Liquidity Provision in the Limit Order Book

Adverse Selection, Iceberg Orders and the Opening Auction

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Wirtschaftswissenschaftlichen Fakultät
der Eberhards-Karls-Universität Tübingen

vorgelegt von

Stefan Frey
aus Frankfurt am Main

2009

Dekan:	Professorin Dr. Kerstin Pull
Erstkorrektor:	Professor Dr. Joachim Grammig
Zweitkorrektor:	Professor Dr. Rainer Schöbel
Tag der mündlichen Prüfung:	20. Dezember 2007

Acknowledgments

> Bernard of Chartres used to say that we are like dwarfs on the shoulders of giants, so that we can see more than they, and things at a greater distance, not by virtue of any sharpness of sight on our part, or any physical distinction, but because we are carried high and raised up by their giant size.
>
> Johannes von Salisbury (1159): *Metalogicon 3,4,46-50*

Scientific work is a challenge. It is a journey into the unknown and many dangers lurk in the shadows. Often you realize after days of work that have only returned to your point of departure. Skill and luck often have to unite before you can achieve your goals. Without the help of a lot of colleagues and friends (often both) I would not have succeeded. This is the place to thank them all.

The most important person to encourage me to return to academia and write this thesis was my advisor Joachim Grammig. I am greatly thankful to his advice and support throughout that time. I always enjoyed and admired his deep knowledge, his enthusiasm, and his wit. I believe he applied the right combination of pushing and pulling to overcome the obstacles. Without him this work would have been impossible.

Research involves the exchange of ideas. Combining forces often helps to excel the capabilities of everyone involved. In this respect I was lucky to work together with excellent colleagues. The process of discussing, disputing, and finally (dis)agreeing was a marvelous experience. I would first like to thank Helena Beltran-Lopez. She implemented the important steps of the reconstruction of the XETRA order book and without her the tasks would have never been completed. It was a pleasure to work with her on the auction topics and it is a great loss for the field that she preferred to work for the private sector. Not less is my gratitude to Patrik Sandås. What I have

Financial support from the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged.

learned from him is invaluable. I really appreciate his trust to start with someone unknown and continually work over such a distance. There have been others project outside of my thesis and I would thank those colleagues too: Andreas Heinen, Patrick Herbst, Erick Rengifo, and Andreas Walter. I hope I did not disappoint you too often when I had to focus on my thesis. Unfortunately there are projects that get never really started or stuck somewhere on the journey. I want to thank those too: Thomas Dimpfl, Ingmar Nolte, and Alejandro Veen. Maybe we get things done in the future.

Many people helped with resources, opportunities, and valuable discussions. I would name only a few here: Winfried Pohlmeier, who allowed me to participate in the Konstanz summer school free of charge. I am grateful to Alexander Kempf who invited me to join the CFR in Cologne. I am heavily indebted to the people at Deutsche Börse namely Uwe Schweickert and Miroslav Budimir, who prepared the XETRA data and were always helpful to answer our - sometimes naive - questions. I want to thank all discussants and the audience at conferences and seminars. I will highlight three of them: Thierry Focault who encouraged me to focus on the liquidity provision at the Konstanz summer school. Bruno Biais inspired me not only with his papers, but also with his market microstructure course in Mannheim. And finally Gideon Saar for his comments at the NBER meeting and his discussion at the AFA which strongly improved the paper.

There are many people here in Tübingen I would like to thank. Too many to name them all, if you feel you belong here simply add yourself. I am grateful to Werner Neuss and Rainer Schöbel to serve on my thesis committee. I enjoyed working with those of you share(d) the same goal: Luis Huergo, Kerstin Kehrle, Franziska Peter, Stephan Jank, Miriam Sperl, and Oliver Wünsche. I relied on the knowledge from Sylvia Bürger and Angelika Hutt to struggle with the administrative tasks and I thank for that. Sven Bauer was a great help to tackle all issues related with IT infrastructure - without him we would still use pen and paper. I experienced all students to be very interested and most of them were a great pleasure to work with. I name one here for all of them: Peter Schmidt for his excellent team play during the organisation of the econometric lab.

Without family and friends I had never got that far. I am very glad to thank Anette Farr for all the support and love she gives to me. She and our Laurenz make me happy every day.

Contents

List of Tables

List of Figures

Prologue

Markets are fascinating.

Larry Harris (2003)
Trading and Exchanges, p.3

Motivation

Throughout my professional years working at an investment management company I was impressed by the intricate process from the trading decision to the trade execution. The investment process itself requires many sophisticated people to evaluate the wide array of available information and opinion. Analysts who focus on quarterly earning reports and balance sheet informations, building huge models to derive a DCF price target. Funds managers who decide on the assset allocation and at the same time proficiently pick stocks. Both groups meet with corporate leaders and sell-side analysts to discuss company strategies and scrutinize their opinions. But in the end there is a trade ticket submitted to the trading desk of the institution - though it exists only on computer networks nowadays.

One trader compared the impact of those large orders to the bow wave of a mighty ocean liner cruising the sea. All the other vessels try to avoid it by crossing earlier or waiting for the sea to calm down. The traders job is to minimize this effect by disguising the large order. It is broken up into several smaller packages submitted over the trading day to various investment brokers, sometimes to two or three simultaneously. For stocks with low liquidity the fullfillment requires up to several days. The broker will first try to match the order internally or against an order from another client. If he does not succeed, he slices and dices it in even smaller units before submitting them to the exchange - usually with the help of automated trading programs.

The science of trading - market microstructure - has to deal with those issues (and many more). The example shows that this field is closely connected to several other important issues in finance. Asset pricing, asset allocation and corporate valuation are important issues that help to understand the trading decision. All of them strongly draw from macroeconomics, monetary economics, and accounting and those again effect the nature of trading and their markets. The recent financial crises beginning with the October 1987 stock market crash up to the current financial crises prove that market liquidity itself is an important topic to consider for the other areas of finance and economics.

Every trader submitting orders to an exchange considers the motives of the other market participants. This leads to the well known problem of adverse selection: One might trade with people who have better knowledge about the fundamental value of the stock. Or - to avoid the fiction of a fundamental value - at least about the future price movement in a time period that is a relevant holding period. Solving those problems requires both input from industrial organization and behavioral economics. In that respect the importance of market microstructure is diametrically opposed to the efficiency of markets. A perfectly efficient market allows everyone to trade with each other without any cost using the true price. Unfortunately those markets do not exist in reality. Quite the contrary is true. With the knowledge about the imperfections market microstructure suggests trading systems that maximize market efficency and minimize trading costs. The output of such research is then highly relevant for exchanges and their regulators.

This thesis collects my work on empirical market microstructure up to the first half of 2007. In 2004 the German exchange operator Deutsche Börse granted the opportunity to my advisor Joachim Grammig and some other researchers to study the XETRA limit order book. The dataset is an extraction of the XETRA logfiles that contained all events of every order during the trading day. It offers a very detailed view on the trading process, but is still very far away from what anyone would recognize as a dataset of an actual limit order book. Thus the extended first year of my work was dedicated to the reconstruction of the order book. Fortunately there was an existing code base from a previous undertaking in 2001 and we had the opportunity to build up an IT infrastructure powerful enough for the task on hand.

Agenda

In the following section I outline the research agenda. The common theme of all chapters is the question how traders participate in the limit order book. The first chapter as an introduction looks at profits from trading both at the level of traders and the exchange. In the second the focus moves on to a model based estimation of profits for limit order suppliers and confirms the surprising result that the marginal limit order is in the red. The third discusses the usage of iceberg orders, an order type which allows trader to enter hidden liquidity into the book. In one respect it violates the characteristic of full transparency. In another it extends the strategy space of traders and may allow them to keep the trade on the exchange instead of falling back to an upstairs markets or a crossing network. The results provide some interesting insights in the interaction between those submitting the iceberg orders and the other traders who try to detect the hidden liquidity. The last chapter then concludes with a specific trading period of the XETRA system, namely the auctions. It employs the limit order book, but in essence is a two-sided uniform price auction with multiple units. The main results are a possible answer to the prediction puzzle of the indicative price price and to outline the effect of the random end phase.

Liquidity provision on Xetra

The first chapter serves as an introduction to the trading system of XETRA. It discusses the profitability of the limit order book for the system provider Deutsche Börse in terms of EBIT contributions and the development of trading costs in the period from 1999 to 2006. It describes the market model of XETRA with its order types and matching rules. The main analysis is an implementation of the methodology of Harris and Hasbrouck (1996) and an adaptation thereof to the auction period.

The main results are the following. (i) One of the reasons of the economic success for Deutsche Börse is the introduction of the XETRA system for their equity trading. The EBIT margin topped 50% plus in 2006. (ii) During continuous trading ex-ante performance increases with aggressiveness, order size and prevailing spread. Ex-post performance on the other hand decreases with aggressiveness. (iii) Ex-ante performance is best for orders submitted early during the opening auction. For the ex-post performance the pattern reverses but the statistical significance is weak.

Liquidity supply and adverse selection

The second chapter concentrates on the profit conditions of liquidity suppliers in the limit order book. It extends the framework of Sandås (2001) to estimate an empirical version of Glosten's (1994) limit order book model in several dimensions. Average profit conditions for the full volume of an individual quote replace those for the marginal units. The assumption of an exponential distribution for volumes of the market order submissions is relaxed with a nonparametric specification. Finally a cross-sectional analysis studies the relation between liquidity supply and adverse selection costs.

The conclusions of the chapter are in summary. (i) The average profit conditions provide a better fit to the observed schedule and imply positive per unit costs. The fixed costs are still estimated to be negative. This implies that at least for quotes close to the spread there is an incentive to trade. (ii) All results are robust for the choice of the exponential distribution compared to the nonparamateric approach. Especially the marginal transaction costs remain significantly negative. (iii) The empirical results support one of the main hypothesis of the theory of limit order markets, namely that liquidity and adverse selection effects are inversely related.

Iceberg orders and the compensation for liquidity provision

The previous chapter studied the conditions to supply visible liquidity in the context of competition between several traders. The focus of the third chapter is on the effects from hidden liquidity that stems from iceberg orders. They are used to reduce the price impact of large orders, thus the other traders have to consider the potential hidden liquidity in their profit calculations. An extension of the model from Sandås (2001) allows to address those questions. The important change is that the order flow and the price impact changes according to the existence of hidden liquidity in the order book. An important assumption is the possibility to detect the iceberg order due to the characteristics of the replenishment process.

The analyses can be summarized as follows. (i) The price impact of a market order is reduced if an iceberg order is on the opposite side of the order book ("buffer effect"). (ii) The order flow is skewed in the direction of the iceberg in the order book. Larger market orders hit the iceberg order more often. (iii) Liquidity provision only partially adjusts to the crowding out by the hidden liquidity. Liquidity provision on the opposite side is strongly negative, whereas on the side of the iceberg there are still profit opportunities.

Auctions

The last chapter analyzes a topic which is less covered in the research literature. At the beginning and end of the trading sessions there are scheduled auctions. There are two major differences to continuous trading. Firstly it is uniform price auction and secondly the execution is delayed until a specific point in time. XETRA is specific in two aspects here. The exact time of execution is triggered randomly, thus after the first time where execution is possible, the duration gets extended by a random amount of time. Before that orders can be submitted without any commitment, cancellation is possible by will. XETRA closes the order book throughout the auction process and publishes only an indicative price. Biais, Hillion, and Spatt (1999) discussed the opening auction for the Paris exchange and spotted a bias in the indicative price if used as a forecast for the true auction price. We extend his model using a specification that allows for (microstructure) noise.

The results are in short summary: (i) The indicative price is an unbiased forecast for the true value of the stock right after the beginning of the morning auction, if one corrects for market microstructure noise. (ii) The executable volume is informative about the amount of noise in the indicative price. Additionally, the order book contains information about future prices that is not incorporated in the indicative price. (iii) The length of the random end phase has no prolonged impact on traded volume for that stock on that day.

One

Liquidity provision on Xetra

The chapter introduces the Xetra trading system. It provides a short history and discusses the economic contribution of Xetra for Deutsche Börse . It outlines the function of the different order types and their contribution to the trading process. To illustrate the success of potential strategies for liquidity suppliers, we use the approach of Harris and Hasbrouck (1996) for the DAX 30 stocks. The methodology is then adopted to the case of the opening auction.

1.1 The Xetra trading system

The exchange Deutsche Börse

The listed company Deutsche Börse is the major exchange in Germany. It supports Xetra as the electronic trading platform for the equity spot market. Xetra is one of the business segments in the diversified set-up of Deutsche Börse that covers the whole value chain from trading and clearing to settlement for a broad range of asset classes. This includes both spot and derivative markets.

The offering for equity markets is twofold: First the Xetra system and second the Frankfurt Parkett (floor) trading. The latter differs from Xetra by the existence of a closed order book that is collected by Maklern (specialists) only and targets private investors. For Xetra and all other German exchanges, Deutsche Börse provides centralized services for clearing and settlement by its subsidaries Eurex Clearing and Clearstream. In addition Deutsche Börse is the majority owner of the derivative exchange Eurex and the fixed-income platforms Eurex Bonds and Eurex Repo. In April

2007 it announced the expansion to the US with the takeover of the equity derivative exchange ISE. The previous attempt to expand in North America through Eurex US had no success.

Nowadays, the networking properties of an exchange is of utmost importance. Market participants can connect to Xetra via access points in the main European financial centers, either by Xetra clients, indirect with the Xentric systems or by implementing the Value API into applications of the market participants. This facilitates the increasing amount of trading volume generated by algorithmic trading programs.

A short history of Xetra

Xetra was introduced by Deutsche Börse in 1997. With release 3.0 in 1998 the functionality offered was already very similar to the system now. Major changes were introduced by release 5.0 in 2000, with iceberg orders, support of multiple exchanges and block trading facility Xetra XXL. Release 7.0 allowed internalization on Xetra BEST. Central counterparty service followed in 2003. The latest release 8.0 focused on improvements of the system's performance including a heartbeat system. The next release 8.1 in late October 2007 will introduce those changes required by the European market directive MIFID. But Xetra is not only the trading platform for the German equity market. Other instruments trading on Xetra in Germany include foreign equities and exchange traded funds (ETFs). Starting April 2008 SCOACH, a joint-venture with the Swiss Exchange SWX for structured products, will run on Xetra. Furthermore the Vienna stock exchange (since 1999), the Irish stock exchanges (since 2000) and the Bulgarian stock exchange (planned for 2008) use Xetra with Deutsche Börse as outsourcing provider. The Shanghai stock exchange will built their next generation market system on the technology of Xetra. Moreover the subsidiaries Eurex Bonds and Eurex Repo and the European Energy Exchange in Leipzig apply it for their markets.

The economic dimension of Xetra

Since the initial public offering on 5th February 2001 shares of Deutsche Börse proved to be a very successful investment. The most recent share price of €105.90 (close 12th October 2007) implies a performance of more than 500 percent compared to an issue price of €16.75 (adjusted for stock splits). In the same time the DAX 30 price index (excluding dividends) increased only by 3.5 percent from 5084 to 5260 points. Can we specify the value of Xetra? The total market capitalization of Deutsche Börse was 14.6 Mrd Euro at the end of 2006. Xetra contributes 17.4 % of the total EBITA (179 Mio Xetra / 1029 Mio Deutsche Börse). Thus a simple multiple calculation derives a market value of somewhere around 2.5 Mrd Euro for Xetra alone. This would rank Xetra among the 50th largest publicly traded companies in Germany.

Figure 1.1: Revenue from Xetra

Figure 1.1.highlights the surge in revenues and profits from the XETRA segment in recent years. The improvement of the EBIT can be partly explained by the doubling of revenues from 1999 to 2006 and the expansion of the EBIT margin (the shaded area) from around five percent in 1999 to more than fifty in 2006.

Figure 1.2: Transaction costs on Xetra

From market participant's point of view strong earnings might indicate that the exchange was slow to pass on lower costs. Figure 1.1 shows that turnover in euro volume (the shaded area) tripled since 1999, whereas volume related transaction costs only decreased by twenty percent. On the other hand per transaction costs more than halved, but this is mainly explained by a simultaneous decrease of order sizes. Table 1.1 summarizes important financial figures related to Xetra, as reported in the annual reports of Deutsche Börse . Another interesting fact is the ongoing concentration within the group of market participants.

The Xetra market model

In chapter five Harris (2003) defines market structure by two components: Trading rules and trading system. Below we discuss Xetra in that respect. Nevertheless, two trading systems may share the same trading rules, and still can behave very differently. The major European Exchanges share – with minor non-substantial exceptions – the same trading rules for their equity cash markets. And still it is very different to execute

Reporting Year	Capital Investment Mio €	Revenue Mio €	EBIT Mio €	EBIT Margin in %
1999	31.1	139.3	9.7	5.9
2000	31.8	277.9	82.5	29.7
2001	28.5	243.1	38.4	15.8
2002	38.5	210.8	17.4	8.3
2003	2.7	212.3	57.6	27.1
2004	9.8	216.3	86.2	39.9
2005	4.0	247.7	112.6	45.5
2006	5.2	314.1	179.0	57.0

Reporting Year	Number Participants	Turnover Mrd €	Number Transactions Mio	Average Tradsize 1000 €	Average Transaction Costs €	in BP
1999	404	515.0	15.8	32.6	4.4	1.4
2000	431	980.6	39.0	25.1	3.6	1.4
2001	413	958.4	49.7	19.3	2.4	1.3
2002	359	876.2	60.0	14.6	1.8	1.2
2003	308	833.1	71.4	11.7	1.5	1.3
2004	283	902.7	69.4	13.0	1.6	1.2
2005	267	1125.5	81.3	13.8	1.5	1.1
2006	262	1592.9	107.7	14.8	1.5	1.0

Table 1.1: Xetra - key financial figures

a large order for a French corporation on Xetra or Euronext in regards to execution costs and speed. On the other hand, exchanges with very different or even opposing rules excel at the same time, as in the example of American exchanges, where Nasdaq and NYSE both compete successfully.

The two archetypes of on-exchange execution systems, see Harris (2003), are the quote-driven dealer market and the order-driven markets. Xetra is a typical example of the latter. It provides electronic, non-discriminatory access to a centralized order book. Xetra supports both single-price auctions and continuous two-sided auctions and combinations thereof. Depending on the instrument, the trading day consists of a single auction, a sequence of multiple auctions, or of continuous trading with auctions at the start, middle, and closing of trading. The last is by far the dominating trading type on Xetra in terms of transaction volume.

Xetra displays the limit order book during continuous trading to all market participants, thus implements ex-ante transparency (for the exception during auctions see chapter 4). Xetra itself is ex-post transparent, but there is no requirement in Germany to report off-exchange OTC trades, which will change after the implementation of European market directive MIFID. It provides traders anonymity, ex-ante for all instruments, ex-post for those where Deutsche Börse has a central counterparty service, which includes all DAX instruments.

The order matching algorithm is the core of the Xetra system. The ruleset is well defined and publicly available. During normal operation it does not require human intervention and allows nearly instantaneous executions. The precedence rules are limit price - order book display - submission time. Price priority is the primary order precedence rule: Any order which improves the price has priority to all other orders which offer a worse execution. For orders with the same price limit, visible volume is preferred to hidden volume, which entered the order book by the usage of the iceberg order type (for more details on iceberg orders refer to chapter 3). Orders with identical price limit and display condition are executed in the sequence of the submission time. The two subsidiary rules encourage traders to submit their liquidity supply visible and early, and this again increases the marketplace's attractiveness for traders searching immediate execution.

The Xetra Q1 2004 dataset

All of the following analyses rely on an event dataset supplied directly by the exchange provider Deutsche Börse . It comprises all activity on Xetra for the constituents of the German blue chip index DAX 30. The dataset allows us to recreate the order book and all transactions during the first three months of 2004. The excellent quality and coverage of the data facilitates the studies, which follow in the next chapters of this work.

Some characteristics of the dataset are shown in table 1.2. The daily number of transactions varies between 600 to 4,300. Liquidity supplying orders exceed the number of transactions by a factor of six and the number of events is on average fifteen times the transaction figure. The average number of events per order is around 2.1 which is close to the minimum of two (entry, followed by either cancel or execution).

Xetra order types and instructions

The main ingredient of an order driven market is the limit order. In turn this manifests in the definition of those exchanges as pure limit order book markets. Orders, which demand or supply liquidity, only differ in their limit price relative to the current state of limit order book. In conclusion market orders are only a short-cut for a limit order with an infinite (for a buy) or zero (for a sell) limit price. In fact, the existence of market orders complicates the ruleset of a market substantially, for example 21 of the 27 matching rules for Xetra involve that order type. Accordingly, the classification into liquidity demanding or supplying is not the one choosing market versus limit for the order type, but rather how the chosen limit relates to the current order book: All orders demanding immediate execution constitute the class of marketable orders. Orders are partially marketable, if the limit allows only a fraction of the order to be executed against standing limit orders. The unexecuted part then enters the order book. Those and all orders without immediate execution provide liquidity to the market.

Table 1.3 refers to the distinction made above: 11% of all orders are marketable orders that execute against the orderbook immediately. The remaining 87% of all submitted orders are non-marketable and contribute to the order book. Only 2% of the orders are partially marketable, but they are responsible for both 17% of the immediate and non-immediate executed volume. A similar result for iceberg orders: 0.5% of orders, 10% of submission volume to the book and 16% of volume executed from the book. Iceberg orders will be discussed in detail in chapter 3. Nonetheless, the dominating order type remains the limit order with 97.4% of all orders.

The table also highlights, that market orders mainly stem from agency trading and that liquidity provision is dominated by proprietary accounts. Traders can assign further specifications to each order using order restrictions like immediate-or-cancel, fill-or-kill and stop orders. Orders can be restricted for execution in auctions only and decide on the expiration date. Table 1.4 reports that the typical Xetra order is the order valid for a day without any restriction. Only immediate-or-cancel is a frequently used restriction and provides an alternative to the market order, which still provides some price protection against adverse market movements.

1.2 Liquidity provision during continuous trading

In the following we focus on the continuous trading phase and the possible strategies of limit order suppliers. It is difficult to evaluate the success of order submission strategies: Institutional investors usually buy shares in blocks of several millions Euro. The typical order size however is in the range of several thousands Euro (see table 1.14). The same holds for professional liquidity suppliers that will execute a substantial amount of trades that have to be considered in combination. The event dataset from the exchange does not allow to follow individual strategies as it does not include any identification of the traders or its institution. Then again, the same reasoning supports the aggregation of a large sample of executions to derive results for submission strategies by statistical means.

To improve tractability, we categorize all order submissions in three dimensions. This allows to identify the strategies chosen by the traders and creates cross-sectional standardization. The first is the order size in euro volume, the second the spread prevailing at the time of submission and third the relative position to the bestquote. For order sizes, each order is assigned to one of three quantiles for each stock, defining the categories small, middle, and large orders. The same for spreads, which defines small, middle and large spreads for each stock individually. The categories relative to the best quote are behind the market (inside the order book), at the market (submission at the price level of the best quote), inside the market (inside the spread, between the two best quotes) and partially marketable. Table 1.14 provides details on the classification for each stock.

Table 1.5 describes the distribution of order submissions. Overall the activity at and inside the market plays an important role, but the majority of submissions occurs inside the order book. It is not surprising that large spreads discourage partially marketable orders due to the adverse execution conditions. With the notable exception for large orders, which already provide the bulk of partially marketable orders. Generally speaking, large orders are more frequent observed at large spreads than expected. Considering the execution probabilities adds to the picture: Not much a surprise is that more aggressive submissions and smaller spreads increase execution probabilities. The exception are partially marketable orders, where larger spreads coincide with more executions. The explanation is, that a larger spread increases the aggressiveness of the partially marketable order itself. Furthermore, larger orders execute more often than the smaller and medium sized orders. Table 1.7 withorder durations illustrates that . Larger orders nearly stay twice as long in the order book compared to the medium sized orders. A surprising result is that orders submitted inside the spread are by far the most persistent limit orders. We conjecture that those orders are submitted by value traders, which do not adjust their orders to market movements as strongly as pure liquidity sponsors do.

To provide more insight to the success of the different order types we apply a methodology by Harris and Hasbrouck (1996) used for a data sample for the NYSE in from 1990/1991. They propose two different measures to capture the performance of an order submission. The ex-ante performance is defined as the difference between the potential execution price at submission minus the execution price for filled trades or the potential execution price at the time of cancellation. The ex-post performance is the difference between the execution price and the same side quote five minutes after the fill. For the Xetra dataset we propose the following changes: The reconstruction of the order book allows the calculation of potential execution prices that consider the order size, as large orders might walk up the book. In contrast Harris and Hasbrouck (1996) had to use the prevailing best quote, assuming that the order will be fully filled at that price. On NYSE an incoming order might receive a price improvement, something that cannot occur on Xetra by definition. We change the ex-post performance to that of the realized spread (a more frequently kind of measure, see for example SEC (2001)), thus the comparison price in the future is the midquote instead of the same side best quote. And third the time horizon is increase to ten instead of five minutes. The following equations summarize the definitions of the ex-ante and ex-post performance.

$$p^{\text{ex-ante}} = \begin{cases} p^{\text{potential}}_{\text{at submission}} - p^{\text{fill}} & \text{for submitted buy orders} \\ p^{\text{fill}} - p^{\text{potential}}_{\text{at submission}} & \text{for submitted sell orders} \end{cases} \tag{1.1}$$

$$p^{\text{fill}} = \begin{cases} p^{\text{execution}} & \text{for executed orders} \\ p^{\text{potential}}_{\text{at cancellation}} & \text{for cancelled orders} \\ p^{\text{auction}}_{\text{at closing}} & \text{orders still in the book at closing} \end{cases} \tag{1.2}$$

$$p^{\text{ex-post}} = \begin{cases} p^{\text{midquote}}_{\text{execution+10min}} - p^{\text{execution}} & \text{for executed buy orders} \\ p^{\text{execution}} - p^{\text{midquote}}_{\text{execution+10min}} & \text{for executed sell orders} \end{cases} \tag{1.3}$$

We focus on liquidity provision only here, thus results for market orders and fully marketable limit orders are omitted. The results in table 1.8 show clear evidence for an improved ex-ante performance with increasing level of aggressiveness, order size, and prevailing spread. The latter two can be explained by increased opportunity costs for an immediate execution of larger orders or because of an increased spread. The worse ex-ante performance of submissions far from the opposite side can be explained by the worse execution by filling the trade with a market order at the time of cancellation.

The ex-post performance measured by the realized spreads in table 1.9 tell a different story though. Performance deteriorates with an increasing level of aggressiveness. Most likely, the terms at execution are more favorable for the orders submitted deeper in the book, and they earn a substantial profit, if they get executed at all. The more aggressive orders are closer to the other side of the book with less a spread to earn. Large spreads at submission provide an increased ex-post performance. This implies that larger spreads attract new liquidity, thus lowering the spread until a new equilibrium is reached. For order sizes, the picture is mixed, there is only a slight evidence, that favors medium order sizes.

1.3 Liquidity Provision during the Opening Auction

Submitting to an auction with a uniform pricing rule requires a different behavior compared to the continuous case. The main decision is whether to participate in the crossing by submitting aggressively enough or not. Only the trader with the marginal order decides the auction price. In the extreme case of infinitely many traders, any individual cannot influence the auction outcome by its own behavior. Both the number of market participants in Xetra (see table 1.1 and the number of orders submitted to the opening auction (see table 1.15) do not fully support the assumption of no influence..

Similar to the previous section, each order submission is categorized in three dimensions: Order size, the time of submission during the auction and the position compared to the indicative price (see chapter 4 for more details about the opening auction). The median order volume separates the two orders sizes (small and large) for each stock. The division of the auction call into the three periods is based on equal order numbers in each period. The auction mechanism itself defines pretrading and random end as specific periods. The position compared to the indicative price is either behind (order would not execute at the current indicative price), marketable (order would execute at the indicative price) or the order is a market order (order executes in any case).

Table 1.10 reports frequencies of order submissions. More than a quarter of all orders are already submitted during pretrading, where no indicative information is published by the exchange. Those are mainly orders of the smaller (3.5 times that of the larger) category and are less aggressive than those submitted later. The closer we get to the auction resolution, the fewer are unlimited market orders. The proportion between small and large order remains mainly constant. The speed of order submissions increases to the end of the auction, the third period of the auction has the same number of orders as the first, but is only 90 instead of 335 seconds long. The random end, with an average duration of 15 seconds, contributes an additional 7% of all submissions. An interesting indication of the traders motive can be derived from table 1.11. Submissions during the pretrading have execution probabilities of combined, auction and later continuous trading, 0.74 (small orders) and 0.6 (larger orders). Orders submitted during the auction call are cancelled much more frequently. The extreme case are large marketable orders submitted at the beginning of the auction. 71% of those are cancelled during the auction. The number increases slightly over time, but reaches the same level of commitment as in the pretrading only during the random end.

We adapt the approach used for continuous trading above for orders submitted during the opening auction. A crucial difference is how to measure the ex-ante performance. Unlike in continuous trading, immediate execution using a market order is not possible. But at any rate, the obvious choice for a reference price, is the auction price itself. Using that, the ex-ante performance of orders executed in the auction, is zero by definition. For those orders which are executed or cancelled in continuous trading after the opening we can compare the actual or potential execution price to the auction price. All orders that are cancelled during the auction call are excluded from the analysis. The ex-post performance is identical to that in continuous trading: We consider all executed orders (both in the auction and later), and compare that to the midquote ten minutes later. Below we summarize the definitions from above..

$$P^{\text{ex-ante}} = \begin{cases} p^{\text{auction}} - p^{\text{fill}} & \text{for buy orders that still live at auction execution} \\ p^{\text{fill}} - p^{\text{auction}} & \text{for sell orders that still live at auction execution} \end{cases} \tag{1.4}$$

$$p^{\text{fill}} = \begin{cases} p^{\text{auction}} & \text{for orders executed in the auction} \\ p^{\text{execution}} & \text{for executed orders in continuous trading} \\ p^{\text{potential}}_{\text{at cancellation}} & \text{for cancelled orders in continuous trading} \end{cases} \tag{1.5}$$

$$p^{\text{ex-post}} = \begin{cases} p^{\text{midquote}}_{\text{execution+10min}} - p^{\text{execution}} & \text{for executed buy orders} \\ p^{\text{execution}} - p^{\text{midquote}}_{\text{execution+10min}} & \text{for executed sell orders} \end{cases} \tag{1.6}$$

Ex-ante and ex-post performance of order submissions during the opening auction are reported in tables 1.12 and 1.13. The ex-ante performance decreases throughout the auction (except for submissions in the random end phase) and larger orders. On average large orders have a worse ex-ante performance than smaller orders and for market orders it is zero by definition. The results for the ex-post performance should be interpreted with care as the cross-sectional variation between stocks, shown by the large standard errors. Still there is an increase in performance for submissions at the end of the auction and during the random end. The group of orders performing worst are the large market orders submitted during pretrading. Those orders might be submitted by liquidity traders which built or close an open position.

1.4 Outlook

From Glosten (1994a) we learned that eventually the non-discriminatory limit order market would prevail over the competition of trading systems with dedicated market makers. They have to be replaced by voluntary liquidity suppliers. This work is dedicated to investigate the environment in which those agents operate. In chapter 2 we discuss the profit conditions during continuous trading, where liquidity supplier must be cautious not too lose against informed traders due to adverse selection effects. An interesting twist in the otherwise straight-forward market model of Xetra are iceberg orders, which allow to hide a substantial share of an order. In chapter 3 we analyze the resulting changes in the market dynamics and the consequences for the liquidity providers' profit conditions . Chapter 4 sees our research extended to the opening auction. Of special interest in this scenario is how to forecast the auction price and those issues that arise from the absence of transparency of the order book.

Table 1.2: Xetra Q1 2004 sample

Trading Activity Group	Ticker Symbol	Stock Name	Market Cap. [€ billions]	Trading Volume [€ billions]	Daily Averages [in 1,000] Trans-actions	Liquidity Orders	Xetra Events
1	ALV	ALLIANZ AG VNA	33.8	18.6	4.1	29.7	71.2
(High)	DBK	DEUTSCHE BANK AG NA	38.2	19.8	3.6	23.1	56.5
	DCX	DAIMLERCHRYSLER AG NA O.N	30.3	12.0	3.1	18.8	46.6
	DTE	DT.TELEKOM AG NA	34.9	22.4	4.3	14.6	42.0
	MUV2	MUENCH.RUECKVERS.VNA	16.4	13.3	3.1	20.1	48.9
	SAP	SAP AG ST	27.4	11.8	2.6	19.7	46.7
	SIE	SIEMENS AG NA	52.9	20.6	4.1	23.7	59.0
	Mean		33.4	16.9	3.6	21.4	53.0
2	BAS	BASF AG	25.4	8.0	2.4	18.3	43.3
	BAY	BAYER AG	15.9	5.7	2.3	15.3	37.1
	BMW	BAY.MOTOREN WERKE AG ST	12.2	5.6	2.0	14.8	35.2
	EOA	E.ON AG	33.8	10.3	2.6	18.9	45.4
	HVM	BAY.HYPO-VEREINSBK.	6.6	6.3	1.8	10.2	25.7
	IFX	INFINEON TECH.AG NA	4.8	9.4	2.8	10.3	28.8
	RWE	RWE AG ST	12.7	6.3	2.1	14.5	35.0
	VOW	VOLKSWAGEN AG ST	9.7	6.7	2.4	13.5	33.8
	Mean		15.1	7.3	2.3	14.5	35.5
3	ADS	ADIDAS-SALOMON AG	4.1	2.0	0.9	8.1	18.6
	CBK	COMMERZBANK AG	7.6	3.4	1.4	12.0	28.0
	DB1	DEUTSCHE BOERSE NA	4.8	2.3	1.0	6.6	16.0
	DPW	DEUTSCHE POST AG NA	6.8	2.8	1.2	6.9	17.4
	LHA	LUFTHANSA AG VNA	4.5	2.8	1.3	8.1	20.0
	MEO	METRO AG ST	5.0	2.5	1.1	8.0	19.2
	SCH	SCHERING AG	7.1	3.3	1.4	9.1	22.2
	TKA	THYSSENKRUPP AG	6.4	2.4	1.2	7.9	19.2
	Mean		5.8	2.7	1.2	8.3	20.1
4	ALT	ALTANA AG	3.3	2.0	1.0	7.7	18.3
(Low)	CONT	CONTINENTAL AG	4.1	1.6	1.0	8.1	18.8
	FME	FRESEN.MED.CARE AG	1.9	0.8	0.6	5.8	13.1
	HEN3	HENKEL KGAA VZO	3.7	1.2	0.7	8.0	17.9
	LIN	LINDE AG	3.4	1.4	0.8	8.3	19.0
	MAN	MAN AG ST	2.4	1.8	0.9	7.2	17.2
	TUI	TUI AG	2.0	1.7	1.0	6.8	16.5
	Mean		3.0	1.5	0.9	7.4	17.2
All	Mean		14.1	7.0	2.0	12.8	31.2

Table 1.2 reports the market capitalization and the trading volume for each stock in the sample. The trading activity groups are formed by sorting the stocks according to the total trading volume in the first quarter of 2004; group 1 has the highest and group 4 the lowest trading volume. The market capitalization is free-float adjusted, and measured in billions of euros as of December 31st, 2003. The last section reports the daily averages for transactions, submission of orders which provide liquidity to the order book and the total number of rows in the event dataset provided by Deutsche Boerse. The last row for each group reports the cross-sectional average of the variables. The last row reports a cross-sectional average of the means with a standard error reported below each mean.

Table 1.3: Breakdown of Order Types

Order Type	Order Numbers	Execution Volume Immediate	Execution Volume from the Order Book	Submission Volume	Account Agency	Account Proprietary
	in Column Percents				in Row Percents	
FULLY MARKETABLE						
Market	2.0	8.8	.	.	76.1	23.9
	(1.1)	(3.0)	.	.	(10.3)	
Limit	9.3	73.4	.	.	28.5	71.5
	(2.8)	(4.3)	.	.	(7.3)	
All	11.4	83.0	.	.	36.9	63.1
	(3.6)	(3.5)	.	.	(7.6)	
PARTIALLY MARKETABLE						
Limit	1.8	14.4	11.7	3.4	34.2	65.8
	(0.5)	(2.9)	(3.0)	(1.3)	(10.6)	
Iceberg	0.1	2.6	5.1	2.1	32.2	63.0
	(0.1)	(2.2)	(3.7)	(1.7)	(26.7)	
All	1.9	17.0	16.7	5.5	34.2	65.8
	(0.6)	(3.5)	(4.7)	(2.6)	(10.6)	
SUBMISSION TO THE ORDER BOOK ONLY						
Limit	86.3	.	72.8	86.5	14.7	85.3
	(4.3)	.	(8.5)	(7.0)	(5.2)	
Iceberg	0.4	.	10.5	7.5	39.9	59.8
	(0.3)	.	(5.9)	(4.9)	(24.7)	
All	86.7	.	83.3	94.0	14.8	85.2
	(4.1)	.	(4.7)	(2.9)	(5.2)	

The contribution of the different order types is shown in Table 1.3. The middle columns presents from left to right the percentages of order numbers, execution volumes (immediate and non-immediate), and submission volumes. The columns on the right differentiate for each order type the usage by the proprietary and agency accounts of the market participants. The main order types of Xetra (Market, Limit, and Iceberg Orders) are divided into three different subgroups, depending on execution or non-execution at the time of submission. The upper panel shows the orders with full execution, the middle panel those with partial execution (the remaining part enters the order book) and the lower panel those which enter the order book. Limitorders with trade restriction Immediate-Or-Cancel and partial execution due to their limit are counted as fully marketable. All figures are averages and standard deviations (in parentheses) based on the results for each stock and trading day combination.

Table 1.4: Order Restrictions and Expiration Instructions

Order Instruction	Submission Numbers	Submission Volume	Account Agency	Account Proprietary
	in Panel Percents		in Row Percents	
PANEL A: ORDER RESTRICTION				
Immediate or Cancel (IOC)	18.5	15.3	14.1	85.9
	(5.3)	(6.7)	(9.8)	
Fill or Kill (FOK)	0.1	0.1	96.7	3.3
	(0.1)	(0.2)	(16.2)	
Stop Order	1.0	0.5	76.7	23.3
	(1.2)	(0.5)	(24.2)	
No Restriction	80.4	84.2	41.0	59.0
	(5.1)	(6.6)	(8.2)	
PANEL B: TRADE RESTRICTION				
Auction only	1.1	1.2	100.0	0.0
	(0.6)	(0.7)	(0.0)	
No Restriction	98.9	98.8	17.7	82.3
	(0.6)	(0.7)	(5.6)	
PANEL C: EXPIRE INSTRUCTIONS				
Day Order	99.2	99.3	15.7	84.3
	(0.6)	(0.6)	(5.2)	
Good until	0.7	0.6	93.2	6.8
	(0.6)	(0.6)	(8.2)	
Good till cancel (GTC)	0.1	0.1	60.5	39.5
	(0.1)	(0.1)	(37.2)	

Table 1.4 shows the usage of various order instructions. The second column reports the percentages in numbers of orders submitted, the third the percentages of the order volume submitted. On the right column four and five differentiate the usage of each instruction between agency and proprietary accounts. The upper panel summarizes the Order Restrictions, which include Stop Orders. Trade restricitions follow in the middle panel, where we summarize all restrictions for executions in auctions only. It comprises both the restriction to any type of auction and for specific auctions. The lower part reports the frequencies of the different expiration instructions. All figures are averages and standard deviations (in parentheses) based on the averages for each stock and trading day combination.

Table 1.5: Order Submission during Continuous Trading

	Small Spreads		Medium Spreads		Large Spreads		Whole Sample	
	Differences in Percent to Expected Count						Column Percent	
PANEL A: SMALL ORDERS								
Behind the Market	3.5	(0.5)	0.9	(0.4)	-9.0	(1.3)	18.8	(0.3)
At the Market	1.8	(1.3)	4.5	(0.8)	-7.6	(1.4)	9.7	(0.3)
In the Market	-2.3	(1.8)	0.6	(0.5)	1.1	(1.1)	4.7	(0.3)
Marketable (partially)	11.0	(0.7)	-37.2	(2.3)	-48.9	(3.0)	0.2	(0.0)
All	0.3	(0.5)	2.5	(0.4)	-3.8	(0.8)		
PANEL B: MEDIUM ORDERS								
Behind the Market	-0.8	(0.5)	-0.3	(0.2)	1.9	(1.1)	22.3	(0.5)
At the Market	-4.9	(1.4)	-0.1	(0.7)	7.6	(1.9)	7.1	(0.4)
In the Market	0.5	(1.5)	-4.2	(0.7)	2.6	(1.1)	3.6	(0.3)
Marketable (partially)	5.7	(0.5)	-19.4	(1.5)	-22.3	(2.1)	0.3	(0.0)
All	0.2	(0.4)	-0.9	(0.3)	0.6	(0.6)		
PANEL C: LARGE ORDERS								
Behind the Market	-2.2	(0.6)	-0.6	(0.4)	6.0	(1.4)	20.2	(0.5)
At the Market	0.3	(1.3)	-6.6	(0.8)	6.1	(1.6)	6.6	(0.3)
In the Market	1.0	(1.1)	1.9	(0.4)	-2.0	(0.6)	5.1	(0.4)
Marketable (partially)	-2.6	(0.3)	8.5	(0.6)	10.8	(1.0)	1.5	(0.1)
All	-0.5	(0.5)	-1.7	(0.4)	3.3	(0.7)		

Table 1.5 reports submission frequencies. Each limitorder submission — excluding fully marketable — is classified by order size (panels), spreads (columns) and relative limit position (rows). The right side (the last two columns) shows absolute percentages. Differences in percent (not percentage point!) to the expected counts for that combination of categories are on the left side (the first six columns). For the definition of the order size quantiles and spreads for each stock refer to table 1.14. Behind the market are limit orders going inside the book, at the market those adding to the best quote, and in the market submissions inside the spread. All figures are means and standard errors (in parentheses) based on the results for each stock.

Table 1.6: Execution Probabilities during Continuous Trading

	Small Spreads		Medium Spreads		Large Spreads		Whole Sample	
PANEL A: SMALL ORDERS								
Behind the Market	5.3	(0.4)	3.9	(0.3)	3.7	(0.4)	4.7	(0.4)
At the Market	27.2	(2.1)	15.4	(1.5)	9.9	(0.9)	19.6	(1.7)
In the Market	38.2	(2.1)	30.2	(2.2)	23.1	(2.3)	29.9	(2.2)
Marketable (partially)	75.8	(0.9)	82.9	(0.8)	89.0	(1.6)	77.0	(0.8)
All	14.6	(1.0)	11.4	(0.9)	10.0	(0.9)	12.6	(1.0)
PANEL B: MEDIUM ORDERS								
Behind the Market	3.8	(0.2)	2.8	(0.1)	2.6	(0.2)	3.3	(0.2)
At the Market	31.6	(0.9)	16.8	(0.6)	10.0	(0.6)	21.1	(0.9)
In the Market	42.6	(1.8)	34.7	(1.7)	25.2	(2.0)	33.2	(1.8)
Marketable (partially)	79.4	(0.8)	84.9	(0.7)	89.9	(0.7)	80.7	(0.7)
All	12.7	(0.6)	9.7	(0.3)	8.3	(0.6)	10.8	(0.5)
PANEL C: LARGE ORDERS								
Behind the Market	8.2	(0.5)	6.3	(0.3)	5.5	(0.4)	7.1	(0.4)
At the Market	40.4	(0.9)	26.6	(0.8)	16.0	(0.8)	29.8	(1.0)
In the Market	46.5	(2.4)	41.5	(1.9)	35.2	(2.1)	40.9	(2.1)
Marketable (partially)	81.6	(0.7)	87.2	(0.5)	90.2	(0.4)	83.3	(0.6)
All	22.7	(0.9)	18.3	(0.8)	15.3	(0.8)	19.6	(0.9)
ALL ORDERS								
Behind the Market	5.6	(0.3)	4.2	(0.2)	3.8	(0.2)	4.9	(0.3)
At the Market	32.9	(1.0)	18.8	(0.9)	11.8	(0.6)	23.2	(0.9)
In the Market	42.6	(1.8)	35.5	(1.6)	28.1	(2.1)	34.9	(1.9)
Marketable (partially)	80.7	(0.7)	86.6	(0.5)	90.1	(0.4)	82.3	(0.6)
All	16.7	(0.5)	13.1	(0.4)	11.2	(0.6)	14.3	(0.5)

Table 1.6 shows the execution probabilities differentiated by the limit order categories. Each limitorder submission — excluding fully marketable — is classified by order size (panels), spreads (columns), and relative limit position (rows). Based on the outcome each order is assigned an execution percentage ranging from zero for no execution to one for a full execution. For the definition of the order size quantiles and spreads for each stock refer to table 1.14. Behind the market are limit orders going inside the book, at the market those adding to the best quote, and in the market submissions inside the spread. All figures are means and standard errors (in parentheses) based on the results for each stock.

Table 1.7: Order Durations during Continuous Trading

	Small Spreads		Medium Spreads		Large Spreads		Whole Sample	
PANEL A: SMALL ORDERS								
Behind the Market	6.0	(1.1)	4.3	(0.4)	6.0	(1.1)	5.9	(1.1)
At the Market	2.6	(0.3)	1.7	(0.2)	1.6	(0.1)	2.1	(0.2)
In the Market	8.0	(0.7)	8.3	(0.8)	8.8	(1.0)	8.7	(0.9)
Marketable (partially)	1.0	(0.1)	1.4	(0.3)	0.7	(0.2)	1.1	(0.1)
All	4.9	(0.6)	4.1	(0.3)	5.2	(0.6)	5.0	(0.6)
PANEL B: MEDIUM ORDERS								
Behind the Market	4.2	(0.6)	3.3	(0.2)	4.0	(0.6)	4.1	(0.6)
At the Market	4.2	(0.4)	2.5	(0.2)	2.0	(0.1)	2.9	(0.2)
In the Market	11.6	(0.8)	12.3	(0.7)	11.8	(1.0)	12.6	(1.0)
Marketable (partially)	1.2	(0.1)	1.0	(0.2)	0.6	(0.1)	1.1	(0.1)
All	4.3	(0.4)	4.0	(0.2)	4.7	(0.4)	4.4	(0.4)
PANEL C: LARGE ORDERS								
Behind the Market	8.9	(0.8)	8.3	(0.7)	9.6	(0.9)	9.1	(0.8)
At the Market	7.0	(0.4)	5.5	(0.2)	4.2	(0.2)	5.7	(0.2)
In the Market	13.1	(1.0)	14.0	(1.0)	12.9	(0.7)	13.5	(0.9)
Marketable (partially)	2.7	(0.2)	2.2	(0.2)	1.7	(0.2)	2.6	(0.2)
All	8.0	(0.5)	8.4	(0.5)	8.7	(0.5)	8.4	(0.5)
ALL ORDERS								
Behind the Market	6.2	(0.7)	5.1	(0.3)	6.4	(0.7)	6.2	(0.7)
At the Market	4.2	(0.3)	2.9	(0.1)	2.5	(0.1)	3.3	(0.1)
In the Market	10.9	(0.7)	11.5	(0.7)	11.1	(0.8)	11.6	(0.8)
Marketable (partially)	2.3	(0.2)	2.0	(0.2)	1.5	(0.2)	2.2	(0.2)
All	5.7	(0.4)	5.5	(0.2)	6.3	(0.4)	5.9	(0.4)

Table 1.7 provides the results on the duration in Minutes of limit order submissions until execution or cancellation. Each limitorder submission, except full marketable, is classified by order size (panels), spreads (columns), and relative limit position (rows). For the definition of the order size quantiles and spreads for each stock refer to table 1.14. Behind the market are limit orders going inside the book, at the market those adding to the best quote, and in the market submissions inside the spread. All figures are means and standard errors (in parentheses) based on the results for each stock.

Table 1.8: Ex-ante Performance during Continuous Trading

	Small Spreads		Medium Spreads		Large Spreads		Whole Sample	
PANEL A: SMALL ORDERS								
Behind the Market	-0.7	(0.1)	0.5	(0.1)	1.8	(0.1)	0.1	(0.0)
At the Market	-0.2	(0.1)	0.7	(0.1)	1.5	(0.1)	0.5	(0.0)
In the Market	0.3	(0.1)	1.2	(0.1)	2.6	(0.2)	1.8	(0.2)
Marketable (partially)	4.0	(0.3)	4.3	(0.3)	5.2	(0.5)	4.1	(0.3)
All	-0.5	(0.1)	0.7	(0.1)	1.9	(0.1)	0.5	(0.0)
PANEL B: MEDIUM ORDERS								
Behind the Market	-0.5	(0.1)	0.7	(0.1)	2.1	(0.2)	0.4	(0.1)
At the Market	-0.2	(0.1)	0.9	(0.1)	1.7	(0.1)	0.7	(0.1)
In the Market	0.6	(0.2)	1.8	(0.2)	3.0	(0.2)	2.2	(0.2)
Marketable (partially)	4.0	(0.3)	4.3	(0.4)	5.6	(0.5)	4.1	(0.3)
All	-0.3	(0.1)	0.9	(0.1)	2.1	(0.2)	0.7	(0.1)
PANEL C: LARGE ORDERS								
Behind the Market	-0.7	(0.1)	0.7	(0.1)	2.5	(0.2)	0.5	(0.1)
At the Market	-0.1	(0.1)	1.2	(0.1)	2.4	(0.2)	0.9	(0.1)
In the Market	0.5	(0.2)	1.7	(0.2)	3.5	(0.3)	2.4	(0.3)
Marketable (partially)	4.3	(0.3)	4.8	(0.4)	6.2	(0.5)	4.5	(0.3)
All	-0.1	(0.1)	1.1	(0.1)	2.8	(0.2)	1.0	(0.1)
ALL ORDERS								
Behind the Market	-0.6	(0.1)	0.6	(0.1)	2.2	(0.2)	0.3	(0.0)
At the Market	-0.2	(0.1)	0.9	(0.1)	1.8	(0.1)	0.7	(0.1)
In the Market	0.5	(0.2)	1.5	(0.2)	3.1	(0.2)	2.1	(0.2)
Marketable (partially)	4.2	(0.3)	4.7	(0.4)	6.1	(0.5)	4.4	(0.3)
All	-0.3	(0.1)	0.9	(0.1)	2.3	(0.2)	0.7	(0.1)

Table 1.8 reports the ex-ante performance in basis point. Each limit order submission, excluding those which are fully marketable, is classified by order size (panels), spreads (columns), and relative limit position (rows). The ex-ante performance is defined as

$$p^{\text{ex-ante}} = \begin{cases} p^{\text{potential}}_{\text{at submission}} - p^{\text{fill}} & \text{for submitted buy orders} \\ p^{\text{fill}} - p^{\text{potential}}_{\text{at submission}} & \text{for submitted sell orders} \end{cases}$$

$$p^{\text{fill}} = \begin{cases} p^{\text{execution}} & \text{for executed orders} \\ p^{\text{potential}}_{\text{at cancellation}} & \text{for cancelled orders} \\ p^{\text{auction}}_{\text{at closing}} & \text{orders still in the book at closing} \end{cases}$$

All figures are means and standard errors (in parentheses) based on the results for each stock.

Table 1.9: Ex-post Performance during Continuous Trading

	Small Spreads		Medium Spreads		Large Spreads		Whole Sample	
PANEL A: SMALL ORDERS								
Behind the Market	0.2	(0.3)	0.1	(0.4)	2.0	(0.7)	0.5	(0.3)
At the Market	-1.1	(0.2)	-0.1	(0.3)	1.5	(0.4)	-0.6	(0.2)
In the Market	-1.2	(0.2)	-1.0	(0.2)	0.0	(0.2)	-0.7	(0.2)
Marketable (partially)	-0.5	(0.6)	-5.5	(2.1)	-4.8	(1.6)	-1.3	(0.5)
All	-0.8	(0.2)	-0.5	(0.2)	0.8	(0.3)	-0.4	(0.2)
PANEL B: MEDIUM ORDERS								
Behind the Market	0.4	(0.3)	0.6	(0.3)	2.1	(0.6)	0.7	(0.3)
At the Market	-0.7	(0.2)	0.4	(0.2)	1.6	(0.3)	-0.2	(0.2)
In the Market	-1.0	(0.2)	-0.7	(0.2)	0.3	(0.2)	-0.4	(0.1)
Marketable (partially)	-1.1	(0.3)	-3.6	(1.1)	-5.7	(1.0)	-1.7	(0.3)
All	-0.5	(0.2)	-0.1	(0.1)	0.8	(0.2)	-0.1	(0.1)
PANEL C: LARGE ORDERS								
Behind the Market	-0.9	(0.3)	-0.4	(0.3)	0.9	(0.5)	-0.4	(0.3)
At the Market	-0.8	(0.1)	-0.1	(0.2)	0.9	(0.3)	-0.4	(0.1)
In the Market	-1.2	(0.2)	-0.7	(0.2)	-0.2	(0.2)	-0.6	(0.1)
Marketable (partially)	-1.2	(0.2)	-2.8	(0.3)	-5.6	(0.5)	-1.9	(0.2)
All	-1.0	(0.1)	-0.7	(0.1)	-0.1	(0.2)	-0.7	(0.1)
ALL ORDERS								
Behind the Market	-0.2	(0.3)	0.0	(0.3)	1.6	(0.5)	0.2	(0.3)
At the Market	-0.9	(0.2)	0.1	(0.2)	1.4	(0.2)	-0.3	(0.1)
In the Market	-1.1	(0.1)	-0.8	(0.1)	0.0	(0.1)	-0.5	(0.1)
Marketable (partially)	-1.2	(0.2)	-2.9	(0.3)	-5.7	(0.4)	-1.8	(0.2)
All	-0.8	(0.2)	-0.4	(0.1)	0.4	(0.2)	-0.4	(0.1)

Table 1.9 summarizes the ex-post performance in basis point for each limit order execution. Each submission, that is fully but not immediate executed, is classified by order size (panels), spreads (columns), and relative limit position (rows). The ex-post performance is defined as

$$p^{\text{ex-post}} = \begin{cases} p^{\text{midquote}}_{\text{execution+10min}} - p^{\text{execution}} & \text{for executed buy orders} \\ p^{\text{execution}} - p^{\text{midquote}}_{\text{execution+10min}} & \text{for executed sell orders} \end{cases}$$

All figures are means and standard errors (in parentheses) based on the results for each stock.

Table 1.10: Order Submissions in the Opening Auction

	Pretrading		Auction Call								Overall	
			Start		Middle		End		Random			
PANEL A: SMALL ORDERS												
Behind the Market	9.0	(0.7)	2.6	(0.2)	2.1	(0.1)	3.4	(0.2)	0.5	(0.0)	17.7	(0.6)
Marketable	3.9	(0.3)	1.8	(0.1)	1.9	(0.1)	2.1	(0.1)	1.2	(0.1)	11.1	(0.2)
Market Order	9.1	(0.7)	4.3	(0.3)	4.1	(0.4)	2.6	(0.2)	1.1	(0.1)	21.2	(0.6)
All	22.1	(1.6)	8.8	(0.5)	8.1	(0.5)	8.1	(0.5)	2.8	(0.2)	49.9	(0.0)
PANEL B: LARGE ORDERS												
Behind IndicPrice	3.1	(0.2)	3.8	(0.1)	4.6	(0.1)	5.4	(0.2)	0.9	(0.1)	17.8	(0.5)
Marketable	0.9	(0.0)	5.0	(0.1)	4.7	(0.1)	4.9	(0.1)	2.7	(0.1)	18.2	(0.4)
Market Order	2.0	(0.1)	3.9	(0.1)	4.1	(0.1)	3.1	(0.2)	1.0	(0.1)	14.1	(0.4)
All	6.0	(0.3)	12.7	(0.1)	13.4	(0.1)	13.4	(0.2)	4.6	(0.2)	50.1	(0.0)
ALL ORDERS												
Behind IndicPrice	12.1	(0.8)	6.5	(0.2)	6.7	(0.1)	8.8	(0.3)	1.4	(0.1)	35.5	(0.7)
Marketable	4.9	(0.3)	6.9	(0.2)	6.6	(0.2)	7.0	(0.2)	3.8	(0.2)	29.3	(0.5)
Market Order	11.1	(0.7)	8.2	(0.3)	8.2	(0.3)	5.7	(0.1)	2.1	(0.1)	35.2	(0.5)
All	28.1	(1.7)	21.5	(0.5)	21.5	(0.5)	21.5	(0.5)	7.4	(0.3)	.	.

Table 1.10 summarizes the order submissions during the opening auction in Xetra. The figures shown are percentages of the overall order numbers. We report the mean and standard error (in parentheses) treating the results for each stock as one observation. Every order submission is categorized by three dimensions: The order size is differentiated in the two panels A and B, the columns indicate the time of submission and the rows in each panel define the price relative to the prevailing indicative price. The order sizes are separated by the median order volume for each stock. The subdivision of the auction call into the three periods start, middle, end are based on equal order numbers for each period. Pretrading and random end are defined by the auction mechanism. The position compared to the indicative price is either behind (order would not execute at the current indicative price), marketable (order would execute at the indicative price) or the order is a market order (order executes in any case). The exchange does not display an indicative price during Pretrading, thus traders do not know their position relative to the indicative price at submission.

Table 1.11: Executions and Cancellations in the Opening Auction

	Trading Phase	Pretrading	Auction Call Start	Auction Call Middle	Auction Call End	Auction Call Random	Overall
PANEL A: SMALL ORDERS							
Behind the Market	Auction	18/29	7/53	10/40	4/20	3/18	11/34
	Cont. Time	33/20	17/22	19/30	11/64	23/56	23/30
Marketable	Auction	66/10	33/44	34/46	40/38	58/35	48/31
	Cont. Time	19/4	10/12	8/11	7/14	4/3	11/8
Market Order	Auction	94/6	60/40	54/46	63/37	90/10	73/27
	Cont. Time	0/0	0/0	0/0	0/0	0/0	0/0
All	Auction	58/17	37/45	36/45	31/31	59/23	45/31
	Cont. Time	16/9	8/10	7/11	6/30	6/12	11/12
PANEL B: LARGE ORDERS							
Behind the Market	Auction	9/27	6/56	9/39	4/26	3/27	7/36
	Cont. Time	32/32	13/25	12/34	9/38	21/48	15/34
Marketable	Auction	39/27	16/71	23/59	31/47	47/41	28/55
	Cont. Time	22/12	5/8	5/11	5/14	6/6	6/10
Market Order	Auction	86/14	35/65	32/68	51/49	85/15	49/51
	Cont. Time	0/0	0/0	0/0	0/0	0/0	0/0
All	Auction	41/22	19/65	21/55	25/39	46/33	26/48
	Cont. Time	19/18	6/11	6/16	5/20	8/13	7/16

Remark: Execution Percent / Cancel Percent

Table 1.11 shows the execution probabilities differentiated by the categories defined for auction submissions. Each order of the opening auction is classified by order size (panels), time of submission (columns), and relative limit position (rows). For every combination of the categories four percentages are reported in 2x2 quadrant, clockwise starting at the upper left: execution in auction, cancellation in auction, cancellation in continuous time, execution in continuous time. For the definition of the order size quantiles and submission time for each stock refer to table 1.15. Behind the market are limit orders that would not execute given the prevailing indicative price, marketable are those that would execute and market orders (usually) have a guaranteed execution. All figures are means and standard errors (in parentheses) based on the results for each stock.

Table 1.12: Ex-ante Performance Opening Auction

	Pretrading		Auction Call								Overall	
			Start		Middle		End		Random			
PANEL A: SMALL ORDERS												
Behind the Market	43.5	(2.4)	25.0	(2.3)	15.3	(2.3)	-4.0	(1.9)	10.0	(3.5)	25.8	(2.5)
Marketable	10.0	(0.7)	2.2	(1.4)	-1.8	(1.4)	-5.4	(1.0)	-0.9	(0.5)	3.1	(0.6)
Market Order	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)
All	16.8	(1.1)	6.8	(0.9)	4.2	(0.8)	-3.2	(1.0)	2.3	(0.8)	9.4	(1.1)
PANEL B: LARGE ORDERS												
Behind the Market	45.1	(3.4)	21.5	(2.2)	5.9	(1.5)	-1.1	(1.7)	3.8	(2.7)	14.5	(1.7)
Marketable	11.6	(1.4)	2.8	(1.4)	-3.7	(1.4)	-7.5	(1.1)	-3.9	(0.6)	-2.4	(0.8)
Market Order	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)
All	22.6	(1.7)	9.0	(1.1)	1.3	(0.9)	-3.6	(1.1)	-1.0	(0.8)	5.1	(0.9)
ALL ORDERS												
Behind the Market	43.5	(2.4)	22.7	(1.8)	8.7	(1.4)	-3.0	(1.6)	6.4	(2.0)	20.5	(2.0)
Marketable	10.1	(0.6)	2.7	(1.2)	-3.2	(1.3)	-6.8	(1.0)	-2.9	(0.5)	0.2	(0.7)
Market Order	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)
All	17.8	(1.1)	7.9	(0.9)	2.5	(0.8)	-3.5	(1.0)	0.2	(0.5)	7.6	(1.0)

Table 1.12 reports the ex-ante performance in basis point. Each order submitted to the opening auction (except those which are cancelled during the auction call) is classified by order size (panels), time of submission (columns), and relative limit position (rows). The ex-ante performance in the opening auction is defined as

$$p^{\text{ex-ante}} = \begin{cases} p^{\text{auction}} - p^{\text{fill}} & \text{for buy orders that still live at auction execution} \\ p^{\text{fill}} - p^{\text{auction}} & \text{for sell orders that still live at auction execution} \end{cases}$$

$$p^{\text{fill}} = \begin{cases} p^{\text{auction}} & \text{for orders executed in the auction} \\ p^{\text{execution}} & \text{for executed orders in continuous trading} \\ p^{\text{potential}}_{\text{at cancellation}} & \text{for cancelled orders in continuous trading} \end{cases}$$

For the definition of the order size quantiles and submission time for each stock refer to table 1.15. Behind the market are limit orders that would not execute given the prevailing indicative price, marketable are those that would execute and market orders (usually) have a guaranteed execution. All figures are means and standard errors (in parentheses) based on the results for each stock.

Table 1.13: Ex-post Performance Opening Auction

	Pretrading		Auction Call								Overall	
			Start		Middle		End		Random			
PANEL A: SMALL ORDERS												
Behind the Market	6.5	(3.8)	6.8	(2.8)	4.8	(4.5)	16.2	(4.0)	10.6	(15.9)	5.5	(2.9)
Marketable	6.4	(2.0)	2.9	(2.1)	6.7	(3.0)	9.2	(1.7)	5.2	(2.7)	6.1	(1.7)
Market Order	-0.9	(1.1)	-1.8	(1.5)	1.4	(1.0)	1.2	(1.2)	2.9	(1.5)	-1.2	(0.9)
All	1.5	(1.2)	-0.4	(1.3)	2.9	(0.9)	4.6	(1.2)	4.2	(1.3)	1.0	(0.8)
PANEL B: LARGE ORDERS												
Behind the Market	4.1	(5.0)	-6.3	(5.2)	1.9	(4.5)	11.9	(3.1)	4.8	(11.6)	2.5	(2.7)
Marketable	3.6	(3.1)	1.3	(2.2)	0.5	(1.8)	10.6	(2.2)	8.9	(3.2)	5.5	(1.7)
Market Order	-4.9	(1.7)	1.6	(1.3)	4.4	(1.3)	3.3	(1.4)	4.5	(2.6)	1.1	(1.1)
All	-3.1	(1.4)	0.7	(1.0)	2.5	(1.1)	7.3	(1.2)	7.5	(2.1)	2.8	(0.7)
ALL ORDERS												
Behind the Market	6.4	(3.8)	-0.9	(3.2)	2.2	(3.1)	13.0	(2.6)	14.3	(10.5)	4.4	(2.6)
Marketable	5.9	(2.0)	2.2	(1.4)	2.6	(1.4)	10.5	(1.7)	7.3	(2.5)	6.0	(1.5)
Market Order	-2.3	(0.9)	-0.7	(1.3)	2.7	(0.8)	2.2	(1.1)	3.7	(1.6)	-0.5	(0.8)
All	0.3	(0.9)	-0.0	(1.0)	2.7	(0.7)	6.0	(1.0)	5.2	(1.3)	1.6	(0.6)

The ex-post performance in basis points is reported in Table 1.13 above. Each order submitted to the opening auction and executed either in the auction itself or in continuous trading after the end of the auction is classified by order size (panels), time of submission (columns), and relative limit position (rows). The ex-post performance in the opening auction is defined as

$$p^{\text{ex-post}} = \begin{cases} p^{\text{midquote}}_{\text{execution+10min}} - p^{\text{execution}} & \text{for executed buy orders} \\ p^{\text{execution}} - p^{\text{midquote}}_{\text{execution+10min}} & \text{for executed sell orders} \end{cases}$$

For the definition of the order size quantiles and submission time for each stock refer to table 1.15. Behind the market are limit orders that would not execute given the prevailing indicative price, marketable are those that would execute and market orders (usually) have a guaranteed execution. All figures are means and standard errors (in parentheses) based on the results for each stock.

Table 1.14: Quantiles Continuous Trading

Trading Activity Group	Ticker Symbol	Order Size Small	Order Size Middle [in 1000 €]	Order Size Large	Spread Small	Spread Middle [in Cent]	Spread Large
1	ALV	3-32	32-52	52-291	1-4	5-7	8-17
(High)	DBK	3-32	32-60	60-355	1-2	3-5	6-11
	DCX	3-31	31-56	56-259	1-2	3-3	4-7
	DTE	1-36	36-81	81-638	1-1	.-.	2-3
	MUV2	3-20	20-42	42-278	1-4	5-7	8-19
	SAP	3-29	29-55	55-328	1-6	7-10	11-27
	SIE	3-36	36-66	66-334	1-2	3-4	5-9
	Mean	3-31	31-59	59-355	1-3	4-6	6-13
2	BAS	3-25	25-45	45-209	1-2	3-3	4-9
	BAY	2-17	17-33	33-171	1-1	2-2	3-6
	BMW	2-15	15-27	27-178	1-2	3-3	4-8
	EOA	3-32	32-47	47-262	1-2	3-4	5-10
	HVM	2-10	10-19	19-250	1-2	3-3	4-7
	IFX	1-9	9-31	31-305	1-1	.-.	2-3
	RWE	3-15	15-29	29-180	1-2	3-3	4-8
	VOW	2-14	14-29	29-198	1-2	3-3	4-9
	Mean	2-17	17-33	33-219	1-2	3-3	4-8
3	ADS	2-9	9-24	24-140	1-6	7-11	12-35
	CBK	2-10	10-19	19-156	1-1	2-2	3-5
	DB1	2-5	5-14	14-152	1-3	4-6	7-18
	DPW	2-9	9-19	19-182	1-1	2-3	4-7
	LHA	1-7	7-15	15-144	1-1	2-2	3-6
	MEO	2-7	7-17	17-164	1-3	4-5	6-17
	SCH	2-9	9-21	21-176	1-3	4-5	6-14
	TKA	2-8	8-17	17-174	1-2	3-3	4-7
	Mean	2-8	8-18	18-161	1-3	4-5	6-14
4	ALT	2-5	5-15	15-134	1-4	5-7	8-22
(Low)	CONT	2-6	6-16	16-100	1-3	4-5	6-16
	FME	1-5	5-12	12-85	1-5	6-10	11-33
	HEN3	2-7	7-13	13-98	1-5	6-9	10-26
	LIN	1-5	5-13	13-152	1-3	4-6	7-18
	MAN	2-3	3-13	13-144	1-3	4-5	6-15
	TUI	2-4	4-10	10-185	1-2	3-3	4-9
	Mean	2-5	5-13	13-128	1-4	5-6	7-20
All	Mean	2-15	15-30	30-214	1-3	4-5	6-13

Table 1.14 above details the creation of quantiles to categorize the limit order submissions during continuous trading. The stocks are divided by turnover into four activity groups. For each stock the order sizes are divided in three groups, after trimming the upper and lower one percentiles of order sizes. Columns three to five show the resulting ranges in Euro. The spreads are treated similar, but only the upper one percentiles is trimmed. Results are given in columns six to eight. DTE and IFX have that many observations at spreads of one and two cents, that the middle quantile is omitted.

Table 1.15: Quantiles Opening Auction

Trading Activity Group	Ticker Symbol	Daily Order Numbers	P 5	Order Size Median [in 1000 €]	P 95	Call duration Start [in Minutes]	Middle	End
1	ALV	219	1.2	45.2	433.1	5:21	3:03	1:35
(High)	DBK	186	1.5	70.8	505.6	5:23	3:06	1:31
	DCX	179	1.1	56.1	359.4	5:33	2:59	1:28
	DTE	289	0.5	21.5	442.2	5:06	3:05	1:49
	MUV2	182	1.7	43.8	359.6	5:25	3:03	1:32
	SAP	141	1.9	64.2	392.6	5:34	2:52	1:33
	SIE	208	1.3	59.4	610.3	5:27	2:58	1:35
	Mean	201	1.3	51.6	443.3	5:24	3:01	1:35
2	BAS	139	2.1	56.4	288.5	5:26	3:04	1:30
	BAY	163	1.1	38.4	234.1	5:17	2:55	1:47
	BMW	143	0.8	26.1	180.0	5:40	2:50	1:30
	EOA	147	1.6	53.1	404.1	5:33	3:01	1:25
	HVM	116	0.9	24.2	204.4	5:47	2:43	1:29
	IFX	245	0.8	16.3	292.3	5:29	2:54	1:37
	RWE	106	0.9	34.2	194.4	5:23	3:08	1:29
	VOW	124	1.7	32.2	205.3	5:37	2:56	1:27
	Mean	148	1.2	35.1	250.4	5:31	2:56	1:32
3	ADS	79	1.3	19.4	140.3	5:46	2:55	1:18
	CBK	97	1.4	28.2	159.5	5:33	2:52	1:35
	DB1	66	2.4	24.1	199.2	5:47	2:52	1:20
	DPW	94	0.6	18.9	173.6	5:45	2:45	1:29
	LHA	104	0.8	21.8	154.9	5:49	2:49	1:22
	MEO	86	2.1	23.9	165.0	5:48	2:41	1:30
	SCH	92	1.7	32.5	200.0	5:44	2:53	1:23
	TKA	87	1.0	16.9	142.7	5:40	2:48	1:32
	Mean	88	1.4	23.2	166.9	5:44	2:49	1:26
4	ALT	75	2.1	17.5	128.5	5:23	3:10	1:27
(Low)	CONT	70	1.6	19.7	129.6	5:48	2:54	1:17
	FME	55	2.5	9.1	105.4	5:47	2:57	1:16
	HEN3	58	3.4	18.6	132.7	5:14	3:10	1:35
	LIN	63	2.2	21.8	129.6	5:56	2:49	1:14
	MAN	64	1.9	13.7	135.5	5:46	2:51	1:22
	TUI	104	1.0	11.0	96.4	5:31	2:51	1:38
	Mean	70	2.1	15.9	122.5	5:38	2:57	1:24
All	Mean	126	1.5	31.3	243.3	5:35	2:56	1:29

Table 1.15 provides a breakdown of the quantiles created to group the order submissions of the opening auction. The stocks are divided by turnover into four activity groups. For each stock the order sizes are divided in two groups after trimming the upper and lower five percentiles of order sizes. Columns three to five show the resulting ranges in Euro. The ten minutes duration of the auction call is divided into periods of equal numbers of orders submitted. The durations for the three phases are detailed in columns six to eight.

Two

Liquidity supply and adverse selection

This chapter analyzes adverse selection costs and liquidity supply in a pure open limit order book market. We relax assumptions of the Glosten/Såndas modeling framework regarding marginal zero profit order book equilibrium and the parametric market order size distribution. We show that using average zero profit conditions considerably increases the empirical performance while a nonparametric specification for market order size combined with marginal zero profit conditions does not. A cross sectional analysis corroborates the finding that adverse selection costs are more severe for smaller capitalized stocks. We also find additional support for one of the central hypothesis put forth by the theory of limit order book markets, which states that liquidity supply and adverse selection costs are inversely related. Furthermore, adverse selection cost estimates based on the structural model and those obtained using popular model-free methods are strongly correlated. This indicates the robustness of the theory-based approach.

2.1 Introduction

Ten years after the question phrased in Glosten's (1994) celebrated paper: 'Is the electronic order book inevitable?' seems to be answered, given the triumphal procession of open order book systems in Continental Europe and recent developments in US stock markets.[1] A central feature of a pure limit order book market is the absence of dedicated market makers. Liquidity is supplied voluntarily by patient market participants

[1] In January 2002 the New York Stock Exchange (NYSE), known as a hybrid specialist market, adopted the key feature of electronic order book markets, namely the public display of all limit orders (NYSE open book program).

who provide an inflow of limit buy and sell orders, the lifeblood of the trading process. The non-executed orders constitute the limit order book, the consolidated source of liquidity. As the viability and resiliency of such a market structure is in the interest of regulators, operators and individual investors it is not surprising that theoretical and empirical studies of limit order markets abound in the literature.[2] However, theoretical models explaining liquidity supply and demand in limit order book markets have not been very successful when confronted with real world order book data. Sandås (2001) extends the methodology proposed by De Jong et al. (1996) and estimates a version of Glosten's (1994) limit order book model allowing for real world features like discrete price ticks and time priority rules. The empirical results obtained using data from the Swedish stock exchange were not encouraging. Formal specification tests reject the model, transaction costs estimates are significantly negative, and book depth is systematically overestimated.

This chapter shows how some potentially restrictive assumptions in the Glosten/ Såndas framework can be relaxed, while retaining suitable moment conditions for GMM estimation. We show that the revised econometric methodology considerably improves the empirical performance. The alternative approach is employed in a cross sectional analysis of adverse selection costs and liquidity supply in a limit order market.

Given the discontenting results reported in the previous literature, it is not surprising that many recent empirical papers analyzing limit order book market data have severed the close connection to the theoretical framework. Extending the approach of the early papers by Biais et al. (1995), Hall et al. (2003), Coppejans et al. (2003), Cao et al. (2004), Grammig et al. (2004), Pascual and Veredas (2004) and Ranaldo (2004) employ discrete choice and count data models to analyze the determinants of order submission activity and the interaction of liquidity supply and demand processes in limit order markets. Beltran et al. (2004) advocate a principal components aproach to extract latent factors that explain the state of the order book. Gomber et al. (2004) and Degryse et al. (2003) conduct intra-day event studies to analyze the resiliency of limit order markets. These papers interpret the empirical results in the light of predic-

[2]Traditionally, market microstructure theory focussed on quote driven markets with one or more market makers (see O'Hara (1995) for an overview). Recent papers by Parlour (1998), Seppi (1997) Foucault (1999) and Foucault et al. (2003) have changed the focus to the analysis of price and liquidity processes in order book markets.

tions of microstructure models. However, a structural interpretation of the parameter estimates cannot be delivered.

This chapter returns to the theoretical basis for the empirical analysis of limit order book markets. We hypothesize that the discontenting empirical model performance is due to the following problems. First, the real world trading process might be organized in a way that deviates too much from the theoretical framework. Second, some of the underlying theoretical model's assumptions might be too restrictive. The Glosten/Såndas model imposes a zero expected profit condition for order book equilibrium which may not hold in a very active order market with discrete price ticks and time priority rules. Furthermore, the parametric distribution of market order sizes assumed by Såndas (2001), though leading to convenient closed form liquidity supply equations and GMM moment conditions, might be misspecified. Hasbrouck (2004) conjectures that the latter is responsible for the empirical failure of the model.

The original methodological contribution of this chapter is to propose alternative estimation strategies which relax some allegedly restrictive assumptions in the Glosten/ Såndas framework. First, we show that the parametric distributional assumption about market order sizes can be abandoned in favor of a straightforward nonparametric alternative that still delivers convenient closed form unconditional moment restrictions that can be used for GMM estimation. Second, we motivate a set of alternative set of moment conditions which replace the zero expected marginal profit conditions used by Såndas (2001). These moment conditions, referred to as average break even conditions, are derived from the assumption that the expected profit of the orders placed on a specific quote is zero.

We estimate the model using both the standard and the revised methodology based on reconstructed order book data from the Xetra electronic order book system which operates at various European exchanges. The data are tailor-made for the purpose of this chapter since the trading protocol closely corresponds to the theoretical trading process from which the moment conditions used for the empirical methodology are derived.

We show that using average break even conditions instead of marginal break even conditions delivers a much better empirical performance. Encouraged by this result, we employ the methodology in a cross sectional analysis of adverse selection effects and liquidity in the Xetra limit order market. This is the original empirical contribution of this chapter. The main results can be summarized as follows. First, we provide new evidence, from a limit order market, that adverse selection effects are more severe for smaller capitalized, less frequently traded stocks. This corroborates the results of previous papers dealing with different theoretical backgrounds, empirical methodologies, and market structures. Second, the empirical results support one of the main hypothesis of the theory of limit order markets, namely that book liquidity and adverse selection effects are inversely related. Finally, we compare the adverse selection components implied by the structural model estimates with popular ad hoc measures which are based on a comparison of effective and realized spreads. The latter approach is model-free, frequently used in practice and academia (see e.g. Boehmer (2004) and SEC (2001)) and requires publicly available trade and quote data only. The first approach is based on a structural model and permits an economic interpretation of the structural parameters, but the demand on the data is higher as reconstructed order books are needed. We show that both methodologies lead to quite similar conclusions. This result indicates the robustness of the structural model approach. It also provides a theoretical underpinning for using the ad-hoc method for the analysis of limit order data.

The remainder of the chapter is organized as follows. Section 2.2 describes the market structure and data. Section 2.3 discusses the theoretical background and develops the empirical methodology. The empirical results are discussed in section 2.4. Section 2.5 concludes with a summary and an outlook for further research.

2.2 Market structure and data

The Xetra open limit order book system

In the empirical analysis We use data from the automated auction system Xetra which operates at various European trading venues, like the Vienna Stock Exchange, the Irish Stock Exchange, the Frankfurt Stock Exchange (FSE) and the European Energy Exchange.[3] Xetra is a pure open order book system developed and maintained by the German Stock Exchange. It has operated since 1997 as the main trading platform for

[3]The Xetra technology was recently licensed to the Shanghai Stock Exchange, China's largest stock exchange.

German blue chip stocks at the FSE. Since the Xetra/FSE trading protocol is the data generating process for this study we will briefly describe its important features.[4]

Between an opening and a closing call auction - and interrupted by another mid-day call auction - Xetra/FSE trading is based on a continuous double auction mechanism with automatic matching of orders based on the usual rules of price and time priority. During pre- and post-trading hours it is possible to enter, revise and cancel orders, but order executions are not conducted, even if possible. During the year 2004, the Xetra/FSE hours extended from 9 a.m. C.E.T to 5.30 p.m. C.E.T. For blue chip stocks there are no dedicated market makers like the Specialists at the New York Stock Exchange (NYSE) or the Tokyo Stock Exchange's Saitori. For some small capitalized stocks listed in Xetra there may exist so-called Designated Sponsors - typically large banks - who are required to provide a minimum liquidity level by simultaneously submitting competitive buy and sell limit orders. In addition to the traditional limit and market orders, traders can submit so-called iceberg (or hidden) orders. An iceberg order is similar to a limit order in that it has pre-specified limit price and volume. The difference is that a portion of the volume is kept hidden from the other traders and is not visible in the open book.

Market orders and marketable limit orders which exceed the volume at the best quote are allowed to 'walk up the book'.[5] In other words, market orders are guaranteed immediate full execution, at the cost of incurring a higher price impact on the trades. This is one of the key features of the stylized theoretical trading environment upon which the econometric modeling is based, but which may not necessarily be found in the real world trading process.[6]

Xetra/FSE faces some local, regional and international competition for order flow. The FSE maintains a parallel floor trading system, which bears some similarities with the NYSE, and, like in the US, some regional exchanges participate in the hunt for

[4]The Xetra trading system resembles in many features other important limit order book markets around the world like Euronext, the joint trading platform of the Amsterdam, Brussels, Lisbon and Paris stock exchanges, the Hong Kong stock exchange described in Ahn et al. (2001), and the Australian stock exchange, described in Cao et al. (2004).

[5]A marketable limit order is a limit order with a limit price that makes it immediately executable against the current book. In my study, 'real' market orders (i.e. orders submitted without a upper or lower price limit) and marketable limit orders are treated alike. Henceforth, both real market orders and marketable limit orders are referred to as market orders.

[6]For example, Bauwens and Giot (2001) describe how the Paris Bourse's trading protocol converted the volume of a market order in excess of the depth at the best quote into a limit order at that price which enters the opposite side of the order book.

liquidity. Furthermore, eleven out of the thirty stocks we analyze in the empirical study are also cross listed at the NYSE, as an ADR or, in the case of Daimler/Chrysler, as a globally registered share. However, the electronic trading platform clearly dominates the regional and international competitors in terms of market shares, at least for the blue chip stocks that we study in the present chapter.

Data and descriptive analyses

The Frankfurt Stock Exchange granted access to a database containing complete information about Xetra open order book events (entries, cancelations, revisions, expirations, partial-fills and full-fills of market, limit and iceberg orders) which occurred during the first three months of 2004 (January, 2nd - March, 31st). The sample comprises the thirty German blue chip stocks constituting the DAX30 index. Based on the event histories we perform a real time reconstruction of the order book sequences. Starting from an initial state of the order book (supplied by the exchange), we track each change in the order book implied by entry, partial or full fill, cancelation and expiration of market, limit and iceberg orders in order to re-construct the order book at each point in time. The reconstruction procedure permits distinguishing the visible and the hidden part of the order book. The latter consists of the hidden part of the non-executed iceberg orders. To implement the empirical methodology outlined below, we take snapshots of the visible order book entries whenever a market order triggers an execution against the book.

Table 2.5 reports descriptive statistics of the cross section of stocks. The activity indicators show an active market. Averaged across stocks, about 13,000 non-marketable limit orders per stock are submitted each day. Among those, almost 11,000 get canceled before execution. This indicates that the limit order traders closely monitor the book for profit opportunities which is in fact one of the core assumptions of the underlying theoretical model. The large trade sizes (on average over 40,000 euros per trade) indicate that Xetra/FSE is a trading venue for institutional traders and not a retail market. Averaged across stocks, 2,100 trades are executed per day. Table 2.5 also reports average effective and realized spreads. Following Huang and Stoll (1996) the average effective spread is computed by taking two times the absolute difference of the transaction price of a trade (computed as average price per share) and the prevailing midquote and averaging over all trades of a stock. Realized spreads are computed similarly, but instead of taking the prevailing midquote, the midquote five minutes after the trade is used.[7] Note that in an open order book market like Xetra, there is no possibility to trade inside the bid-ask spread. Orders are either executed at the best quote or they walk up the book until they are completely filled. Table 2.5 shows that on average 15% of the order volume walks up the book, i.e. part of the order is matched by standing limit orders beyond the best bid and ask. This implies that the effective spread is then, by definition, larger than or equal to than the quoted spread. To ensure comparability across stocks, we compute effective and realized spreads relative to the midquote prevailing at the time of the trade. Analyzing effective and realized spreads is a straightforward way to assess and compare transaction costs and adverse selection effects across stocks or trading venues. The realized spread can be viewed as a transaction costs measure that is purged of informational effects while the difference of effective and realized spread (referred to as price impact) is a natural measure for the amount of informational content of the order flow.[8] Average effective spreads range from 0.04 % to 0.13%. Realized spreads are considerably smaller. This implies that price impacts, computed as the difference between effective and realized spreads, are relatively large. In other words, a large fraction of the spread is due to informational order flow. This is not an unexpected result. In an open automated auction market there is no justification for inventory costs associated with market making or monopolistic power of a market maker, the other factors that may explain the spread. Furthermore, order submission fees, i.e. operational costs, are very small.

Table 2.5 shows that there is a considerable variation of price impacts, market capitalization and trading activity across stocks. The Spearman rank correlation between market capitalization and price impacts is -0.88 ($p-$value < 0.001) and the correlation between price impacts and daily number of trades is -0.87 ($p-$value < 0.001). Price impacts thus tend to be larger for smaller capitalized, less frequently traded stocks. We will come back to this result when discussing the empirical results based on the structural model.

[7] By choosing a five minutes lag we follow the previous literature, see e.g. SEC (2001)

[8] Boehmer (2004) and SEC (2001) conduct exhaustive comparisons of transaction costs and adverse selection effects in US exchanges based on effective and realized spread analyses.

2.3 Methodology

Såndas' basic framework

Sandås (2001) develops a variant of Glosten's (1994) limit order book model with discrete price ticks and time priority rules. The model delivers equations which predict that order book depth and adverse selection effects are inversely related. The associated empirical methodology is rooted in economic theory, and delivers structural parameter estimates of transaction costs and adverse selection effects in a limit order book market. Below we will briefly describe the assumptions of the basic model and the estimation strategy proposed in Sandås (2001). The fundamental asset value X_t is described by a random walk with innovations depending on an adverse selection parameter α, which gives the informational content of a signed market order of size m_t,

$$X_{t+1} = \mu + X_t + \alpha m_t + \eta_{X,t+1}. \tag{2.1}$$

Negative values of m_t denote sell orders, positive values buy orders. Furthermore, it is assumed that $E(X_t) = 0$. $\eta_{X,t+1}$ is an innovation orthogonal to X_t. μ gives the expected change in the fundamental value. Market buy and sell orders are assumed to arrive with equal probability with a two-sided exponential density describing the distribution of order sizes m_t: [9]

$$f(m_t) = \begin{cases} \frac{1}{2\lambda} e^{\frac{-m_t}{\lambda}} & \text{if } \; m_t > 0 \;\; \text{(market buy)} \\ \frac{1}{2\lambda} e^{\frac{m_t}{\lambda}} & \text{if } \; m_t < 0 \;\; \text{(market sell).} \end{cases} \tag{2.2}$$

Risk neutral limit order traders face a order processing cost γ (per share) and have knowledge about the distribution of market order size and the adverse selection component α, but not about the true asset price. They choose limit order prices and quantities such that their expected profit is maximized. If the last unit at any discrete price tick exactly breaks even, i.e. has expected profit equal to zero, the order book is in equilibrium.

Denote the ordered discrete price ticks on the ask (bid) side by p_{+k} (p_{-k}) with $k = 1, 2, \ldots$ and the associated volumes at these prices by q_{+k} (q_{-k}). Given these as-

[9] In an alternative specification we allowed for additional flexibility by allowing the expected buy and sell market order sizes to be different. However, the parameter estimates and diagnostics changed only marginally. We therefore decided to stick to the specification in equation (2.2) which is more appealing both from a methodological and theoretical point of view.

sumptions and setting $q_{0,t} \equiv 0$, the equilibrium order book at time t can recursively be constructed as follows:

$$
\begin{aligned}
q_{+k,t} &= \frac{p_{+k,t} - X_t - \mu - \gamma}{\alpha} - Q_{+k-1,t} - \lambda \quad k = 1, 2, \ldots \quad \text{(ask side)} \\
q_{-k,t} &= \frac{X_t + \mu - p_{-k,t} - \gamma}{\alpha} - Q_{-k+1,t} - \lambda \quad k = 1, 2, \ldots \quad \text{(bid side)},
\end{aligned} \tag{2.3}
$$

where $Q_{+k,t} = \sum_{i=+1}^{+k} q_{i,t}$ and $Q_{-k,t} = \sum_{i=-1}^{-k} q_{i,t}$. Equation (2.3) contains the model's key message. Order book depth and informativeness of the order flow are inversely related. If the model provides a good description of the real world trading process, and if consistent estimates of the model parameters can be provided, one can use equation (2.3) to predict the evolution of the order book for a given stock and quantify adverse selection costs and their effect on order book depth.

Såndas (2001) proposes to employ GMM for parameter estimation and specification testing. Assuming mean zero random deviations from order book equilibrium at each price tick, and eliminating the unobserved fundamental asset value X_t by adding the resulting bid and ask side equations for quote $+k$ and $-k$, the following unconditional moment restrictions can be used for GMM estimation,

$$
E\left(p_{+k,t} - p_{-k,t} - 2\gamma - \alpha\left(Q_{k,t} + 2\lambda + Q_{-k,t}\right)\right) = 0 \quad k = 1, 2, \ldots. \tag{2.4}
$$

Since equation (2.4) follows from the assumption that the last (marginal) limit order at the respective quote has zero expected profit, it is referred to as 'marginal break even condition'. A second set of moment conditions results from eliminating X_t by subtracting the deviations from equilibrium depths at the kth quote at time $t+1$ and t and taking expectations which yields

$$
\begin{aligned}
E\left(\Delta p_{+k,t+1} - \alpha\left(Q_{k,t+1} - Q_{k,t}\right) - \mu - \alpha m_t\right) &= 0 \qquad k = 1, 2, \ldots \\
E\left(\Delta p_{-k,t+1} + \alpha\left(Q_{-k,t+1} - Q_{-k,t}\right) - \mu - \alpha m_t\right) &= 0 \qquad k = 1, 2, \ldots,
\end{aligned} \tag{2.5}
$$

where $\Delta p_{j,t+1} = p_{j,t+1} - p_{j,t}$. We refer to the equations in (2.5) as 'marginal update conditions'. They relate the expected changes in the order book to the market order flow. An obvious additional moment condition to identify the expected market order size is given by

$$
E(|X_t| - \lambda) = 0. \tag{2.6}
$$

Moment conditions (2.4), (2.5) and (2.6) can conveniently be exploited for GMM estimation a la Hansen (1982).

Såndas (2001) derives the moment conditions from the basic model setup outlined by Glosten (1994). Both Glosten's framework and Såndas' empirical implementation entail a set of potentially restrictive assumptions that may be problematic when confronting the model with real world data. Maybe the most crucial assumption of the Glosten framework is that limit order traders are assumed to be uninformed and that private information is only revealed through the arrival of market orders. Recent literature, however, suggests that limit orders may also be information-motivated Seppi (1997); Kaniel and Liu (2001); Cheung, de Jong, and Rindi (2003). Bloomfield, O'Hara, and Saar (2003) observe in an experimental limit order market that informed traders use more limit orders than liquidity traders. Since both break even and update conditions are derived from the assumption of uniformed limit order traders, the rejection of the model when confronted with real world data might be a result from a violation of this fundamental assumption.[10] Another important consideration is the number of active liquidity providers. Glosten(1994) assumes perfect competition. Biais, Martimort, and Rochet (2000) propose solutions for oligopolstic competition.

The following section proposes a revised set of moment conditions which are derived from a relaxation of the expected marginal profit condition and the parametric assumption of the market order distribution. However, we leave the basic assumption of uninformed limit order traders intact. Its relaxation would entail a fundamental revision of the theoretical base model. This is left for further research.

Revised moment conditions

Alternatives to the distributional assumption on market order sizes

Reviewing the Såndas/Glosten framework Hasbrouck (1994) conjectures that the parametric specification for the market order size distribution (2.2) may be incorrect.[11] Indeed, the plot of the empirical market order distribution against the fitted exponential densities depicted in figure 3 in Såndas (2001) sheds some doubt on this distributional assumption. To provide a formal assessment, we have employed the nonparametric

[10] We are grateful to a referee for pointing this out.

[11] It should be noted that the exponential assumption in DeJong et al.'s (1996) implementation of the Glosten model did not seem to be a restrictive assumption.

testing framework proposed by Fernandes and Grammig (2005) and found that the exponential distribution is rejected on any conventional level of significance for the sample of stocks. Hasbrouck (2004) argues that the misspecification of the exponential distribution could be responsible for the discontenting empirical results which have been reported when the model is confronted with real world data.

Of course, the exponential assumption is convenient both from a theoretical and an econometric perspective. It yields the closed form conditions for order book equilibrium (2.3) which, in turn, lend itself conveniently to GMM estimation. However, the parametric assumption can easily dispensed with and a straightforward nonparametric approach can be pursued for GMM estimation. In the appendix we show that the zero expected profit condition for the marginal unit at ask price p_{+k} can be written as

$$p_{+k} - \gamma - \alpha E\left[m|m \geq Q_{+k}\right] - X - \mu = 0.^{12} \tag{2.7}$$

Assuming of exponentially distributed market orders as in equation (2.2) We have $E\left[m|m \geq Q_{+k}\right] = Q_{+k} + \lambda$. Hence, equation (2.7) becomes

$$Q_{+k} = \frac{p_{+k} - X - \gamma - \mu}{\alpha}. \tag{2.8}$$

This is an alternative to equation (2.3) to describe order book equilibrium. Although the closed form expression implied by the parametric distributional assumption is convenient, it is not necessary for the econometric methodology to rely on it. Instead, we can rewrite equation (2.7) to obtain

$$E\left[m|m \geq Q_{+k}\right] = \frac{p_{+k} - X - \gamma - \mu}{\alpha}. \tag{2.9}$$

In order to utilize equation (2.9) for GMM estimation, one can simply replace $E\left[m|m \geq Q_{+k}\right]$ by the conditional sample means $\widehat{E}\left[m|m \geq Q_{+k}\right]$. Since the number of observations will be large for frequently traded stocks (which is the case in my application), conditional expectations can be precisely estimated by the conditional sample means. Nonparametric equivalents of the marginal break even and update conditions (2.4) and (2.5) can be derived in the same fashion as described in the previous section. GMM estimation is more computer intensive since evaluating the GMM objective function

[12] For notational brevity we omit the subscripts. Market order size m and fundamental price X are observed at time t, and the equation holds for any price tick $p_{+k,t}$ with associated cumulative volume $Q_{+k,t}$, $k = 1, 2....$

involves computation of the conditional sample means, but it is a straightforward exercise.

Empirical evidence suggests that market orders are timed in that market order traders closely monitor the state of the book when deciding on the size of the submitted market order (see e.g. Biais et al. (1995), Ranaldo (2004) and Gomber et al. (2004)). To account for state dependency, Såndas (2001) proposed using a set of instruments which scale the value of the λ parameter in equation (2.2). The nonparametric strategy developed here can be easily adapted to account for a market order distribution that changes with the state of the book. One only has to base the computation of the conditional upper tail expectation on a vector of state variables F, i.e. calculate $\widehat{E}\left[m|m \geq Q_{+k}, F\right]$. For the purpose of this study we focus on the unconditional market order distribution and leave modeling the conditional market order distribution as a topic for further research.

Average profit conditions

To justify the marginal zero expected profit assumption, one implicitly assumes a repetitive two phase trading process. In phase one, agents submit and cancel limit orders until the book is free of (expected) profit opportunities and no agent wants to submit, revise or cancel her order. Limit orders are sorted by price priority and, within the same price tick, by time priority. When the book is such an equilibrium the order book should display no 'holes', i.e. zero volumes in between two price ticks. In phase two, a single market order of a given size arrives and is executed against the equilibrium order book. After this event we go back to phase one, during which the book is replenished again until equilibrium is reached and another market order arrives and so forth. Can this be a reasonable description of a real world trading process? The descriptive statistics on the trading and order submission activity reported in table 1 indicate a dynamic trading environment. For a large stock, like Daimler Chrysler, we have on average over 3,000 trade events per day, about 19,000 submissions of limit orders, of which over 80 % are canceled before execution. One could argue that such an active limit order trader behavior indicates a thorough monitoring of the book which eliminates any profit opportunities. This is quite in line with the theoretical framework. However, with on average 10 seconds duration between trade events (for Daimler Chrysler) the time to reach the new equilibrium after a market order hits the book and before a new order arrives, seems a short span.

The marginal break even conditions can also be challenged by the following reasoning. The conditions imply nonzero expected profits for limit order units that do not occupy the last position of the respective price ticks. On the other hand, this implies that the whole book offers positive expected profits for traders acting as market makers. If market making provides nonzero expected profit opportunities, then this would attract new entrant and the competition between these would-be market makers ultimately eliminate any profit opportunities.

These considerations lead us to consider an alternative to the marginal profit conditions which does not rely on the assumption that limit order traders immediately cancel or adjust all their orders which show negative expected profit on a marginal unit, and that also acknowledges the effect of market maker competition on expected profits. For this purpose we retain most of the assumptions of the Glosten/ Såndas framework. However, instead of evaluating the expected profit of the marginal profit for the last unit at each quote k, we assume that the expected profit of the whole block of limit orders at any quote is zero. The marginal zero profit condition is thus replaced by an 'average zero profit condition'. This assumption allows to differentiate between two types of costs associated with the submission of a limit order, a fixed cost component, like order submission and surveillance costs, and marginal costs (per share), like execution or clearing fees and opportunity costs of market making. In the appendix we show that the liquidity supply equations which are implied by the zero expected profit condition can be written as

$$\begin{aligned} q_{+k,t} &= 2\left(\frac{p_{+k,t} - X_t - \gamma - \frac{\xi}{q_{+k,t}}}{\alpha} - \lambda - \mu\right) - Q_{+k-1,t} \quad k = 1,2,\ldots \quad \text{(ask side)} \\ q_{-k,t} &= 2\left(\frac{X_t - p_{-k,t} - \gamma - \frac{\xi}{q_{-k,t}}}{\alpha} - \lambda + \mu\right) - Q_{-k+1,t} \quad k = 1,2,\ldots \quad \text{(bid side)} \end{aligned} \tag{2.10}$$

ξ denotes the fixed cost component which is assumed to be identical for each price tick in the order book. To derive the equations in (2.10), we have retained the parametric assumption about the distribution of trade sizes. Considering a nonparametric alternative along the lines described in the previous subsection is also feasible. Proceeding as above, i.e. by eliminating the unobserved fundamental asset value X_t by adding the bid and ask side equations for quote $+k$ and $-k$ yields the following unconditional moment restrictions which we refer to as average break even conditions,

$$E\left(\Delta p_{\pm k,t} - 2\gamma - \frac{\xi}{q_{+k,t}} - \frac{\xi}{q_{-k,t}} - \alpha\left(\frac{1}{2}Q_{+k,t} + 2\lambda + \frac{1}{2}Q_{-k,t}\right)\right) = 0 \quad k = 1, 2, \ldots, \tag{2.11}$$

where $\Delta p_{\pm k,t} = p_{+k,t} - p_{-k,t}$. Subtracting deviations from the implied depths at the kth quote at time $t+1$ and t and taking expectation yields the following equations which we refer to as average update conditions,

$$\begin{aligned} E\left(\Delta p_{+k,t+1} - \frac{\xi}{q_{+k,t}} + \frac{\xi}{q_{+k,t-1}} - \frac{\alpha}{2}\left(Q_{+k,t+1} - Q_{+k,t}\right) - \mu - \alpha m_t\right) &= 0 \qquad k = 1, 2, \ldots \\ E\left(\Delta p_{-k,t+1} + \frac{\xi}{q_{+k,t}} - \frac{\xi}{q_{+k,t-1}} + \frac{\alpha}{2}\left(Q_{-k,t+1} - Q_{-k,t}\right) - \mu - \alpha m_t\right) &= 0 \qquad k = 1, 2, \ldots, \end{aligned} \tag{2.12}$$

where $\Delta p_{j,t+1} = p_{j,t+1} - p_{j,t}$. The average break even and update conditions replace the marginal break even and update conditions of equation (2.4) and equation (2.5).

2.4 Empirical results

Performance comparisons

Using the DAX30 order book data we follow Såndas (2001) and estimate the model parameters exploiting the marginal break even conditions (2.4) and the marginal updating conditions (2.5) along with (2.6). To construct the moment conditions we use the respective first four best quotes, i.e. $k = 1, \ldots, 4$ on the bid and the ask side of the visible order book. This yields thirteen moment conditions: four break even conditions, eight update conditions, and the moment condition (2.6). Order sizes X_t are expressed in 1000 shares.

Table 2.5 contains the first stage GMM results.[13]

We report parameter estimates, $t-$statistics and the value of the GMM $J-$statistic with associated $p-$values. Under the null hypothesis that the moment conditions are correctly specified, the $J-$statistic is asymptotically χ^2 with degrees of freedom equal

[13] Two stage and iterated GMM estimates are similar and therefore not reported to conserve space. To compute the parameter standard errors and the J-statistic we employ the Bartlett Kernel with bandwidth equal to 10 lags when computing the spectral density matrix. We have tested various lags and the results are robust with respect to bandwidth choice.

to the number of moment conditions minus the number of estimated parameters. The estimation results based on the Xetra data are in line with the central findings reported by Såndas (2001). Only for two out of thirty stocks the model is not rejected at 1 % significance level. In Såndas' (2001) application the model was rejected for all stocks. Like in Såndas' (2001) application, the transaction cost estimates (γ) are significantly negative, a result that is difficult to reconcile with the underlying theoretical model. Hence, even with a data generating process that corresponds very close to the theoretical framework, the model does not seem to fit the data very well.

Tables and 2.5 and 2.5 report the results that are obtained when the modified moment conditions suggested in the previous section are used. As before, the first four quotes on each market side are used for the construction of break even and update conditions. Table 2.5 reports the estimation results for a specification that does not rely on a parametric assumption on the distribution of the market order size when constructing the marginal break even and the update conditions as described in section 2.3. The results reported in table 2.5 are obtained when using the average break even and update conditions (2.11) and (2.12) for GMM estimation, while maintaining the parametric assumption (2.2) about the trade size distribution. For each parametric specification, the moment condition of equation (2.6) is employed, too. The full set of eight update conditions is exploited.

The estimates reported in table 2.5 show that abandoning the parametric assumption concerning the market order size distribution improves the results only marginally. For four of the thirty stocks the model is not rejected at the 1 % significance level. Hasbrouck's (2004) conjecture that the distributional assumption might be responsible for the model's empirical failure is therefore not supported. Generally, the estimates of the adverse selection components, transaction costs and drift parameters do not change dramatically compared to the baseline specification. The transaction cost estimates remain negative.

Maintaining the distributional assumption, but using average break even conditions instead of marginal break even conditions, considerably improves the empirical performance. Table 2.5 shows that we have model non-rejection for 22 out of the 30 stocks at the 1% significance level. With a single exception the estimates of the marginal transaction cost parameter γ are positive for those stocks for which the model is not rejected at 1 % significance level. The size of the implied transaction cost estimates are broadly comparable with the relative realized spreads figures reported in table 2.5. For example, the estimation results imply that transaction costs account for 0.013% of the euro value of a median sized DaimlerChrsyler trade. This value is quite comparable with the average relative realized spread which amounts to 0.010% (see table 1).

Figure 2.5 shows graphically the improved empirical performance delivered by the revised methodology. The figure depicts means an medians of implied and observed ask side price schedules of four selected stocks. The results obtained from the baseline estimation which uses marginal moment conditions confirm the disturbing findings reported in Såndas (2001). The price schedules implied by the model estimates are below the observed price schedules at all relevant volumes. The economically implausible negative price discount at small volumes is caused by the negative transaction costs estimates. This suggests that the model is not only rejected on the grounds of statistical significance, but that fundamentally fails to explain the data. The model does a bad job even in describing the 'average' state of the order book. However, figure 2.5 shows that when working with average break even and update conditions the empirical performance of the model is considerably improved. Especially the median observed price schedules correspond closely to those implied by the model.[14]

Cross sectional analyses

Encouraged by the improved empirical performance of the revised methodology, this section uses the estimation results reported in table 2.5 to conduct a cross sectional analysis of liquidity supply and adverse selection costs in the Xetra limit order book market. To ensure comparability across stocks, we follow a suggestion by Hasbrouck (1991) and standardize the adverse selection component α by computing

$$\tau = \frac{\alpha \cdot \overline{m}}{\overline{P}}, \tag{2.13}$$

where $\overline{m}$ is the average (non-signed), stock specific trade size expressed in number of shares. $\overline{P}$ is the sample average of the midquote of the respective stock. τ (times 100) approximates the percentage change of the stock price caused by a trade of (stock specific) 'average' size. This is a relative measure which is comparable across stocks. The

[14] We have also estimated a specification that combines the nonparametric approach towards trades sizes and average moment conditions, but the results (not reported) are not improved compared to the parametric version. In this version, the model is not rejected for 16 out of 30 stocks. The following analysis therefore focuses on the parametric specification using average moment conditions.

τ estimates are reported in the last column of table 2.5. In the following subsections we study the relation of τ and market capitalization, trading frequency, liquidity supply and alternative adverse selections measures.

Adverse selection effects, market capitalization and frequency of trading

In their seminal papers Hasbrouck (1991) and Easley et al. (1996) have reported empirical evidence that adverse selection effects are more severe for smaller capitalized stocks. Easley et al. (1996) use a formal model assuming a Bayesian market maker who updates quotes according to the arrival of trades while Hasbrouck (1991) estimates a vector autoregression (VAR) involving trade and midquote returns. Both methodologies have modest data requirements. To estimate the model by Easley et al. (1996) one only needs to count the number of buyer and seller initiated trades per trading day to estimate the probability of informed trading (PIN), the central adverse selection measure in this framework. As it allows a structural interpretation of the model parameter estimates the methodology is quite popular in empirical research. Hasbrouck's VAR methodology is not based on a formal model, but the reduced form VAR equations are compatible with a general class of microstructure models. The adverse selection measure is given by the cumulative effect of a trade innovation on the midquote return. To estimate the model, standard trade and quote data are sufficient.

Both methodologies are not specifically designed for limit order markets, but rather for market maker systems. Accordingly, their main applications have been to analyze NYSE and NASDAQ stocks. In the present work, the data generating process, the theoretical background and the empirical methodology are quite different. However, we reach the same conclusion as Hasbrouck (1991) and Easley et al. (1996). The Spearman rank correlation of the market capitalization and the estimated standardized adverse selection component τ (using only the results for those 22 out of 30 stocks for which the model is not rejected at 1 % significance level) is -0.928 ($p-$value < 0.001). The correlation of τ and the daily number of trades is -0.946 ($p-$value < 0.001), and the correlation of τ and the daily turnover is -0.966 ($p-$value < 0.001). The estimation results thus confirm the previous evidence also for an open limit order market: Adverse selection effects are more severe for smaller capitalized stocks.

Adverse selection and liquidity supply

Many theoretical market microstructure models predict that liquidity supply and informed order flow are inversely related. As the standard framework for microstructure models is a stylized NYSE trading process with a single market maker quoting best bid and ask prices and associated depths (the 'inside market'), liquidity in those models is usually measured by the inside spread set by the specialist/market maker. Sequential trading models like Easley et al. (1996) and spread decomposition models like Glosten and Harris (1988) predict that liquidity (as measured by the spread) and informed order flow are inversely related. In the presence of informed order flow, the market maker widens the spread in order to balance the losses that occur when trading with superiorly informed agents. More informed order flow thus implies reduced liquidity. Empirical analyses of specialist markets have confirmed this prediction. The results reported in table 2.5 provide evidence that the inverse relation of inside spread and informed order flow also holds for open limit order book markets in which limit order traders, instead of specialists, determine the inside market. The table reports the cross sectional correlation (Spearman rank correlation) of the standardized adverse selection component τ and the effective, quoted, and the realized spreads. Effective and quoted spreads and τ are strongly positively correlated while the correlation with the realized spread and τ is not significantly different from zero. Given the interpretation of the realized spreads as a transaction costs measure which is purged of any informational effects, this is an expected result.[15]

Ad hoc versus model-based estimates of adverse selection effects in a limit order book market

In this subsection we investigate whether the adverse selection estimates obtained from the formal model and those delivered by the simple analysis of effective and realized spreads (see section 2.2) point in the same direction. The two methodologies differ in two main aspects. First, the estimation of adverse selection components by taking the difference of effective and realized spread is not based on a specific theoretical model. The economic intuition behind the methodology, however, is quite clear, which explains the popularity of the approach. A large difference between effective

[15] Huang and Stoll (1996) and DeJong et al. (1996) provide evidence for a negative correlation of realized spread and adverse selection costs.

and realized spread indicates informational content of the order flow as the midquote tends to move in the direction of the trade. If a market buy (sell) order initiates a trade at time t, then the midquote five minutes after the trade is on average above (below) the time t midquote. By contrast, the estimates of the standardized adverse selection component reported in table 2.5 are based on a formal model assuming rational limit order traders who place their order submissions explicitly taking into account the amount of informational content of the order flow. Second, computation of price impacts by taking the difference of effective and realized spread only requires publicly available trade and (best) quote data. To obtain the estimates in the formal framework considered in this chapter, reconstructed order book data are needed. The latter methodology thus uses richer data, which are, however, more difficult to obtain.

But do the two different methodologies lead to the same conclusions? To address this question we compute the Spearman rank correlation between the standardized adverse selection components (τ) reported in table 2.5 and the difference of effective and realized spread. The cross sectional correlation (using the 22 out of 30 stocks for which the model is not rejected at 1 % significance level) is 0.95. The two different methodologies thus point in the same direction. This result indicates the robustness of the estimation results of the formal model and also provides a theoretical justification to use the popular ad hoc method for the analysis of adverse selection effects in limit order book markets.

2.5 Conclusion and outlook

An increasing number of financial assets trade in limit order markets. These markets can be characterized by the following keywords: Transparency, anonymity and endogenous liquidity supply. They are transparent, because a more or less unobstructed view on the liquidity supply is possible and anonymous, because prior to a trade the identity of none of the agents participating in the transaction is revealed. Liquidity supply is endogenous, because typically there are no dedicated market makers responsible for quoting bid and ask prices. The question how liquidity quality and price formation in such a trading design is affected by informed order flow is a crucial one, both from a theoretical and a practical point of view. Glosten (1994) has put forth a formal model that describes how an equilibrium order book emerges in the presence of potentially informed order flow. Såndas (2001) has confronted the Glosten model with real world data and reported quite discouraging results. His findings suggest that Glosten's model contains too many simplifying assumptions in order to provide a valid description of the intricate real world trading processes in limit order markets.

This chapter shows that the ability of Glosten's basic framework to explain real world order book formation is greater than previously thought. We estimate the model using data produced by a DGP that closely corresponds to the Glosten's theoretical framework and confirm the previous finding that the baseline specification put forth by Såndas (2001) is generally rejected. However, relaxing the assumption about marginal zero profit order book equilibrium in favor of a weaker equilibrium condition, considerably improves the empirical performance. The equilibrium condition proposed in this chapter does not assume that traders immediately cancel a marginal order that shows non-positive expected profit. It also acknowledges the fact that competition between potential market makers will render the expected profit offered by the whole book ultimately to zero (after accounting for opportunity costs). Employing the revised econometric methodology, formal specification tests now accept the model in the vast majority of cases at conventional significance levels. A comparison of implied and observed order book schedules shows that the model estimated on the revised set of moment conditions fits the data quite well. We conclude that Glosten's theoretical framework can also be transferred into a quite useful empirical model.

On the other hand, the conjecture put forth by Hasbrouck (2004), which states that the distributional assumption regarding the market order sizes is responsible for the empirical model failure is not supported. The chapter has developed a straightforward way to circumvent the restrictive distributional assumption and proposes a nonparametric alternative. However, this modification does not deliver an improved empirical performance.

Given the overall encouraging results, the empirical methodology is employed for an analysis of liquidity supply and adverse selection costs in a cross section of stocks traded in one of the largest European equity markets. The main results can be summarized as follows:

- We have provided new evidence, from a limit order market, that adverse selection effects are more severe for smaller capitalized, less frequently traded stocks.

This corroborates the results of previous papers dealing with a quite different theoretical background, empirical methodology and market structure.

- The empirical results support one of the main hypothesis of the theory of limit order markets, namely that liquidity and adverse selection effects are inversely related.

- The adverse selection component estimates implied by the structural model and ad hoc measures of informed order flow which are based on a comparison of effective and realized spreads point in the same direction. This is a useful result, because is not always possible to estimate the structural model, most often because of the lack of suitable data. The result also points towards the robustness of the structural model.

Avenues for further research stretch in various directions. The results reported in this chapter have vindicated the empirical relevance the Glosten type market order model. Practical issues in market design can thus be empirically addressed based on a sound theoretical framework. The revised methodology could be employed to evaluate changes in trading design on liquidity quality, with the advantage that the results can be interpreted on a sound theoretical basis. A comparison of (internationally) cross listed stocks seems also promising, especially after the NYSE's move towards adopting the key feature of a open limit order market, the public display of the limit order book. An interesting question would be to investigate whether the recently reported failures of cross listings (in terms of insufficient trading volume in the foreign markets) are due to market design features that aggravate potential adverse selection effects.

Second, a variety of methodological extensions could be considered. Såndas (2001) has already addressed the issue of state dependence of the model parameters. He used a set of plausible instruments to scale the model parameters. Recent papers on price impacts of trades point to alternative, powerful instruments that could be used, and which might improve the empirical performance and explanatory power. For example, Dufour and Engle (2000) have emphasized the role of time between trades within Hasbrouck's (1991) VAR framework. As the Glosten/Såndas type model considered in this chapter is also estimated on irregularly spaced data, it seems natural to utilize their findings. Furthermore, the exogeneity of the market order flow is a restrictive assumption that should be relaxed. Gomber et al. (2004) and Coppejans et al (2003) show that market order traders time their trades by submitting larger trade sizes at times when the book is relatively liquid. Hence, using the liquidity state of the book as a scaling instrument for the expected order size parameter seems a promising strategy. As in many GMM applications, the number of moment conditions that are available is large, and the difficult task is to pick both relevant and correct moment conditions. Recent contributions by Andrews (1999) and Hall and Peixe (2003) could be utilized to base the selection of moment conditions on a sound methodological basis. Another direction of future research points to a further relaxation of the model's parametric assumptions. Specifically, the linear updating function 2.14 could be replaced by a nonlinear relation of asset price and market order size. Combined with a conditional nonparametric distribution for the market order sizes this would provide a quite flexible modeling framework.

2.A Appendix: Derivation of revised moment conditions

This section outlines the background for the revised set of moment conditions describing order book equilibrium. We start by writing the zero expected profit condition for one unit of a limit sell order as

$$E(R_t - X_{t+1}) = 0, \tag{2.14}$$

where R_t denotes the net revenue (minus transaction costs) received from selling one unit of a limit order at price p_t to a market order trader who submitted a market buy order of size m_t.[16] X_{t+1} denotes the fundamental value of the stock after the arrival the (buy) market order. X_{t+1} depends on the current value X_t and the signed market order size m_t, i.e. $X_{t+1} = g(m_t, X_t)$. For brevity of notation we henceforth omit the time t subscripts whenever it is unambiguous to do so.

The expected profit of the market order depends on the position of the limit order in the order queue and the distribution of market orders, i.e. we can write equation (2.14) as

$$\int_Q^\infty (R - g(m, X)) f(m) dm = 0. \tag{2.15}$$

Q is the cumulated sell order volume standing in the book before the considered limit order unit and $f(m)$ denotes the probability density function of m. Alternatively, equation (2.15) can be written as

$$(R - E\left[g(m, X) | m \geq Q\right]) \cdot P\left(m \geq Q\right) = 0. \tag{2.16}$$

Assuming the linear specification in equation (2.1) for $g(m, X)$, and dividing by the unconditional probability, $P\left(m \geq Q\right)$, equation (2.16) simplifies to

$$R - \alpha E\left[m | m \geq Q\right] - X - \mu \cdot = 0. \tag{2.17}$$

Equation (2.17) highlights that the expected profit of a limit order trader depends on the upper tail expectation of the market order distribution.

Assuming exponentially distributed market order sizes as in equation (2.2) we have

$$E\left[m | m \geq Q\right] = Q + \lambda \tag{2.18}$$

[16] The exercise is analogous for the bid side, but to conserve space, we focus on the sell side of the book.

Using $R = p - \gamma$ this yields

$$Q = \frac{p - X - \gamma - \mu}{\alpha} - \lambda, \tag{2.19}$$

which is a generalized form of equation (2.3). Without the distributional assumption, the equivalent of equation (2.19) is

$$E\left[m|m \geq Q\right] = \frac{p - X - \gamma - \mu}{\alpha} \tag{2.20}$$

Replacing $E\left[m|m \geq Q\right]$ by the conditional sample mean $\widehat{E}\left[m|m \geq Q\right]$, i.e. the observed upper tail market order distribution in the sample, one can construct update and break even moment conditions for GMM estimation which do not require a parametric assumption of market order sizes.

So far, the results are valid for an order book with a continuous price grid. We now focus on a specific offer side quote with price p_{+k} and corresponding limit order volume q_{+k}. Abstracting from the discreteness of limit order size shares and assuming that the execution probabilities for all units at the quote tick p_{+k} are identical, we calculate the expected profit of all limit orders with identical limit price p_{+k} by integrating the left hand side of equation, (2.17), [17]

$$\int_{Q_{+k-1}}^{Q_{+k}} \left(p_{+k} - \gamma - \alpha E\left[m|m \geq Q\right] - X - \mu\right) dQ \cdot P\left(m \geq Q_{+k-1}\right). \tag{2.21}$$

Assuming exponentially distributed order sizes and subtracting quote specific fixed execution costs ξ yields the total expected profit of the limit order volume at price p_{+k}. Dividing by the volume at quote q_{+k}, yields the average expected profit per share at the $+k$th quote,

$$\left(p_{+k} - X - \mu - \gamma - \frac{\xi}{q_{+k}} - \alpha(Q_{+k} + \lambda - \frac{q_{+k}}{2})\right) \cdot P\left(m \geq Q_{+k-1}\right) \tag{2.22}$$

. In the main text we discuss the implications of the situation that the average profit equals zero. This implies that

$$p_{+k} - X - \mu - \gamma - \frac{\xi}{q_{+k}} - \alpha(Q_{+k} + \lambda - \frac{q_{+k}}{2}) = 0. \tag{2.23}$$

Reordering equation (2.23) and replacing Q_{+k} by $Q_{+k-1} + q_k$ yields the average profit conditions (2.10) from which average break even and update conditions can be derived again.

[17] The same result can be derived using the precise probabilities and a first-order Taylor approximation for the emerging exponential terms.

Table 2.1: Xetra 2004 - Sample Descriptives.

company name	ticker symbol	turnover	mkt. cap.	$\overline{m}$	% aggr. trades	trades per day	LO sub.	LO canc.	$\overline{P}$	eff. spread	real. spread
TUI	TUI	26281175	2025	24723	17.6	1063	6767	5714	18.7	0.125	0.015
CONTINENTAL	CONT	25627638	4060	25574	13.5	1002	8036	7052	31.6	0.092	-0.011
MAN	MAN	27685031	2434	26189	13.0	1057	7214	6235	27.7	0.096	0.003
METRO	MEO	38874669	5018	31480	15.7	1235	7975	6702	35.0	0.089	0.000
LINDE	LIN	22378772	3448	24971	15.8	896	8342	7454	43.6	0.080	-0.009
LUFTHANSA	LHA	43946809	4548	32504	11.9	1352	8079	6780	14.2	0.111	0.022
FRESENIUS	FME	12850947	1944	20680	16.7	621	5764	5195	54.0	0.098	0.010
THYSSEN-KRUPP	TKA	37892493	6450	30017	11.3	1262	7864	6672	15.9	0.111	0.029
DEUTSCHE POST	DPW	43836617	6806	33330	11.0	1315	6861	5666	18.2	0.097	0.018
HYPO-VEREINSB.	HVM	98351090	6629	50783	15.0	1937	10204	8293	18.7	0.098	0.019
COMMERZBANK	CBK	53171668	7569	36659	12.6	1450	11922	10476	15.4	0.100	0.023
ADIDAS-SALOMON	ADS	31976047	4104	32635	20.1	980	8057	7105	92.6	0.070	-0.002
DEUTSCHE BOERSE	DB1	35696903	4847	36359	18.4	982	6598	5698	46.9	0.075	0.003
HENKEL	HEN3	18174548	3682	25904	16.6	702	7989	7306	65.9	0.077	0.005
ALTANA	ALT	30985416	3338	28310	18.9	1095	7718	6609	48.6	0.079	0.008
SCHERING	SCH	51413053	7055	33756	16.2	1523	9111	7669	40.8	0.071	0.004
INFINEON	IFX	146462315	4790	52331	8.6	2799	10320	7744	11.6	0.104	0.040
BAYER	BAY	88776121	15911	36994	12.4	2400	15258	12988	23.1	0.076	0.012
RWE	RWE	97655566	12653	42203	13.0	2314	14438	12355	33.8	0.062	0.002
BMW	BMW	87854358	12211	41639	14.4	2110	14736	12764	34.7	0.060	0.003
VOLKSWAGEN	VOW	104249843	9688	40963	16.0	2545	13474	11273	39.2	0.056	0.004
BASF	BAS	124434537	25425	48236	13.8	2580	18211	15898	43.3	0.051	0.002
SAP	SAP	184628162	27412	65795	21.9	2806	19733	17095	131.5	0.049	0.001
E.ON	EOA	160625983	33753	55950	13.6	2871	18899	16468	52.5	0.048	0.003
MUENCH.RUECK	MUV2	207353230	16396	60534	20.7	3425	20154	16894	93.9	0.049	0.005
DAIMLERCHRYSLER	DCX	187737846	30316	56736	14.5	3309	18722	15919	36.4	0.055	0.010
DEUTSCHE TELEKOM	DTE	350627866	34858	78884	5.0	4445	14498	11009	15.7	0.072	0.031
DEUTSCHE BANK	DBK	309282831	38228	78083	19.3	3961	23169	19772	67.2	0.044	0.004
ALLIANZ	ALV	289980556	33805	64114	21.4	4523	29791	25882	100.1	0.049	0.010
SIEMENS	SIE	321704299	52893	72831	16.7	4418	23659	19920	64.0	0.041	0.006
Average		108683880	14076	42972	15.2	2099	12785	10887	44.5		

Mkt. cap. is the market capitalization in million euros at the end of December 2003, $\overline{m}$ is the average trade size (in euros). *%Aggr. trades* gives the percentage of total trading volume that has not been executed at the best prices (that is, the order walked up the book). *Turnover* is the average trading volume in euros per trading day, *trades per day* is the average number of trades per day, *LO sub.* and *LO canc.*, respectively, denote the average number of non-marketable limit order order submissions and cancelations per day. $\overline{P}$, *eff. spread* and *real. spread* refer to the sample averages of midquote, effective spread and realized spread, respectively. The average effective spread is computed by taking two times the absolute difference of the transaction price of a trade and the prevailing midquote and averaging over all trades of a stock. The average realized spread is computed similarly, but instead of taking the prevailing midquote, we use the midquote five minutes after the trade. To ensure comparability across stocks, we compute effective and realized spreads relative to the midquote prevailing at the time of the trade and multiply by 100 to obtain a % figure. The table is sorted in descending order by the difference of effective and realized spread.

Table 2.2: First stage GMM results baseline specification.

ticker	α	γ	λ	μ	$J(9)$	p-value
LIN	0.0228	-0.0185	0.5728	0.0003	0.5	1.000
	(118.1)	(55.5)	(134.6)	(3.1)		
DPW	0.0025	-0.0063	1.8362	-0.0001	21.0	0.013
	(149.8)	(47.0)	(150.0)	(3.3)		
HEN3	0.0415	-0.0178	0.3937	-0.0003	23.2	0.006
	(95.8)	(40.1)	(114.0)	(1.9)		
MEO	0.0132	-0.0163	0.9066	0.0000	25.3	0.003
	(135.1)	(66.1)	(166.7)	(0.4)		
LHA	0.0019	-0.0072	2.3210	0.0001	31.2	0.000
	(159.7)	(63.6)	(150.8)	(4.4)		
MAN	0.0104	-0.0151	0.9445	0.0004	35.2	0.000
	(121.9)	(61.8)	(134.5)	(5.2)		
DB1	0.0162	-0.0158	0.7739	0.0001	45.0	0.000
	(114.5)	(46.7)	(114.1)	(1.1)		
FME	0.0456	-0.0210	0.3839	0.0000	53.2	0.000
	(85.0)	(34.8)	(96.4)	(0.0)		
TUI	0.0054	-0.0095	1.3215	-0.0002	57.4	0.000
	(130.3)	(50.6)	(127.0)	(2.7)		
ALT	0.0224	-0.0144	0.5785	-0.0002	79.8	0.000
	(121.9)	(50.5)	(142.4)	(2.1)		
CBK	0.0016	-0.0048	2.4055	-0.0001	81.1	0.000
	(164.0)	(50.4)	(152.0)	(2.0)		
CONT	0.0131	-0.0168	0.8166	0.0002	85.8	0.000
	(116.6)	(60.4)	(139.6)	(2.2)		
ADS	0.0549	-0.0183	0.3528	-0.0002	118.2	0.000
	(113.2)	(38.1)	(141.0)	(1.5)		
BMW	0.0053	-0.0087	1.2029	-0.0001	173.7	0.000
	(173.8)	(69.6)	(203.0)	(2.9)		
TKA	0.0024	-0.0075	1.9075	0.0000	206.6	0.000
	(148.6)	(61.0)	(158.9)	(1.3)		
SCH	0.0106	-0.0101	0.8250	0.0000	232.6	0.000
	(135.3)	(49.6)	(168.8)	(0.4)		
RWE	0.0053	-0.0095	1.2460	0.0001	239.4	0.000
	(212.2)	(86.2)	(210.3)	(3.3)		
DTE	0.0002	-0.0010	5.0499	0.0000	292.8	0.000
	(303.0)	(32.1)	(232.7)	(0.0)		
IFX	0.0004	-0.0023	4.5335	0.0000	360.9	0.000
	(196.7)	(45.6)	(170.6)	(0.4)		
HVM	0.0015	-0.0043	2.8391	0.0000	363.8	0.000
	(109.0)	(40.1)	(130.9)	(1.0)		
VOW	0.0065	-0.0099	1.0472	0.0001	429.9	0.000
	(21.2)	(17.7)	(195.8)	(0.2)		
BAY	0.0024	-0.0046	1.6352	0.0000	458.2	0.000
	(216.9)	(59.2)	(225.8)	(1.2)		
BAS	0.0056	-0.0077	1.1206	0.0000	683.1	0.000
	(219.9)	(77.4)	(244.1)	(1.1)		
EOA	0.0060	-0.0070	1.0663	0.0000	1011.3	0.000
	(219.2)	(65.0)	(252.7)	(1.0)		
DCX	0.0031	-0.0049	1.5638	0.0002	1376.9	0.000
	(258.2)	(65.4)	(254.9)	(7.2)		
SAP	0.0370	-0.0147	0.5030	0.0006	1609.9	0.000
	(212.6)	(49.5)	(237.4)	(5.9)		
MUV2	0.0196	-0.0106	0.6476	0.0001	2101.9	0.000
	(212.1)	(60.4)	(246.9)	(1.0)		
DBK	0.0065	-0.0061	1.1517	0.0000	2584.6	0.000
	(248.7)	(57.7)	(256.1)	(0.9)		
ALV	0.0187	-0.0080	0.6453	-0.0002	2701.8	0.000
	(232.7)	(35.9)	(294.4)	(4.5)		
SIE	0.0052	-0.0039	1.1442	0.0001	3827.8	0.000
	(273.3)	(36.4)	(297.3)	(2.9)		

2×4 quotes from the bid and ask side of the visible book are used to construct update and break even conditions derived from the zero marginal expected profit condition as in Såndas (2001). The numbers in parentheses are $t-$values. The fifth and sixth column report the GMM J statistic and the associated $p-$value. The stocks are sorted by ascending order of the $J-$statistic.

Table 2.3: First stage GMM results for the nonparametric specification.

ticker	α	γ	μ	$J(9)$	$p-$value
TUI	0.0040	-0.0091	-0.0001	0.7	1.000
	(140.1)	(52.9)	(2.3)		
LIN	0.0169	-0.0142	0.0003	3.4	0.945
	(122.1)	(49.8)	(2.8)		
DB1	0.0111	-0.0123	0.0001	11.1	0.270
	(110.2)	(39.0)	(1.1)		
HEN3	0.0301	-0.0120	-0.0002	19.7	0.020
	(95.2)	(31.5)	(1.5)		
ALT	0.0169	-0.0117	-0.0002	33.5	0.000
	(129.8)	(47.9)	(1.8)		
HVM	0.0012	-0.0058	0.0000	41.9	0.000
	(134.2)	(60.8)	(0.8)		
ADS	0.0403	-0.0125	-0.0002	68.3	0.000
	(119.3)	(30.1)	(1.2)		
MEO	0.0103	-0.0139	0.0000	72.8	0.000
	(141.2)	(64.3)	(0.3)		
FME	0.0299	-0.0119	0.0000	84.5	0.000
	(79.8)	(22.3)	(0.2)		
CONT	0.0094	-0.0127	0.0001	87.5	0.000
	(120.6)	(52.6)	(2.1)		
IFX	0.0003	-0.0032	0.0000	97.9	0.000
	(235.2)	(69.9)	(0.2)		
MAN	0.0076	-0.0125	0.0003	101.3	0.000
	(127.5)	(58.8)	(5.0)		
BMW	0.0039	-0.0072	-0.0001	112.9	0.000
	(188.1)	(66.8)	(2.7)		
LHA	0.0015	-0.0075	0.0001	165.3	0.000
	(162.9)	(68.5)	(4.1)		
VOW	0.0048	-0.0092	0.0001	169.4	0.000
	(28.4)	(21.9)	(0.2)		
SCH	0.0078	-0.0080	0.0000	173.0	0.000
	(147.3)	(45.9)	(0.4)		
DPW	0.0019	-0.0064	-0.0001	176.8	0.000
	(168.3)	(53.0)	(2.9)		
RWE	0.0038	-0.0073	0.0001	189.2	0.000
	(209.4)	(74.0)	(3.2)		
BAY	0.0018	-0.0034	0.0000	349.3	0.000
	(196.6)	(44.4)	(1.3)		
CBK	0.0013	-0.0060	0.0000	427.8	0.000
	(196.2)	(68.0)	(2.0)		
BAS	0.0041	-0.0057	0.0000	574.9	0.000
	(222.3)	(62.7)	(0.8)		
TKA	0.0019	-0.0077	0.0000	721.3	0.000
	(164.7)	(68.7)	(1.1)		
DCX	0.0022	-0.0036	0.0001	760.9	0.000
	(235.6)	(49.8)	(5.8)		
EOA	0.0043	-0.0047	0.0000	1050.8	0.000
	(227.8)	(50.3)	(0.7)		
SAP	0.0274	-0.0099	0.0004	1070.5	0.000
	(228.6)	(39.5)	(4.9)		
MUV2	0.0140	-0.0074	0.0000	1508.2	0.000
	(219.1)	(47.7)	(0.7)		
DTE	0.0001	-0.0022	0.0000	1514.1	0.000
	(380.6)	(73.5)	(0.0)		
DBK	0.0046	-0.0046	0.0000	1633.5	0.000
	(265.3)	(49.6)	(0.7)		
ALV	0.0134	-0.0048	-0.0002	2152.3	0.000
	(241.2)	(22.7)	(4.0)		
SIE	0.0039	-0.0032	0.0001	2912.9	0.000
	(312.6)	(35.5)	(2.6)		

2×4 quotes from the bid and ask side of the visible book are used to construct update and break even conditions derived from the zero marginal expected profit condition as in Såndas (2001). For the construction of the moment conditions, the empirical distribution of the market order sizes is used instead of the exponential distribution. The stocks are sorted by ascending order of the $J-$statistic.

Table 2.4: First stage GMM results based on average profit conditions.

ticker	α	γ	λ	ξ	μ	$J(8)$	$p-$value	$\tau(\%)$
IFX	0.0004	0.0042	4.5335	-0.0051	0.0000	0.1	1.000	0.0157
	(151.6)	(24.3)	(170.6)	(14.9)	(0.2)			
DBK	0.0057	0.0137	1.1517	-0.0031	0.0000	0.3	1.000	0.0099
	(102.7)	(20.4)	(256.1)	(19.4)	(0.7)			
SAP	0.0286	0.0532	0.5030	-0.0058	0.0004	0.7	1.000	0.0109
	(32.8)	(9.7)	(237.4)	(9.6)	(4.1)			
DCX	0.0029	0.0059	1.5638	-0.0023	0.0002	0.8	0.999	0.0124
	(145.8)	(17.8)	(254.9)	(17.9)	(6.1)			
DB1	0.0152	0.0049	0.7739	-0.0016	0.0001	0.8	0.999	0.0251
	(56.2)	(2.0)	(114.1)	(4.5)	(1.1)			
SIE	0.0045	0.0133	1.1442	-0.0030	0.0001	0.8	0.999	0.0080
	(117.0)	(27.8)	(297.3)	(23.7)	(2.6)			
TUI	0.0053	0.0021	1.3215	-0.0016	-0.0002	0.9	0.999	0.0374
	(103.0)	(2.8)	(127.0)	(7.0)	(2.6)			
MUV2	0.0171	0.0234	0.6476	-0.0036	0.0001	1.1	0.998	0.0117
	(89.0)	(15.4)	(246.9)	(15.7)	(0.8)			
FME	0.0428	0.0101	0.3839	-0.0016	0.0000	1.5	0.992	0.0303
	(24.1)	(1.0)	(96.4)	(1.7)	(0.0)			
HVM	0.0015	0.0055	2.8391	-0.0032	0.0000	1.8	0.986	0.0218
	(83.7)	(11.7)	(130.9)	(12.2)	(0.9)			
ALT	0.0216	0.0055	0.5785	-0.0015	-0.0002	1.9	0.985	0.0259
	(83.1)	(3.2)	(142.4)	(6.2)	(2.1)			
BAS	0.0053	0.0037	1.1206	-0.0015	0.0000	3.1	0.930	0.0136
	(161.4)	(11.4)	(244.1)	(16.6)	(1.0)			
DTE	0.0002	0.0046	5.0499	-0.0061	0.0000	3.8	0.877	0.0064
	(232.1)	(49.9)	(232.7)	(18.8)	(0.0)			
ALV	0.0155	0.0280	0.6453	-0.0039	-0.0002	3.8	0.872	0.0099
	(80.3)	(19.8)	(294.4)	(18.9)	(3.8)			
ADS	0.0503	0.0252	0.3528	-0.0028	-0.0002	5.1	0.750	0.0191
	(31.8)	(3.0)	(141.0)	(3.4)	(1.4)			
VOW	0.0062	0.0020	1.0472	-0.0011	0.0001	6.5	0.595	0.0166
	(21.8)	(1.4)	(195.8)	(3.7)	(0.2)			
EOA	0.0053	0.0099	1.0663	-0.0027	0.0000	7.5	0.488	0.0107
	(79.1)	(12.0)	(252.7)	(12.9)	(0.8)			
CBK	0.0016	0.0020	2.4055	-0.0018	-0.0001	9.7	0.283	0.0248
	(139.6)	(10.5)	(152.0)	(13.5)	(1.8)			
BMW	0.0051	0.0018	1.2029	-0.0013	-0.0001	9.7	0.284	0.0176
	(143.8)	(5.2)	(203.0)	(13.4)	(2.8)			
BAY	0.0024	0.0037	1.6352	-0.0017	0.0000	12.9	0.117	0.0167
	(168.5)	(17.4)	(225.8)	(17.6)	(1.2)			
LIN	0.0232	-0.0107	0.5728	0.0004	0.0004	15.0	0.059	0.0305
	(93.8)	(9.8)	(134.6)	(3.1)	(3.1)			
DPW	0.0025	0.0012	1.8362	-0.0015	-0.0001	16.0	0.042	0.0252
	(131.8)	(4.0)	(150.0)	(8.7)	(3.2)			
HEN3	0.0405	-0.0004	0.3937	-0.0005	-0.0003	26.9	0.001	0.0241
	(27.5)	(0.1)	(114.0)	(0.6)	(1.9)			
SCH	0.0099	0.0052	0.8250	-0.0017	0.0000	35.1	0.000	0.0201
	(92.5)	(6.1)	(168.8)	(9.4)	(0.4)			
CONT	0.0132	-0.0105	0.8166	0.0009	0.0002	51.0	0.000	0.0338
	(108.1)	(12.7)	(139.6)	(5.0)	(2.2)			
MAN	0.0105	-0.0086	0.9445	0.0006	0.0004	53.0	0.000	0.0359
	(105.0)	(10.2)	(134.5)	(3.4)	(5.2)			
LHA	0.0019	-0.0006	2.3210	-0.0008	0.0001	63.0	0.000	0.0307
	(137.1)	(2.0)	(150.8)	(6.4)	(4.3)			
MEO	0.0134	-0.0071	0.9066	0.0000	0.0000	99.1	0.000	0.0345
	(115.8)	(9.3)	(166.7)	(0.1)	(0.4)			
RWE	0.0051	0.0000	1.2460	-0.0009	0.0001	112.6	0.000	0.0188
	(178.5)	(0.1)	(210.3)	(11.5)	(3.3)			
TKA	0.0024	-0.0012	1.9075	-0.0002	0.0000	362.6	0.000	0.0285
	(139.4)	(5.0)	(158.9)	(2.2)	(1.3)			

2×4 quotes from the bid and ask side of the visible book are used to construct average update and average break even conditions. The exponential assumption on the distribution of the trade size is maintained. $\tau = \frac{\alpha \cdot \overline{m}}{\overline{P}}$, where $\overline{m}$ and $\overline{P}$ denote stock specific sample averages of the non-signed trade sizes (number of shares) and the midquotes, respectively. The stocks are sorted by ascending order of the $J-$statistic.

Table 2.5: Correlation of standardized adverse selection component τ with liquidity indicators.

liquidity variable	correlation	$p-$value
quoted spread (%)	0.873	$<$ .0001
effective spread (%)	0.794	$<$ .0001
realized spread (%)	0.050	0.824

The table reports the cross sectional Spearman rank correlations of the standardized adverse selection component τ reported in table 2.5 with average quoted, effective and realized spread reported in table 2.5. To compute the correlations we include the stocks for which the model is not rejected at 1 % significance level (22 out of 30 stocks). To obtain stock specific measures we take averages over all order book snapshots.

Figure 2.1: Comparison of implied and observed price schedules (visible book)

The figure depicts means an medians of implied and observed ask side price schedules of four selected stocks. In each figure the values on the horizontal axis show trade volumes (number of shares) up to the 0.9 quantile of the respective stock. The vertical axis show the per share price decrease that a sell trade of a given volume would incur if it were executed against the current book. The solid line depicts sample means and the short dashed lines sample medians computed by using all order book snapshots during the three month period. The bold long-dashed lines depict the mean slope implied by the estimation results reported in table 2.5 (baseline model that uses marginal break even and update conditions). The dash-dot lines and the long-dashed lines are the mean and the median of the book slope as implied by the estimation results reported in table 2.5 (revised specification which uses average break even and update restrictions). The stock in the left upper panel is Daimler Chrysler (DCX, from the largest trade volume quartile), the stock in the right upper panel is Bay. Hypo Vereinsbank (HVM, second volume quartile), the stock in the left lower panel is Altana (ALT, third volume quartile) and the stock in the right lower panel is Deutsche Boerse (DB1, fourth volume quartile).

Three

Iceberg orders and the compensation for liquidity provision

Limit order books in many markets contain hidden liquidity because traders are able to submit iceberg orders. We study the interaction between hidden liquidity and overall liquidity provision using a sample from a limit order market that includes both iceberg and limit orders. We report evidence that iceberg orders can be detected using public information and that market participants follow state-dependent order submission strategies. We show that iceberg orders influence the flow and price impact of market orders. After adjusting for those effects the marginal compensation for liquidity provision changes.

3.1 Introduction

Most limit order markets allow traders to submit iceberg orders and as a result the total liquidity in the order book consists of both hidden and displayed liquidity. Traders submit iceberg orders to try to reduce the price impact of large orders. The potential hidden liquidity in the order book is also relevant for traders who submit visible orders. Traders who submit limit orders try to detect hidden liquidity because it has an impact on expected payoffs for limit orders. Traders who submit market orders try to detect hidden liquidity because it has an impact on the cost of immediate execution. What is the impact of iceberg orders on the strategies followed by traders who demand liquidity? What is the impact of iceberg orders on the strategies of traders who supply liquidity?

We address these questions using a sample from a limit order market that includes both limit and iceberg orders. An iceberg order specifies a price, a total order size, and a visible peak size. The peak size is the maximum number of shares that is displayed to the market at any time. The remainder of the iceberg order is not displayed in the order book. When the first peak size has been fully executed, the displayed part is immediately replenished by a size equal to the peak size. At a given price level in the order book all displayed order volume has time priority relative to any undisplayed order volume, irrespective of the order entry times.

We document the impact that iceberg orders have on the order books and the price dynamics using a sample that includes iceberg and limit orders. Our sample is from the Xetra system of the German Stock Exchange and it includes all orders in the DAX-30 stocks for a three-month period. We document that iceberg orders have a significant impact on the price impact of market orders and on the market order flow. The price impact of a buy market order, measured after 10 minutes or 30 trades, is 2-4 basis points higher when there is a buy iceberg order in the book. Conversely, a market buy order submitted when a sell iceberg order is on the book has, on average, a price impact that is 2-5 basis points lower. The average size of buy market order increases by 13 to 27% when a buy iceberg order is present, and by 40 to 59% when a sell iceberg order is present. Both the price impact and the market order effect are approximately symmetric for buy and sell orders and imply that iceberg orders have an impact on the price dynamics and on the order books.

The replenishment rules for iceberg orders implies that an iceberg order is not truly hidden to a trader who observes the order book. The immediate replenishment with a new peak size after a transaction provides the market with a signal that the price level is likely to contain hidden liquidity. We construct an iceberg detection algorithm that uses the order book dynamics to make predictions of whether or not the bid or the ask side has an iceberg order. For a typical stock, the detection algorithm correctly detects approximately 70% of the books with iceberg orders and when the algorithm generates an iceberg signal it is correct for approximately 90% of the books. These results suggest that marginal liquidity providers can, on average, respond to iceberg orders by changing their order submission strategies.

Based on the insight that traders partly anticipate iceberg orders, we develop a state-dependent model of liquidity provision. Our model framework is similar to the one used in Frey and Grammig (2006) and Sandås (2001) and builds on the Glosten model for the order book (Glosten (1994a)). We use the model to decompose marginal payoff to liquidity provision to determine whether the compensation for liquidity provision is systematically different for order books with or without iceberg orders. We measure the discount or premia earned by marginal liquidity providers when iceberg orders are present. Our findings suggest that, after accounting for the changes in order flow and the price impact of the order flow, the compensation for liquidity provision is different in iceberg versus no iceberg order books. Overall, the marginal compensation for liquidity provision is higher than normal when there is at least one iceberg order on the same side of the book, and is lower than normal when there is at least one iceberg order on the opposite side of the book. The higher compensation for liquidity provision when an iceberg order is on the same side of the book is consistent with traders placing orders less aggressively when they anticipate their order competing with hidden liquidity.

The iceberg order type illustrates the trend towards a common prototypical implementation of the popular limit order market mechanism. The London Stock Exchange's SETS limit order market system did not feature iceberg orders originally—all limit orders were fully displayed—but in 2003 iceberg orders were introduced. The Toronto Stock Exchange reintroduced iceberg orders in 2002 after a six-year period without iceberg orders. The Australian Stock Exchange replaced its unique undisclosed order category with iceberg orders of the kind used elsewhere in 2006. The European Union's "Markets in Financial Instruments Directive" (MiFID), which is to be implemented by November 2007, includes iceberg orders as one exception from the goal of full pre-trade transparency. Given the growing popularity of iceberg orders it is important to understand the interaction between hidden liquidity and overall liquidity provision.

Several studies provide evidence on the use of iceberg or hidden orders in various limit order markets; Aitken, Berkman, and Mak (2001) for the Australian Stock Exchange; Bessembinder, Panayides, and Venkataraman (2007) and De Winne and D'Hondt (2007) for Euronext; Hasbrouck and Saar (2004) for INET; Harris and Hasbrouck (1996) for the Paris Bourse and Toronto Stock Exchange; Næs and Skjeltorp (2006) for the Oslo Stock Exchange; Pardo and Pascual (2006) for the Madrid Stock Exchange; and Tuttle (2006) for Nasdaq. Esser and Mönch (2005) present a framework for analyzing the optimal use of iceberg orders in a multi-period liquidation strategy. Our study add to the existing literature by focusing on the impact that iceberg order have on the traders demanding and supplying liquidity.

Many traders are interested in detecting iceberg orders and predicting their size since that will help them make better order placement decisions. Applications such as Credit Suisse First Boston's Guerilla, Instinet's Nighthawk, and Citigroup's Dagger are all design to detect hidden liquidity. Bongiovanni, Borkvec, and Sinclair (2006) and De Winne and D'Hondt (2007) report empirical evidence consistent with traders being able to predict the existence and location of iceberg and hidden orders. Our results also suggest that in the Xetra system a straightforward algorithm generates informative predictions of whether an iceberg order is present at the best quote.

The rest of the paper is organized as follows. We discuss the market we study and our sample in Section 2. We report empirical results for the price impact of order flow and the order flow distribution in Section 3. We present a framework for liquidity provision with iceberg orders in section 4. We present our empirical results in Section 5. Section 6 concludes. Appendix A reports more details on our sample and our estimation results.

3.2 The Xetra Trading System and the Sample

Our sample includes all order entries, trades, and cancellations in the thirty stocks that comprise the DAX 30 German blue chip index for the period January 2nd to March, 31st, 2004. The sample is from the Xetra trading system, which is operated by the Frankfurt Stock Exchange. During the first quarter of 2004 trading in the Xetra system accounted for approximately 96% of all trading in Germany of the DAX 30 stocks (Source: Deutsche Börse Group, Factbook 2004). Xetra is a fully electronic trading platform that matches buy and sell orders from licensed traders in a limit order book. There were 302 licensed traders during our sample period, and there were no designated market makers for our sample stocks.

Traders in the Xetra system can, in addition to market and limit orders, submit iceberg orders.[1] An iceberg order specifies a price, a total size, and a peak size. The peak size is the maximum displayed volume of the order. When the iceberg order is submit-

[1]Iceberg orders were introduced with Xetra release 5.0, in October 2000.

ted, the first peak size is displayed in the order book. At that time, the hidden volume of the order is equal to the order's total size minus its peak size. When the first peak size has been fully executed, the system automatically replenishes the visible part of the iceberg order by a number of shares equal to the peak size, and reduces the hidden part of the order by the corresponding number of shares. The system continues to replenish the iceberg order until no hidden volume remains or the order is cancelled.[2]

Orders in the order book are given priority according to price, display condition and time. Sell orders at lower prices have priority relative to sell orders at higher prices, irrespective of the orders' time of submission or display condition. At the same price level, orders that were submitted earlier usually have priority relative to order submitted later. The exception is that all displayed orders have priority relative to all hidden volume, irrespective of the submission time. The peak sizes of iceberg orders are considered as displayed volume, indistinguishable from regular limit orders. Both limit orders and iceberg peaks are then executed according to time priority. All visible volume must be executed at a given price level before a market order or a marketable limit order can execute against hidden volume. After the execution of all visible volume, iceberg orders replenish the order book by displaying a new peak size in the order book. In that case the iceberg order (both the hidden and the visible part) is assigned a new time stamp. In the case of multiple iceberg orders at the same price, the combined peak size volumes are to be displayed on the book simultaneously. Iceberg orders are valid for the current trading day only.

We reconstruct the sequence of order books from the event histories in our sample. The order records include a flag for iceberg orders and we use that flag to construct complete histories for all limit and iceberg orders. From these histories we reconstruct snapshots of the visible and hidden order books before every transaction. In addition, we construct individual order histories that we use to examine the placement, execution, cancellation, and duration of limit and iceberg orders.

The continuous trading in the Xetra system begins at 9:00 after the opening auction and ends at 17:30. A closing auction takes place between 17:30 and 17:35. In addition, continuous trading is interrupted by a mid-day auction that takes place between 13:00 and 13:02. We focus on the order submissions during the continuous trading period; 9:00 to 13:00 and 13:02 to 17:30. Our reconstruction procedure takes into account any

[2]The iceberg order's last peak size may be smaller than its initial peak size.

effect that order submissions, transactions, or cancellations that occur in the auctions have on the state of the order book during continuous trading.

Table 3.1 reports descriptive statistics for the stocks in our sample. We divide the sample stocks into four trading activity groups based on the total trading volume during the sample period—stocks in group 1 have the highest and stocks in group 4 the lowest trading trading volume. The second column lists the ticker symbols for the stocks.[3] The next five columns report the free-float market capitalization measured in billions of euros as of the end of 2003; the trading volume for the sample period measured in billions of euros; the average transaction price; the average number of trades per day measured in thousands of trades;and the average trade size in thousands of shares. The last two columns report the average relative and absolute bid-ask spreads.

Table 3.2 reports descriptive statistics for iceberg and limit orders. Columns three and four report the percentage of all submitted and executed shares that are iceberg orders; market orders and marketable limit or iceberg orders are excluded. On average, iceberg orders represent 8% of all shares submitted, but they represent 16% of all shares executed implying a higher execution rate for iceberg orders than for regular limit orders. The higher execution rate may reflect a difference in the average placement of iceberg and limit orders, or a difference in the time that iceberg and limit orders remain outstanding. The last two columns on the right report the median distance between the same-side best quotes and the iceberg and limit order prices. While there are some stocks, for example, DBK, DCX and BMW for which icebergs tend to be placed closer to the best quotes there are other stocks such as SAP, ADS or ALT for which the opposite is true. Overall the median distance is 3.6 basis points for icebergs and 3.9 basis points for limit orders and we conclude that there is little evidence that the higher execution rate for iceberg orders is driven by difference in the order placement.

The middle four columns (columns five through eight) of Table 3.2 provide descriptive statistics on order sizes. The fifth column reports the average size of limit orders, which ranges from 1,400 shares in group 1 to 400 shares in group 4. The next column reports the average peak size of iceberg orders, which ranges from 3,200 for group 1 to 1,500 for group 4 implying that, on average, an iceberg order's visible portion is approximately two to three times the average limit order size. The sixth column, labeled 'Total Size/Peak Size,' reports the average of the ratio of the iceberg order's total size

[3]Table 3.12 in the appendix provides the full company names and the associated ticker symbols.

to its peak size. This ratio is often clustered at even multiples such as five or ten times the peak size and that explains why the average ratio is with one exception between 5 and 10 for all stocks. The column labeled 'Executed Shares/Peak Size' reports the ratio of executed shares to peak size for all iceberg orders whose first peak size was executed. The average ratio is 4.6 compared with an average size of 7.6 implying that on average iceberg orders are replenished almost four times conditional on the first peak size being executed. This implies that approximately 80% of the executed shares for iceberg orders come from the (initially) hidden parts of the iceberg orders.

Table 3.6 provides more details on the execution and duration of iceberg and limit orders. Panel A of the table, reports the percentage of limit and iceberg order with at least one execution. Almost one-half (47%) of all iceberg orders receive at least one execution whereas only 14% of all limit orders receive at least one execution. The difference is somewhat surprising given that, on average , iceberg and limit orders are submitted at approximately the same distance from the best quotes. The last column of Panel A reports the ratio of the median duration for iceberg orders to the median duration of limit orders. The ratio of median durations of 7.1 is evidence that iceberg orders may be more likely to received executions because they spend a longer time in the order book. The longer time that iceberg orders on average spend in the order book also implies that the chance that a randomly selected order book contains at least one iceberg orders is substantially greater than what one would expect based on the iceberg orders' 8% share of all submitted shares.

How different are iceberg orders from regular limit orders? Does it matter whether the iceberg order is likely to be undetected or not? We examine two sub-samples to gain a better understanding of possible differences. Panel B of Table 3.6 report, for two sub-samples, the execution rates and the ratio of median time-to-fill for iceberg and limit orders.

The sub-sample labeled 'First Peak' consists of iceberg orders that have not been executed yet. In addition, for every stock the iceberg orders are selected to have a peak size that differs by 10% or less from the modal peak size and have an order price relative to the same-side best quote that falls between the 30th and 70th percentile for all iceberg orders. The limit order sample includes all limit orders with order sizes and quantities that fall within the price and size cut-offs for the iceberg orders. The iceberg orders in this sub-sample are special in the sense that there is little information that can help identify whether a given order originates from a limit or an iceberg order.

The execution frequencies for the First Peak sample are comparable with an average of 86% for limits and 90% for iceberg orders. The ratio of median time-to-fill is 1.0. These figures suggest that iceberg orders that are likely to be undetected behave very similarly to regular limit orders with a comparable order price and size.

The Second Peak includes only iceberg orders whose first peak was executed. The replenishment of these iceberg orders provides the market with a signal implying that these iceberg orders are likely to have been detected. The execution of the first peak implies that at the time of the execution the iceberg order was at the front of the order queue. Accordingly, we keep only the limit orders that undercut the best quote. The execution frequencies are comparable although the relative ordering is reversed relative to the first peak case. The time to fill, however, is shorter for iceberg orders with an average ratio of 0.7; 23 out of 30 ratios of the median times-to-fill that are below 1 and only one ratio is above one (TUI). Iceberg orders that may have been detected appear to attract order flow and execute more rapidly when they execute.

Table 3.4 reports descriptive statistics for the average limit order books observed before transactions when the order book contains no iceberg orders and when the order book contains iceberg orders on either side of the book. The order book snapshots are created 1/100th second before every transaction. For each stock we compute an average spread and depth for each constellation of iceberg orders. The table reports the mean values of these averages across the stock in each group.

The top part of the table reports average spreads and differences of spreads. The average bid-ask spread for stocks in group 1 is 4.9 basis points for books without iceberg orders and 4.2 basis points for books with at least one iceberg order on either the bid or the ask side. Across all trading activity groups, the average spread without iceberg orders is 0.8 to 1.4 basis points wider than the spread with one or more iceberg orders in the book. For all stock, we reject the null hypothesis of the difference in spreads being zero.

The spread between the best and second best price levels in the order book is narrower when there is an iceberg order on the opposite side of the book, but it is wider when there is an iceberg order on the same side of the book. The spread is 0.2 to 0.5 basis points tighter with the iceberg order on the opposite side and we reject the null of no difference for 27 out of the 30 stocks. The difference is greater in magnitude when the iceberg order is on the same side; the difference ranges from 0.6 to 1.7 basis points. The average magnitudes of the difference suggest that the tighter inside spread when iceberg orders are present are approximately offset by a wider spread between the best and the second best quotes.

The bottom part of Table 3.4 report statistics on the *visible* order book depths at the best and the second best price levels with or without iceberg orders at the same or opposite sides of the order book. The visible depth at the best quote is greater when an iceberg order is present at the best quote at either the opposite or the same side of the book. The depth increases by more when the iceberg order is on the same side. The null hypothesis of no change is rejected for 29 out of the 30 stocks in the same side case and for 23 out of the 30 stocks for the opposite side case. The magnitude of the increase in the same side case corresponds to approximately one average-size limit order and thus is less than the peak size of a typical iceberg order. This implies that traders submit fewer or smaller limit orders when an iceberg order is present, but that the drop is small enough to make the net depth higher. At the second best quote level, the depths are are fairly similar. We reject the null of no difference for approximately half of the cases, but with the exception of group 2 and opposite side, the differences are small in magnitude.

Overall, the presence of iceberg orders in the book is associated with narrower inside spreads and greater depths at the best quote levels. The presence of iceberg orders appear to increase the amount of liquidity provided and to lower the average price of liquidity. We next examine to what extent the presence of iceberg orders also influence the price impact and the order flow.

3.3 Price Impact and Order Flow with Iceberg Orders

Iceberg orders may influence the short-term price dynamics in a number of ways. We start by examining whether the price impact of the order flow is different when iceberg orders are present in the order book. We then turn to the order flow itself and examining whether the distribution of market orders is different when iceberg orders are present. For tractability we restrict our focus to iceberg orders at the best bid or ask quotes. We will later address the issue of the extent to which traders can detect iceberg orders. Here we examine the problem from the perspective of an observer who perfectly observes whether there is an iceberg order at the best bid or ask quotes but who does not observe the size of the iceberg order.

Price Impact

An iceberg order may have a direct and an indirect effect on the price impact. A direct effect may arise because iceberg orders add significant depth and have longer durations than regular orders. Incoming market orders therefore need to 'eat through' the iceberg orders for prices to move unless the iceberg order is cancelled. An indirect effect may arise because other traders change their order submission strategies when they believe an iceberg order is present, see, Harris (1997). Traders may speed up their own buying or selling by submitting more aggressive orders. Some traders may follow front running strategies.

We estimate price impact regressions that allow for different price impact of order flow when iceberg orders are present in the order book. Let mq_t denote the mid-quote that is observed immediately prior to the time t transaction. Let $\Delta mq_{t+\tau,t}$ denote the relative change in the mid-quote from time t to time $t+\tau$ measured in basis points. We denote the size of the time t market order by m_t; $m_t > 0$. The market order size is normalized for each stock so that m_t is measured in units of the average market order size. Let d_t denote the sign of the market order at time t, $d_t = 1$ for a market buy order and $d_t = -1$ for a market sell order. Let $\tilde{m}_t = m_t d_t$ denote the signed market order quantity.

Let I_t^{own} and I_t^{opp} denote iceberg indicators. An order book with an iceberg order at the best bid quote and no iceberg orders at the best ask quote at time t when a market buy order is submitted has $I_t^{own} = 1$ and $I_t^{opp} = 0$. The same order book when a market sell order arrives has $I_t^{own} = 0$ and $I_t^{opp} = 1$. The indicators are defined from the perspective of the direction of the market order so that either there is an iceberg order in the 'opposite' direction or there is an iceberg in the 'own' direction or both or no iceberg orders at all.

We consider three time horizons τ for measuring the mid-quote change. The shortest horizon is the next trade; trade to trade mid-quote changes. The next trade horizon may be too short since iceberg orders are long-lived. We therefore also estimate the regressions for a 30 trade horizon and a 10 minute horizon. The longer time horizons add more noise but they also allow us to detect effects that are revealed with a lag. If the time horizon goes beyond the closing time we use the closing price as the revised mid-quote.

We estimate the following regression:

$$\Delta mq_{t+\tau,t} = c + a_0 d_t + b_0 \tilde{m}_t + (a_1 d_t + b_1 \tilde{m}_t) I_t^{own} + (a_2 d_t + b_2 \tilde{m}_t) I_t^{opp} + \epsilon_t. \quad (3.1)$$

The baseline price impact without iceberg orders is given by a fixed component a_0 and a variable component b_0 that depends on the order size m_t. The change in the price impact for the case with iceberg orders at the same side as the market orders—a market buy order when there is an iceberg at the best bid—is capture by a_1 and b_1. The corresponding change in the price impact when there is an iceberg order in the order book at the opposite side of the market order—a market buy order when there is an iceberg at the best ask—is capture by a_2 and b_2.

Table 3.5 report the average parameter estimates and standard errors by trading activity group and by time horizon τ. The top panel reports the results for a time horizon of 10 minutes. The middle panel reports results for a time horizon of 30 trades and the bottom panel reports results for a next trade horizon, i.e., trade to trade price impact.

The average parameter estimates for the baseline price impact, a_0 and b_0, are monotonically increasing as we go from the most active trading group to the least active trading group across all three time horizons. The two longer time horizons, 10 minutes and 30 trades, have larger average a_0 and b_0 estimates implying that the next trade horizon may not capture the full price impact of an order.

The change in the price impact when an iceberg order is present on the same side is zero for the next trade horizon but for the 10- minute and 30-trade horizons it is positive and comparable in magnitude of the baseline fixed component. The variable component, b_1, is own average negative but often not significantly different from zero.

The corresponding estimate for the opposite side effect are both negative for all three time horizons with estimates that are larger in magnitude for the longer horizons. A negative a_2 is consistent with a sort of buffer effect whereby the iceberg order absorb a lot of market order flow and reduced the impact of the market order flow relative to the baseline case. The negative estimates of b_2 suggest that the per unit price impact is also reduced. This effect may partly reflect the greater visible depth in the order book when iceberg order are present (see Table 3.4).

Order Flow

A trader who anticipates an iceberg order at the best ask quote may choose to submit a larger buy order to take advantage of the greater total depth. We regress the market order size on iceberg indicators to determine if there is a such an effect in our sample. We estimate the following regression:

$$m_t = c + b_1 I_t^{own} + b_2 I_t^{opp} + \epsilon_t \tag{3.2}$$

Panel A of Table 3.6 reports the average parameter estimates and standard errors by trading activity group. The average estimate of c is slightly below one which means that when there are no iceberg orders we observe market orders that are 5-8% smaller than the overall average market order size. The average estimates of the parameter on I_t^{own} are 0.13 for the first group and increase to 0.27 for the fourth group. The implication is that when there is an iceberg order at the best bid or ask then even market orders in the same direction tend to be larger. The corresponding average parameters on I_t^{opp} are also positive and range from 0.37 to 0.59 implying that when an iceberg orders is present at the opposite side of the book the average market order size is 37% to 59% larger.

The bottom half of the Panel A provides the average F-statistic for the null hypothesis that the parameters b_1 and b_2 are jointly equal to zero and the number of stocks for which the hypothesis is rejected. The null is rejected for all thirty stocks. The next row reports the t-statistic and the number of rejections of the null hypothesis that the market sell order sizes are different from the market buy order sizes. For the most actively traded stocks the null is rejected in six out of seven cases but for the other activity groups the evidence is mixed with a total of 10 rejections for 23 stocks. The last row report the average F-statistics and the number of rejections of the null that the own and opposite effects are identical for buy and sell market orders. The evidence for different effects is mixed with 14 rejections with the strongest evidence against the null for the most actively traded stocks. The parameter estimates, which are not reported to save space, are in general small, especially compared to the baseline model estimates, and we therefore focus on the symmetric case below.

In order to determine whether the existence of iceberg orders also affect the skewness of the order flow we estimate a logit model of the probability of a market buy order. Panel B of Table 3.6 reports the average marginal effects and standard errors for the logit model. The marginal effects are negative for I^{own} and positive for I^{opp}. The negative marginal effect for the own side ranges from -0.13 to -0.15 implying that if the baseline probability is one-half the probability of a market buy is 0.35 to 0.37 when there is an iceberg at the best bid. The positive opposite side marginal effect ranges from 0.12 to 0.16 and implies that the probability of observing a market buy when there is an iceberg at the best ask is 0.62 to 0.66 assuming again that the baseline probability of one-half.

Overall, the results provide evidence of an opposite and an own side effect of iceberg orders. The opposite side effect is that iceberg orders on the opposite side of market orders tend to act as buffers that reduce the price impact of market orders and the price impact per unit traded. These iceberg orders tend to attract market orders and they tend to attract larger market orders. The own side effect is that iceberg orders tend to magnify the price impact of market orders in the same direction as the iceberg order. The market orders in the same direction tend to be larger but the probability of a market order in the same direction is actually smaller.

The results provide some evidence that consistent with the iceberg order submitters being perceived as large uninformed traders by the market. The larger market order sizes when iceberg orders are present is consistent with that. The higher execution frequencies for iceberg orders are also consistent with that view.

We now turn to the problem faced by liquidity providers when there may be iceberg orders in the order book. The liquidity provider we have in mind is not submitting iceberg order but instead submit regular limit orders that make up the visible limit order book. The expected marginal payoffs for a limit order depends on the order book, the price impact, and the order flow. We will incorporate the evidence presented above into a framework that allows us to determine empirically the discount or premium to liquidity provision in a setting with iceberg orders.

3.4 A Model of Liquidity Provision with Iceberg Orders

We characterizes the compensation for liquidity provision in a limit order book that may contain hidden liquidity. We start by characterizing the limit order book and the compensation for liquidity provision with no hidden liquidity. We then consider the case in which traders anticipate the existence of hidden liquidity, but face uncertainty about the amount of hidden liquidity. Finally, we present a set of moment conditions implied by our model.

The Order Book With No Hidden Liquidity

We focus on how the ask side of the book is determined, the bid side is determined analogously.

All liquidity providers agree on a fundamental value of X_t for the stock at time t; X_t may be interpreted as the liquidity providers' time t expectation of the liquidation value of the stock.

The time t market order is submitted by a trader who may be informed about the future value of the stock. Let $m_t > 0$ denote the size of the time t market order. Let d_t be an indicator for the trade direction and let $\tilde{m}_t$ denote the signed market order size, $\tilde{m}_t = d_t m_t$. We assume that the market order size is exponentially distributed with a mean order size of λ.

The following three components determine the change in the fundamental value between t and $t+\tau$: a drift term μ, new public information, and private information revealed by the market order flow. The new fundamental value at $t+\tau$, $X_{t+\tau}$, is given by:

$$X_{t+\tau} = X_t + \mu + \alpha_0 d_t + \alpha_1 \tilde{m}_t + \epsilon_{t+\tau}, \tag{3.3}$$

in which, α_0 and α_1 are parameters that measure the information content of the market order flow, and $\epsilon_{t+\tau}$ reflects the public news arrival between t and $t+\tau$.

The ask side of the limit order book at time t is characterized by a series of ask quotes $p_t^{ask,1} < p_t^{ask,2} <, \ldots, < p_t^{ask,K}$. The total number of shares offered at the kth best ask is denoted by $q_t^{ask,k}$. The cumulative number of shares offered at all ask quotes lower than or equal to the kth highest ask, $p_t^{ask,k}$, is denoted by $Q_t^{ask,k}$, and is determined as, $Q_t^{ask,k} = \sum_{i \leq k} q_t^{ask,i}$. The bid quotes, depths, and cumulative depths are defined analogously and denoted, $p_t^{bid,k}$, $q_t^{bid,k}$, and $Q_t^{bid,k}$.

Limit orders are executed in a discriminatory fashion. A limit sell order that is executed if a market buy order of size m arrives will also be executed by any market buy order that is greater than m. Upper- and lower-tail expectations of the fundamental value summarize the expected price impact conditional on execution. Let $x_{t+\tau}^U(m;\theta)$ denote the upper-tail expectation. It is defined as the expected value of $X_{t+\tau}$ conditional on a market buy order at time t of size m or greater:

$$\begin{aligned} x_{t+\tau}^U(m;\theta) &= E[X_{t+\tau}|m_t \geq m, d_t = 1] \\ &= X_t + \mu + \alpha_0 + \alpha_1 E[m|m_t \geq m, d_t = 1] \\ &= X_t + \mu + \alpha_0 + \alpha_1(m + \lambda), \end{aligned} \tag{3.4}$$

with θ denoting the vector of parameters, $\theta = \{\lambda, \mu, \alpha_0, \alpha_1\}$. The corresponding lower-tail expectation is denoted $x_{t+\tau}^L(m;\theta)$.

The expected payoff for a marginal limit order at the kth best ask quote that is executed at time t is denoted by $\delta^{ask,k}$ and is determined as:

$$\begin{aligned} \delta^{ask,k} &= E[p_t^{ask,k} - X_{t+\tau} - |\text{marginal limit at } p_t^{ask,k} \text{ executes at } t], \\ &= E[p_t^{ask,k} - X_{t+\tau}|m_t \geq Q_t^{ask,k}, d_t = 1], \\ &= E[p_t^{ask,k} - x_{t+\tau}^U(Q_t^{ask,k};\theta)]. \end{aligned} \tag{3.5}$$

The corresponding expected payoff for a marginal limit at the kth best bid level is denoted by $\delta^{bid,k}$ and is determined analogously.

Perfect competition among liquidity providers would imply that the expected payoffs, $\{\delta^{ask,k}, \delta^{bid,k}\}_{k=1}^K$, at each bid and ask level are driven to the per share cost of processing orders. In the next section, we consider how the possibility of hidden liquidity modifies the expected payoffs for the marginal limit orders.

The Order Book With Hidden Liquidity

We consider a special case in which all traders know whether or not the book contains hidden liquidity at the best bid and ask quotes. Traders do not, however, know the amount of hidden liquidity. We show below, in section 3.5, that it is reasonable to assume that traders often can detect the existence of hidden liquidity even if the total amount of hidden liquidity is uncertain.

An iceberg has a direct and an indirect effect on limit orders in the order book. The direct effect arises because any hidden volume at higher bids or lower asks have price priority relative to all limit orders at a lower bids or higher asks. For example, if there is hidden volume at the best bid level, then the length of the queue of limit orders ahead of the marginal limit order at the second best bid level includes the visible volume at the two levels plus the hidden volume at the best level. One implication is that the upper- and lower-tail expectations must be revised to take into account the any market order that will execute the marginal limit order is strictly greater than the total visible volume ahead of the order.

Liquidity demanders may adjust their strategies, when there is hidden liquidity, causing the market order flow and the information content of the market order flow to change. This is the indirect effect. The indirect effect changes the payoffs to limit orders at the best bid and ask levels despite the fact that these limit order have price priority with respect to any hidden volume. The indirect effect changes the payoffs on limit orders at the bid side of the book when the only hidden liquidity is on the ask side, and vice versa.

Let I_t^{bid} be an indicator that takes on value one if there is at least one iceberg order at the best bid level in the limit order book at time t. Let I_t^{ask} be the corresponding indicator order for the best ask level. Let $h_t^{ask} = [1 \; I_t^{ask} \; I_t^{bid}]$ be a vector that captures the state of the order book from the perspective of liquidity providers on the ask side. When there are no iceberg orders in the order book $h_t^{ask} = [1 \; 0 \; 0]'$. When there is an iceberg order on the ask side and no iceberg on the bid then $h_t^{ask} = [1 \; 1 \; 0]'$. When there is an iceberg order on the bid side and no iceberg on the ask side then $h_t^{ask} = [1 \; 0 \; 1]'$. When there are iceberg orders on both sides then $h_t^{ask} = [1 \; 1 \; 1]'$. A vector h_t^{bid} is defined analogously; $h_t^{bid} = [1 \; I_t^{bid} \; I_t^{ask}]$.

The hidden volume at the kth best ask quote is denoted by $\hat{q}_t^{ask,k}$ and the corresponding hidden volume at the kth best bid quote is denoted $\hat{q}_t^{bid,k}$.

We define a vector h_t^{mkt} that captures the iceberg state from the market order submitters perspective; an iceberg order at the ask side is one the ‘own’ side for a market sell order but on the ’opposite’ side for a market buy order.

If $d_t = 1$ then $h_t^{mkt} = [1 \; I_t^{bid} \; I_t^{ask}]$, and if $d_t = -1$ then $h_t^{mkt} = [1 \; I_t^{ask} \; I_t^{bid}]$. Consider a market buy order, $\tilde{m}_t > 0$. If there are no iceberg orders, then $h_t^{mkt} = [1 \; 0 \; 0]'$; if there is an iceberg order on the bid side then $h_t^{mkt} = [1 \; 1 \; 0]'$; if there is an iceberg order on the ask side then $h_t^{mkt} = [1\ 0\ 1]'$; and, finally, if there are iceberg orders on both sides we have $h_t^{mkt} = [1\ 1\ 1]'$.

While traders observe the existence of an iceberg order they do not observe the exact quantity. We denote the expected hidden size of an iceberg order by η.

There are now three parameters that capture the market order size and we denoted the vector of parameters by $\bar{\lambda} = [\lambda_0\ \lambda^{own}\ \lambda^{opp}]$. Similarly, we have three parameters for the fixed and the variable components of the price impact function. We denote them $\bar{\alpha}_0 = [\alpha_0\ \alpha_0^{own}\ \alpha_0^{opp}]$ and $\bar{\alpha}_1 = [\alpha_1\ \alpha_1^{own}\ \alpha_1^{opp}]$. Let θ denote the vector of parameters with $\theta = \{\bar{\lambda}\ \mu\ \bar{\alpha}_0\ \bar{\alpha}_1\}$.

The payoffs for liquidity provision are also state dependent. But for liquidity provision the state-dependence is defined from the perspective of the liquidity provider. Consider a marginal limit order at the best ask quote. The parameter δ_1^{own} measures the difference in the compensation for liquidity provision in the presence of an iceberg order at the best ask quote. The parameter δ_1^{opp} measures the difference in the compensation for liquidity provision in the presence of an iceberg order at the best bid quote. The parameters are symmetrically defined for the bid side. Again there is three-parameter vector denoted by $\bar{\delta}_1 = [\delta_1\ \delta_1^{own}\ \delta_1^{opp}]$. The corresponding vector for the second-best level is denoted $\bar{\delta}_2 = [\delta_2\ \delta_2^{own}\ \delta_2^{opp}]$.

The state-dependent upper-tail expectations for the best ask level $x_{t+\tau,i}^U(q_t^{ask,1};\theta)$ take the following form:

$$x_{t+\tau,0}^U(q_t^{ask,1};\theta) = X_t + \mu + \bar{\alpha}_0 h_t^{ask} + \left(\bar{\alpha}_1 h_t^{ask}\right)\left(q_t^{ask,1} + \bar{\lambda} h_t^{ask}\right), \quad (3.6)$$

in which $\bar{\alpha}_0 h_t^{ask}$ captures the shift in the fixed price impact when iceberg orders are present, and $\bar{\alpha}_1 h_t^{ask}$ captures the effect on the variable component. The interaction between the effect on the price impact and the effect on the market order flow is captured by the product of $\bar{\alpha}_1 h_t^{ask}$ and $\bar{\lambda} h_t^{ask}$.

For the second-best ask level the upper-tail expectations for the take the following form:

$$x_{t+\tau,0}^U(Q_t^{ask,2};\theta) = X_t + \mu + \bar{\alpha}_0 h_t^{ask} + \left(\bar{\alpha}_1 h_t^{ask}\right)\left(Q_t^{ask,2} + \bar{\lambda} h_t^{ask} + I_t^{ask}\eta\right), \quad (3.7)$$

in which the main difference relative to equation 3.6 is that expected market order size conditional on execution is greater both because of the visible order quantity, $Q_t^{ask,2}$, and because of the hidden quantity, $I_t^{ask}\eta$.

Moment Conditions

We use the restrictions implied by the model to specify a set of moment conditions.

The state-dependent marginal payoff conditions generate a set of moment conditions that involve the lower-tail expectations, the δ parameters, and the iceberg state vector for the bid side of the order book:

$$h^1(\theta) = E\left[\left[\begin{array}{l} x^L_{t+\tau}(q^{bid,1}_t;\theta) - p^{bid,1}_t - \bar{\delta}_1 h^{bid}_t \\ x^L_{t+\tau}(Q^{bid,2}_t;\theta) - p^{bid,2}_t - \bar{\delta}_2 h^{bid}_t \end{array}\right] \otimes h^{bid}_t\right] = 0. \tag{3.8}$$

Similarly, for the ask side we get a set of moment conditions that involve the upper-tail expectations, the δ parameters, and the iceberg state vector for the ask side of the order book:

$$h^2(\theta) = \left[\left[\begin{array}{l} p^{ask,1}_t - x^U_{t+\tau}(q^{ask,1}_t;\theta) - \bar{\delta}_1 h^{ask}_t \\ p^{ask,2}_t - x^U_{t+\tau}(Q^{ask,2}_t;\theta) - \bar{\delta}_2 h^{ask}_t \end{array}\right] \otimes h^{ask}_t\right] = 0. \tag{3.9}$$

The state-dependent price impact and market order flow imply the following set of moment conditions:

$$h^3(\theta) = \left[\left[\begin{array}{l} m_t - \bar{\lambda} h^{mkt}_t \\ X_{t+\tau} - X_t - \mu - \left(\bar{\alpha}_0 h^{mkt}_t\right) d_t - \left(\bar{\alpha}_1 h^{mkt}_t\right) \tilde{m}_t \\ \left(X_{t+\tau} - X_t - \mu - \left(\bar{\alpha}_0 h^{mkt}_t\right) d_t - \left(\bar{\alpha}_1 h^{mkt}_t\right) \tilde{m}_t\right) d_t \\ \left(X_{t+\tau} - X_t - \mu - \left(\bar{\alpha}_0 h^{mkt}_t\right) d_t - \left(\bar{\alpha}_1 h^{mkt}_t\right) \tilde{m}_t\right) \tilde{m}_t \end{array}\right] \otimes h^{mkt}_t\right] = 0 \tag{3.10}$$

Finally, the size of the iceberg orders, η, is identified from the following moment conditions:

$$h^4(\theta) = \left[\begin{array}{l} \left(\hat{q}^{ask}_t - \eta\right) 1(\hat{q}^{ask}_t > 0) \\ \left(\hat{q}^{bid}_t - \eta\right) 1(\hat{q}^{bid}_t > 0) \end{array}\right] = 0. \tag{3.11}$$

We replace X_t and $X_{t+\tau}$ by the mid-quotes observed immediately before the transactions at times t and $t+\tau$. The time horizon, τ, is set to 30 trades.

We stack the moment conditions, $h^1(\cdot), ..., h^4(\cdot)$ and estimate the model parameters using GMM in two stages. Table 7 summarizes the moment conditions. In the second-stage estimation we use a Newey-West type weighting matrix with a 10 lags Bartlett kernel.

3.5 Empirical Results

We start by presenting our parameter estimates for the model of liquidity provision with iceberg orders. We then study the robustness of the main findings to uncertainty about the hidden liquidity.

Parameter Estimates for the Liquidity Provision Model

Table 3.7 summarizes the parameters and the moment conditions of our model. We estimate the model separately for each stock and report average parameter estimates by group. Table 3.8 report the average parameter estimates for the iceberg order book model by trading activity group. We provide the parameter estimates for each stock in the appendix in table 3.13.

Panel A of Table 3.8 reports the estimates for the price impact function. Overall, the parameter estimates for the price impact function are very close to the results for the price impact regression reported in table 3.5; time-horizon 30 trades.

Panel B reports the estimates for the discount or premium for marginal liquidity provision at the best and second best order book levels. The δ_1 estimates are on average negative and range from 0.8 to 1.5 basis points. One interpretation of the negative estimates is that, on average, traders who determine the marginal prices at the best bid or ask level have some intrinsic reason for trading. Accepting a negative payoff of this magnitude is rational if the alternative is to pay one-half of the bid-ask spread, which ranges from 4 to 7 basis points. The estimates of δ_2 are positive and range from 1.2 to 2.4 basis points implying that liquidity providers a the second best level expect to have a positive net payoff after accounting for the adverse selection cost.

The state-dependent δ_i^{own} parameters ($i = 1, 2$) are positive for both the best and the second best levels albeit that for the second best level the estimates are typically not significantly different from zero. For the best price level, however, the negative average marginal payoffs turns in to a positive payoff when there is an iceberg order at the best price level. One interpretation is that limit orders that undercut an iceberg order earn a positive marginal payoff perhaps because the iceberg order limits the price impact of incoming market orders.

Panel C of Table 3.8 reports the estimates for the market order flow and iceberg order size parameters. Again, the estimates for the market order size—the λ's—are close to the regression results reported above in table 3.6. The average hidden size of iceberg orders ranges from 9 to 14 times the normal market order size. There is no clear pattern across the activity groups.

Table 3.9 reports the average value for the net marginal compensation for liquidity provision by order book level and by iceberg state. The top part reports the average values for δ_1 which is the baseline compensation for liquidity provision. Our results here are in line with the findings in Frey and Grammig (2006) and Sandås (2001) and suggest that the marginal orders at the best quotes are not submitted by value traders but instead by patient liquidity traders. The positive estimate for δ_2 may be interpreted as the cost of providing liquidity, if value traders are the marginal providers and liquidity supply is competitive. To identify the proportion of δ_2 that reflects order processing cost versus any rents one would need some additional restriction like the one used in Biais, Bisiere, and Spatt (2002).

The marginal compensation when iceberg orders are on the own and opposite side of the order book show two main regularities. First, when iceberg orders are on the same side of the book, then the marginal compensation for liquidity provision is positive both for the best and second-best quotes. Second, when iceberg orders are on the opposite side the marginal compensation for liquidity provision is negative both for the best quotes.

The positive compensation for liquidity provision at the best level when iceberg orders are present on the same side suggests that traders are not bidding aggressively enough when it is likely that their limit order compete with hidden liquidity. Similarly, at the second best quotes the results provide some support for the idea that after controlling for other factors traders bid less aggressively when they are likely to compete with hidden liquidity.

Conversely, when iceberg orders are on the opposite side we observe that traders are bidding too aggressively. Of course, one may also interpret this as evidence that when iceberg orders are present on the opposite side then liquidity traders tend to determine the marginal prices. Our evidence on the order flow suggests that when iceberg orders are present there is more frequent and larger market orders making trades more likely.

Robustness

In this section we address the robustness of the our findings. Traders in the market do not observe the true iceberg states. Instead they have to use available information to try to detect or predict the existence of iceberg orders. Even if we assume that market participants know the iceberg states, our model still captures the uncertainty of the precise size of the hidden volume. To improve from that we develop a reasonable approximation of the possible predictions used by the traders.

An Iceberg Detection Algorithm

The limit order book changes provides signals to traders about the existence of iceberg orders. When a trade occurs that involves the execution of volume in excess of the displayed volume at given price, it is very likely that the price contains an iceberg order. The ratio between total size to peak size in Table 3.2 implies that after the replenishing of one peak, additional hidden volume is to typically expected.[4] This implies that one can detect iceberg orders comparing the recent history of transaction prices and volume with the transition of the visible order book.

We construct an iceberg detection algorithm that works as follows. Every time an iceberg order is replenished—new visible order volume is added—the algorithm sets an indicator for hidden volume at that price. The algorithm resets the indicator for that price to zero only when an event occurs that could not have occurred had the iceberg order remained at that price. The indicator and the volume until the next replenishment is stored specific for the price level. The indicator remains unchanged unless a potential replenishment is omitted. For the iceberg state variable we apply the prevailing indicators at the best price levels. A detailed example how the algorithm works is given in the appendix 3.A.

The algorithm will make both type I and type II errors (assuming that the null is to predict no iceberg state), but given the properties of iceberg order documented above, see, table 3.2, it is likely to provide informative signals. Table 3.10 reports the average percentages of correct and in-correct detections. False detections are in the range between a half and two percents. True iceberg states that are missed by the algorithm occur between two and five percents. Around a third of all iceberg states

[4] De Winne and D'Hondt (2007) propose an algorithm for detecting iceberg orders that use the type of information and the methodology applied in Pardo and Pascual (2006) is also based on the same idea.

remain undiscovered, as the iceberg was not yet executed at least once. Comparing our results to De Winne and D'Hondt (2007) the algorithm displays similar size (false detection) for Euronext, but slightly less power with around 40 percent undiscovered states or somewhat more than ten percents in absolute terms.

We do not know, of course, to what extent the predictions from our algorithm closely approximates the predictions of the market participants. Conversations with market participants suggest that it is reasonable to assume that active participants are able to collect this type of information. Of course, it may well be the case that market participants apply algorithms that generate even more accurate predictions.

Robustness of Model Parameter Estimates

We re-estimate the model parameters using the same moment conditions as above, but by letting the indicators generated by the detection algorithm determine the values of the indicator variables. Table 3.6 provides a comparison of the δ estimates obtained in the baseline case—labeled 'True'—and in the case of the detection algorithm labeled 'Algorithm.' Overall, the parameter estimates are fairly close. The price impact parameters and the order flow parameters are also fairly similar suggestion that our main findings for the compensation for liquidity provision are robust to uncertainty about the iceberg orders.

3.6 Conclusions

We study the interaction between hidden and visible liquidity in a limit order market. We show that the hidden liquidity that is supplied by iceberg orders influences that strategies followed by traders supplying visible liquidity using limit orders. We report evidence that iceberg orders have an economically significant impact on the order book and short term price dynamics despite representing a relatively small fraction of all submitted orders. We show that while it may be hard to predict the amount of hidden liquidity, a trader can, using the history of visible order books, successfully predict whether the book contains hidden liquidity or not. Based on these findings we develop a state-dependent model of liquidity provision in which traders who submit limit orders to the order book follow different strategies depending on whether there is hidden liquidity or not.

We use our model to extract estimates of the net marginal compensation for liquidity provision. Overall, our results show that the marginal compensation varies with the state of the order book. The higher compensation for liquidity provision when an iceberg order is on the same side of the book is consistent with traders placing orders less aggressively when they anticipate their order competing with hidden liquidity.

Our approach focuses on the response of traders submitting market and limit orders to the possibility of hidden liquidity. But indirectly the results also offer some insights into the motives of the traders submitting iceberg orders. The fact, that iceberg orders are partly detectable and we observe more frequent and larger market orders suggest that iceberg orders are used successfully as a tool for coordinating trading.

3.A Appendix: The iceberg detection algorithm

The following example illustrates how the iceberg detection algorithm operates. The algorithm maintains and updates a detection vector with the following four elements for all price levels in the order book: price; detection flag; visible volume; volume until next replenishment. In the example, the best bid is initially 9.70 and there are no iceberg orders at this price level, i.e., the detection flag is zero.

1. A buy iceberg order is submitted at 9.77 with a total size of 9000, a peak size of 1000, implying a hidden volume of 8000 and a visible volume of 1,000 shares. The detection flag is currently zero. Detection vector: [9.77, 0, 1000, 0].

2. A sell market order for 1200 shares is submitted. The iceberg order's first peak size of 1000 shares is completely executed and another 200 share are automatically executed from the iceberg order's second peak size of 1000 shares. The remaining 800 shares of the iceberg order's second peak size are displayed in the book. The algorithm sets the detection flag to one and sets the volume until next replenishment to 800. Detection vector: [9.77, 1, 800, 800].

3. A buy limit order is submitted at 9.79 for 200 shares. The flag for 9.79 is zero. Detection vector: (9.79, 0, 200, 0). There is no change at 9.77

4. A sell market order for 500 shares is submitted. 200 shares are executed at 9.79, and 300 shares are executed against the second peak size of the iceberg order at 9.77. The detection flag at 9.77 remains at one, and the volume until next replenishment is revised to 500. Detection vector: [9.77, 1, 500,500].

5. A buy limit order is submitted at 9.77 for 5000 shares. The flag at 9.77 remains at one and volume until replenishment is unchanged because the new limit order is behind the visible 500 share of the iceberg order's second peak size. Detection vector: [9.77, 1, 5500, 500].

6. The iceberg order at 9.77 is cancelled. The detection flag remains at one.[5] Detec-

[5]In principle, the algorithm could make use of the fact that the remaining volume of 5,000 at 9.77 exactly matches the size of the previously submitted limit order to infer that it is very likely that the drop in the visible volume at 9.77 was caused by the iceberg being cancelled. However, if the limit order in question was for 500 shares instead of 5,000 shares, it would be a fifty-fifty chance that the cancellation was due to the iceberg order. Our algorithm has not been optimized with respect to these scenarios so it is possible that its performance could be enhanced.

tion vector: [9.77, 1, 5000, 500].

7. A sell market order is submitted for 600 shares. The flag is reset to zero as the volume until next replenishment is exceeded without the expected replenishment. Detection vector: [9.77, 0, 4500, 0].

Table 3.1: Descriptive Statistics: Sample Stocks

Trading Activity Group	Ticker Symbol	Market Cap. [€ billions]	Trading Volume [€ billions]	Transaction Price [€]	Trades Per Day [1000]	Trade Size [1000 shrs]	Bid-Ask Spread [b.p.]	Bid-Ask Spread [€ cents]
1	ALV	33.8	18.6	99.5	4.4	0.6	4.6	4.5
(High)	DBK	38.2	19.9	67.6	4.0	1.2	4.0	2.7
	DCX	30.3	12.1	36.3	3.3	1.6	5.2	1.9
	DTE	34.9	22.3	15.6	4.4	5.1	7.1	1.1
	MUV2	16.4	13.3	93.6	3.4	0.7	4.6	4.3
	SAP	27.4	11.8	131.0	2.8	0.5	4.6	6.0
	SIE	52.9	20.6	63.7	4.3	1.1	3.9	2.5
	Mean	33.4	16.9	72.5	3.8	1.5	4.8	3.3
2	BAS	25.4	7.9	43.1	2.6	1.1	4.7	2.0
	BAY	15.9	5.7	22.8	2.4	1.6	7.1	1.6
	BMW	12.2	5.6	34.7	2.1	1.2	5.6	1.9
	EOA	33.8	10.5	52.5	2.9	1.1	4.5	2.4
	HVM	6.6	6.3	18.5	1.9	2.8	9.2	1.7
	IFX	4.8	9.4	11.6	2.8	4.5	10.0	1.2
	RWE	12.7	6.2	33.9	2.3	1.2	5.9	2.0
	VOW	9.7	6.7	39.2	2.5	1.1	5.2	2.0
	Mean	15.1	7.3	32.0	2.4	1.8	6.5	1.9
3	ADS	4.1	2.1	92.5	1.0	0.4	6.7	6.2
	CBK	7.6	3.4	15.3	1.5	2.4	9.6	1.5
	DB1	4.8	2.3	47.0	1.0	0.8	7.0	3.3
	DPW	6.8	2.8	18.2	1.3	1.8	9.3	1.7
	LHA	4.5	2.8	14.1	1.4	2.3	10.6	1.5
	MEO	5.0	2.5	34.8	1.2	0.9	8.4	2.9
	SCH	7.1	3.3	41.0	1.5	0.8	6.7	2.7
	TKA	6.5	2.5	15.8	1.3	1.9	10.7	1.7
	Mean	5.8	2.7	34.8	1.3	1.4	8.6	2.7
4	ALT	3.3	1.9	48.8	1.1	0.6	7.4	3.6
(Low)	CONT	4.1	1.6	31.5	1.0	0.8	8.8	2.8
	FME	1.9	0.8	53.8	0.6	0.4	9.4	5.1
	HEN3	3.7	1.2	65.8	0.7	0.4	7.3	4.8
	LIN	3.4	1.4	43.6	0.9	0.6	7.5	3.3
	MAN	2.4	1.8	27.7	1.1	0.9	9.1	2.5
	TUI	2.0	1.7	18.7	1.1	1.3	11.5	2.2
	Mean	3.0	1.5	41.4	0.9	0.7	8.7	3.5
All	Mean	14.1	7.9	44.4	2.1	1.4	7.2	2.8
	s.e.	2.6	1.2	5.4	0.2	0.2	0.4	0.3

Table 3.1 reports the market capitalization, the trading volume, the average transaction price, the average number of trades per day, the average trade size, and the average relative and absolute bid-ask spreads for the sample stocks. The trading activity groups are formed by sorting the stocks according to the total trading volume; group 1 has the highest and group 4 the lowest trading volume. The company names and their associated ticker symbols are listed in Table 3.12 in the appendix. The market capitalization is calculated using a free-float methodology. It is measured in billions of euros as of December 31st, 2003. The average bid-ask spread is reported in basis points and in cents. The last row for each group reports the average of the mean values across the stocks. The last two rows report the cross-sectional average of the means and cross-sectional standard error of the means.

Table 3.2: Descriptive Statistics: Iceberg and Limit Orders

Trading Activity Group	Ticker Symbol	Iceberg Orders as % Shares Submitted	Iceberg Orders as % Shares Executed	Limit Size [1000 shrs]	Iceberg Order Size: Peak Size [1000 shrs]	Iceberg Order Size: Total Size/ Peak Size	Iceberg Order Size: Executed Shares/ Peak Size	Distance: Order Price To Best Quote [b.p.] Iceberg	Distance: Order Price To Best Quote [b.p.] Limit
1 (High)	ALV	5%	15%	0.5	1.5	8.5	5.5	2.8	3.7
	DBK	7%	17%	0.9	2.2	7.5	5.2	1.6	3.2
	DCX	8%	22%	1.4	2.7	6.8	5.2	2.7	5.1
	DTE	7%	12%	5.4	10.8	4.9	3.7	6.3	6.2
	MUV2	8%	17%	0.5	1.5	8.5	5.2	2.2	3.2
	SAP	6%	11%	0.4	1.3	8.2	4.2	3.9	3.1
	SIE	7%	17%	1.1	2.4	6.9	4.8	3.0	3.2
	Mean	7%	16%	1.4	3.2	7.3	4.8	3.2	4.0
2	BAS	7%	17%	0.9	2.0	8.0	5.1	2.5	4.4
	BAY	6%	14%	1.4	2.6	7.1	4.8	4.7	4.3
	BMW	9%	20%	0.8	2.1	7.3	4.8	2.8	3.0
	EOA	6%	16%	1.0	2.1	7.4	4.9	1.9	3.9
	HVM	18%	26%	1.4	4.2	6.9	4.8	5.6	5.3
	IFX	18%	22%	3.1	7.2	5.7	4.6	8.6	8.6
	RWE	9%	19%	0.9	2.2	7.3	4.9	3.0	5.3
	VOW	12%	24%	0.8	2.1	7.7	5.1	2.4	2.9
	Mean	10%	20%	1.3	3.1	7.2	4.9	3.9	4.7
3	ADS	7%	7%	0.2	1.3	8.3	3.0	2.1	0.0
	CBK	8%	20%	1.4	3.8	6.9	4.8	6.1	6.5
	DB1	16%	23%	0.4	1.7	7.7	4.4	2.1	2.2
	DPW	14%	24%	1.2	3.1	6.8	4.7	5.1	5.5
	LHA	13%	21%	1.3	3.4	7.1	5.2	6.7	6.9
	MEO	11%	15%	0.5	1.6	8.4	4.9	3.0	2.8
	SCH	9%	15%	0.5	1.6	8.3	4.9	2.5	2.6
	TKA	10%	13%	1.3	3.3	6.6	4.0	5.9	6.2
	Mean	11%	17%	0.9	2.5	7.5	4.5	4.2	4.1
4 (Low)	ALT	10%	11%	0.4	1.2	9.0	4.1	3.7	1.9
	CONT	7%	12%	0.5	1.9	8.1	4.3	3.3	3.1
	FME	7%	10%	0.3	1.2	9.3	4.1	2.0	1.8
	HEN3	4%	9%	0.2	1.2	9.1	4.8	0.0	1.6
	LIN	7%	10%	0.4	1.3	8.5	4.0	2.4	2.3
	MAN	12%	18%	0.5	1.9	7.6	4.8	3.5	3.6
	TUI	12%	16%	0.8	2.2	7.4	4.2	5.3	5.4
	Mean	8%	12%	0.4	1.5	8.4	4.3	2.9	2.8
All	Mean	9%	16%	1.0	2.6	7.6	4.6	3.6	3.9
	s.e.	1%	1%	0.2	0.4	0.2	0.1	0.3	0.3

Table 3.2 reports descriptive statistics for iceberg and limit orders. Columns three and four report the iceberg orders' share of the total number of shares submitted and executed, excluding any orders that are immediately executed. The next column reports the average size of limit orders, followed by the average of the iceberg orders' peak size, ratio of total size to peak size, and ratio of executed shares to peak size for iceberg orders whose first peak size was executed. The last two columns report the median distance between the same-side best quote and the order price of iceberg and limit orders. The last row for each group reports the average of the mean values across the stocks. The last two rows report the cross-sectional average of the means and cross-sectional standard error of the means.

Table 3.3: Order Execution and Duration

		A: Unconditional			B: Matched Group: Modal Iceberg Order					
					First Peak			Second Peak		
Trading Activity	Ticker	Execution		Duration (dur)	Execution		Time-to-Fill (ttf)	Execution		Time-to-Fill (ttf)
Group	Symbol	Limit	Iceberg	$\overline{dur_{iceberg}}/\overline{dur_{limit}}$	Limit	Iceberg	$\overline{ttf_{iceberg}}/\overline{ttf_{limit}}$	Limit	Iceberg	$\overline{ttf_{iceberg}}/\overline{ttf_{limit}}$
1	ALV	13%	49%	8.5	88%	90%	1.1	84%	81%	0.9
(High)	DBK	14%	54%	7.2	87%	92%	1.2	86%	89%	1.0
	DCX	15%	54%	6.8	82%	90%	1.0	85%	86%	1.0
	DTE	22%	42%	3.0	91%	95%	0.9	90%	81%	0.9
	MUV2	16%	51%	8.9	80%	85%	0.8	85%	83%	0.8
	SAP	13%	40%	8.4	89%	86%	0.8	87%	72%	0.7
	SIE	15%	50%	6.9	88%	95%	1.4	87%	85%	0.9
2	BAS	13%	47%	8.0	81%	93%	1.4	87%	81%	0.6
	BAY	14%	48%	5.9	90%	97%	1.0	87%	87%	1.0
	BMW	13%	50%	6.0	86%	92%	1.5	87%	83%	0.9
	EOA	13%	51%	8.0	84%	93%	1.3	87%	82%	1.0
	HVM	17%	51%	5.9	91%	94%	0.7	89%	80%	0.8
	IFX	21%	44%	2.6	91%	93%	0.7	88%	80%	1.0
	RWE	14%	45%	6.1	90%	97%	1.1	89%	83%	1.0
	VOW	16%	55%	6.1	86%	90%	0.7	84%	83%	0.6
3	ADS	12%	43%	7.6	77%	80%	1.6	80%	66%	0.4
	CBK	11%	54%	8.1	90%	85%	1.2	87%	81%	0.6
	DB1	13%	46%	7.6	85%	86%	0.8	86%	77%	0.4
	DPW	16%	51%	6.4	88%	96%	1.3	90%	87%	0.9
	LHA	15%	52%	7.3	87%	91%	1.3	84%	81%	0.7
	MEO	16%	44%	7.0	90%	91%	0.6	87%	79%	0.5
	SCH	15%	49%	6.3	89%	90%	1.0	86%	83%	0.6
	TKA	15%	48%	6.5	88%	89%	0.5	83%	67%	0.8
4	ALT	14%	40%	5.0	87%	85%	0.9	84%	75%	0.4
(Low)	CONT	12%	48%	9.4	79%	92%	1.4	81%	84%	0.3
	FME	10%	39%	8.2	77%	82%	1.4	79%	74%	0.5
	HEN3	9%	48%	11.3	80%	94%	1.1	81%	75%	0.3
	LIN	11%	37%	8.7	84%	85%	0.5	86%	67%	0.4
	MAN	13%	47%	8.6	82%	88%	1.0	83%	78%	0.6
	TUI	15%	46%	6.3	81%	88%	1.1	85%	68%	1.1
All	Mean	14%	47%	7.1	86%	90%	1.0	85%	79%	0.7
	s.e.	0.5%	1%	0.3	1%	1%	0.1	1%	1%	0.04

Panel A of Table 3.6 reports the percentage of limit and iceberg orders with at least one execution and the ratio of the median durations of iceberg and limit orders. Order duration is measured in from the submission until the time of the last order execution or cancellation. Group B is comprised of two sub-groups; First Peak and Second Peak. First Peak includes all iceberg orders with relative order prices—measured from the best same-side quote—and order sizes close to the median values. A matching limit order sample is constructed by matching on size and relative order price. Second Peak includes iceberg orders whose first peak size was executed. The matching limit order sample includes all limit order submissions that undercut the best same-side quote and that have a size that closely matches the modal iceberg peak size. The execution frequency reported for First Peak is the percent of executed first peaks. The time-to-fill ratio is the ratio of median time-to-fill for iceberg and limit orders. For the Second Peak, the median time-to-fill refers to the time it takes for the second peak to be executed.

Table 3.4: Limit Order Books and Iceberg Orders

	Trading Activity Group			
	1 (High)	2	3	4 (Low)
Spreads				
Bid-Ask Spread [basis points]				
(a) No Iceberg	4.9	6.7	8.7	8.7
(b) Iceberg Opposite/Same	4.2	5.9	7.7	7.3
Difference (b)-(a) (# p-values < 0.001)	-0.8 (7)	-0.7 (8)	-1.1 (8)	-1.4 (7)
2nd Best Quote - Best Quote [basis points]				
(a) No Iceberg	3.3	5.0	6.0	5.4
(b) Iceberg Opposite	3.0	4.7	5.7	4.9
(c) Iceberg Same	4.0	5.5	7.1	7.1
Difference (b)-(a) (# p-values < 0.001)	-0.3 (6)	-0.2 (8)	-0.3 (7)	-0.5 (6)
Difference (c)-(a) (# p-values < 0.001)	0.7 (6)	0.6 (8)	1.1 (8)	1.7 (7)
Depths				
Visible Depth at Best Quote [1,000 shares]				
(a) No Iceberg	6.0	4.4	2.9	1.2
(b) Iceberg Opposite	6.4	5.2	3.2	1.3
(c) Iceberg Same	7.3	5.4	3.5	1.8
Difference (b)-(a) (# p-values < 0.001)	0.4 (4)	0.8 (7)	0.3 (7)	0.1 (5)
Difference (c)-(a) (# p-values < 0.001)	1.3 (7)	1.0 (7)	0.6 (8)	0.6 (7)
Visible Depth at 2nd Best Quote [1,000 shares]				
(a) No Iceberg	9.7	6.0	3.5	1.3
(b) Iceberg Opposite	9.6	6.8	3.8	1.3
(c) Iceberg Same	9.8	6.0	3.4	1.1
Difference (b)-(a) (# p-values < 0.001)	-0.1 (1)	0.8 (7)	0.3 (4)	0.1 (2)
Difference (c)-(a) (# p-values < 0.001)	0.1 (1)	-0.0 (6)	-0.1 (5)	-0.1 (6)

Table 3.4 reports average spreads and depths in the order books observed before transactions stratified by the presence of one or more iceberg orders at the best bid or ask quotes. The No Iceberg strata includes all order books with no iceberg orders at either best quotes. The Iceberg Opposite Side strata includes the bid side of all order books with an iceberg order at the best ask side, and vice versa. The Iceberg Same Side strata includes the bid side of all order books with an iceberg order at the best bid side, and vice versa. All averages are first computed by stock and then averaged across stocks within each trading activity group. Next to the mean differences, in parenthesis, is the number of stocks within each group that have a mean difference that has the same sign as the overall mean difference and a p-value of 0.001 or less for a test of the null that the difference is zero.

Table 3.5: Price Impact and Iceberg Orders

Panel A: Price Impact Regression:

$$\Delta mq_{t+\tau,t} = c + a_0 d_t + b_0 \tilde{m}_t + (a_1 d_t + b_1 \tilde{m}_t) I_t^{own} + (a_2 d_t + b_2 \tilde{m}_t) I_t^{opp} + \epsilon_t$$

	Trading Activity Groups							
	1 (High)		2		3		4 (Low)	
	Time Horizon: $\tau = 10$ Minutes							
intercept	-0.07	(0.04)	-0.51	(0.07)	-0.06	(0.09)	0.29	(0.11)
d_t	1.66	(0.06)	2.57	(0.09)	3.03	(0.12)	3.38	(0.15)
$\tilde{m}_t$	0.43	(0.03)	0.55	(0.05)	0.83	(0.07)	1.12	(0.09)
$d_t I_t^{own}$	1.62	(0.22)	2.54	(0.28)	3.51	(0.42)	4.26	(0.66)
$\tilde{m}_t I_t^{own}$	-0.03	(0.10)	-0.12	(0.13)	-0.26	(0.20)	-0.16	(0.30)
$d_t I_t^{opp}$	-1.61	(0.17)	-2.42	(0.23)	-2.81	(0.34)	-4.73	(0.50)
$\tilde{m}_t I_t^{opp}$	-0.13	(0.07)	-0.12	(0.10)	-0.28	(0.15)	-0.61	(0.20)
	Time Horizon: $\tau = 30$ Trades							
intercept	-0.16	(0.02)	-0.25	(0.04)	-0.26	(0.08)	-0.11	(0.11)
d_t	1.83	(0.03)	2.57	(0.05)	3.04	(0.10)	3.41	(0.14)
$\tilde{m}_t$	0.45	(0.01)	0.62	(0.03)	0.89	(0.06)	1.10	(0.08)
$d_t I_t^{own}$	1.38	(0.11)	2.13	(0.17)	3.64	(0.37)	5.05	(0.66)
$\tilde{m}_t I_t^{own}$	-0.12	(0.05)	-0.13	(0.08)	-0.28	(0.18)	-0.36	(0.30)
$d_t I_t^{opp}$	-1.74	(0.08)	-2.20	(0.14)	-3.27	(0.31)	-5.13	(0.50)
$\tilde{m}_t I_t^{opp}$	-0.19	(0.03)	-0.23	(0.06)	-0.29	(0.14)	-0.50	(0.20)
	Time Horizon: $\tau =$ Next Trade							
intercept	-0.00	(0.00)	-0.02	(0.00)	0.01	(0.01)	0.01	(0.01)
d_t	0.98	(0.00)	1.55	(0.01)	1.97	(0.02)	2.08	(0.02)
$\tilde{m}_t$	0.29	(0.00)	0.42	(0.00)	0.59	(0.01)	0.58	(0.01)
$d_t I_t^{own}$	-0.07	(0.02)	-0.04	(0.03)	-0.11	(0.06)	-0.22	(0.11)
$\tilde{m}_t I_t^{own}$	-0.02	(0.00)	-0.05	(0.01)	-0.06	(0.03)	-0.07	(0.05)
$d_t I_t^{opp}$	-1.00	(0.01)	-1.56	(0.02)	-1.98	(0.05)	-2.08	(0.08)
$\tilde{m}_t I_t^{opp}$	-0.18	(0.00)	-0.22	(0.01)	-0.33	(0.02)	-0.38	(0.03)

Panel B: Asymmetric Price Impact Regression:

$$\Delta mq_{t+\tau,t} = [\text{baseline model as above}] \ldots$$
$$+ \left(\hat{a}_1 I_t^{own} + \hat{a}_2 I_t^{opp} + \hat{b}_0 \tilde{m}_t + \hat{b}_1 \tilde{m}_t I_t^{own} + \hat{b}_2 \tilde{m}_t I_t^{opp} \right) 1_{d_t=-1} + e_t$$

Time Horizon	# Rejections: F-test of $H_0 : \hat{a}_1 = \hat{a}_2 = 0$			
τ=10 minutes	6	3	3	3
τ=30 trades	2	1	1	3
τ= next trade	3	2	2	0
	# Rejections: F-test of $H_0 : \hat{b}_0 = \hat{b}_1 = \hat{b}_2 = 0$			
τ=10 minutes	4	2	2	1
τ=30 trades	1	1	0	0
τ= next trade	4	6	7	4

Panel A of Table 3.5 reports results for price impact regressions for the following three different time horizons: 10 minutes; 30 trades; trade to trade. The mid-quote change over the horizon is regressed on a constant (c), and on the trade direction indicator (market buy $d_t = 1$, market sell $d_t = -1$) and the signed normalized market order size ($\tilde{m}_t = m_t \times d_t$) by themselves as well as interacted with indicators for iceberg orders on the same side as the market order (I^{own}) and opposite side (I^{opp}), for example, $I^{own} = 1$ for a market buy order if there is an iceberg at the best bid quote. The mid-quote changes are measured in basis points and the market order sizes are measured in units of the average market order size. Panel B report the number of rejections for a regression that allows for asymmetric price impact for market buy versus sell orders.

Table 3.6: Market Order Flow and Iceberg Orders

	Panel A: Market Order Size Regression: $m_t = c + b_1 I_t^{own} + b_2 I_t^{opp} + \epsilon_t$							
	Trading Activity Groups							
	1 (High)		2		3		4 (Low)	
Intercept	0.95	(0.00)	0.92	(0.00)	0.93	(0.00)	0.94	(0.00)
I_t^{own}	0.13	(0.01)	0.17	(0.01)	0.17	(0.02)	0.27	(0.03)
I_t^{opp}	0.40	(0.01)	0.37	(0.01)	0.43	(0.01)	0.59	(0.02)
F-test $I_t^{own} = I_t^{opp} = 0$	803.3	(7)	633.3	(8)	386.9	(8)	387.1	(7)
T-test $c \times$Sell = 0	44.6	(6)	21.2	(5)	5.2	(2)	8.5	(3)
F-test $I_t^{own} \times Sell = I_t^{opp} \times Sell = 0$	6.9	(5)	9.3	(4)	4.0	(2)	8.6	(3)

Marginal effects	Panel B: Logit for the Probability of a Market Buy Order: $Prob(\text{market buy}\mid I_t^{opp}, I_t^{own}) = \frac{e^{c+b_1 I_t^{own}+b_2 I_t^{opp}}}{1+e^{c+b_1 I_t^{own}+b_2 I_t^{opp}}}$							
I_t^{own}	−0.15	(0.00)	−0.14	(0.00)	−0.13	(0.01)	−0.15	(0.01)
I_t^{opp}	0.16	(0.00)	0.14	(0.00)	0.12	(0.01)	0.15	(0.01)

Table 3.6 reports results for a regression of the market order size on a constant and indicators for iceberg orders on the opposite and same side as the incoming market order. The regressions are estimated for each stock and the average parameters estimates by trading activity group are reported. Average standard errors are given in parenthesis. The market order size is normalized by the average market order size. The indicators for iceberg orders I^{opp} and I^{own} are defined so that for a market buy order $I^{opp} = 1$, if there is one or more iceberg orders at the best ask quote, and $I^{own} = 1$, if there is one or more iceberg orders at the best bid quote. The fourth row report the average F-statistics and, in parentheses, the number of rejections at the 1% level for a test of the null of constant market order size. The next two rows report average T-statistics and F-statistics for tests that allow for asymmetric effects between buy and sell market orders. Panel B reports the mean average marginal effects from a logit model of the probability of a market buy order as a function of the indicators for iceberg orders on the opposite and same side as the incoming market order. The mean average marginal effects are computed for a value of one for the indicator versus a value of zero, and averaged across stocks within each group. The standard errors reported in parenthesis are computed for the average marginal effects for each stock, and are then averaged across stocks within each group.

Table 3.7: The Order Book Model

ORDERBOOK CONDITIONS - ASK SIDE

$$E\left[\left[\begin{array}{l} p_t^{ask,1} - x_{t+\tau}^U(q_t^{ask,1};\theta) - \bar{\delta}_1 h_t^{ask} \\ p_t^{ask,2} - x_{t+\tau}^U(Q_t^{ask,2};\theta) - \bar{\delta}_2 h_t^{ask} \end{array}\right] \otimes h_t^{ask}\right] = 0$$

ORDERBOOK CONDITIONS - BID SIDE

$$E\left[\left[\begin{array}{l} x_{t+\tau}^L(q_t^{bid,1};\theta) - p_t^{bid,1} - \bar{\delta}_1 h_t^{bid} \\ x_{t+\tau}^L(Q_t^{bid,2};\theta) - p_t^{bid,2} - \bar{\delta}_2 h_t^{bid} \end{array}\right] \otimes h_t^{bid}\right] = 0$$

UPPER/LOWER-TAIL EXPECTATIONS - 1ST QUOTE

$$\begin{aligned} x_{t+\tau,0}^U(q_t^{ask,1};\theta) &= X_t + \mu + \left[\bar{\alpha}_0 h_t^{ask} + \left(\bar{\alpha}_1 h_t^{ask}\right)\left(q_t^{ask,1} + \bar{\lambda} h_t^{ask}\right)\right] \\ x_{t+\tau,0}^L(q_t^{bid,1};\theta) &= X_t + \mu - \left[\bar{\alpha}_0 h_t^{bid} + \left(\bar{\alpha}_1 h_t^{bid}\right)\left(q_t^{bid,1} + \bar{\lambda} h_t^{bid}\right)\right] \end{aligned}$$

UPPER/LOWER-TAIL EXPECTATIONS - 2ND QUOTE

$$\begin{aligned} x_{t+\tau}^U(Q_t^{ask,2};\theta) &= X_t + \mu + \left[\bar{\alpha}_0 h_t^{ask} + \left(\bar{\alpha}_1 h_t^{ask}\right)\left(Q_t^{ask,2} + \bar{\lambda} h_t^{ask} + I_t^{ask}\eta\right)\right] \\ x_{t+\tau}^L(Q_t^{bid,2};\theta) &= X_t + \mu - \left[\bar{\alpha}_0 h_t^{bid} + \left(\bar{\alpha}_1 h_t^{bid}\right)\left(Q_t^{bid,2} + \bar{\lambda} h_t^{bid} + I_t^{bid}\eta\right)\right] \end{aligned}$$

PRICE IMPACT AND ORDER FLOW CONDITIONS

$$E\left[\left[\begin{array}{l} m_t - \bar{\lambda} h_t^{mkt} \\ X_{t+\tau} - X_t - \mu - \left(\bar{\alpha}_0 h_t^{mkt}\right) d_t - \left(\bar{\alpha}_1 h_t^{mkt}\right) \tilde{m}_t \\ \left(X_{t+\tau} - X_t - \mu - \left(\bar{\alpha}_0 h_t^{mkt}\right) d_t - \left(\bar{\alpha}_1 h_t^{mkt}\right) \tilde{m}_t\right) d_t \\ \left(X_{t+\tau} - X_t - \mu - \left(\bar{\alpha}_0 h_t^{mkt}\right) d_t - \left(\bar{\alpha}_1 h_t^{mkt}\right) \tilde{m}_t\right) \tilde{m}_t \end{array}\right] \otimes h_t^{mkt}\right] = 0$$

HIDDEN ORDER SIZE CONDITIONS

$$E\left[\begin{array}{l} \left(\hat{q}_t^{ask} - \eta\right) 1(\hat{q}_t^{ask} > 0) \\ \left(\hat{q}_t^{bid} - \eta\right) 1(\hat{q}_t^{bid} > 0) \end{array}\right] = 0$$

Continued on next page

The Order Book Model - continued

MODEL VARIABLES

$p^{\cdot,1}$ $\quad p^{\cdot,2}$	price 1st and 2nd quotes
$q^{\cdot,1}$ $\quad Q^{\cdot,2}$	volume 1st and cumulated volume 2nd quotes
$\hat{q}^{ask}$ $\quad \hat{q}^{ask}$	hidden volume at 1st quotes
X_t $\quad X_{t+\tau}$	share price - current and at $t+\tau$
d_t	sign of market orders (buy = 1, sell = -1)
m_t $\quad \tilde{m}_t = d_t m_t$	volume, signed volume of market orders
$h_t^{bid} = [1\ I_t^{bid}\ I_t^{ask}]$	iceberg states - bid side
$h_t^{ask} = [1\ I_t^{ask}\ I_t^{bid}]$	iceberg states - ask side
$h_t^{mkt} = \begin{cases} [1 I_t^{bid} I_t^{ask}] & \text{if } d_t = 1 \\ [1 I_t^{ask} I_t^{bid}] & \text{if } d_t = -1 \end{cases}$	iceberg states - market orders

MODEL PARAMETERS

$\bar{\alpha}_0 = [\alpha_0\ \alpha_0^{own}\ \alpha_0^{opp}]$	fixed component of the price impact
$\bar{\alpha}_1 = [\alpha_1\ \alpha_1^{own}\ \alpha_1^{opp}]$	variable component of the price impact
μ	drift of the share price
$\bar{\delta}_1 = [\delta_1\ \delta_1^{own}\ \delta_1^{opp}]$	payoff for liquidity provision at the 1st quotes
$\bar{\delta}_2 = [\delta_2\ \delta_2^{own}\ \delta_2^{opp}]$	payoff for liquidity provision at the 2nd quotes
$\bar{\lambda} = [\lambda\ \lambda^{own}\ \lambda^{opp}]$	expected market order volume
η	expected hidden order volume at the 1st quotes

Table 3.8: Parameter Estimates for Order Book Model

	Trading Activity Group							
	1 (High)		2		3		4 (Low)	
A: Price Impact Function								
α_0	1.92	(0.08)	2.60	(0.24)	3.30	(0.33)	3.64	(0.36)
α_0^{own}	1.40	(0.35)	2.04	(0.39)	2.47	(1.12)	4.77	(1.09)
α_0^{opp}	-1.97	(0.30)	-1.95	(0.35)	-3.04	(0.68)	-4.65	(1.09)
α_1	0.41	(0.03)	0.54	(0.06)	0.82	(0.14)	0.86	(0.13)
α_1^{own}	0.01	(0.08)	-0.05	(0.08)	-0.01	(0.25)	-0.10	(0.29)
α_1^{opp}	-0.18	(0.04)	-0.29	(0.06)	-0.41	(0.14)	-0.54	(0.14)
μ	-0.21	(0.10)	-0.24	(0.20)	0.02	(0.37)	0.09	(0.44)
B: Compensation for Liquidity Provision								
δ_1	-0.79	(0.09)	-0.93	(0.21)	-1.25	(0.46)	-1.51	(0.36)
δ_1^{own}	1.75	(0.23)	2.30	(0.34)	3.34	(0.77)	4.85	(1.11)
δ_1^{opp}	-2.22	(0.42)	-2.37	(0.38)	-3.23	(1.00)	-5.53	(1.18)
δ_2	1.21	(0.17)	2.58	(0.32)	2.91	(0.72)	2.41	(0.49)
δ_2^{own}	0.19	(0.70)	1.52	(0.60)	1.70	(1.50)	3.30	(1.79)
δ_2^{opp}	-3.29	(0.78)	-2.57	(0.56)	-3.85	(1.31)	-5.88	(1.64)
C: Market Order Flow and Iceberg Order Volume								
λ	0.95	(0.00)	0.92	(0.01)	0.93	(0.01)	0.94	(0.01)
λ^{own}	0.11	(0.02)	0.17	(0.02)	0.16	(0.07)	0.37	(0.08)
λ^{opp}	0.40	(0.02)	0.37	(0.02)	0.43	(0.04)	0.59	(0.05)
η	11.05	(0.13)	9.20	(0.12)	10.70	(0.21)	14.48	(0.37)

Table 3.8 report the average parameter estimates for the order book model developed in section 3.4. The model parameters are estimated using GMM. The second stage estimates are computed using a Newey-West 10-lag weighting matrix. Table 3.7 lists the moment conditions. The model is estimated for each stock and the average parameter estimates are reported with the average standard errors in parenthesis. Panel A provides the parameters for the price impact function, the αs, and the drift (μ). Panel B provides the parameters for the marginal premium or discount for liquidity provision. Panel C provides the parameters for the market order flow, the λs, and for the mean iceberg order volume, (η).

Table 3.9: Net Compensation for Liquidity Provision

		Trading Activity Group											
		1			2			3			4		
Hypothesis = 0	Method	Mean	+	-	Mean	+	-	Mean	+	-	Mean	+	-
δ_1	true	-0.8	0	7	-0.9	0	6	-1.2	0	5	-1.5	0	6
	algo.	-0.6	0	7	-0.7	0	6	-0.7	1	4	-1.3	0	4
$\delta_1 + \delta_1^{own}$	true	1.0	6	0	1.4	8	0	2.1	5	0	3.3	5	0
	algo.	0.6	5	0	1.0	4	0	1.5	4	0	2.4	3	0
$\delta_1 + \delta_1^{opp}$	true	-3.0	0	7	-3.3	0	8	-4.5	0	7	-7.0	0	7
	algo.	-2.7	0	7	-3.5	0	8	-4.9	0	7	-7.7	0	7
$\delta_1 + \delta_1^{own} + \delta_1^{opp}$	true	-1.3	0	4	-1.0	0	6	-1.1	0	4	-2.2	0	2
	algo.	-1.5	0	6	-1.8	0	8	-2.7	0	6	-4.0	0	4
δ_2	true	1.2	7	0	2.6	8	0	2.9	6	0	2.4	7	0
	algo.	1.3	7	0	2.9	8	0	3.6	7	0	2.8	7	0
$\delta_2 + \delta_2^{own}$	true	1.4	6	0	4.1	7	0	4.6	6	0	5.7	6	0
	algo.	2.1	6	0	3.8	7	0	4.7	6	0	4.6	5	0
$\delta_2 + \delta_2^{opp}$	true	-2.1	0	3	0.0	2	2	-0.9	1	2	-3.5	0	3
	algo.	-1.2	0	4	-0.4	2	4	-1.4	0	3	-4.7	0	4
$\delta_2 + \delta_2^{own} + \delta_2^{opp}$	true	-1.9	2	1	1.5	2	0	0.8	1	0	-0.2	0	0
	algo.	-0.4	1	0	0.5	1	1	-0.3	0	0	-2.9	0	1

Table 3.9 reports by trading activity group the mean values of the net δs for different states of the order book; no iceberg orders, iceberg order on the same side as the marginal limit (own), iceberg order on the opposite side, and iceberg orders on both sides (own+opp). The mean value of the estimated parameter values or sums of parameter values are reported and the number of mean values that are significantly different from zero at the 1% level are reported separately for positive and negative values in the columns labeled '+' and '-.'

Table 3.10: Iceberg Detection Algorithm

	Trading Activity Group											
	1 (High)			2			3			4 (Low)		
	Algorithm			Algorithm			Algorithm			Algorithm		
True State	No Ice	Ice	Sum	No Ice	Ice	Sum	No Ice	Ice	Sum	No Ice	Ice	Sum
No Iceberg	89.8	1.0	90.8	85.5	1.3	86.9	87.5	1.1	88.5	92.7	0.5	93.2
Iceberg	3.0	6.2	9.2	4.3	8.8	13.1	3.9	7.5	11.5	2.2	4.6	6.8
Sum	92.8	7.2	100.0	89.9	10.1	100.0	91.4	8.6	100.0	94.9	5.1	100.0

Table 3.10 reports the average distribution of the prediction generated by the iceberg detection algorithm and the true iceberg state. The results are averaged for the bid and ask side for each stock and averaged across stock within each trading activity group. The row entries in each 2×2 table refers to the true state of the book, and the columns refer to the predictions made by the algorithm.

Table 3.11: Robustness of Liquidity Compensation Parameter Estimates

δ_1	True	-0.79	(0.09)	-0.93	(0.21)	-1.25	(0.46)	-1.51	(0.36)
	Algorithm	-0.65	(0.08)	-0.70	(0.20)	-0.68	(0.38)	-1.27	(0.35)
δ_1^{own}	True	1.75	(0.23)	2.30	(0.34)	3.34	(0.77)	4.85	(1.11)
	Algorithm	1.25	(0.22)	1.67	(0.33)	2.17	(0.71)	3.68	(1.07)
δ_1^{opp}	True	-2.22	(0.42)	-2.37	(0.38)	-3.23	(1.00)	-5.53	(1.18)
	Algorithm	-2.07	(0.26)	-2.77	(0.38)	-4.20	(0.90)	-6.41	(1.16)
δ_2	True	1.21	(0.17)	2.58	(0.32)	2.91	(0.72)	2.41	(0.49)
	Algorithm	1.31	(0.17)	2.88	(0.31)	3.63	(0.61)	2.85	(0.47)
δ_2^{own}	True	0.19	(0.70)	1.52	(0.60)	1.70	(1.50)	3.30	(1.79)
	Algorithm	0.81	(0.59)	0.91	(0.56)	1.08	(1.38)	1.78	(1.65)
δ_2^{opp}	True	-3.29	(0.78)	-2.57	(0.56)	-3.85	(1.31)	-5.88	(1.64)
	Algorithm	-2.49	(0.55)	-3.30	(0.54)	-5.01	(1.31)	-7.50	(1.57)

Table 3.6 reports the parameter estimates for the δ parameters of the iceberg order book model for the case in which iceberg states are observed without error—rows labeled True—and the case in which iceberg states are detected using the algorithm—rows labeled Algorithm. The average parameters estimates for each trading activity group is report with average standard errors in parenthesis.

Table 3.12: Sample Firms

Company Name	Ticker Symbol	Trading Activity Group
Adidas-Salomon	ADS	3
Allianz	ALV	1
Altana	ALT	4
BASF	BAS	2
Bayer	BAY	2
Bayerische Motoren Werke	BMW	2
Commerzbank	CBK	3
Continental	CONT	4
DaimlerChrysler	DCX	1
Deutsche Bank	DBK	1
Deutsche Börse	DB1	3
Deutsche Post	DPW	3
Deutsche Telekom	DTE	1
E.ON	EOA	2
Fresenius Medical Care	FME	4
Henkel	HEN3	4
Hypo-Vereinsbank	HVM	2
Infineon Technologies	IFX	2
Linde	LIN	4
Lufthansa	LHA	3
Man	MAN	4
Metro	MEO	3
Münchener Rück	MUV2	1
RWE	RWE	2
SAP	SAP	1
Schering	SCH	3
Siemens	SIE	1
ThyssenKrupp	TKA	3
Touristik Union International	TUI	4
Volkswagen	VOW	2

Table 3.13: Stock-by-Stock Model Parameter Estimate

Group/Ticker	α_0		α_0^{own}		α_0^{opp}		α_1		α_1^{own}		α_1^{opp}		μ	
1/ALV	1.88	(0.05)	1.24	(0.19)	-1.82	(0.17)	0.45	(0.02)	-0.07	(0.06)	-0.28	(0.03)	-0.32	(0.08)
1/DBK	1.79	(0.07)	1.00	(0.16)	-1.53	(0.15)	0.37	(0.02)	-0.08	(0.05)	-0.21	(0.03)	-0.04	(0.08)
1/DCX	1.76	(0.16)	1.82	(0.23)	-1.11	(0.31)	0.57	(0.06)	-0.10	(0.07)	-0.32	(0.06)	-0.38	(0.14)
1/DTE	2.06	(0.09)	0.83	(1.16)	-3.62	(0.81)	0.36	(0.03)	0.49	(0.17)	0.14	(0.10)	0.00	(0.13)
1/MUV2	2.11	(0.07)	1.60	(0.20)	-2.06	(0.23)	0.40	(0.03)	-0.28	(0.04)	-0.30	(0.03)	-0.35	(0.10)
1/SAP	2.17	(0.07)	2.19	(0.34)	-2.55	(0.29)	0.38	(0.03)	0.20	(0.14)	-0.13	(0.04)	-0.20	(0.11)
1/SIE	1.66	(0.05)	1.12	(0.15)	-1.12	(0.15)	0.36	(0.02)	-0.07	(0.05)	-0.19	(0.03)	-0.15	(0.07)
Panel	1.92	(0.08)	1.40	(0.35)	-1.97	(0.30)	0.41	(0.03)	0.01	(0.08)	-0.18	(0.04)	-0.21	(0.10)
2/BAS	2.53	(0.14)	1.38	(0.27)	-1.77	(0.24)	0.31	(0.05)	-0.05	(0.09)	-0.15	(0.06)	-0.06	(0.13)
2/BAY	3.17	(0.24)	2.16	(0.38)	-2.05	(0.39)	0.46	(0.08)	-0.01	(0.11)	-0.20	(0.09)	-0.20	(0.21)
2/BMW	2.03	(0.21)	2.23	(0.32)	-1.29	(0.31)	0.79	(0.08)	-0.21	(0.09)	-0.51	(0.08)	-0.47	(0.18)
2/EOA	2.07	(0.10)	1.34	(0.23)	-1.53	(0.21)	0.36	(0.05)	0.02	(0.08)	-0.17	(0.06)	0.12	(0.12)
2/HVM	3.32	(0.63)	2.91	(0.66)	-2.78	(0.63)	0.46	(0.06)	-0.06	(0.06)	-0.24	(0.05)	-1.20	(0.43)
2/IFX	2.86	(0.27)	2.63	(0.33)	-1.61	(0.45)	0.62	(0.05)	-0.16	(0.05)	-0.32	(0.04)	-0.22	(0.23)
2/RWE	2.54	(0.16)	1.42	(0.70)	-2.32	(0.29)	0.71	(0.06)	0.14	(0.13)	-0.34	(0.07)	0.51	(0.17)
2/VOW	2.29	(0.14)	2.26	(0.27)	-2.23	(0.27)	0.59	(0.04)	-0.11	(0.07)	-0.34	(0.04)	-0.38	(0.15)
Panel	2.60	(0.24)	2.04	(0.39)	-1.95	(0.35)	0.54	(0.06)	-0.05	(0.08)	-0.29	(0.06)	-0.24	(0.20)
3/ADS	2.62	(0.19)	4.33	(0.98)	-4.04	(0.78)	0.83	(0.09)	-0.64	(0.32)	-0.73	(0.10)	-0.01	(0.28)
3/CBK	3.93	(0.50)	-2.50	(1.72)	-2.63	(0.63)	0.65	(0.07)	0.25	(0.15)	-0.22	(0.09)	0.01	(0.47)
3/DB1	2.86	(0.32)	3.38	(0.65)	-2.47	(0.71)	0.60	(0.11)	0.14	(0.20)	-0.29	(0.11)	-0.07	(0.34)
3/DPW	3.48	(0.41)	2.26	(0.52)	-2.74	(0.84)	0.84	(0.13)	-0.18	(0.12)	-0.50	(0.11)	-0.14	(0.34)
3/LHA	3.44	(0.21)	1.32	(2.67)	-2.99	(0.67)	1.69	(0.49)	-0.10	(0.60)	-0.89	(0.46)	0.52	(0.53)
3/MEO	3.71	(0.29)	3.37	(0.73)	-2.85	(0.65)	0.98	(0.09)	0.09	(0.26)	-0.44	(0.12)	-0.16	(0.35)
3/SCH	2.60	(0.23)	3.35	(0.92)	-3.10	(0.43)	0.81	(0.08)	0.01	(0.21)	-0.46	(0.10)	-0.17	(0.23)
3/TKA	3.74	(0.52)	4.28	(0.74)	-3.49	(0.73)	0.11	(0.04)	0.39	(0.11)	0.22	(0.06)	0.22	(0.38)
Panel	3.30	(0.33)	2.47	(1.12)	-3.04	(0.68)	0.82	(0.14)	-0.01	(0.25)	-0.41	(0.14)	0.02	(0.37)
4/ALT	2.57	(0.25)	3.71	(1.34)	-2.91	(0.74)	0.96	(0.09)	-0.27	(0.20)	-0.67	(0.09)	-0.42	(0.30)
4/CONT	3.90	(0.28)	2.41	(0.87)	-1.53	(0.86)	0.67	(0.15)	0.10	(0.27)	-0.32	(0.17)	0.33	(0.44)
4/FME	2.93	(0.43)	8.88	(1.58)	-8.24	(1.36)	1.26	(0.16)	-0.49	(0.54)	-1.06	(0.18)	-0.54	(0.54)
4/HEN3	3.11	(0.29)	5.34	(1.25)	-6.40	(0.97)	0.42	(0.08)	-0.31	(0.30)	-0.32	(0.10)	0.90	(0.39)
4/LIN	3.77	(0.33)	3.95	(0.89)	-4.11	(0.84)	0.66	(0.12)	-0.05	(0.21)	-0.39	(0.12)	-0.74	(0.39)
4/MAN	3.79	(0.42)	4.13	(0.74)	-5.00	(0.75)	1.07	(0.11)	0.32	(0.24)	-0.56	(0.12)	0.04	(0.44)
4/TUI	5.38	(0.56)	4.97	(0.97)	-4.34	(2.13)	0.96	(0.16)	-0.02	(0.27)	-0.45	(0.18)	1.04	(0.55)
Panel	3.64	(0.36)	4.77	(1.09)	-4.65	(1.09)	0.86	(0.13)	-0.10	(0.29)	-0.54	(0.14)	0.09	(0.44)
Sample	2.87	(0.26)	2.64	(0.74)	-2.87	(0.60)	0.66	(0.09)	-0.04	(0.18)	-0.35	(0.10)	-0.08	(0.28)

Table 3.13: Stock-by-Stock Model Parameter Estimates (continued)

Group/Ticker	δ_1		δ_1^{own}		δ_1^{opp}		δ_2		δ_2^{own}		δ_2^{opp}	
1/ALV	-0.70	(0.06)	1.75	(0.17)	-1.74	(0.21)	1.02	(0.08)	1.15	(0.33)	-2.19	(0.31)
1/DBK	-0.74	(0.07)	1.71	(0.16)	-1.08	(0.17)	1.11	(0.10)	1.31	(0.28)	-1.17	(0.26)
1/DCX	-0.76	(0.15)	1.55	(0.30)	-2.01	(0.24)	1.49	(0.28)	0.86	(0.48)	-2.00	(0.43)
1/DTE	-1.10	(0.15)	1.74	(0.34)	-4.81	(1.53)	1.48	(0.43)	-6.07	(2.60)	-10.90	(3.33)
1/MUV2	-0.76	(0.07)	2.27	(0.23)	-1.36	(0.21)	1.20	(0.11)	2.62	(0.30)	-1.24	(0.26)
1/SAP	-0.76	(0.07)	2.07	(0.29)	-3.28	(0.40)	1.20	(0.11)	0.59	(0.65)	-4.09	(0.63)
1/SIE	-0.71	(0.05)	1.16	(0.14)	-1.27	(0.16)	0.98	(0.09)	0.91	(0.25)	-1.44	(0.25)
Panel	-0.79	(0.09)	1.75	(0.23)	-2.22	(0.42)	1.21	(0.17)	0.19	(0.70)	-3.29	(0.78)
2/BAS	-1.01	(0.10)	1.80	(0.23)	-1.61	(0.28)	1.68	(0.19)	1.48	(0.49)	-1.90	(0.47)
2/BAY	-1.06	(0.14)	2.22	(0.34)	-2.66	(0.39)	2.92	(0.33)	1.35	(0.72)	-2.94	(0.69)
2/BMW	-1.35	(0.14)	2.13	(0.28)	-2.17	(0.30)	0.99	(0.27)	1.66	(0.49)	-2.03	(0.47)
2/EOA	-0.82	(0.15)	1.56	(0.24)	-1.86	(0.28)	1.41	(0.26)	0.66	(0.49)	-2.28	(0.47)
2/HVM	0.13	(0.68)	2.82	(0.65)	-3.32	(0.66)	5.84	(0.75)	2.05	(0.75)	-3.67	(0.74)
2/IFX	-0.51	(0.22)	2.72	(0.40)	-2.41	(0.35)	5.19	(0.36)	2.55	(0.57)	-1.85	(0.60)
2/RWE	-1.55	(0.12)	2.54	(0.30)	-2.53	(0.48)	1.21	(0.23)	0.24	(0.83)	-3.54	(0.59)
2/VOW	-1.25	(0.13)	2.64	(0.26)	-2.38	(0.28)	1.37	(0.18)	2.17	(0.44)	-2.38	(0.41)
Panel	-0.93	(0.21)	2.30	(0.34)	-2.37	(0.38)	2.58	(0.32)	1.52	(0.60)	-2.57	(0.56)
3/ADS	-1.41	(0.21)	4.90	(0.82)	-3.54	(0.98)	0.70	(0.32)	5.46	(1.61)	-2.92	(1.33)
3/CBK	-1.14	(0.52)	2.70	(0.68)	1.04	(1.45)	4.52	(0.62)	-0.54	(1.34)	-0.60	(1.33)
3/DB1	-0.80	(0.26)	2.39	(0.66)	-4.13	(0.68)	2.53	(0.36)	1.11	(1.04)	-4.75	(1.01)
3/DPW	-1.27	(0.23)	3.46	(0.75)	-2.41	(0.51)	3.16	(0.47)	2.83	(0.93)	-2.02	(0.84)
3/LHA	-3.31	(1.42)	4.97	(1.42)	-1.79	(1.96)	0.43	(2.79)	2.67	(3.56)	-1.99	(2.65)
3/MEO	-1.91	(0.31)	2.86	(0.68)	-4.49	(0.85)	1.91	(0.41)	1.04	(1.35)	-5.54	(1.32)
3/SCH	-1.40	(0.17)	3.36	(0.43)	-4.20	(0.80)	1.30	(0.26)	1.58	(1.11)	-4.92	(1.05)
3/TKA	1.27	(0.53)	2.08	(0.73)	-6.35	(0.78)	8.73	(0.56)	-0.50	(1.03)	-8.10	(0.98)
Panel	-1.25	(0.46)	3.34	(0.77)	-3.23	(1.00)	2.91	(0.72)	1.70	(1.50)	-3.85	(1.31)
4/ALT	-1.23	(0.21)	3.76	(0.73)	-3.75	(1.55)	1.60	(0.29)	2.78	(1.12)	-3.73	(1.84)
4/CONT	-1.38	(0.49)	1.18	(0.97)	-3.65	(0.99)	3.12	(0.75)	-0.86	(1.65)	-4.02	(1.44)
4/FME	-1.50	(0.43)	8.88	(1.45)	-9.48	(1.68)	1.79	(0.59)	7.43	(3.17)	-8.59	(2.62)
4/HEN3	-0.57	(0.29)	6.41	(0.96)	-5.19	(1.32)	2.54	(0.36)	6.57	(1.69)	-5.10	(1.66)
4/LIN	-1.73	(0.34)	4.13	(0.84)	-4.71	(0.90)	1.91	(0.46)	3.00	(1.18)	-5.50	(1.12)
4/MAN	-1.99	(0.40)	5.16	(0.72)	-6.06	(0.85)	2.16	(0.50)	1.33	(1.30)	-7.34	(1.32)
4/TUI	-2.19	(0.36)	4.44	(2.07)	-5.88	(1.00)	3.73	(0.51)	2.83	(2.43)	-6.89	(1.47)
Panel	-1.51	(0.36)	4.85	(1.11)	-5.53	(1.18)	2.41	(0.49)	3.30	(1.79)	-5.88	(1.64)
Sample	-1.12	(0.28)	3.05	(0.61)	-3.30	(0.74)	2.31	(0.43)	1.67	(1.14)	-3.85	(1.06)

Table 3.13: Stock-by-Stock Model Parameter Estimates (continued)

Group/Ticker	λ		λ^{own}		λ^{opp}		η		# Obs	J-Stat	P-value
1/ALV	0.95	(0.00)	0.14	(0.02)	0.46	(0.01)	12.69	(0.13)	284	135.3	0.00
1/DBK	0.96	(0.00)	0.06	(0.02)	0.35	(0.01)	9.46	(0.11)	248	30.3	0.00
1/DCX	0.94	(0.00)	0.07	(0.01)	0.28	(0.01)	8.42	(0.11)	207	125.5	0.00
1/DTE	0.96	(0.01)	0.00	(0.06)	0.30	(0.03)	10.82	(0.18)	279	167.3	0.00
1/MUV2	0.93	(0.00)	0.22	(0.02)	0.51	(0.02)	12.84	(0.15)	214	120.4	0.00
1/SAP	0.96	(0.00)	0.21	(0.03)	0.55	(0.02)	13.81	(0.17)	174	66.0	0.00
1/SIE	0.95	(0.00)	0.08	(0.01)	0.34	(0.01)	9.33	(0.09)	277	36.8	0.00
Panel	0.95	(0.00)	0.11	(0.02)	0.40	(0.02)	11.05	(0.13)	240	97.4	(7/7)
2/BAS	0.95	(0.00)	0.12	(0.02)	0.33	(0.01)	9.75	(0.13)	160	105.5	0.00
2/BAY	0.95	(0.00)	0.17	(0.02)	0.30	(0.01)	8.01	(0.12)	148	60.6	0.00
2/BMW	0.92	(0.01)	0.17	(0.02)	0.40	(0.01)	9.77	(0.12)	130	18.1	0.01
2/EOA	0.95	(0.00)	0.10	(0.02)	0.36	(0.01)	10.25	(0.15)	179	48.8	0.00
2/HVM	0.87	(0.01)	0.26	(0.02)	0.46	(0.02)	7.37	(0.12)	119	183.0	0.00
2/IFX	0.89	(0.01)	0.21	(0.02)	0.30	(0.01)	8.21	(0.08)	174	86.1	0.00
2/RWE	0.93	(0.00)	0.13	(0.02)	0.38	(0.02)	9.47	(0.14)	143	22.0	0.00
2/VOW	0.91	(0.01)	0.17	(0.02)	0.42	(0.02)	10.75	(0.11)	158	24.1	0.00
Panel	0.92	(0.01)	0.17	(0.02)	0.37	(0.02)	9.20	(0.12)	151	68.5	(7/8)
3/ADS	0.97	(0.01)	0.16	(0.05)	0.58	(0.05)	21.16	(0.55)	58	8.8	0.27
3/CBK	0.92	(0.01)	0.19	(0.02)	0.35	(0.02)	8.93	(0.15)	88	18.1	0.01
3/DB1	0.88	(0.01)	0.24	(0.03)	0.67	(0.03)	11.95	(0.21)	58	46.7	0.00
3/DPW	0.90	(0.01)	0.13	(0.02)	0.39	(0.02)	8.41	(0.12)	79	34.4	0.00
3/LHA	0.92	(0.02)	0.08	(0.32)	0.31	(0.10)	7.21	(0.11)	81	25.1	0.00
3/MEO	0.94	(0.01)	0.17	(0.03)	0.46	(0.02)	9.55	(0.17)	74	13.8	0.05
3/SCH	0.94	(0.01)	0.15	(0.02)	0.41	(0.02)	10.96	(0.21)	92	21.7	0.00
3/TKA	0.95	(0.01)	0.17	(0.02)	0.27	(0.02)	7.42	(0.15)	76	22.7	0.00
Panel	0.93	(0.01)	0.16	(0.07)	0.43	(0.04)	10.70	(0.21)	76	23.9	(5/8)
4/ALT	0.94	(0.01)	0.41	(0.05)	0.51	(0.03)	12.75	(0.22)	65	59.6	0.00
4/CONT	0.94	(0.01)	0.24	(0.04)	0.50	(0.04)	12.17	(0.27)	59	46.7	0.00
4/FME	0.92	(0.01)	0.92	(0.29)	0.81	(0.09)	23.23	(0.93)	35	41.0	0.00
4/HEN3	0.95	(0.01)	0.27	(0.06)	0.94	(0.07)	18.83	(0.48)	40	44.4	0.00
4/LIN	0.96	(0.01)	0.31	(0.05)	0.47	(0.04)	13.87	(0.31)	53	43.5	0.00
4/MAN	0.91	(0.01)	0.21	(0.03)	0.56	(0.03)	11.84	(0.22)	63	40.5	0.00
4/TUI	0.94	(0.01)	0.19	(0.03)	0.34	(0.03)	8.70	(0.19)	63	63.9	0.00
Panel	0.94	(0.01)	0.37	(0.08)	0.59	(0.05)	14.48	(0.37)	54	48.5	(7/7)
Sample	0.93	(0.01)	0.20	(0.05)	0.44	(0.03)	11.26	(0.21)	129	58.7	(26/30)

Number of observations (# Obs) is measured in 1,000. For the last row for each group the p-value column reports the number of stocks for which the test of over-identifying restrictions rejects at the 0.001 level over the total number of stocks.

Four

Auction design in order book markets

The chapter analyzes the impact of opening auction design on the quality of the auction price. The exchange provides throughout the auction a market clearing price (called indicative price) together with the executable volume which would result from an immediate close of the call phase. By getting rid of microstructure noise, we find that the indicative price becomes informative about the true value at the very beginning of the call phase, and that traders estimate the extent of microstructure noise by means of the executable volume. Moreover, the (undisclosed) order book provides information about future prices that is not incorporated in the auction price. Finally, the auction ends at a random time between 9 and 9:00:30 a.m. The random phase duration bears no impact on traded volume in the morning.

4.1 Introduction

Platform breakdowns at the Tokyo Stock Exchange, internalization of orders by banks, collusion between market makers at the NASDAQ, or the role of the specialist in New York: market design is now more than ever a prime concern for many of the actors of exchanges around the world. This led to numerous new regulations, either pushed by the exchange itself or imposed by regulators (for instance the European Investment Services Directive).

Although most of the attention has focused on the continuous phases of the trading process, auctions are of great importance for market participants. Indeed, most exchanges start the day with an opening auction that provides the market with its first price, and, thus, its first important piece of information of the day. A well-designed morning auction ensures that this first price is informative, therefore limiting volatility and improving liquidity during at least the morning trading phase, if not the entire day. Similarly, many investors, e.g. index funds, rely on a closing auction to ensure that their assets are traded at the closing price. Auctions may represent a small portion of daily trading (around 10% in our sample), but this is a portion crucial for the market.

In this chapter we question the design of auctions in stock exchanges, particularly in limit order book markets. Most auctions in an order book platform follow the same pricing rule: the auction price results from the maximization of the executable volume, and, when several prices lead to the same executable volume, of the minimization of the surplus between the buy and sell sides. Differences between exchanges in the auction design generally stem either from the information disclosed to the market during the auction, or from the setting that ends the call phase. Tables 4.1 and 4.2 report similarities and discrepancies in auction design at the London Stock Exchange (LSE), Euronext, and the German Stock Exchange (Deustche Börse).

Most exchanges provide throughout the auction a theoretical, market clearing price (called indicative price) together with the executable volume which would result from an immediate close of the call phase. What is exchange specific is the release of information about the order book and order flow. Euronext and the LSE, two of the leading European exchanges, disclose the order book during the call phase. Traders can see the entire supply and demand schedule. On the other hand, their main competitor, the Deutsche Börse, closes the book during the call phase. The rationale for closing the book is not clear, since a lack of transparency may harm the dissemination of information. The main argument is that by hiding the book, the exchange protects large orders (therefore fostering traded volume) while limiting price manipulation by traders during the call phase.

Another remarkable difference between Euronext/LSE and the Deutsche Börse is the setting of the close of the call phase. Euronext and the LSE end the auction at a fixed, pre-determined time; market participants know exactly when their orders are matched. Therefore, traders don't bear any execution risk during an auction. At the Deutsche Börse, the call phase is followed by a random phase. During that random phase, traders can submit and cancel orders as during the call, but matching might take place at any time. The maximum time span for the random phase is 30 seconds. Again, random phases were set to avoid price manipulation. Without the random phase, traders can hide their position either by submitting orders that are cancelled just before the matching or by submitting their orders just before the matching takes place.

We focus on opening auctions. For exchanges, the opening auction aims at processing overnight information in a short amount of time and without price disruption. The auction has been successful if the opening price incorporates all overnight information. To address these questions, we analyze first the (informational) quality of the auction price. Second, we investigate whether additional information, namely the executable volume (disclosed) and the order book (undisclosed) speeds price discovery. Finally, we study the impact of the random phase on traded volumes in the early morning.

There are two dimensions in the quality of opening prices. The auction (or indicative) price is said to be informative if it is the expectation at time i of the price at 9:10 a.m. (taken as the proxy for the true value once the overnight information has been processed). Additionally, traders may improve the precision of their estimate of the value (learning hypothesis). When both dimensions hold, prices are said to be efficient. To test for price efficiency, we estimate a model similar to Biais, Hillion, and Spatt (1999) on the Xetra platform, a pure order book system operating at the Deutsche Börse. Our analysis focuses on the 30 stocks that compose the German leading index called DAX30. However, we depart from Biais, Hillion, and Spatt (1999) and propose a new estimation method, the consistent adjusted least squares (CALS). This new methodology is not a mere econometric game: the estimation of the structural model in Biais, Hillion, and Spatt (1999) is subject to the well-known problem of measurement error due to the presence of microstructure noise. CALS enables to get estimates of the structural parameters that are free of microstructure noise.

Our results read as follows. First, we reproduce the methodology in Biais, Hillion, and Spatt (1999) and find that the indicative prices become informative only close to the end of the auction. Nevertheless, once we get rid of the microstructure noise by means of CALS, the estimates indicate that the indicative price is indeed informative quickly after the start of the opening auction. Traders take advantage of disclosed information, namely the executable volume, to assess the amount of microstructure noise in the indicative price. Second, we show that the order book contains information relevant for future prices that is not incorporated in the indicative price. Finally, the random phase bears no impact on the volumes traded in the morning.

The remainder of the chapter is organized as follows. The next section introduces market design and the data. Price efficiency and the CALS estimator are the focus of section 4.3. Section 4.4 sheds light on the informational content of the executable volume and of the order book. Section 4.5 analyzes the impact of the random phase. The last section concludes.

4.2 Market design

The Xetra platform

The Xetra platform carries out orders and trades at the Deutsche Börse. Xetra is a pure order book[1] that follows the usual rules applying to order books around the world. Trading is enforced based on the price/time priority - the order with the best price is executed first, and at this price execution is granted to the order submitted first. Traders have a large range of orders at their disposal. The most common ones are the well-known limit and market orders. Additionally, a portion of the order's volume can be hidden from other traders. These orders are referred to as hidden orders or iceberg orders. The visible portion of the hidden order, called the peak, benefits from full price and time priorities. The hidden portion, however, gets only the price priority. When all orders ahead in the queue are consumed by an incoming marker order, a new portion of the hidden order, equal to the peak size, is revealed to the market. Disclosed shares are then granted time priority over subsequent order submissions.

Trading on the Xetra platform starts around 9:00 a.m. and ends around 5:30 p.m.[2] Trading is not allowed outside these hours. Three scheduled auctions take place at the beginning, in the middle and at the end of the trading day. Additionally, unscheduled auctions are launched when trading prices leave a defined price range; these are the so-called "volatility interruptions" or "trading halts". All auctions follow a similar setting: the auction starts with a call of fixed duration, followed by a call with a random end ("random phase"). At the end of the second phase, the auction price is set and executable orders are matched. Figure 4.1 shows the time frame of the different trading phases for the DAX30 stocks.

[1] The Xetra platform allows the presence of a committed liquidity provider for less liquid stocks. This is not the case for the stocks studied in this work.

[2] All times refer to the Continental European Time, which is GMT time plus one hour.

Supply and demand determine the Xetra auction price.[3] Two quantities are of special importance for the understanding of the auction price. The executable volume at time i is the volume that can be traded according to the current state of the book. The surplus is the exceeding volume that cannot be executed due to an overhang of volume on one side. Both the executable volume and the surplus are computed based on all standing orders (limit, market, iceberg orders). When the auction ends, the auction price is set such that it maximizes the executable volume and minimizes the surplus.

The order book remains closed during the auction. Nevertheless, traders are not left without information. At each moment in time during the call phase, the Xetra computer applies the price setting rules but no trading takes place. If there is a positive executable volume (crossed order book), a hypothetical auction price, the executable volume, and the size and side of the surplus are displayed on traders' screens. Since none of the standing orders are executed, the hypothetical auction price is referred to as the indicative price. If the buy and sell sides of the book do not cross, the indicative price cannot be computed and traders are provided with the best bid and ask price limits together with their depths.

The data

The German Stock Exchange granted access to a database containing all order events for each of the 30 DAX stocks, over a 3-month period ranging from January 2 to March 31, 2004. Additionally, the exchange provided us with the state of the book as of January 1, 2004. The database enables to reconstruct the entire life of an order, from its submission to its death by (partial) execution and/or cancellation. Moreover, all the trading phases (scheduled auctions, volatility interruptions, continuous trading) are perfectly identified in the database.

Based on this information and on the initial book, we reconstructed the entire order book as available to market participants at every point in time during the three-month period. This is done by implementing the rules of the Xetra trading protocol.[4] Similarly, we are able to reconstruct the series of indicative prices, volumes and surpluses as displayed to traders during auctions.

[3] It is worth noting that the hidden portion of the iceberg orders take part in the price setting

[4] We run an exhaustive battery of consistency checks. We checked for example that the spread is always positive and that market orders are executed against volumes available in the book. No errors occur during the reconstruction process.

Table 4.6 displays the main characteristics of the 30 stocks in our sample. Those stocks are quite actively traded (on average 2,000 trades per day) and liquid (the average percentage spread is 0.09%).

4.3 Auctions and Price efficiency

In this section we analyze whether the indicative prices displayed during the auction are efficient. Price efficiency refers to two dimensions. First, an efficient price has to be informative, which means it must reflect the underlying value of the stock. Second, the price should exhibit a low variance. Traders learn about the true value of the stock during the auction if the variance of the indicative price return decreases throughout the auction (learning hypothesis). The first dimension will be referred to as price informativeness. If both price informativeness and learning hold during an auction, then the indicative prices are said to be efficient.

Price informativeness: the CALS estimator

We propose a model similar to Biais, Hillion, and Spatt (1999). We assume that the overnight change in the value ρ_t^N is on average incorporated fully into the stock price after 10 minutes of continuous trading:

$$r_t^O = \rho_t^N + \omega_t^O, \tag{4.1}$$

where r_t^O is the return of the stock price (at 9:10) on the previous day closing price, and ω_t^O is an error term uncorrelated with the overnight return ρ_t^N. The index t distinguishes observations for different days or stocks.

The opening auction allows market participants to discover the overnight change in the fundamental value of the traded asset without the risk of immediate execution. The indicative price of the auction at each time i can partially or fully reflect that overnight change:

$$r_t^i = \frac{1}{\beta}\rho_t^N + \omega_t^i. \tag{4.2}$$

The parameter $\frac{1}{\beta}$ measures the amount of price discovery. The parameter ω_t^i captures microstructure effects and is uncorrelated to the pricing error after 10 minutes of trading. Prices are fully informative if β is equal to one. Partial informativeness would result in a parameter β greater than one, and the parameter β is zero if prices are not at all informative.[5]

The variance of the indicative price at time i is then

$$s_i^2 = \frac{1}{\beta^2}\sigma_N^2 + \sigma_i^2, \tag{4.3}$$

where σ_N^2 and σ_i^2 are the variances of the overnight return ρ_t^N and the microstructure noise w_t^i respectively. The indicative price variance can therefore be decomposed into the part stemming from changes in the value and the part originating in microstructure effects. The latter might stem e.g. from short-term mispricing, price manipulation, or from submitted orders that aim at scanning the book and are cancelled before the auction is called. Both components of the indicative price variance are unobserved.

From equation (4.2), the estimated equation reads:

$$r_t^O = br_t^i + \epsilon_t. \tag{4.4}$$

Our main deviation from the model in Biais, Hillion, and Spatt (1999) is that we consider a different proxy for the value of the asset. Information spills over the market during the trading day, due to, e.g., informed trades or the opening of foreign markets (mostly the American ones). Therefore, the close-to-close return, while a good proxy for the change in the value during that day, may depart dramatically from the actual change in the value overnight. Moreover, the opening price might not incorporate all the overnight information if the auction fails at its goal. This consideration leads us to pick the price at 9:10 a.m. instead of the closing price. In the following, the overnight return refers to the return of the price at 9:10 a.m. on the previous day closing price.

Biais, Hillion, and Spatt (1999) estimate equation (4.4) by means of standard OLS. However, as pointed out in Barclay and Hendershott (2003), the observed returns r_t^O and r_t^i are equal to the true change in the value plus some noise. Measurement errors in the variables imply that the OLS estimate $\hat{b}_{OLS}$ is a biased estimate of the parameter β.

There exists a variety of approaches to diminish or circumvent the effects of measurement errors. A list, by no means exhaustive, starts with weighted regression, instrumental variables, and LIML, and goes on but does not end with GMM and factor models. Books by Fuller (1987) and more recently Wansbeek and Meijer (2000)

[5] See Biais, Hillion, and Spatt (1999) for a discussion of the possible values for β.

provide a comprehensive review of the methods at hand. We use an adjustment of the OLS regression that can be seen as a special case of the consistent adjusted least squares (CALS) estimator. The main advantage of CALS lies in its ability to estimate the amount of measurement error σ_i^2. We simply need to provide reasonable identification assumptions. The CALS estimator is described in the appendix.

In our framework, we need to identify value from microstructure noise. In our model, the covariance $s_{EOC,C}$ of the return of the close-to-9 a.m. indicative price return, and the close-to-close return is equal to the variance of the overnight return adjusted by the parameter of the amount of price discovery:

$$s_{EOC,C} = \frac{1}{\beta}\sigma_N^2. \tag{4.5}$$

All remaining components of the two prices vanish as they are assumed to be mutually uncorrelated to the others. If we additionally assume that the magnitude of price discovery is stable throughout the auction, then we are able to identify the two components in the indicative price variance s_i^2. This turns out to be sufficient to obtain unbiased, microstructure-free estimates of the β parameter.

The CALS estimator then reads (see the appendix for the exact derivation):

$$\hat{b}_{CALS} = \frac{s_{i,O}}{s_{EOC,C}}, \tag{4.6}$$

$$\hat{\sigma}^2_{\epsilon,CALS} = s_O^2 - \frac{s_{i,O}^2}{s_{EOC,C}}. \tag{4.7}$$

where $s_{i,O}$ is the covariance of the return of the indicative price and the return until 9:10 a.m.

By comparison, the OLS estimator for β reads:

$$\hat{b}_{OLS} = \frac{s_{i,O}}{s_i^2}. \tag{4.8}$$

Both estimators for β share the same numerator, but the denominator is different. Instead of the variance of the regressor s_i^2, which includes microstructure noise, the CALS estimator denominator is the covariance of the closing return and the auction return at the end of the opening, which measures the amount of price discovery. The difference between the two estimators corresponds to the amount of measurement error (microstructure noise) in the observed returns. The asymptotic variance/covariance-matrix for the estimated parameter vector $(\hat{b}_{CALS}, \hat{\sigma}^2_{\epsilon,CALS})$ is derived in the appendix.

Results

Equation (4.4) is estimated across the 65 days for each stock and each time i. We estimate the model every 5 minutes between 7:30 and 8:50 a.m., every 15 seconds between 8:50 and 8:59 a.m., and finally every second until 9:00:30 a.m.[6] For reporting purposes, and to improve the quality of the estimation, we adopt a classification of the thirty DAX stocks in four groups based on trading activity.[7] The first group is then composed of the seven most frequently traded stocks whereas the eight least frequently traded stocks are in group four. The second and third group contain respectively eight and seven stocks each (see Table 4.6). Results based on a single stock or on the pooled regression (one regression for each quartile) lead to similar results. In the following we present results based on the pooled regressions.

Figure 4.6 reports the estimates for the b coefficient with both OLS and CALS estimation methods. The OLS estimate starts from zero at the beginning of the pre-trading phase (7.30 a.m.), and linearly increases to reach the value of one during the random phase. The CALS estimate displays a markedly different pattern: according to CALS, prices are already fully informative ($\beta = 1$) around 8:50, which is the launch of the opening call phase. Interestingly, although the estimated value reaches the value of one early, the variance of the estimator decreases sharply as we get closer to the opening. Our interpretation is that the indicative price incorporates most of the overnight information quite early, but that the quality of the indicative price improves as we head toward the end of the call phase. This is confirmed by the size of the microstructure noise (measurement error) reported in Figure 4.6. Microstructure noise is very large at the beginning of the call phase, and decreases towards zero during the auction. Since the OLS estimate does incorporate microstructure effects, it reaches the value of one only at the very end of the auction (around 9:00). Therefore, contrary to Biais, Hillion, and Spatt (1999), we conclude that the indicative price during the auction becomes informative very early, but is then very noisy due to the importance of microstructure effects.

[6] During the random phase, all stocks will progressively open to trading. For those stocks, the indicative price is replaced by the auction price until 9:00:30.

[7] An analysis conducted on a classification based on market capitalization leads to similar conclusions.

Do traders learn during the auction?

We showed that the indicative price becomes informative quite early during the morning auction. But how much do traders learn from the auction? Figure 4.6 reports the variance of the indicative price return r_t^i: the indicative price becomes more precise when the auction gets closer to the call. The figure also reports $var(r_t^{9:10} - r_t^i)$ and $var(r_t^C - r_t^i)$, which are the variances of the difference between the indicative price return and two returns taken after full price discovery. Again, both variances decrease as the opening gets closer. This suggest that traders do learn during the auction.

The regressions in the previous subsection help interpreting this finding. We showed that the indicative price becomes informative quite early during the morning auction. Even when we incorporate microstructure effects (OLS estimates), prices reflect most of the overnight change in the value at the opening. The comparison between the OLS and CALS estimates suggest that traders learn mainly about the extent of microstructure noise during the auction phase.

4.4 Speed of price discovery

Opening auctions are designed primarily for the processing of overnight information about the underlying value of the stock. If the auction fails to provide efficient prices, it failed to its main goal. But the exchange has other purposes, e.g. to provide a starting book for the rest of the trading day. In the following section, we analyze what pieces of information, whether they are disclosed or not by the exchange, speed price discovery in the market. We are not interested in estimating a structural model as in the previous section. Observed prices reflect a microstructure factor throughout the day, and traders are also interested in forecasting the observed price, not only the value component of the price as in the previous section. Therefore we focus our attention on the ability of different pieces of information to forecast the price at 9:10. Since we do not desire to disentangle value from microstructure effects, the OLS estimates are perfectly suited. Our starting point will be the OLS estimates displayed in the previous section, equation (4.4). We showed that the indicative price helps forecasting the value of the asset ($\hat{b}_{OLS}$ is positive and significant). Hereinafter, we analyze whether additional pieces of information improve the forecasting of the price at 9:10.

Surplus and executable volume

At every moment during the call phase, the exchange provides traders with three pieces of information: the indicative price, the executable volume and the surplus. We have already discussed the interest of the indicative price. The surplus provides information on a very small portion of the book, since it reveals the signed volume a trader has to submit to move the indicative price. Although that might be useful for traders scanning the book, or manipulating the indicative price, this does not provide much information about the true value. Indeed, surplus tend to be rather constant during the auction phase, as shown in Figure 4.6.

Similarly, traders cannot infer the underlying value of the asset by observing the executable volume at times i. However, it indicates the amount of agreement on the indicative price among traders. Consequently, traders might use the executable volume as a proxy for the amount of microstructure noise in the indicative price. If this is indeed the case, the inclusion of volumes in our regression should increase our estimate for $\hat{b}_{OLS}$. Figure 4.6 reports the evolution of the executable volume throughout the auction. After 8:50, volume levels off linearly, while microstructure noise, as estimated by CALS and displayed in Figure 4.6, goes to zero. We test more formally our hypothesis by estimating the following equation:

$$r_t^O = a^v + b^v r_t^i + c^v (r_t^i * v^i) + \nu_t, \tag{4.9}$$

where a^v is a constant, v^i stands for executable volume at time i, normalized by the average opening volume for each stock, and ν_t is an error term. If the parameter c^v is positive and significant at time i, and if the estimated $(b^v + c^v)$ is larger than $\hat{b}_{OLS}$ in equation 4.4, the executable volume helps traders to disentangle value from noise in the indicative price at time i.

Figure 4.6 reports the estimates for b^v and c^v. The parameter c^v is clearly positive and significant. Moreover, the sum of the two parameters $(b^v + c^v)$ is larger than the b estimates in equation 4.4. This confirms our intuition that executable volumes are a good indicator of the quality of the indicative price for market participants.

The order book

There is an important piece of information that is not disclosed to the market: the order book. In recent years, many papers have shown that the order book contains information not incorporated in the price, see for instance Cao, Hansch, and Wang (2004b), Irvine, Benston, and Kandel (2000), Kalay, Sade, and Wohl (2004), or Beltran-Lopez, Giot, and Grammig (2006). Although those papers focus on the continuous phase of the trading process, the book during the auction might be informative about future prices.

The state of the book during continuous trading is typically described by the slopes of the buy and sell schedules. When the buy and sell schedules cross, trading occurs. During an auction, the buy and sell schedules cross at the indicative price. We consider a measure of the slopes of the book similar to the elasticities defined in Kalay, Sade, and Wohl (2004). The ask elasticity x quotes away from the indicative price at time i for a stock is:

$$ask^i_{t,x} = \frac{\Delta v^i_{t,x} / v^i_t}{\Delta p^i_{t,x} / p^i_t}, \tag{4.10}$$

where v^i_t is the executable volume and p^i_t the indicative price. $\Delta v^i_{t,x}$ ($\Delta p^i_{t,x}$) is the change in the available ask volume (price) when we move x ask quotes away from the indicative price. The same formula applies for the bid elasticity. We have at our disposal 4 elasticities: the ask elasticity below the indicative price, the bid elasticity above the indicative price, the ask elasticity above the indicative price, and the bid elasticity below the indicative price. The first two elasticities measure the slopes of the executable portion of the book. The last two measure the slopes of the non-executable portion of the book. Some of the information contained in the executable portions of the book is already captured by the executable volume. Therefore, we focus hereinafter on the non-executable portion of the book. Figure 4.6 displays the average of the 3-quote ask and bid elasticities during the opening auction phase. Liquidity improves dramatically during the first 5 minutes of the opening auction, and remain quite stable afterwards. Nevertheless, elasticities rise slightly close to 9:00. This might be due to order cancellations since both value and the extent of microstructure noise have been estimated by traders by then.

We estimate the following regression:

$$r^O_t = a^{OB} + b^{OB} r^i_t + c^{OB}_1 ask^i_{t,3} + c^{OB}_2 bid^i_{t,-3} + \mu_t. \tag{4.11}$$

A large elasticity on the ask side implies that the order book is deep on the sell side. Thus, an increase in the ask elasticity should drive prices down: the parameter c_1^{OB} is expected to be negative. Similarly, a deeper book on the buy side will be reflected by a higher bid elasticity and a positive value for c_2^{OB}. We estimated the above equation with two elasticities, one taken at one quote from the indicative price, one taken at 3 quotes. Both led to similar results, although coefficients tend to be slightly larger and more significant in the latter case. We present here the estimation based on the 3-quote elasticity.

Results are reported in Figure 4.6. The parameters exhibit the expected signs. Both elasticity parameters become significant around 8:59 a.m. This confirms that the order book helps forecasting the price at 9:10 a.m. Nevertheless, the book loses its relevance as the opening gets closer: the absolute value of both coefficients decreases as we head toward 9:00:30 a.m. while still being significant. This is consistent with the evidence presented above that the indicative price becomes more efficient as we get closer to the random phase. When fully efficient, there is consequently little additional information contained in the order book.

4.5 The random phase

As in London, the German Stock Exchange uses a 30-second random phase to end auctions. This means that traders do not know when the auction is called. The impact of the random duration on traded volume is unclear. On one hand, traders have additional time to submit orders at what is likely to be the opening price: if so, we may observe a higher volume executed at the end of the auction when durations are longer. On the other hand, the random phase delays the start of continuous trading. It might simply lead to a transfer of traded volume from the continuous trading phase to the opening call, but it could also dampen traded volume in the first minutes of the day by preventing trades. We estimate by OLS the following regression:

$$log(v_t) = a + bd_t + \epsilon_t, \tag{4.12}$$

where d_t is the random duration observed at day t. v_t stands for volume. We additionally include appropriate dummies to control for stock effects.

We first consider as dependant variable the volume executed at the opening. The estimated b is not significant at 1%, which means that the random phase does not foster additional trading during the auction. When we turn to the volume executed before 9:00:40 (opening auction volume included), b shows up significant (at 1%) and negative (-0.0046). This means that not only the random phase does not foster trading at the opening: it additionally decrease trading in the very first seconds of the continuous phase. For the exchange, this means that shortening the random phase by one second increases trading volume in the first 40 seconds by 0.46%. This might seem small, but for the longest random phase (30 seconds), it means a drop of execution by nearly 14%. However, this portion of the daily trading is not lost for the exchange. Indeed, the regression above with traded volume before 9:10 (opening auction volume included) indicates that there is reversal: b is again not significant. We conclude that the random phase has no impact on traded volume in the first 10 minutes of the day.

4.6 Conclusion

In this chapter we analyze how auction design impacts the quality of the auction price. We show that the indicative price is indeed informative about the true value of the stock right after the beginning of the morning auction. Nevertheless, traders still learn about the amount of microstructure noise in the indicative price by means of the executable volume. Additionally, the order book contains information about future prices that is not incorporated in the indicative price. Finally, the random phase bears no impact on traded volume.

4.A Appendix: Derivation of the CALS estimator

The CALS estimator was originally developed by Kapteyn and Wansbeek (1984). We follow here the presentation in Wansbeek and Meijer (2000) and Meijer and Wansbeek (2000). In a linear model with measurement error, there are three groups of parameters to be estimated: the regression parameters $\hat{b}$, the variance of the residuals of the true model $\hat{\sigma_\epsilon^2}$ and the measurement error variance matrix $\hat{\Omega}$. The CALS estimator allows a consistent estimation of all parameters and their variance/covariance matrix.

A model for CALS requires all the assumptions of the standard OLS linear model, and additionally, arbitrary restrictions on the three parameter groups. The number of restrictions has to be equal to the number of independent elements of the measurement error variance matrix (just-identification). Otherwise the system is not properly identified.

With the restrictions, the CALS estimator is defined by the three following equations

$$(I - S_X^{-1}\hat{\Omega})\hat{b}_{CALS} - \hat{b}_{OLS} = 0 \quad (4.13)$$

$$\hat{\sigma}^2_{\epsilon,CALS} + \hat{b}'_{CALS}\hat{\Omega}\hat{b}_{OLS} - \hat{\sigma}^2_{\epsilon,OLS} = 0 \quad (4.14)$$

$$r(\hat{b}_{CALS}, \hat{\sigma}^2_{\epsilon,CALS}, \hat{\Omega}) - r_0 = 0, \quad (4.15)$$

where $\hat{b}_{OLS}$ and $\hat{\sigma}^2_{\epsilon,OLS}$ are the estimated parameters and the variance of the residuals of the according OLS regression. S_X is the covariance matrix of the exogenous variable, including measurement errors.

The first equation provides a bias-corrected estimator for b, the second a bias-corrected estimator for the variance of the residuals, and the third provides the necessary restrictions to identify the measurement error: r(.) is a differentiable function that equals to r_0. In the general case this system of equations has to be solved simultaneously.

Our identifying assumption for microstructure effects enables us to derive an estimator in the case of a univariate and known measurement error. The estimator for the case of a known measurement error is:

$$\hat{b}_{CALS} = (S_X - \hat{\Omega})^{-1}\hat{b}_{OLS},$$
$$\hat{\sigma}^2_{\epsilon,CALS} = \hat{\sigma}^2_{\epsilon,OLS} - \hat{b}'_{CALS}\hat{\Omega}\,\hat{b}_{OLS}, \quad (4.16)$$

with asymptotic distribution

$$\sqrt{N}\begin{bmatrix}\hat{b}_{CALS}-\beta\\ \hat{\sigma}^2_{\epsilon,CALS}-\sigma^2_\epsilon\end{bmatrix}\to^d \mathcal{N}\left(0,\begin{bmatrix}\gamma\, S_K^{-1}S_XS_K^{-1}+\omega\omega' & -2\gamma\omega\\ -2\gamma\omega' & 2\gamma^2\end{bmatrix}\right),\tag{4.17}$$

where $S_K = S_X - \Omega$, $\gamma = \sigma^2_\epsilon + \beta'\Omega\beta$ and $\omega = S_K^{-1}\Omega\beta$, c.f. Kapteyn and Wansbeek (1984). In our univariate model, S_X equals s_i^2 and we identify the measurement error by

$$\hat{\Omega} = s_i^2 - s_{EOC,C}.\tag{4.18}$$

The OLS estimator reads in our notation:

$$\hat{b}_{OLS} = \frac{s_{i,O}}{s_i^2},\tag{4.19}$$

$$\hat{\sigma}^2_{\epsilon,OLS} = s_O^2 - \frac{s_{i,O}}{s_i^2}.\tag{4.20}$$

Combining those identities with equations (4.16) results in

$$\hat{b}_{CALS} = \frac{s_{i,O}}{s_{EOC,C}},\tag{4.21}$$

$$\hat{\sigma}^2_{\epsilon,CALS} = s_O^2 - \frac{s_{i,O}^2}{s_{EOC,C}}.\tag{4.22}$$

Similarly, the asymptotic distribution reads:

$$\sqrt{N}\begin{bmatrix}\hat{b}_{CALS}-\beta\\ \hat{\sigma}^2_{\epsilon,CALS}-\sigma^2_\epsilon\end{bmatrix}\to^d \mathcal{N}\left(0, s_{EOC,C}^{-4}\begin{bmatrix}s_{i,O}^2(2s_i^4+s_{EOC,C}^2)+s_i^2 s_{EOC,C}(s_O^2 s_{EOC,C}-4s_{i,O}^2)\\ \frac{1}{2}\left(s_O^2 s_{EOC,C}^2+s_{i,O}^2(s_i^2-2s_{EOC,C})\right)^2\end{bmatrix}\right).\tag{4.23}$$

Figure 4.1: Xetra trading phases for DAX30 stocks.

07:30	08:50	09:00:xx	13:00	13:02:xx	17:30	17:35:xx
Pretrading	Opening Auction	Continuous Trading	Mid Auc- tion	Continuous Trading	Closing Auction	Post- trading

The figure reports the organization of a trading day at the German Stock Exchange. The exact times are valid for the Dax30 companies and can vary for other instruments.

Figure 4.2: The auction price mechanism at Xetra.

Highest executable volume
$V = max(V_{ASK}, V_{BID}) \quad \forall p$

None: no AP
Display best bid & ask

Only one ⇒ AP

Multiple: Lowest surplus
$S = V - min(V_{ASK}, V_{BID})$

Surplus on which side?

Only one ⇒ AP

Bid side only
highest ⇒ AP

Ask side only
lowest ⇒ AP

Both sides
reference ⇒ AP

The figure reports how prices are set at the end of the auction under the Xetra platform. V denotes executable volume, p the indicative price, S the surplus, and AP the auction price. V_{ASK} and V_{BID} are the executable volumes at the bid and the ask side of the book respectively.

Table 4.1: Comparison of Auction Designs.

Exchange	Opening Auction	Closing Auction	Intraday Auction	Volatility Interruption	Random End
Deutsche Börse	10 min	5 min	2/5 min [1]	2 min s+d[2]	30 sec
London Stock Exchange	10 min	5 min	-/5 min [1]	5 min d:± 5%	30 sec
Euronext	105 min	5 min	-	4 min d:± 2% s:± 10%	-

[1] Only on expiration days for derivative contracts.
[2] Price ranges are not disclosed and vary depending on the market conditions and the instrument.

The table reports the duration of the main auction phases at 3 of the largest exchanges operating with an order book, namely the Deutsche Börse, the London Stock Exchange (LSE), and Euronext. Columns report the duration of each type of auction. Volatility interruptions are launched when prices reach the static (s) or dynamic (d) price bounds. The bounds are defined as $\pm x\%$ of a reference price, typically the last auction price for the static bounds, and the last transaction price for the dynamic bounds. In all panels, the sign '-' indicates that the auction phase is not part of the exchange platform.

Table 4.2: Comparison of Auction Designs - continued.

Exchange	Extension due to Market Order Imbalance	Extension due to Price Control	Volume Control	Visibility of Indicative Price and Volume	Visibility of Order Book
Deutsche Börse	2 min [1,2]	A:2 min[1] s+d [3] B: manual termination, 2d	-	x	-
London Stock Exchange	2 min [1]	A: 5 min [1]d:±1/3/5%[4] B:5 min [1]d:±1/3/5%[4]	Closing auction only	x	x
Euronext	-	4 min d:± 2% s:± 10%	-	x	x [6]

[1] There is an additional random phase of 30 seconds maximum.
[2] The phase terminates as soon as market orders can be executed.
[3] Price ranges are undisclosed, and vary depending on the market conditions and the instrument. Volatility interruptions exhibit only one extension (B).
[4] The dynamic ranges are stock- and auction- specific. Opening auctions can be extended only once, volatility interruptions cannot be extended. Trading is not allowed in-between the auctions if a volatility interruption is triggered within 10 minutes of the closing or intraday auctions.
[5] The executable volume at the end of the auction is compared to $max(0.5 * NMS, 2500)$, where NMS stands for the "Nominal Market Size".
[6] Executable limit orders are reported with the indicative price as the limit price. Non-members see only the best five quotes.

The table reports characteristics of the auction phases at 3 of the largest exchanges operating with an order book, namely the Deutsche Börse, the London Stock Exchange (LSE), and Euronext. The left panel reports characteristics of the auction extensions: the duration of the market imbalance phase, and the duration and new price ranges of a price control extension. The latter is triggered when the auction price is set outside the static or dynamic price ranges defined in Table 4.1. A (resp. B) refers to the first (resp. second) extension. The middle panel reports auctions where no orders are executed when the executable volume does not reach a minimum threshold. In both panels, the sign '-' indicates that the phase is not part of the exchange platform. In the right panel, the sign 'x' (resp. '-') indicates that the piece of information is disclosed to the market (resp. undisclosed).

Table 4.3: Characteristics of the stocks in the sample.

Company name	Market cap	Turnover	Nb trades	Spread	Grp
DT.TELEKOM AG NA	34858	351	4445	0.07	1
SIEMENS AG NA	52893	322	4418	0.05	1
DEUTSCHE BANK AG NA O.N.	38228	309	3961	0.05	1
ALLIANZ AG VNA O.N.	33805	290	4523	0.05	1
MUENCH.RUECKVERS.VNA O.N.	16396	207	3425	0.06	1
DAIMLERCHRYSLER AG NA O.N	30316	188	3309	0.06	1
SAP AG ST O.N.	27412	185	2806	0.06	1
E.ON AG O.N.	33753	161	2871	0.06	2
INFINEON TECH.AG NA O.N.	4790	146	2799	0.10	2
BASF AG O.N.	25425	124	2580	0.06	2
VOLKSWAGEN AG ST O.N.	9688	104	2545	0.07	2
BAY.HYPO-VEREINSBK.O.N.	6629	98	1937	0.11	2
RWE AG ST O.N.	12653	98	2314	0.08	2
BAYER AG O.N.	15911	89	2400	0.08	2
BAY.MOTOREN WERKE AG ST	12211	88	2110	0.07	2
COMMERZBANK AG O.N.	7569	53	1450	0.11	3
SCHERING AG O.N.	7055	51	1523	0.09	3
LUFTHANSA AG VNA O.N.	4548	44	1352	0.12	3
DEUTSCHE POST AG NA O.N.	6806	44	1315	0.11	3
METRO AG ST O.N.	5018	39	1235	0.12	3
THYSSENKRUPP AG O.N.	6450	38	1262	0.13	3
DEUTSCHE BOERSE NA O.N.	4847	36	982	0.10	3
ADIDAS-SALOMON AG O.N.	4104	32	980	0.09	4
ALTANA AG O.N.	3338	31	1095	0.10	4
MAN AG ST O.N.	2434	28	1057	0.12	4
TUI AG O.N.	2025	26	1063	0.14	4
CONTINENTAL AG O.N.	4060	26	1002	0.11	4
LINDE AG O.N.	3448	22	896	0.11	4
HENKEL KGAA VZO O.N.	3682	18	702	0.10	4
FRESEN.MED.CARE AG O.N.	1944	13	621	0.13	4
Average	14076	109	2099	0.09	

The table reports characteristics of the 30 stocks in our sample. All statistics report daily averages for the sample period January 2, 2004 to March 31, 2004 except for the column *Market cap.* which gives the market capitalization of the respective stock in million euros at the end of December 2003. *Turnover* is the total average turnover (in millions of euros) per trading day. *Nb. trades* is the average daily number of trades. *Price* and *Spread (%)* denote the average midquote and relative spread over the 3 months sample period. The stocks are sorted into four groups according to traded volume, i.e. by the column *Turnover*.

Figure 4.3: Price informativeness, OLS vs CALS.

The graphs display the estimated parameters from equation (4.4). The triangle (resp. circle) refers to the OLS (resp. CALS) estimates of b. The horizontal bars are the heteroskedastic-consistent confidence intervals at 95% for each coefficient estimate. The top left corner displays the results for the group of most frequently traded stocks. The top right panel shows the results for the second and the lower left panel depicts the result for the third trading activity quartile. The lower right panel presents the results for the least frequently traded stocks.

Figure 4.4: Price informativeness, microstructure noise and residual variances.

The graphs display results from equation (4.4). The triangle (resp. the plus sign) refers to the OLS (resp. CALS) residual variance. The circle reproduces the estimated measurement error (microstructure noise). The top left corner displays the results for the group of most frequently traded stocks. The top right panel shows the results for the second and the lower left panel depicts the result for the third trading activity quartile. The lower right panel presents the results for the least frequently traded stocks.

Figure 4.5: Learning hypothesis.

The graphs display the variance of three variables: $var(r_t^i)$ (dashed-and-dotted line), $var(r_t^{9:10} - r_t^i)$ (solid line) and $var(r_t^C - r_t^i)$ (dashed line). r_t^i is the return of the indicative price at time i on the closing price the previous day, $r_t^{9:10}$ is the return of the midquote price at 9:10 on the same closing price, and r_t^C is the close-to-close return. The top left corner displays the results for the group of most frequently traded stocks. The top right panel shows the results for the second and the lower left panel depicts the result for the third trading activity quartile. The lower right panel presents the results for the least frequently traded stocks.

Figure 4.6: Volume, surplus and elasticities throughout the auction.

The graphs display the average for each time *i* of the following 4 variables: executable volume (represented by a triangle), surplus (represented by a star), the 3-quote bid elasticity (dashed line) and the 3-quote ask elasticity (solid line). The top left corner displays the results for the group of most frequently traded stocks. The top right panel shows the results for the second and the lower left panel depicts the result for the third trading activity quartile. The lower right panel presents the results for the least frequently traded stocks.

Figure 4.7: Biais model augmented by volume.

The graphs display results from the Biais model, augmented by an interaction term, eq. (4.9). The triangle reports the estimate for the return of the indicative price on the previous day closing price. The circle reports the estimate for the interaction term. The horizontal bars are the heteroskedastic-consistent confidence intervals at 95% for each coefficient estimate. The top left corner displays the results for the group of most frequently traded stocks. The top right panel shows the results for the second and the lower left panel depicts the result for the third trading activity quartile. The lower right panel presents the results for the least frequently traded stocks.

Figure 4.8: Biais model augmented by book elasticities.

The graphs display results from the Biais model, augmented by book elasticities, eq. (4.11). The triangle reports the estimate for the return of the indicative price on the previous day closing price. The star (resp. the plus sign) reports the estimate for the bid (resp. ask) elasticity. The horizontal bars are the heteroskedastic-consistent confidence intervals at 95% for each coefficient estimate. The top left corner displays the results for the group of most frequently traded stocks. The top right panel shows the results for the second and the lower left panel depicts the result for the third trading activity quartile. The lower right panel presents the results for the least frequently traded stocks.

Epilogue

> Liquidity is created through a give and take process in which multiple counterparties selectively reveal information in exchange for information ultimately leading to a trade.
>
> Hasbrouck (2004): *Empirical Market Microstructure*, p.5

An increasing number of financial assets trade in limit order markets. These markets can be characterized by the following keywords: Transparency, anonymity and endogenous liquidity supply. Endogenous here implies voluntarily participation of market participants to ensure the liquidity of the limit order book.

This is very different to the previously dominating market mechanism, where supply was guaranteed by the market maker. During phases of diminishing outside participation in the order book, the market maker acts as last resort of liquidity, even if he has to suffer some losses on her own positions. To compensate for those potential losses, the exchange grants market makers advantages over the other participants that allow them to earn additional profits during less volatile periods. Due to the rareness of the extreme periods the market maker receives the compensation upfront. Thus exchanges enforce a sufficient level of liquidity provision with rulings about minimum levels of certain liquidity measures.

One can easily identify two substantial problems with the market maker approach. First the details of minimum liquidity provision in volatile phases has to be defined by the regulators ex ante. As those periods occur infrequently and vary in their characteristic, it is at least a demanding task, most likely even an impossible one. And the level of compensation throughout the quiet periods has to be balanced with the expected losses during the more extreme events. Overcompensation implies a welfare loss, undercompensation risks that the market makers go bankrupt at precisely the moment when their services are most needed. The regulator could consider to auction the market maker position but then faces both the problems of adverse selection and moral hazard.

However, the limit order book by construction offers a flexible instrument to regulate the compensation of liquidity provision: Its slope defines exactly price to trade a specific quantity instantaneously. In periods of increased uncertainty or financial distress each liquidity supplier reviews her expected risk and adjust her offering accordingly. Thus the order book allows for both competition between dealers and flexibility of the reward scheme. This in my view is the lesson learned from Glosten (1994b).

In hindsight his paper proved to be a prophecy: The limit order book indeed proved to be inevitable - at least for trading standardized assets in large volumes between a large number of traders. For those kind of markets the limit order book is "... a stable institution and, within the set of economic and trading structures considered, the only stable institution".

Bibliography

Ahn, H.-J., Bae K.-H., and K. Chan, 2001, Limit orders, depth, and volatility: Evidence from the stock exchange of Hong Kong, *Journal of Finance* 54, 767–788.

Aitken, Michael J., Henk Berkman, and Derek Mak, 2001, The use of undisclosed limit orders on the australian stock exchange, *Journal of Banking & Finance* 25, 1589–1603.

Andrews, D. W. K., 1999, Consistent moment selection procedures for generalized method of moments estimation, *Econometrica* 67, 543–564.

Barclay, MJ, and T Hendershott, 2003, Price discovery and trading after hours, *Review of Financial Studies* 16, 1041–1073.

Bauwens, L., and P. Giot, 2001, *Econometric Modelling of Stock Market Intraday Activity* (Kluwer Academic Publishers).

Beltran-Lopez, H., P. Giot, and J. Grammig, 2004, Commonalities in the order book, Mimeo, CORE, Universities of Louvain, Namur, and Tübingen.

——— , 2006, Commonalities in the order book, Mimeo, CORE.

Bessembinder, Hendrik, Marios Panayides, and Kumar Venkataraman, 2007, In search of liquidity: An analysis of order submission strategies in automated markets, Discussion paper University of Utah.

Biais, Bruno, Christophe Bisiere, and Chester Spatt, 2002, Imperfect competition in financial markets: Island versus nasdaq, Discussion paper University dé Toulouse.

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

Biais, Bruno, Pierre Hillion, and Chester Spatt, 1999, Price discovery and learning during the preopening period in the paris bourse, *Journal of Political Economy* 107, 1218–1248.

Biais, B., D. Martimort, and J.-C. Rochet, 2000, Competing mechanisms in a common value environment, *Econometrica* 68, 799–837.

Bloomfield, R., M. O'Hara, and G. Saar, 2003, The make or take decision in an electronic market: Evidence on the evolution of liquidity, *Journal of Financial Economics* 75, 165–199.

Boehmer, E., 2004, Dimensions of execution quality: Recent evidence for u.s. equity markets, EFA 2004 Working Paper Texas A&M University.

Bongiovanni, Steve, Milan Borkvec, and Robert Sinclair, 2006, Let's play hide-and-seek: The location and size of undisclosed limit order volume, Discussion paper Investment Technology Group.

Cao, C., O. Hansch, and X. Wang, 2004a, The informational content of an open limit order book, EFA 2004 Working Paper, Penn State University and Southern Illinois University.

Cao, Charles, Oliver Hansch, and Xiaoxin Wang, 2004b, The informational content of an open limit order book, mimeo, Penn State University.

Cheung, Y.C., F. de Jong, and B. Rindi, 2003, Trading european sovereign bonds: The microstructure of the mts trading platforms, Discussion paper EFA 2003 Annual Conference Paper `www.ssrn.com/abstract=424936`.

Coppejans, M., I. Domowitz, and A. Madhavan, 2003, Dynamics of liquidity in an electronic limit order book market, Discussion paper Department of Economics, Duke University.

De Winne, Rudy, and Catherine D'Hondt, 2007, Hide-and-seek in the market: Placing and detecting hidden orders, Discussion paper EDHEC Business School.

Degryse, H., F. de Jong, M. Ravenswaaij, and G. Wuyts, 2003, Aggressive orders and the resiliency of a limit order market, Working Paper, K.U. Leuven, Department of Economics.

DeJong, F., T. Nijman, and A. Roëll, 1996, Price effects of trading and components of the bid-ask spread on the Paris Bourse, *Journal of Empirical Finance* pp. 193–213.

Dufour, A., and R.F. Engle, 2000, Time and the price impact of a trade, *Journal of Finance* 55(6), 2467–2499.

Easley, D., N.M. Kiefer, M. O'Hara, and J. Paperman, 1996, Liquidity, information, and less-frequently traded stocks, *Journal of Finance* 51, 1405–1436.

Esser, Angelika, and Burkart Mönch, 2005, The navigation of an iceberg: The optimal use of hidden orders, Discussion paper Goethe University.

Fernandes, M., and J. Grammig, 2005, Non-parametric specification tests for conditional duration models, *Journal of Econometrics* 127, 35–68.

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

——— , O. Kadan, and E. Kandel, 2003, Limit order book as a market for liquidity, Mimeo, HEC.

Frey, Stefan, and Joachim Grammig, 2006, Liquidity supply and adverse selection in a pure limit order book market, *Empirical Economics* 30, 1007–1033.

Fuller, Wayne A., 1987, *Measurement Error Models* (New York: Wiley).

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

Glosten, L.R., and L.E. Harris, 1988, Estimating the components of the bid-ask spread, *Journal of Financial Economics* 21, 123–142.

Glosten, Lawrence R., 1994b, Is the electronic limit order book inevitable?, *Journal of Finance* 49, 1127–1161.

Gomber, P., U. Schweickert, and E. Theissen, 2004, Zooming in on liquidity, EFA 2004 Working Paper, University of Bonn.

Grammig, J., A. Heinen, and E. Reginfo, 2004, Trading activity and liquidity supply in a pure limit order book market, CORE Discussion Paper 58 Université Catholique de Louvain.

Hall, A., N. Hautsch, and J. Mcculloch, 2003, Estimating the intensity of buy and sell arrivals in a limit order book market, Mimeo, University of Technology, Sidney and University of Konstanz.

Hall, A., and F. Peixe, 2003, A consistent method for the selection of relevant instruments, *Econometric Reviews* 22, 269–87.

Hansen, L. P., 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1054.

Harris, Lawrence, 1997, Order exposure and parasitic traders, Discussion paper University of Sourthern California.

Harris, L., 2003, *Trading & Exchanges* (Oxford University Press: New York).

———, and J. Hasbrouck, 1996, Market vs. limit orders: The superdot evidence on order submission strategy, *Journal of Financial and Quantitative Analysis* pp. 213–231.

Hasbrouck, J., 1991, Measuring the information content of stock trades, *Journal of Finance* 46, 179–207.

———, 2004, Lecture notes for PhD seminar in empirical market microstructure, Stern School of Business, NYU.

Hasbrouck, Joel, and Gideon Saar, 2004, Technology and liquidity provision: The blurring of traditional definitions, Discussion paper New York University.

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

Irvine, P., G. Benston, and E. Kandel, 2000, Liquidity beyond the inside spread: measuring and using information in the limit order book, Mimeo, Goizueta Business School, Emory University, Atlanta.

Kalay, Avner, Orly Sade, and Avi Wohl, 2004, Measuring stock illiquidity: An investigation of the demand and supply schedules at the tase, *Journal of Financial Economics* 74, 461–486.

Kaniel, R., and H. Liu, 2001, So what orders do informed traders use?, Discussion paper University of Texas at Austin.

Kapteyn, A., and T.J. Wansbeek, 1984, Errors in variables: Consistent Adjusted Least Squares (CALS) estimation, *Communications in Statistics - Theory and Methods* 13, 1811–1837.

Meijer, Erik, and Tom Wansbeek, 2000, Measurement error in a single regressor, *Economics Letters* 69, 277–284.

Næs, Randi, and Johannes Skjeltorp, 2006, Order book charcateristics and the volume-volatility relation: Empirical evidence from a limit order market, *Journal of Financial Markets* 9, 408–432.

O'Hara, M., 1995, *Market Microstructure Theory* (Blackwell Publishers: Malden, MA).

Pardo, Angel, and Roberto Pascual, 2006, On the hidden side of liquidity, Discussion paper University of Valencia.

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

Pascual, R., and D. Veredas, 2004, What pieces of limit order book information are informative? an empirical analysis of a pure order-driven market, Mimeo. Dept. of Business. Universidad de las Islas Baleares.

Ranaldo, A., 2004, Order aggressiveness in limit order book markets, *Journal of Financial Markets* 7(1), 53–74.

Sandås, Patrik, 2001, Adverse selection and competitive market making: Empirical evidence from a limit order market, *Review of Financial Studies* 14, 705–734.

SEC, 2001, Report on the comparison of order executions across equity market structures, Office of Economic Analysis United States Securities and Exchange Commission Working Paper.

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

Tuttle, Laura, 2006, Hidden orders, trading costs and information, Discussion paper University of Kansas.

Wansbeek, Tom, and Erik Meijer, 2000, *Measurement Error and Latent Variables in Econometrics* (North-Holland).