and the adverse selection component of the spread according to the GKN-method. The independent variables are the total holding which is *not* owned by the five largest owners ("free float"), the aggregate holdings of the primary insiders, the aggregate holding of a particular owner group, the logarithm of the market capitalization value, the average closing price, the standard deviation of daily returns, the average daily number of trades, and and dummies for the fixed effect of each company. For each model, we report the estimated coefficients (except the coefficients for the fixed effect dummies). *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, and * denotes significance at the 10 percent level. The sample includes 28 time series observations covering 93 companies. To check the residuals of the panel regression models, we estimate a standard OLS regression model for each company and test, for each company, whether the residuals are autocorrelated. We then re-estimate the panel regression models without the companies with autocorrelated residuals. At the 1 percent level (5 percent level), the number of companies with autocorrelated residuals vary from 1-9 companies (4-19 companies). Removing these companies makes the foreign ownership variable no longer significant in the relative weighted spread model and more significant in the relative effective spread model. For the GKN information component model with individual ownership, the free float variable becomes highly significant when we remove the companies with autocorrelated residuals. For the other models, removing these companies does not seriously change our results.

	Dependent variables				
Independent variables	Rel weighted spread	Rel eff spread	Depth	GH comp.	GKN comp
Free float	−1.1177***	−1.3902***	−0.3254*	−0.0008***	−0.0092
Primary Insiders	0.2677***	0.2034**	−0.1502	−0.0001	0.0013
State ownership	0.2333	−0.2137	−1.3654***	0.0008**	−0.0020
Market cap	−0.2396***	−0.2138***	−0.0465	−0.0003***	0.0011
Price	−0.0008***	−0.0009***	−0.0046***	0.0000	0.0000***
Return volatility	2.9033***	2.5957***	−1.3415***	0.0024***	−0.0131
Trades per day	−0.0018***	−0.0015***	0.0015***	0.0000	0.0000
Free float	−1.1494***	−1.2800***	−0.1039	−0.0009***	−0.0054
Primary Insiders	0.2701***	0.2098**	−0.1600	−0.0001	0.0016
Institutional ownership	0.0344	−0.2190	−0.2850	0.0000	−0.0085
Market cap	−0.2327***	−0.2233***	−0.0881***	−0.0002***	0.0009
Price	−0.0008***	−0.0008***	−0.0044***	0.0000	0.0000***
Return volatility	2.9269***	2.5477***	−1.4912***	0.0024***	−0.0145
Trades per day	−0.0018***	−0.0015***	0.0016***	0.0000*	0.0000
Free float	−1.2268***	−1.4841***	−0.1982	−0.0008***	−0.0091
Primary Insiders	0.2168***	0.1350	−0.1557	0.0000	0.0012
Non-financial ownership	0.5068***	0.5997***	−0.1591	−0.0005**	−0.0001
Market cap	−0.2321***	−0.2181***	−0.0835***	−0.0002***	0.0010
Price	−0.0008***	−0.0008***	−0.0044***	0.0000	0.0000***
Return volatility	2.9818***	2.6480***	−1.4702***	0.0024***	−0.0133
Trades per day	−0.0018***	−0.0014***	0.0016***	0.0000*	0.0000
Free float	−1.1329***	−1.3736***	−0.2307	−0.0009***	−0.0092
Primary Insiders	0.2757***	0.2029**	−0.1823	−0.0001	0.0010
Individual ownership	−0.2953*	−0.2169	0.6719**	0.0011***	0.0199*
Market cap	−0.2460***	−0.2288***	−0.0543	−0.0002***	0.0019
Price	−0.0008***	−0.0008***	−0.0044***	0.0000	0.0000*
Return volatility	2.9174***	2.5753***	−1.4407***	0.0024***	−0.0129
Trades per day	−0.0018***	−0.0015***	0.0016***	0.0000*	0.0000
Free float	−1.3202***	−1.4727***	0.0214	−0.0010***	−0.0101
Primary Insiders	0.2202***	0.1727**	−0.1041	−0.0001	0.0010
Foreign ownership	−0.2909***	−0.1534*	0.3893***	−0.0002	−0.0015
Market cap	−0.2234***	−0.2143***	−0.0965***	−0.0002***	0.0011
Price	−0.0008***	−0.0008***	−0.0044***	0.0000	0.0000***
Return volatility	2.9660***	2.6019***	−1.5101***	0.0025***	−0.0130
Trades per day	−0.0018***	−0.0015***	0.0015***	0.0000	0.0000

Table 5.11: Market liquidity and ownership turnover

The table reports results from estimating a panel regression model with one-way fixed effects (least squares dummy variable estimation) for five measures of liquidity as the dependent variable: log(relative weighted spread), log(relative effective spread), log(weighted depth), adverse selection costs according to the GH-method (variable proportional costs), and the adverse selection component of the spread according to the GKN-method. The independent variables are the aggregate holdings of the five largest owners, the aggregate holdings of the primary insiders, the absolute value of the change in the number of owners for each owner type, the logarithm of the market capitalization value, the average closing price, the standard deviation of daily returns, the average daily number of trades, and dummies for the fixed effect of each company. For each model, we report the estimated coefficients (except the coefficients for the fixed effect dummies), R-squared, and the F-test for no fixed effects. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, and * denotes significance at the 10 percent level. The sample includes 28 time series observations covering 93 companies. To check the residuals of the panel regression models, we estimate a standard OLS regression model for each company and test, for each company, whether the residuals are autocorrelated. We then re-estimate the panel regression models without the companies with autocorrelated residuals. At the 1 percent level (5 percent level), the number of companies with autocorrelated residuals vary from 0-8 companies (2-9 companies). For the effective spread model, removing these companies makes all variables significant. For the GKN information component model, removing these companies makes the five largest variable significant. For the other models, removing these companies does not seriously change our results.

	Dependent variables				
Independent variables	Rel weighted spread	Rel eff spread	Depth	GH comp.	GNK comp
Five largest	1.2452***	1.4885***	−0.1258	0.0010***	0.0098
Primary insiders	0.2819***	0.2139**	−0.2008	−0.0001	0.0014
$\lvert\Delta\rvert$ no of owners, state	−0.0122***	−0.0172***	0.0232***	0.0000	−0.0002
$\lvert\Delta\rvert$ no of owners, institutional	−0.0034***	−0.0027**	0.0098***	0.0000*	0.0000
$\lvert\Delta\rvert$ no of owners, non-financial	−0.0002	−0.0005	0.0008**	0.0000	0.0000
$\lvert\Delta\rvert$ no of owners, individual	0.0000*	0.0000	0.0000	0.0000	0.0000**
$\lvert\Delta\rvert$ no of owners, foreign	−0.0002**	−0.0003	0.0001	0.0000	0.0000
Market cap	−0.2354***	−0.2215***	−0.0722**	−0.0002***	0.0010
Price	−0.0007***	−0.0007***	−0.0047***	0.0000	0.0000***
Return volatility	2.873***	2.5162***	−1.2980***	0.0024***	−0.0134
Trades per day	−0.0017***	−0.0014***	0.0012***	0.0000	0.0000
R square	0.8523	0.8610	0.7845	0.2810	0.1044
F-test no fixed effects	22.35***	24.63***	17.92***	3.20***	2.13***

In this sub-section, we investigate whether ownership variables can forecast liquidity and vice versa by estimating some simple Granger tests of the form $x_{i,t} = f(y_{i,t-1}, x_{i,t-1}, ..)$ and $y_{i,t} = f(x_{i,t-1}, y_{i,t-1}, ..)$, where we use 12 lags to determine the autocorrelation structure of the dependent variables, and include all the significant lags.[34] Table 5.12, Table 5.13, and Table 5.14 show the results from these tests for, respectively, liquidity versus insider holdings and owner concentration, liquidity versus the aggregate holdings of the owner groups, and liquidity versus the number of owners in each group.

There is no significant Granger causality relation between liquidity and insider holdings or between the spread measures and owner concentration measured by the aggregate holdings of the five largest owners. For the owner concentration, this result might seem surprising given the strong relationship found between the variables in the estimation without lags.[35] No significant Granger causality suggests that the variables are determined simultaneously, possibly by the same variables. The aggregate holding of the five largest owners is found to forecast market depth, and vice versa.

Turning to the *type* of the largest owner, the Granger causality is found to go both ways between spreads and owner concentration if the owner is a non-financial company. If the largest owner is the state, the Granger causality is found to go one way only from ownership to liquidity. If the largest owner is foreign, there is a one-way Granger causality from spread to ownership, while the Granger causality against depth goes both ways.

The estimated Granger causality relations between the aggregate holding of different owner types and liquidity show that; (i) the aggregate holding of non-financial companies forecasts the relative spread, (ii) the aggregate holding of foreign investors forecasts both spread measures, (iii) the spread measures forecast the aggregate holding of individual ownership, and (iv) the market depth forecasts the aggregate holding of the state, the institutional investors, and the foreign owners. Finally, when we focus on the number of owners, there are in general one-way Granger causality relations from ownership to spread, and two-ways Granger causality relations between ownership and depth.

[34] Note that we do not address the general causality problem with these tests. A liquidity measure may affect an ownership variable even though it cannot be used to forecast it.

[35] Both variables are highly autocorrelated, and when we estimate a causality without removing the autocorrelation there is a strong two ways relationship.

Table 5.12: Granger causality: liquidity, insider holdings, and owner concentration

The table reports results from estimating a panel regression model of the form $x_t = f(y_{t-1}, x_{t-1}, ..)$ and $y_t = f(x_{t-1}, y_{t-1}, ..)$ with one-way fixed effects (least squares dummy variable estimation). 12 lags are used to determine the autocorrelation structure of the dependent variables, and the significant lags are included in the model. The model is estimated for three liquidity measures: log(relative weighted spread), log(relative effective spread), and log(weighted depth), and seven ownership variables: aggregate holdings of primary insiders, aggregate holdings of the five largest owners, and five variables where we split the holding of the largest owner into one variable for each owner type. The independent variables also include dummies for the fixed effects of each company. For each model, we report the estimated coefficients (except the coefficients for the fixed effect dummies). *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, and * denotes significance at the 10 percent level. The sample includes 17 time series observations covering 87 companies .

	Insider holdings and owner concentration (t-1)						
Liquidity (t)	Prim insid	5 larg	Larg, state	Larg, institut	Larg, non-fin	Larg, individ	Larg, foreign
% spread	0.1907	0.1574	1.0993*	−0.3662	0.8383***	0.8961	−0.2298
% eff spread	0.0352	0.3517	1.1526**	−0.5176	0.7226***	1.1846	−0.2223
Depth	−0.1116	−1.3275***	0.5327	−0.5947	−0.3815	−0.3899	−0.6028**

	Liquidity (t-1)		
Ownership structure (t)	% weighted spread	% eff spread	Weighted depth
Primary insiders	0.0040	0.0038	−0.0024
Five largest	−0.0016	−0.0004	0.0114***
Largest, state	−0.0016	−0.0022	−0.0042
Largest, institut	0.0024**	0.0013	−0.0007
Largest, non-financial	0.0040*	0.0051**	0.0032**
Largest, individ	0.0060*	0.0076**	0.0008
Largest, foreign	−0.0116**	−0.0120*	0.01286***

Table 5.13: Granger causality: liquidity and aggregate holdings of owner types

The table reports results from estimating a panel regression model of the form $x_t = f(y_{t-1}, x_{t-1}, ..)$ and $y_t = f(x_{t-1}, y_{t-1}, ..)$ with one-way fixed effects (least squares dummy variable estimation). 12 lags are used to determine the autocorrelation structure of the dependent variables, and the significant lags are included in the model. The model is estimated for three liquidity measures: log(relative weighted spread), log(relative effective spread), and log(weighted depth), and five ownership variables: the aggregate holdings of each owner group. The independent variables also include dummies for the fixed effects of each company. For each model, we report the estimated coefficients (except the coefficients for the fixed effect dummies). *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, and * denotes significance at the 10 percent level. The sample includes 17 time series observations covering 87 companies.

	Aggregate holding (t-1)				
Liquidity (t)	State	Institutional	Non-financial	Individual	Foreign
% spread	0.0598	0.0896	0.6974***	0.6427	−0.4226**
% eff spread	−0.0071	−0.1877	0.7040	0.4290	−0.2689*
Depth	−0.2007	−0.1250	0.5100	−0.1022	−0.1443

	Liquidity (t-1)		
Aggregate holding (t)	% weighted spread	% eff spread	Weighted depth
State	−0.0012	−0.0020	−0.0044***
Institutional	0.0041	0.0032	−0.0089***
Non-financial	0.0035	0.0035	−0.0009
Individual	0.0090***	0.0117***	−0.0001
Foreign	−0.0082*	−0.0054	0.0172***

Table 5.14: Granger causality: liquidity and the number of owners

The table reports results from estimating a panel regression model of the form $x_t = f(y_{t-1}, x_{t-1}, ..)$ and $y_t = f(x_{t-1}, y_{t-1}, ..)$ with one-way fixed effects (least squares dummy variable estimation). 12 lags are used to determine the autocorrelation structure of the dependent variables, and the significant lags are included in the model. The model is estimated for three liquidity measures: log(relative weighted spread), log(relative effective spread), and log(weighted depth), and five ownership variables: the number of owners in each group of owners. The independent variables also include dummies for the fixed effects of each company. For each model, we report the estimated coefficients (except the coefficients for the fixed effect dummies). *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, and * denotes significance at the 10 percent level. The sample includes 17 time series observations covering 87 companies.

	Number of owners (t-1)				
Liquidity (t)	State	Institutional	Non-financial	Individual	Foreign
% spread	−0.1051***	−0.2984***	−0.1064	−0.0479	−0.2395***
% eff spread	−0.1028***	−0.3923***	−0.1903***	−0.1049*	−0.2306***
Depth	0.2360***	0.1852**	0.3706***	0.3478***	0.4045***

	Liquidity (t-1)		
No of owners (t)	% weighted spread	% eff spread	Weighted depth
State	−0.0413*	−0.0289	−0.0370**
Institutional	0.0142	0.0032	−0.0595***
Non-financial	−0.0040	−0.0242	−0.0507***
Individual	0.0047	−0.0083	−0.0381***
Foreign	0.0070	0.0031	−0.0359***

6 Concluding remarks

In this paper, we analyze the relationship between ownership structure and liquidity in the Norwegian stock market using a panel regression approach. Our main findings are summarized below.

- Both owner concentration, measured by the aggregate holdings of the five largest owners, and insider holdings, measured by the aggregate holdings of primary insiders, are found to increase the spread. These results are in accordance with theoretical predictions. The result for owner concentration is stronger than the result for insider holdings. Moreover, while owner concentration has a negative impact on GH-estimates of adverse selection as well, this is not the case for insider holdings. Our results concerning the effects of insider ownership seem weaker than reported in some other empirical studies.

- The negative effect on spreads from the holding of the largest owner is found to be strongest when the largest owner is a non-financial company, and not significant if the largest owner is an institutional investor.

- In contrast to some other studies, we do not find any special effects on liquidity from the ownership of institutional investors. We do find support for the hypothesis that foreign investors concentrate their holdings in liquid stocks.

- Ownership variables which are found to affect the spread do not in general jointly affect the depth in the predicted way.

- In general, the assumption of a one-way Granger causality from ownership to liquidity is dubious. There are no significant Granger causality relations between owner concentration measured by the five largest owners and the spread measures, suggesting that the these variables are determined simultaneously, possibly by the same variables. We do find a significant one-way Granger causality from the aggregate holdings of non financial companies to the relative spread. Some support is provided for the hypothesis that foreign investors buy stocks with low spreads and high depth, but the Granger causality goes both ways.

The data sample underlying this study is more comprehensive than the data samples used in comparable studies, and the results therefore make a robust contribution to the existing results. The study also represents a natural starting point for some further work.

Firstly, the documented effects on liquidity from ownership structure are interesting inputs to the study of ownership structure and performance. Bøhren and Ødegaard (2003a) find that the negative effects from owner concentration outweight the benefits of monitoring, and that the holdings of insiders are value increasing. Linking these results with the results in this study, suggests that the effects on performance corroborate the liquidity effects; owner concentration is more costly in the form of reduced liquidity than insider holdings. A natural topic for further work on this topic is the link between liquidity, expected returns, and measures of performance.

Secondly, widely used methods for estimating adverse selection costs are based on some form of dealer intermediation, and the relevance of these methods for a fully automated limit order market is unclear. The significant difference in explanatory power for the two estimates of adverse selection chosen in this study suggests that the choice of method makes a difference. Moreover, Næs and Skjeltorp (2003) find evidence that the average trade size explains little of the price changes in this market.[36] Since asymmetric information is of central interest in all studies of liquidity, and given the increasing role of limit order trading arrangements, the validity of the empirical method used for decomposing the spread is an important issue. One possible starting point is to investigate the relationship between ownership and estimates of the probability of informed trading according to Easley et al. (1996). Their method is not based on optimizing behavior of competitive dealers. Rather it estimates the probability of information events by observing the total buys and the sells during each trading day, and combines them with assumed arrival rates of buyer and sellers.

[36] This result applies in a competitive dealer environment as well, according to Jones et al. (1994).

Appendix A

Data appendix

Table A.1: The corporate governance structure in Norway, 1989-1997

The table reports statistics of aggregate holdings of different owner types as well as the holdings of the mean owner, the largest owner and the five largest owners of all companies listed on the OSE in the period 1989-1997. All numbers are taken from the data appendix in Bøhren and Ødegaard (2000). All data except insider holdings are from the VPS. Insider holdings are estimated on the basis of statements given by the insiders to the OSE. The table shows the equally weighted average, the standard deviation, the value weighted average, the median observation, and the number of firms. The value weighted averages are weighted based on the value of the firm's equity. All number are in percent. Averages over years are based on values which are transformed into 1997 kroner.

Ownership structure 1989-1997	EW mean	Std	VW mean	Median	n
Owner types:					
State	5.00	14.00	18.00	0.00	1189
Institutional	17.00	15.00	18.00	15.00	1189
Non-financial	38.00	24.00	24.00	36.00	1189
Individual	18.00	16.00	10.00	12.00	1189
Foreign	22.00	22.00	31.00	15.00	1189
Insiders	14.00	25.00	7.00	1.00	1197
Owner concentration:					
Mean owner	0.15	0.32	0.06	0.08	1189
Largest owner	28.00	19.00	29.00	22.00	1189
Five largest owners	55.00	19.00	53.00	54.00	1189

Table A.2: Monthly returns for liquidity based portfolios - Sub-periods

The table shows monthly returns and the number of securities for four five years periods and one three years period between 1980-2002. Portfolios are grouped at the beginning of each year, using the average relative bid-ask spread in the previous year as the criterion for grouping. The sample include all listed securities at the OSE which comply with the following three filtering criteria: (i) the stock price is above NOK 10, (ii) the total value outstanding of the company is at least NOK 1 million, and (iii) the security is traded at least 20 days during one year. The filtering criteria imply that, on average, a year contains 121 companies. All portfolios are equally weighted.

Panel A	Returns					No of securities		
1980-1984	mean	std	min	median	max	avg	min	max
1 (smallest)	1.64	5.41	-14.14	0.79	17.38	16.1	10	23
2	1.61	5.86	-16.01	1.47	13.2	15.8	10	23
3	3.41	6.16	-9.87	2.25	18.41	16.2	11	23
4	4.41	7.01	-13.08	3.28	21.87	15.8	10	23
5	5.24	9.67	-15.92	2.56	35.04	15.8	10	23

Panel B	Returns					No of securities		
1985-1989	mean	std	min	median	max	avg	min	max
1 (smallest)	1.37	7.09	-24.9	2.45	18.65	27.4	23	32
2	0.95	6.41	-20.04	2.3	18.8	26.8	22	31
3	1.15	5.94	-17.25	2.05	21.45	26.7	21	31
4	1.53	5.13	-11.61	1.76	19.19	26.5	23	31
5	3.08	4.96	-6.66	2.64	16.7	26.5	20	31

Panel C	Returns					No of securities		
1990-1994	mean	std	min	median	max	avg	min	max
1 (smallest)	0.44	7.75	-20.74	1.09	14.61	27.4	24	31
2	0.26	7.56	-19.41	1.24	14.65	26.5	22	31
3	0.65	7.73	-19.29	-0.24	16.78	26.1	22	30
4	1.87	7.37	-12.48	0.75	21.25	25.5	21	31
5	2.61	8.27	-9.51	1.52	30.85	25.7	19	30

Panel D	Returns					No of securities		
1995-1999	mean	std	min	median	max	avg	min	max
1 (smallest)	1.42	5.85	-23.47	1.42	14.35	36.9	30	45
2	1.96	5.56	-20.23	2.02	15.45	35.6	29	44
3	1.56	5.63	-21.24	1.74	14.12	35.2	26	45
4	1.49	5.68	-15.51	1.69	21.85	35.6	26	44
5	2.64	5.07	-10.92	1.57	22.66	35.1	28	44

Panel E	Returns					No of securities		
2000-2002	mean	std	min	median	max	avg	min	max
1 (smallest)	-1.78	8.3	-20.81	-0.71	14.26	35.6	31	42
2	-0.61	7.9	-22.45	-1.37	15.71	34	29	39
3	-0.74	5.57	-13.04	-0.11	10.54	33.8	29	41
4	-0.01	5.36	-11.29	-0.1	11.68	34.7	30	41
5	0.64	5.63	-11.96	0.28	13.78	34.8	30	41

Table A.3: Liquidity over time and across size portfolios

	1999		2000		2001	
	mean	**median**	**mean**	**median**	**mean**	**median**
Largest companies						
Market cap (bill NOK)	15.96	16.04	18.80	18.60	19.98	20.01
Price	139.29	139.03	147.46	147.40	137.92	138.74
Quoted spread	1.49	1.18	1.33	1.07	1.12	0.88
Effective spread	0.98	0.85	0.97	0.81	0.81	0.71
Relative spread (time weighted)	1.11	1.04	1.12	1.06	0.81	0.78
Adverse selection component	0.016	0.015	0.007	0.007	0.005	0.005
Quoted depth (time weighted)	15201	12685	26341	14659	14644	12042
Avg daily trade size	2474	1592	2091	1191	2067	1266
Avg daily no of trades	85	76	123	108	136	113
Avg daily trade volume, in shares	289981	235788	262003	208508	320374	237936
- in 1000 NOK	27855.463	22979.697	30155.068	24503.628	34203.766	25881.279
N	25		24		24	
Medium large companies						
Market cap (bill NOK)	2.58	3.65	3.65	3.65	3.57	3.57
Price	114.12	114.22	125.29	125.94	91.85	92.71
Quoted spread	2.23	1.58	2.57	2.14	2.01	1.74
Effective spread	1.58	1.37	1.95	1.68	1.65	1.52
Relative spread (time weighted)	1.64	1.50	1.81	1.64	1.85	1.75
Adverse selection component	0.023	0.025	0.000	0.017	0.000	0.006
Quoted depth (time weighted)	6523	5741	5010	4401	6835	5776
Avg daily trade size	1550	1035	1166	766	1528	1005
Avg daily no of trades	36	30	57	49	77	66
Avg daily trade volume, in shares	105636	81639	92437	74280	129176	101573
- in 1000 NOK	4907.489	3740.188	7149.697	5700.717	9977.332	7732.947
N	23		22		22	
Medium small companies						
Market cap (bill NOK)	1.09	1.05	1.69	1.57	1.39	1.38
Price	74.38	74.16	106.63	106.79	102.19	102.18
Quoted spread	1.49	1.26	1.98	1.62	1.88	1.56
Effective spread	1.08	0.98	1.51	1.32	1.50	1.28
Relative spread (time weighted)	1.92	1.81	1.72	1.58	1.65	1.57
Adverse selection component	0.046	0.034	0.020	0.020	0.015	0.023
Quoted depth (time weighted)	7953	7049	5254	4443	6654	5989
Avg daily trade size	2295	1647	1342	894	1313	941
Avg daily no of trades	24	20	38	31	27	22
Avg daily trade volume, in shares	68446	52121	61299	44383	46141	35275
- in 1000 NOK	2684.397	2073.844	4277.765	3164.938	2572333.679	1993471.793
N	24		24		23	
Smallest companies						
Market cap	0.32	0.32	0.58	0.57	0.49	0.48
Price	51.22	51.28	60.47	60.68	46.16	46.05
Quoted spread	1.60	1.35	1.77	1.51	1.79	1.60
Effective spread	1.12	1.00	1.33	1.14	1.36	1.24
Relative spread (time weighted)	3.46	3.25	2.84	2.65	3.17	3.00
Adverse selection component	0.110	0.072	0.058	0.045	-0.286	-0.284
Quoted depth (time weighted)	6059	5453	5974	4877	8057	7075
Avg daily trade size	2094	1539	1746	1218	1921	1418
Avg daily no of trades	14	12	38	31	36	29
Avg daily trade volume, in shares	33581	26935	83689	62094	96967	75594
- in 1000 NOK	1023.883	812.440	2638339.398	1862235.877	1936.782	1418.752
N	23		23		24	

Table A.4: Ownership structure over time and across size portfolios

	1999		2000		2001	
	mean	**median**	**mean**	**median**	**mean**	**median**
Largest companies						
Owner types:						
State owners	0.1380	0.1373	0.1475	0.1506	0.1214	0.1224
Institutional owners	0.2525	0.2545	0.2305	0.2329	0.2276	0.2325
Non-financial owners	0.2402	0.2383	0.2018	0.2014	0.2266	0.2272
Individual owners	0.0686	0.0686	0.0682	0.0691	0.0799	0.0805
Foreign owners	0.3012	0.2977	0.3526	0.3459	0.3449	0.3349
Insider holdings	0.0606	0.0.0672	0.0544	0.0630	0.0359	0.0.0341
Owner concentration:						
Largest owner	0.2069	0.2036	0.2499	0.2422	0.2152	0.2119
Five largest owners	0.4596	0.4564	0.4844	0.4816	0.4309	0.4230
N	25		24		24	
Medium large companies						
Owner types:						
State owners	0.0676	0.0662	0.0845	0.0849	0.1000	0.1002
Institutional owners	0.3030	0.3067	0.2841	0.2870	0.2311	0.2322
Non-financial owners	0.2835	0.2860	0.2915	0.2853	0.3080	0.3089
Individual owners	0.1147	0.1157	0.1144	0.1133	0.1117	0.1111
Foreign owners	0.2326	0.2221	0.2278	0.2313	0.2508	0.2469
Insider holdings	0.1014	0.0943	0.0505	0.0447	0.0691	0.0712
Owner concentration:						
Large owners	0.2160	0.2122	0.2346	0.2315	0.2579	0.2577
Five largest owners	0.4476	0.4418	0.4636	0.4597	0.5156	0.5143
N	23		22		22	
Medium small companies						
Owner types:						
State owners	0.0398	0.0324	0.0302	0.0314	0.0431	0.0434
Institutional owners	0.2921	0.2952	0.2763	0.2794	0.2647	0.2707
Non-financial owners	0.3021	0.3046	0.3069	0.3060	0.2907	0.2898
Individual owners	0.1902	0.1911	0.1999	0.1982	0.2222	0.2224
Foreign owners	0.1764	0.1692	0.1906	0.1834	0.1812	0.1759
Insider holdings	0.0486	0.0496	0.0642	0.0663	0.0696	0.0664
Owner concentration:						
Largest owner	0.1633	0.1576	0.1671	0.1624	0.1671	0.1624
Five largest owners	0.4070	0.4019	0.4056	0.3993	0.3957	0.3915
N	24		24		23	
Smallest companies						
Owner types:						
State owners	0.0046	0.0042	0.0044	0.0042	0.0042	0.0042
Institutional owners	0.1727	0.1757	0.1776	0.1768	0.1934	0.3446
Non-financial owners	0.2938	0.1757	0.3123	0.3097	0.3447	0.3446
Individual owners	0.3561	0.3456	0.3058	0.3071	0.2585	0.2579
Foreign owners	0.1742	0.1710	0.2026	0.2017	0.2010	0.2021
Insider holdings	0.1185	0.1038	0.1415	0.1367	0.1166	0.1153
Owner concentration:						
Largest owner	0.1752	0.1743	0.1910	0.1937	0.1937	0.1932
Five largest owners	0.4250	0.4186	0.4193	0.4215	0.4381	0.4377
N	23		23		24	

Table A.5: Descriptive statistics over industry groups

The table provides some descriptive statistics for the data sample split into A and B shares. The A shares are further split into the FTSE global classification system: RESOR = Resources, BASIC = Basic Industries, GENIN = General Industrials, CYCGD = Cyclical Consumer Goods, NCYCG = Non-Cyclical Consumer Goods, CYSER = Cyclical Services, NCYSR = Non-Cyclical Services, UTILS = Utilities, TOTLF = Financials, and ITECH = Information Technology. All stocks were traded at least 400 out of 597 trading days during the period from February 5 1999 to June 30 2001.

Variable	All A-shares	RESOR	BASIC	GENIN	CYCGD	NCYCG	CYSER	NCYSR	UTILS	TOTLF	ITECH	B-shares
N	88	10	3	8	3	7	19	4	3	15	16	6
Firm size (bill NOK):												
-mean	5.97	12.44	8.44	4.73	1.11	13.84	4.96	5.25	2.45	5.76	2.10	5.71
-median	1.85	12.48	8.29	4.16	1.11	13.65	4.72	4.84	2.23	5.66	2.24	4.69
Price:												
-mean	103.15	101.49	192.44	94.39	46.20	75.63	95.02	171.02	69.71	151.22	61.14	79.20
-median	103.27	101.70	193.67	94.40	46.05	75.62	95.14	170.90	69.78	151.13	61.35	79.47
Trades per day:												
-mean	59	101	45	57	24	70	34	49	49	36	96	19
-median	51	88	40	47	19	61	29	41	43	32	80	16
Trade size:												
-mean	1817	1792	1117	1463	2146	2311	1754	1014	1525	1991	2104	1962
-median	1207	1194	777	1064	1508	1456	1227	706	1113	1202	1365	1362
Return volatility:												
-mean	0.0327	0.0293	0.0296	0.0379	0.0340	0.0350	0.319	0.0414	0.0427	0.0196	0.0412	0.0285
-median	0.0274	0.0279	0.0192	0.0345	0.0291	0.0301	0.0284	0.0354	0.0419	0.0176	0.0352	0.0271
Rel spread (weighted):												
-mean	0.0189	0.0114	0.0107	0.0236	0.0238	0.0190	0.0232	0.0220	0.0337	0.0133	0.0190	0.0210
-median	0.0177	0.0107	0.0101	0.0220	0.0223	0.0181	0.0215	0.0204	0.0313	0.0122	0.0180	0.0192
Effective spread (NOK):												
-mean	1.3090	0.6483	1.4538	1.2459	0.9042	0.9114	1.6092	1.6985	1.2446	1.9847	0.8673	1.1749
-median	1.1431	0.5642	1.1953	1.0347	0.8162	0.8149	1.4160	1.3965	1.1112	1.7772	0.7543	1.0013
Depth:												
-mean	11073	9261	7896	5807	7687	14471	12304	3798	6022	16698	11022	18650
-median	9734	7628	4347	5169	6913	11128	12878	3171	5376	15009	8463	5804
Five largest owners:												
-mean	0.4417	0.4688	0.4235	0.4590	0.3496	0.4273	0.5367	0.5220	0.6178	0.3461	0.3673	0.4190
-median	0.4212	0.4716	0.4172	0.4619	0.3503	0.4164	0.5353	0.5189	0.6408	0.3442	0.3684	0.4354
Prim insiders:												
-mean	0.0771	0.0640	0.0001	0.0872	0.0450	0.0587	0.1664	0.1633	0.0047	0.0033	0.0678	0.0898
-median	0.0.0034	0.0684	0.0001	0.1006	0.0460	0.0288	0.1437	0.1926	0.0065	0.0038	0.0660	0.0387
State owners												
-mean	0.0655	0.0684	0.0763	0.0980	0.0027	0.0734	0.0242	0.0876	0.1833	0.0995	0.0431	0.0689
-median	0.0262	0.0699	0.0742	0.1026	0.0014	0.0719	0.0238	0.0879	0.1992	0.0970	0.0499	0.0330
Institutional owners												
-mean	0.2318	0.2199	0.3624	0.2788	0.2937	0.2500	0.2330	0.2330	0.1695	0.1894	0.2159	0.4364
-median	0.2122	0.2232	0.3642	0.2703	0.2995	0.2501	0.2360	0.2307	0.1683	0.1875	0.2132	0.3833
Non-financial owners												
-mean	0.2846	0.3264	0.2920	0.1981	0.3403	0.2509	0.3833	0.2138	0.2792	0.2448	0.2412	0.1996
-median	0.2692	0.3265	0.2989	0.1938	0.3421	0.2526	0.3782	0.2177	0.2697	0.2507	0.2343	0.1828
Individual owners												
-mean	0.1798	0.1079	0.0862	0.2169	0.1602	0.2486	0.0896	0.0740	0.1802	0.2965	0.2314	0.0783
-median	0.1229	0.1067	0.0848	0.2239	0.1546	0.2447	0.0876	0.0671	0.1856	0.2962	0.2177	0.0675
Foreign Owners												
-mean	0.2401	0.2811	0.1833	0.2087	0.2035	0.1779	0.2709	0.3980	0.1892	0.1706	0.2729	0.2177
-median	0.1946	0.2753	0.1827	0.2105	0.1948	0.1762	0.2767	0.3714	0.1897	0.1651	0.2718	0.2007

Table A.6: Monthly returns for portfolios sorted on effective relative spread - Sub-periods

The table shows characteristics of the return distribution of monthly returns for four equally weighted liquidity portfolios. The companies included in the data sample are all firms with price greater than NOK 10 which are traded on at least 400 days out of the 597 trading days from February 5 1999 to June 30 2001. The portfolios are grouped at the beginning of each half year, using the average relative effective spread in the previous half year as the criterion for grouping. The panels show the return characteristics for portfolios which are held one half year.

Sub-period	Effective spread					Return				
1999.2	mean	std	min	median	max	mean	std	min	median	max
Portfolio 1	0.59	0.18	0.21	0.58	0.82	3.95	3.65	-1.27	2.97	15.26
Portfolio 2	1.07	0.12	0.86	1.07	1.28	4.19	4.71	-3.50	3.04	16.52
Portfolio 3	1.84	0.36	1.28	1.89	2.28	5.72	4.81	-5.35	5.01	14.54
Portfolio 4	3.76	1.63	2.30	3.41	9.66	6.68	9.72	-4.91	3.62	32.13

Sub-period	Effective spread					Return				
2000.1	mean	std	min	median	max	mean	std	min	median	max
Portfolio 1	0.53	0.16	0.14	0.55	0.76	0.44	6.64	-11.76	1.64	12.85
Portfolio 2	1.02	0.18	0.77	0.98	1.35	0.41	5.02	-10.88	0.45	8.24
Portfolio 3	1.78	0.22	1.37	1.77	2.15	1.86	8.08	-11.09	-0.70	17.07
Portfolio 4	3.21	0.98	2.16	2.88	5.45	4.79	9.28	-7.52	2.40	38.62

Sub-period	Effective spread					Return				
2000.2	mean	std	min	median	max	mean	std	min	median	max
Portfolio 1	0.58	0.17	0.19	0.59	0.80	-1.86	6.01	-19.77	-1.63	7.60
Portfolio 2	1.13	0.16	0.84	1.13	1.39	-0.66	5.47	-16.38	0.10	7.95
Portfolio 3	1.68	0.21	1.41	1.62	2.03	-0.30	5.35	-12.35	-0.71	9.23
Portfolio 4	2.80	0.55	2.06	2.75	4.17	-1.20	7.41	-17.34	0.49	13.76

Sub-period	Effective spread					Return				
2000.2	mean	std	min	median	max	mean	std	min	median	max
Portfolio 1	0.52	0.15	0.19	0.49	0.72	-1.18	6.16	-19.92	-0.58	8.86
Portfolio 2	1.05	0.21	0.74	1.08	1.39	-1.02	6.87	-25.98	0.11	7.20
Portfolio 3	1.69	0.21	1.40	1.62	2.04	2.00	4.88	-7.43	1.62	14.65
Portfolio 4	3.46	1.35	2.14	3.08	6.73	-1.79	6.40	-14.40	-0.06	6.93

Sub-period	Effective spread					Return				
2001.2	mean	std	min	median	max	mean	std	min	median	max
Portfolio 1	0.51	0.17	0.17	0.52	0.75	-1.53	5.77	-17.05	-0.54	9.70
Portfolio 2	1.02	0.20	0.77	0.98	1.39	-3.08	5.99	-18.64	-1.59	3.67
Portfolio 3	1.71	0.25	1.39	1.73	2.22	0.79	3.00	-4.67	0.62	7.49
Portfolio 4	3.91	1.33	2.26	3.43	6.45	-1.18	4.34	-10.42	-0.14	7.36

Table A.7: A reference model for market liquidity - Sub-periods

The table reports results from estimating the relation between liquidity and four control variables using a panel regression model with one-way fixed effects (least squares dummy variable estimation). The model is estimated using the available observations for each of the years 1999-2001. The control variables are the logarithm of the market capitalization value, the average closing price, the standard deviation of daily returns, and the average daily number of trades. In addition, the model includes dummies for the fixed effect of each company. The model is estimated for five liquidity measures: log(relative weighted spread), log(relative effective spread), log(weighted depth), adverse selection costs according to the GH-method (variable proportional costs), and the adverse selection component of the spread according to the GKN-method. For each model, we report the estimated coefficients (except the coefficients for the fixed effect dummies), R-squared, and the F-test for no fixed effects. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, and * denotes significance at the 10 percent level.

	Dependent variables				
Independent variables	Rel weighted spread	Rel eff spread	Weighted depth	GH info comp	GKN info comp
1999					
Market cap	−0.1542***	−0.1576***	−0.2048***	−0.0003***	−0.0045
Price	−0.0017***	−0.0014**	−0.0038***	0.0000	0.0000
Return volatility	0.7754***	0.6348**	−0.6770	0.0026***	−0.0230
Trades per day	−0.0019***	−0.0018***	0.0024***	0.0000	0.0000
R square	0.9069	0.9046	0.8718	0.4012	0.1525
F-test no fixed effects	20.09***	17.97***	10.62***	3.03***	1.44***
2000					
Market cap	−0.0955**	−0.1057**	0.0817	−0.0001	−0.0010
Price	−0.0017***	−0.0015***	−0.0029***	0.0000	0.0000
Return volatility	3.8889***	3.8297***	−3.0266***	0.0031***	−0.0171
Trades per day	−0.0016***	−0.0016***	0.0021***	0.0000**	0.0000
R square	0.8779	0.8769	0.7923	0.3920	0.1729
F-test no fixed effects	16.94***	16.70***	9.59***	3.79***	1.86***
2001					
Market cap	0.0480	−0.0737	−0.0492	−0.0001	0.0080
Price	−0.0021**	−0.0030***	−0.0049***	0.0000	−0.0001
Return volatility	6.0876***	6.0211***	−0.2987	0.0009	0.0438
Trades per day	−0.0014***	−0.0010***	0.0011***	0.0000	0.0000
R square	0.8686	0.9262	0.8366	0.3992	0.2414
F-test no fixed effects	6.74***	13.28***	8.73***	2.18***	1.32*

Table A.8: Market liquidity, owner concentration and holdings of primary insiders - Sub-periods

The table reports results from estimating the relation between liquidity and ownership structure using a panel regression model with one-way fixed effects (least squares dummy variable estimation). The model is estimated using the available observations for each of the years 1999-2001. The ownership structure variables are the aggregate holdings of the five largest owners, and the aggregate holdings or the primary insiders. The control variables are the logarithm of the market capitalization value, the average closing price, the standard deviation of daily returns, and the average daily number of trades. In addition, the model includes dummies for the fixed effect of each company. The model is estimated for five dependent variables: log(relative weighted spread), log(relative effective spread), log(weighted depth), adverse selection costs according to the GH-method (variable proportional costs), and the adverse selection component of the spread according to the GKN-method. For each model, we report the estimated coefficients (expect the coefficients for the fixed effect dummies), R-squared, and the F-test for no fixed effects. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, and * denotes significance at the 10 percent level.

	Dependent variables				
Independent variables	Rel weighted spread	Rel eff spread	Depth	GH comp.	GKN comp
1999					
Five largest	0.8965***	0.8398***	0.8881**	0.0024***	−0.0333*
Primary Insiders	−0.1122	−0.0307	0.2453	0.0001	−0.0113
Market cap	−0.1889***	−0.1900***	−0.2390***	−0.0004***	−0.0032
Price	−0.0015***	−0.0013**	−0.0036***	0.0000	0.0000
Return volatility	0.5665*	0.4554	−0.8058*	0.0021***	−0.0186
Trades per day	−0.0019***	−0.0018***	0.0024***	0.0000	0.0000
R square	0.9092	0.9062	0.8730	0.4147	0.1571
F-test no fixed effects	16.04***	14.15***	9.48***	2.78***	1.46***
2000					
Five largest	1.0595***	1.3430***	0.7898**	−0.0002	0.0303**
Primary Insiders	−0.0502	−0.0472	0.1062	0.0001	−0.0076
Market cap	−0.1020**	−0.1134***	0.0819	−0.0001	−0.0014
Price	−0.0019***	−0.0017***	−0.0031***	0.0000	0.0000
Return volatility	3.7596***	3.6668***	−3.1159***	0.0032***	−0.0211
Trades per day	−0.0015***	−0.0013***	0.0022***	0.0000**	0.0000
R square	0.8816	0.8826	0.7935	0.3922	0.1796
F-test no fixed effects	15.45***	15.49***	9.61***	3.51***	1.87***
2001					
Five largest	−0.0671	0.0970	0.8005*	0.0001	−0.0031
Primary Insiders	0.4215	0.7562	1.1534	−0.0018*	0.0170
Market cap	0.0558	−0.0696	−0.0740	−0.0001	0.0076
Price	−0.0023**	−0.0033***	−0.0048***	0.0000	0.0000
Return volatility	5.9352***	5.7689***	−0.5071	0.0008	0.0577
Trades per day	−0.0014***	−0.0009***	0.0011***	0.0000	0.0000
R square	0.8683	0.9274	0.8402	0.4060	0.3038
F-test no fixed effects	5.93***	12.00***	8.87***	2.16***	1.75***

Table A.9: Market liquidity and owner types - Sub-periods

The table reports results from estimating the relationship between liquidity and ownership structure using a panel regression model with one-way fixed effects (least squares dummy variable estimation). The model is estimated using the available observations for each of the years 1999-2001. For each model, we report the estimated coefficients (except the coefficients for the fixed effect dummies). *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, and * defines significance at the 10 percent level.

	Dependent variables				
Independent variables	Rel weighted spread	Rel eff spread	Depth	GH comp.	GKN comp
1999					
Primary Insiders	−0.0201	0.0510	0.3210*	0.0002	−0.0152
Largest owner, state	0.6675***	0.4677*	0.7848**	0.0024***	−0.0189
Largest owner, institutional	−0.4955	−0.6184	2.4091**	0.0010	−0.0263
Largest owner, non-financial	0.9856**	1.3274***	1.4814***	0.0003	−0.0556
Largest owner, individual	0.6602	0.4367	2.2566**	0.0108***	−0.0461
Largest owner, foreign	−0.2160	−0.2438	1.4087	0.0018**	−0.0569*
No of state owners	−0.0211	−0.0473	−0.0599	−0.0001	0.0053*
No of institutional owners	−0.0159	−0.0854	0.1224	−0.0005**	0.0110
No of non-financial owners	−0.2136**	−0.1440	−0.0257	0.0007**	−0.0103
No of individual owners	−0.2241***	−0.2030**	−0.1999	−0.0004	0.0017
No of foreign owners	0.1887***	0.1523***	0.1857*	−0.0001	0.0059
Market cap	−0.2789***	−0.2452***	−0.2279**	−0.0002	−0.0084*
Price	−0.0005	−0.0003	−0.0037***	0.0000	0.0000
Return volatility	0.9138***	0.9488***	−1.2179**	0.0017**	−0.0120
Trades per day	−0.0012***	−0.0011***	0.0023***	0.0000	0.0000
2000					
Primary Insiders	−0.0242	−0.0197	0.2818	0.0001	−0.0079
Largest owner, state	0.6072	0.5034	−1.5096	−0.0014	−0.0002
Largest owner, institutional	0.7491	1.1186*	−1.3900	−0.0021	0.0230
Largest owner, non-financial	0.4325	0.4198	−1.2099*	−0.0010	0.0404*
Largest owner, individual	0.5108*	0.7018**	−0.1829	0.0004	−0.0030
Largest owner, foreign	0.3011	0.3780	−1.5851***	−0.0010*	0.0273
No of state owners	0.0311	0.0191	0.0005	0.0002**	0.0010
No of institutional owners	0.1077	0.0154	−0.6074***	−0.0002	−0.0063
No of non-financial owners	−0.2895***	−0.2742**	−0.1293	−0.0002	0.0037
No of individual owners	−0.0828	−0.0624	0.3261*	0.0001	0.0045
No of foreign owners	0.0332	0.0449	−0.3438**	0.0000	−0.0095*
Market cap	−0.0684	−0.0607	0.3536***	−0.0001	0.0005
Price	−0.0025***	−0.0023***	−0.0034***	0.0000	0.0000
Return volatility	3.4113***	3.3924***	−2.9669***	0.0030***	−0.0192
Trades per day	−0.0013***	−0.0012***	0.0026***	0.0000	0.0000
2001					
Primary Insiders	0.3777	0.7916	1.6015*	−0.0020**	0.0136
Largest owner, state	−1.7543	0.1745	−2.4102	0.0006	0.0262
Largest owner, institutional	−3.3955*	−0.6112	−2.5363	0.0002	−0.0004
Largest owner, non-financial	−0.2513	0.8092	−4.0408***	0.0005	−0.0299
Largest owner, individual	−3.4791	−1.1408	−7.2773	0.0022	−0.4419
Largest owner, foreign	−0.5254	0.0553	−1.8028**	0.0004	−0.0268
No of state owners	−0.0822	−0.0170	0.1937	−0.0001	−0.0046
No of institutional owners	−0.3078	−0.1248	−0.5423**	−0.0001	−0.0078
No of non-financial owners	0.6505*	0.4226	−0.7553*	0.0006	−0.0175
No of individual owners	0.1292	−0.1742	−0.5582	−0.0003	0.0182
No of foreign owners	−0.6687*	−0.2975	−0.1410	0.0005	0.0137
Market cap	0.0863	−0.0956	−0.1008	0.0000	0.0127
Price	−0.0024**	−0.0032***	−0.0060***	0.0000	0.0000
Return volatility	6.0250***	5.7900***	−1.5162	0.0008	0.0637
Trades per day	−0.0014***	−0.0009***	0.0013***	0.0000	0.0000

Table A.10: Market liquidity and aggregate holdings of owner groups - Sub-periods

The table reports results from estimating five panel regression model with one-way fixed effects (least squares dummy variable estimation). The model is estimated using the available observations for each of the years 1999-2001 for five measures of liquidity as the dependent variable: log(relative weighted spread), log(relative effective spread), log(weighted depth), adverse selection costs according to the GH-method (variable proportional costs), and the adverse selection component of the spread according to the GKN-method. The independent variables are the total holding which is *not* owned by the five largest owners ("free float"), the aggregate holdings of the primary insiders, the aggregate holding of a particular owner group, the logarithm of the market capitalization value, the average closing price, the standard deviation of daily returns, the average daily number of trades, and fixed effect dummies for each company. For each model, we report the estimated coefficient for the aggregate holding of the particular owner type. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, and * denotes significance at the 10 percent level.

	Dependent variables				
Owner groups	Rel weighted spread	Rel eff spread	Depth	GH comp.	GKN comp
1999					
State	0.1903	−0.0502	−0.2342	0.0004	0.0421*
Institutional	−0.7120***	−0.6308**	0.5162	−0.0015***	0.0223
Non-financial	0.3239	0.5011*	−0.1450	−0.0011*	−0.0258
Individual	−0.2301	−0.1311	−0.0474	0.0045***	0.0130
Foreign	0.2151	0.1450	−0.1007	−0.0003	−0.0229
2000					
State	−0.0501	−0.2157	−2.7078***	−0.0001	−0.0444*
Institutional	0.5522**	0.4332	−0.7774	0.0005	0.0116
Non-financial	0.2207	0.1716	1.0798**	0.0000	0.0011
Individual	−0.6152**	−0.4667	0.8586*	0.0008	−0.0249
Foreign	−0.0558	−0.0531	−0.0448	−0.0004	0.0050
2001					
State	1.0532	−0.8752	−4.9653**	−0.0002	−0.0281
Institutional	−0.5195	−0.1122	−0.2537	0.0007	−0.0585
Non-financial	1.6895***	1.8967***	−2.5697***	0.0008	0.0142
Individual	1.1795	−1.4937	1.5165	−0.0006	0.0539
Foreign	−1.3400***	−1.2433***	2.3548***	−0.0008	0.0072

Table A.11: Market liquidity and changes in ownership - Sub-periods

The table reports results from estimating the relation between liquidity and ownership structure using a panel regression model below with one-way fixed effects (least squares dummy variable estimation). The model is estimated using the available observations for each of the years 1999-2001. The independent variables are the aggregate holdings of the five largest owners, the aggregate holdings or the primary insiders, the absolute value of the change in the number of owners for all owner groups, the logarithm of the market capitalization value, the average closing price, the standard deviation of daily returns, the average daily number of trades, and dummies for the fixed effect of each company. The model is estimated for five dependent variables: log(relative weighted spread), log(relative effective spread), log(weighted depth), adverse selection costs according to the GH-method (variable proportional costs), and the adverse selection component of the spread according to the GKN-method. For each model, we report the estimated coefficients (except the coefficients for the fixed effect dummies). *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, and * denotes significance at the 10 percent level.

	Dependent variables				
Independent variables	Rel weighted spread	Rel eff spread	Depth	GH comp.	GKN comp
1999					
Five largest	0.8142***	0.7476***	0.5616	0.0032***	−0.0210
Primary insiders	−0.1174	−0.0534	0.1941	0.0002	−0.0107
$\|\Delta\|$ no of owners, state	−0.0111**	−0.0115*	0.0126	0.0000	0.0004
$\|\Delta\|$ no of owners, institutional	−0.0001	0.0007	0.0045	0.0000	−0.0003
$\|\Delta\|$ no of owners, non-financial	−0.0001	−0.0001	0.0000	0.0000	0.0000
$\|\Delta\|$ no of owners, individual	0.0000	0.0000	0.0000	0.0000	0.0000
$\|\Delta\|$ no of owners, foreign	0.0000	−0.0001	0.0000	0.0000	0.0000
Market cap	−0.2292***	−0.2351***	−0.2388***	−0.0003***	−0.0031
Price	−0.0006	−0.0002***	−0.0036***	0.0000	0.0000
Return volatility	1.0076***	1.0027	−0.8439	0.0026***	−0.0157
Trades per day	−0.0016***	−0.0014***	0.0019***	0.0000	0.0000
2000					
Five largest	1.3124***	1.5294***	−0.5526	−0.0001	0.0078
Primary insiders	−0.0666	−0.0857	0.1692	0.0001	−0.0070
$\|\Delta\|$ no of owners, state	−0.0130*	−0.0115	0.0360***	0.0000	0.0003
$\|\Delta\|$ no of owners, institutional	−0.0037***	−0.0034***	0.0102***	0.0000	0.0001
$\|\Delta\|$ no of owners, non-financial	−0.0005	−0.0006	0.0003	0.0000	0.0000*
$\|\Delta\|$ no of owners, individual	0.0001***	0.0001***	0.0000	0.0000	0.0000***
$\|\Delta\|$ no of owners, foreign	−0.0015**	−0.0015**	0.0047***	0.0000	0.0000
Market cap	−0.1164***	−0.1295***	0.1175*	−0.0001	−0.0016
Price	−0.0016***	−0.0015***	−0.0036***	0.0000	0.0000
Return volatility	3.6414***	3.5382***	−2.8491***	0.0032***	−0.0246
Trades per day	−0.0013***	−0.0012***	0.0016***	0.0000**	0.0000*
2001					
Five largest	−0.3064	0.1335	0.8155	0.0004	−0.0044
Primary insiders	0.4949	0.8229*	1.1628	−0.0018*	0.0050
$\|\Delta\|$ no of owners, state	−0.0023	−0.0117	0.0197	0.0000	−0.0008
$\|\Delta\|$ no of owners, institutional	0.0029	0.0027	−0.0034	0.0000	0.0000
$\|\Delta\|$ no of owners, non-financial	−0.0014	−0.0027**	0.0042***	0.0000	0.0000
$\|\Delta\|$ no of owners, individual	0.0003	0.0002*	−0.0003	0.0000	0.0000
$\|\Delta\|$ no of owners, foreign	−0.0027**	−0.0008	0.0001	0.0000	0.0002***
Market cap	0.0693	−0.0610	−0.0858	−0.0001	0.0089
Price	−0.0025**	−0.0030***	−0.0051***	0.0000	−0.0001
Return volatility	6.1505***	5.6563***	−0.4675	0.0006	0.0276
Trades per day	−0.0016***	−0.0010***	0.0012**	0.0000	0.0000

Appendix B

Decomposing the spread

We decompose the spread according to a version of the Glosten and Harris (1988) method without inventory costs, and one of the methods suggested in George et al. (1991).

1 The Glosten and Harris (1988)-method

Our description of this method is largely based on the description in Brennan and Subrahmanyam (1996). Let m_t be the expected value of a stock, conditional of the information set at time t. The GH-method is based on a Kyle (1985) type of price formation, i.e. it allows for a linear price adjustment rule to capture the information effect (Kyle's lambda) and a fixed cost of executing a trade[1],

$$m_t = m_{t-1} + \lambda q_t + y_t \tag{B.1}$$

where q_t is the order flow and y_t is an informational signal. Let ΔP_t be the intra-day change in the transaction price P_t from time $t-1$ to t, and let D_t be a dummy variable taking the value $+1/-1$ if the trade at time t was buyer-initiated/seller-initiated. Assuming no inventory costs, the transaction price can be written,

$$P_t = m_t + \psi D_t \tag{B.2}$$

where ψ is a measure of the compensation for per share execution costs and possible costs related to price discreteness and rents. Substituting out m_t using equation B.1, we have

$$P_t = m_{t-1} + \lambda q_t + \psi D_t + y_t \tag{B.3}$$

The change in the the price of a stock from one transaction to the next can then be decomposed in the following way,

[1]Equations 1-4 below are taken from Brennan and Subrahmanyam (1996), page 444.

$$\Delta P_t = \lambda q_t + \psi[D_t - D_{t-1}] + y_t \tag{B.4}$$

The empirical version of the model is

$$\Delta P_t = \beta_0 + \beta_1 q_t + \beta_2 [D_t - D_{t-1}] + e_t \tag{B.5}$$

where e_t is an error term. The information component of the spread will be reflected in the parameter β_1, i.e., we should find that $\beta_1 > 0$.

2 The George et al. (1991)-method

The method of George et al. (1991) (GKN-method) is largely based on the empirical measure of the effective spread introduced by Roll (1984). The underlying assumptions are no inventory costs, no private information, and a probability of trade reversals equal to 0.5). The special feature in the GKN-method is that the "true" expected return of a security is allowed to vary through time. Let E_t be the unobservable expected return for the period between transaction t and $t-1$ conditional on the information set at time $t-1$, let S_q be the quoted spread, and let M_t be the "true" price conditional on the information set immediately following transaction t. The model of transaction prices is given by,

$$P_t = M_t + \psi(S_q/2)D_t \tag{B.6}$$

and

$$M_t = E_t + M_{t-1} + (1-\psi)(S_q/2)D_t + y_t \tag{B.7}$$

It follows that,

$$\Delta P_t = E_t + \psi(S_q/2)[D_t - D_{t-1}] + (1-\psi)(S_q/2)D_t + y_t \tag{B.8}$$

The first-order serial covariance of successive price changes is given by,

$$cov(\Delta P_t, \Delta P_{t-1}) = cov(E_t, E_{t-1}) - \psi(S_q^2/4) \tag{B.9}$$

Thus, the relation between the serial covariance of trade-to-trade returns and the quoted spread is given by

$$\sqrt{\psi} S_q = 2\sqrt{-[cov(\Delta P_t, \Delta P_{t-1}) - cov(E_t, E_{t-1})]} \tag{B.10}$$

This version of the model of the effective spread developed by Roll (1984) includes time varying expected returns. Assuming that the time varying expected returns follow a first-order autoregressive process, George et al. (1991) suggest two techniques for taking the time variation in expected returns into account. In this study, we use the technique for extracting E_t based on bid quotes. Let PB_t be the bid quote subsequent to transaction t. The change in price calculated from bid quotes is given by,

$$\Delta PB_t = E_t + (1-\psi)(S_q/2)D_t + y_t \tag{B.11}$$

Let DP be the difference between ΔP and ΔPB. The serial correlation in DP is given by,

$$cov(DP_t, DP_{t-1}) = -\psi^2(S_q^2/4) \tag{B.12}$$

which gives the relation,

$$2\sqrt{-cov(DP_t, DP_{t-1})} = \psi S_q \tag{B.13}$$

Setting $\hat{S} = 2\sqrt{-cov(DP_t, DP_{t-1})}$, the empirical versions of the model is,

$$\hat{S} = \beta_0 + \beta_1 S_q + e_t \tag{B.14}$$

where e_t is an error term, and the estimate of adverse selection costs is given by $1-\beta_1$.

Bibliography

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

Angel, J. J. (1997). Who gets price improvement on the NYSE? Working Paper FINC-1377-12-694, Georgetown University.

Bacidore, J. M., Ross, K., and Sofianos, G. (1999). Quantifying best execution at the New York Stock Exchange: Market orders. NYSE Working Paper 1999-05.

Bhide, A. (1993). The hidden costs of stock market liquidity. *Journal of Financial Economics*, 34:31–51.

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

Bøhren, Ø. and Ødegaard, B. A. (2000). The ownership structure of Norwegian firms: Characteristics of an outlier. Research report, Norwegian School of Management BI.

Bøhren, Ø. and Ødegaard, B. A. (2001). Patterns of corporate ownership: Insights from a unique data set. *Nordic Journal of Political Economy*, 27:55–86.

Bøhren, Ø. and Ødegaard, B. A. (2003a). Governance and performance revisited. Norwegian School of management BI, unpublished paper.

Bøhren, Ø. and Ødegaard, B. A. (2003b). Norsk eierskap: Særtrekk og sære trekk. *Praktisk Økonomi & Finans*, 2:3–17.

Brennan, M. J. and Subrahmanyam, A. (1996). Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics*, 41:341–364.

Chiang, R. and Venkatesh, P. (1988). Insider holdings and perceptions of information asymmetry: A note. *Journal of Finance*, 43:1041–1048.

Cho, M.-H. (1998). Ownership structure, investments, and the corporate value: An empirical analysis. *Journal of Financial Economics*, 47:103–121.

Chordia, T., Subrahmanyam, A., and Anshuman, V. R. (2001). Trading activity and expected stock returns. *Journal of Financial Economics*, 59:3–32.

Easley, D., Kiefer, N. M., O'Hara, M., and Paperman, J. B. (1996). Liquidity, information, and infrequently traded stocks. *Journal of Finance*, 51:1405–1436.

George, T., Kaul, G., and Nimalendran, M. (1991). Estimation of the bid-ask spread and its components: A new approach. *Review of Financial Studies*, 4:623–656.

Glosten, L. and Harris, L. (1988). Estimating the components of the bid/ask spread. *Journal of Financial Economics*, 21:123–142.

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

Heflin, F. and Shaw, K. W. (2000). Blockholder ownership and market liquidity. *Journal of Financial and Quantitative Analysis*, 35:621–633.

Holmstrøm, B. and Tirole, J. (1993). Market liquidiity and performance monitoring. *Journal of Political Economy*, 101:678–709.

Huang, R. and Stoll, H. (1997). The components of the bid-ask spread: A general approach. *Review of Financial Studies*, 10:995–1034.

Jensen, M. C. and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3:305–360.

Jones, C., Kaul, G., and Lipson, M. (1994). Transactions, volume, and volatility. *Review of Financial Studies*, 7:631–651.

Keim, D. B. and Madhavan, A. (1998). The cost of institutional equity trades. *Financial Analysts Journal*, 54(4):50–69.

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

Madhavan, A. and Smidt, S. (1991). A Bayesian model of intraday specialist pricing. *Journal of Financial Economics*, 30:99–134.

Maug, E. (1998). Large shareholders as monitors: Is there a trade-off between liquidity and control? *Journal of Finance*, 53:65–98.

Næs, R. and Skjeltorp, J. A. (2003). Strategic investor behaviour and the volume-volatility relation in equity markets. Working Paper no 9, Norges Bank.

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

Sarin, A., Shastri, K., and Shastri, K. (2000). Ownership structure and stock market liquidity. Working Paper, University of Pittsburgh.

Shleifer, A. and Vishny, R. W. (1997). A survey of corporate governance. *Journal of Finance*, 52:737–783.

Sjo, H. (1998). Likviditet i aksjene på Oslo Børs. *Praktisk Økonomi og Ledelse*, 14(2):80–89.

Tobiasson, R., Heszlein, J., and Lange, T. (1999). Likviditet og eierstruktur: En empirisk studie av likviditeten, og sammenhengen mellom likviditet og eierstruktur, i aksjene på Oslo Børs. Master's thesis, Norwegian School of Management BI.

ISSN 1503-0741 (printed), 0809-652X (online)
ISBN 82-7553-309-0 (printed), 82-7553-310-4 (online)

Lobo Media 33340