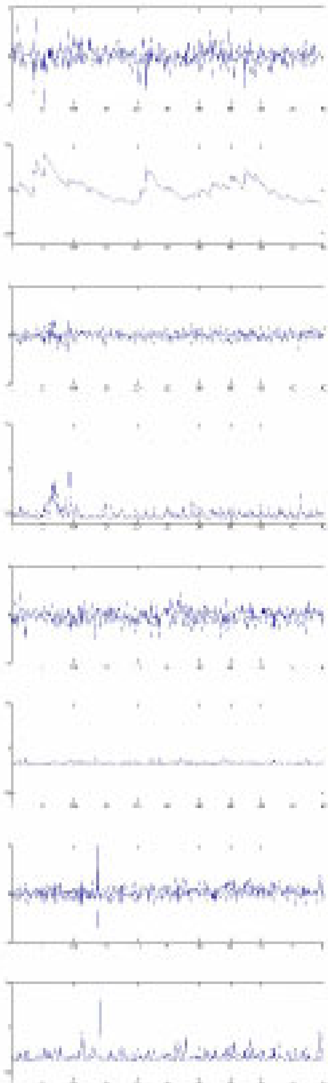
Limit Order Book Dynamics and Asset Liquidity

Georg Pristas



Cuvillier Verlag Göttingen

Limit Order Book Dynamics and Asset Liquidity

Dissertation
for the Faculty of Economics, Business Administration and Information Technology of the University of Zurich

to achieve the title of
Doctor of Economics

presented by

Georg Pristas
from Volketswil (Schweiz)

approved at the request of

Prof. Dr. Markus Leippold
Prof. Dr. Marc Chesney

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar.
1. Aufl. - Göttingen : Cuvillier, 2008
Zugl.: Zürich, Univ., Diss., 2007

978-3-86727-679-5

The Faculty of Economics, Business Administration and Information Technology of the University of Zurich hereby authorises the printing of this Doctoral Thesis, without thereby giving any opinion on the views contained therein.

Zurich, 13 June 2007

the Dean: Prof. Dr. H. P. Wehrli

Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

1. Auflage, 2008
Gedruckt auf säurefreiem Papier

978-3-86727-679-5

Acknowledgements

Research projects require ingenuity, passion, purposefulness and, last but not least, patience. To succeed one has to revise its ideas, thoughts and solution approaches and methods continuously. From this point of view, every instance of guidance, feedback or input can be very helpful and supportive.

I would like to thank my supervisor Prof. Dr. Markus Leippold, who provided me the opportunity to work as a research assistant at his chair and offered me the chance to work in a productive and inspiring environment. He gave me the academic freedom to develop my own ideas and to write this thesis. Moreover, he pushed me repeatedly to high performance, gave me valuable hints and tips to several areas of my work and supported me with the collection of a set of ambitious, challenging doctoral seminars and academic papers. Thank you very much, Markus! Additionally I am also grateful to my co-supervisor Prof. Dr. Marc Chesney, for his help and for agreeing to make the doctoral exam.

For inspiring discussions on finance, econometrics, data matters, Matlab, Latex or administrative issues concerning the writing of a doctoral thesis I am pleased to thank Prof. Dr. Marc Paolela and Prof. Dr. Paolo Vanini. I would like to thank Dr. Carsten Murawski, Dr. Simon Lamprecht, Stephan Jöhri and Dr. Alessia Ortner for excellent proofreading and giving me constructive feedback. Then I would like to thank all the people related to the academic field who are not namely here.

Outside of the academic field I would like to thank all my friends, especially my family for listening to my thoughts and supporting me even during hard times. Moreover, I give thanks to Michèle Rütimann for designing the cover of the book (thank you for avoiding the elves on the front!).

Especially, I would like to thank my girlfriend Simone, for her patience in repeatedly listening to the many details about and around the writing of this thesis. She has been a constant source of love, concern, support and strength all these years. I would like to express my heart-felt gratitude to her!

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Liquidity, in the sense of market liquidity, is an essential characteristic of a well working financial market. In fact, the absence of liquidity can influence the trading process considerably. The simple situation that an investor is not able to sell any given amount of assets at a given point of time can cause a financial distress even up to its insolvency. From this point of view, market liquidity can be seen as the life elixir of financial markets.

Albeit it is easy to circumscribe or to think of intuitively, liquidity is much more difficult to define appropriately. Market liquidity, or, as related to one single asset, asset liquidity, has numerous dimensions. As mentioned above for the buyer side, timing is one of them. An additional dimension is the size of the amount of shares the investor needs to sell. The combination of just these two dimensions influences the price impact of the order in the market. But the price impact depends as well on the trading preparedness of the counterparty, the seller side. Hence, asset liquidity can be considered as a multidimensional problem.

Already a descriptive specification of a liquid asset or a liquid market reveals the complex nature of liquidity considerations. Since liquidity consists of several different dimensions it is difficult to find a definition that accounts for all these attributes. Such a definition has to be bound very general. An example can be found for instance in O'Hara (2004), where she states that a liquid market is one in which investors can trade into and out of positions quickly and without causing large price effects.

The vast majority of academic works suppose the existence of enough asset liquidity implicitly. Most models for valuation assume that the considered asset can be sold or bought immediately and for no costs, i.e. that enough asset liquidity is available. For that reason the aspect of not being able to trade is not included in valuation. However, two stories from financial history should help to highlight the importance of the topic of asset liquidity and in particular, the consequences of the lack of asset liquidity. The first one is the market crash of October 1987 and the second event mentioned here is the financial collapse of the Long Term Capital Management (LTCM) in 1998.

In financial markets history, Black Monday is the name given to Monday, October 19, 1987, when markets were suddenly flooded by sell orders overnight, sweeping away the outstanding buy orders. Because of this imbalance of orders, markets declined extraordinarily strongly and caused the greatest financial distress of modern times. Grossman and Miller (1988) discuss this event in the framework of their liquidity model and stated that in that situation "...markets had become highly illiquid and virtually incapable of supplying immediacy at low cost...". The crash started in New York and Chicago and activated a chain reaction across all financial markets around the world. It is the largest market breakdown reported in financial history. Although the event was entirely, or even primarily, not a matter of liquidity rather than of fundamentals, the large effects were assisted, if not even caused by the liquidity question.

The second event was the crisis of the hedge fund Long-Term Capital Management (LTCM) in 1998. The company had developed complex mathematical models and trading strategies to take advantages of different market arbitrage situations. Starting with fixed income arbitrage deals of different government bonds of U.S., Japanese and European markets, they moved into several riskier market environments. To enhance the high returns additionally, the fund had been extremely leveraged with borrowed capital. Consequently the fund grew so much that it had became the primary supplier in several markets. Hence, to find a counterparty to trade immediately large positions had been a hard task. As the Russian crises started and Russian government bonds defaulted, LTCM had been forced to sell positions. Not least by virtue of few market liquidity LTCM collapsed.[1]

[1]For further readings see for instance Dunbar (2001) or MacKenzie (2003).

These two crises have shown an important aspect of liquidity - the risk of suddenly losing it. However, there is another more subtle aspect of liquidity. It is the day by day costs of transactions that have a significant impact on market prices and that are influenced by the level of liquidity. And of course the impact of dynamically changing liquidity levels. The better understanding of these changes and of course their influence on liquidity risk will be in the center of interest of this thesis.
The two liquidity risk stories of the previous section are a good example of the many faces of liquidity or its absence - illiquidity. Among the named symptoms of these liquidity crises were:

- The inability to find trading partners at market prices
- The inability to execute orders immediately
- High price impacts of trades
- High differences between offered buying and selling prices

These costs also occur in financial markets in the absence of a liquidity crisis but are influenced by the level of market respective asset liquidity.

1.1 Review of Related Literature

The term liquidity is multifaceted and can be used in several financial contexts. On the one hand in corporate finance or in accounting matters the concept of liquidity can be used to describe the ability to fulfill payment obligations at any time. This view is more cash or company based rather than market related. On the other hand an asset is considered liquid if it can be traded quickly, in large quantities and with little impact on the market price[2]. In literature this kind of concept is called asset liquidity or, if refers to the market as a whole, market liquidity.
Asset or market liquidity, or henceforth simply liquidity, is an elusive concept. Because neither is directly nor explicitly identifiable as a risk factor, there exist

[2]See for similar definitions Keynes (1936), Black (1971), Glosten and Harris (1988) or Harris (1990a) among others.

a lot of different interpretations and definitions of it. Even though the existence of liquidity was already mentioned by Keynes (1936), Black (1971) tried to make a first descriptive characterization. Black (1971) established three dimensions to capture liquidity in accordance with today's understanding. He introduced the amount of stocks that can be traded at a given price (depth), the ability to trade across assets without affecting the price (breadth) and the dynamic of how quickly the price returns to its pre-traded price level (resiliency).

From another point of view that trading causes transacting costs, Demsetz (1968) already argued that the difference between the buy order price, so-called bid price, and the sell order price, referred to as ask price, reflects transaction costs. Hence, this difference between the bid price and the ask price, conventionally called the bid-ask spread, may be a proxy candidate for measuring liquidity.

Commenced with these main considerations, a vast amount of researchers started to investigate the aspects of liquidity. West and Tinic (1971), Amihud and Mendelson (1980), S. Phillips and Simth (1980), Amihud and Mendelson (1982), Ho and Stoll (1981), Ho and Stoll (1983) or Copeland and Galai (1983) for example, explored different aspects of transaction costs connected to the bid-ask spread. They conclude that the bid-ask spread may be an appropriate measure for liquidity. Garbade (1982) and Stoll (1985) shows that a negative link between the bid-ask spread and trading volume, number of shares, number of market makers and stock price continuity exists. Analysis of transaction costs in the context of a fixed investment horizon has been made by A. Chen, Kim, and Kon (1975), Levy (1978) or Milne and Smith (1980), among others.

Kyle (1985), Glosten and Milgrom (1985), Easley and O'Hara (1987) and Admati and Pfleiderer (1988) for instance started to theoretically describe the impact of liquidity on stock returns. Beside theoretical works empirical investigations such as Amihud and Mendelson (1986), Constantinides (1986), Grossman and Miller (1988), Heaton and Lucas (1996), Vayanos (1998), M. Huang (2003), O'Hara (2003), O'Hara (2004) or Eisfeldt (2004) followed and emphasized this proposition. Huberman and Halka (2001) analyzed the systematic nature of stock market liquidity. Holmström and Tirole (2001) developed a model where a security's expected return is related to its covariance of aggregate liquidity. Based on the impact of liquidity on stock returns the predictability of stock returns is another direction of sizable amount of investigations. Comparing two equal stocks, which only differ in liquidity aspects, the more liquid one has lower returns but higher prices. Amihud and Mendelson (1986), Brennan and Subrahmanyam (1996) or Brennan, Chordia, and Subrahmanyam (1998) showed that measures of increased liquidity are associated with lower future returns, which seems to evidence this proposition. Among others, Pastor and Stambaugh (2003) find that aggregate liquidity is a priced risk factor in that sense. Amihud (2002) and C. Jones (2002) are even of the opinion that liquidity movements can be used to forecast aggregate returns. Moreover Baker and Stein (2004) stated that time variation in liquidity, on aggregate market or on single asset level, might deliver appropriate estimations of changes in returns and can be used to reveal insider trading.

Research has been extended to other markets than stock markets. The relationship between liquidity and option pricing have been investigated by Longstaff (1995) or Cetin, Jarrow, Protter, and Warachka (2003), for instance. Liquidity in treasury markets have been analyzed by M. J. Fleming (2003) or Strebulaev (n.d.) among others. A variety of different countries have been taken into account. It attracted attention that the kind of market-organization and market-structure, as well as trading rules and regulations, are summarized in market microstructure matters to measure liquidity. O'Hara (1995), Muranaga and Shimizu (1997) or Madhavan (2000) contributed works in that area.

Based oon the knowledge of liquidity so far, some different measures have been tested in different markets. Different authors showed that the market microstructure affects liquidity measures. In the majority of cases two reasons for that are cogitable. First, liquidity measures cannot be easily compared with each other. And second, the efficiency of market microstructure can also be measured using liquidity measures. Ceteris paribus measuring lower transaction costs indicates a more efficient market microstructure.

1.2 Research Question

Economic theory assumes the existence of liquidity in financial markets. Asset liquidity is an essential component for a well working financial market because in the absence of liquidity no trading is possible. Since this liquidity restriction is so elementary it affects not only trading but also several other disciplines of

economic theory such as asset pricing or risk management. To measure liquidity accurately is therefore fundamental for the assumptions of a vast amount of economic concepts and financial models. It delivers the basics of the existence of all these models. The two stories at the beginning of this chapter highlight the effect of a sudden liquidity shock.
Lately, caused through these and other different events, researches into market and asset liquidity matters have grown. A huge amount of different measurement approaches have been developed. Different new figures and alternative concepts have been proposed. But all these concepts and figures consider the time proceeding dynamical behavior only implicitly.
In this thesis the main concern is to deal with the dynamical behavior of asset liquidity as time proceeds. Moreover we define the measure of the dynamical changes as the measure of asset liquidity and show that this measure reflects asset liquidity most accurately. From this point of view several questions demand to be answered. In particular these are:

- Is there a possibility to reflect the limit order book and its movements?
- Are there characteristic structures in this movements?
- Is it possible to capture the time series of these movements with a dynamical model?
- Is the measure of these dynamical changes an appropriate measure for asset liquidity?

These goals will be attained with both a sound mathematical background and on an economic fundament. The detailed structure of this thesis will now be described in the next paragraph.

1.3 Structure

The thesis is structured as follows. The first section is separated into two parts. In the first part of the section, market microstructure and some elements of it will be introduced. It will be shown how market liquidity depends on the structure and on the environment of financial markets. In connection with market microstructure the concept of the limit order book will be introduced and explained in more detail. The second part of this section gives an overview about liquidity concepts in general. A definition of liquidity is given, liquidity dimensions are provided and several basic liquidity measures are introduced. The next section gives a short overview about specific econometric models which are used in later sections. The subsequent section introduces our new concept of an asset liquidity measure. In this section the basic ideas behind the concept are derived. We show the additional information which can be captured by the model and introduces the measurement approach in more detail. In particular, the new measurement approaches will be considered under an economic point of view. The next section is the main part of the thesis. It presents the empirical application of the model and delivers the results. The chapter starts with the reconstruction of the limit order book and shows evidence of characteristic structures in the dynamic of the limit order book. The limit order book will be examined under several points of view while these results build the fundament of the new measurement concept. Subsequently, an extensive econometric analysis follows and finally, the new concept of measuring asset liquidity based on the dynamic of the limit order book will be applied. A liquidity ranking based on this new measurement approach is presented. The chapter closes with the presentation of the liquidity premium computed on the liquidity ranking from the previous section. The last chapter summarizes the findings and gives some concluding remarks and an outlook for additional researches.

Chapter 2

Market Microstructure and Liquidity Measures

Liquidity and particularly liquidity concepts are affected by the structure of the market. Since liquidity is commonly considered as the ability to trade, the organizational form of the market is an integral part of the concept. Different market structures, respectively market microstructures, provide liquidity in different ways. Hence, the market environment is essential to explain and understand liquidity. However, beside the market microstructure, the measurement approach itself is explicitly relevant to identify how liquid an asset is. Moreover the measurement approach is closely related to the market microstructure since the market environment determines the ways how liquidity can be provided. From this point of view this chapter is divided into a section about market microstructure in the light of liquidity and a section about liquidity measures.

2.1 Market Microstructure

Market microstructure is an area of finance that is concerned with the processes and effects of trading assets under explicit trading rules. While most of the scientific economic researchers simplify or neglect the mechanisms and influences of practical execution of trading, market microstructure literature analyzes how these trading mechanisms affect price behavior and price formation processes. According to Madhavan (2000) these analysis range from investigations of the market structure and design over price formation and price discovery to information arrivals and regulation of disclosures. Aspects as strategic trading, market performance or liquidity measurement concepts are also covered by market microstructure theory as pointed out by O'Hara (1995). Market structure and design addresses to the architecture of the market. It refers to the set of rules governing the trading process. Rules like the degree of automation and continuity, the existence of dealer intermediation, order placement procedure, market open and close regulations or information transparency are some elements of it. Price formation and price discovery includes aspects of asset pricing theory. Models that typically assume complete markets, no taxes and transaction costs, perfect competition or free market entry are extended in this field of market microstructure research. It also concerns the process by which prices absorb new information. Additionally researches in market microstructure theory submit recommendations for regulation to handle the publications of price relevant information and disclosures of listed companies.

These mentioned components of market microstructure explicitly or implicitly affect liquidity measurement concepts, market performance and trading strategies. The following section presents an overview about selected disciplines of market microstructure, focused on particular components necessary in order to understand the Swiss market microstructure. Hence, in order to establish an appropriate framework, the section covers mainly elements of market structure and design. In particular, elements are highlighted which, if they are different in other markets, influence trading and related activities.

The section's structure and the terms are mainly oriented at Madhavan (2000). It starts with describing markets with and without dealer intermediation, i.e. order driven and quote driven markets will be introduced. Afterwards the degree of automation will be explained. The next part presents order placement procedure and how an order can be placed. The limit order book will be established. The section closes with a part about an artificially constructed limit order book.

2.1.1 Market Structures and Design

A market can be structured in several ways. The choice of the market structure, or sometimes also referred to as the market architecture, depends basically on the asset that is traded. Different assets require different market architecture. Commodities, for instance, have other trading requirements than financial derivative instruments have, while securities like stocks are traded differently than real estate is. The market architecture mainly defines the type of auction how assets are traded. Beside other diverse market forms, our main focus in the next following paragraphs forming the market architecture is on stock markets, respectively the Swiss stock market.

Auction Type, Degree of Continuity and Automation

Many kinds of trading can take place at a market. The most established type of a market is an auction market or also called double auction market.[1] In an auction market, an investor who is willing to buy an asset hands in a competitive bid. An investor who aims to sell an asset presents an offer to sale. Both supply and demand offers are collected in an order book and may take place at different times or simultaneously. Depending on the degree of continuity the offers will be matched in a continuous system immediately and in a periodic system (call auction) at a specific point in time.[2] How the orders arrive to the market place, or more precisely are submitted to the market place system, is governed by the automation of the market. On a floor-based market exchange, for example, trading takes place via open outcry. Screen-based markets collect orders electronically.

Intermediary

An intermediary dealer is a designated third party that offers intermediation services between two trading parties. The intermediary institution, also referred

[1] As opposition to an auction market trading, an over-the-counter market can be considered. At an over-the-counter market, trades are negotiated mostly one by one. Some tailor-made derivative instruments or unique assets like real estates are traded in that manner. Stock markets are commonly organized as auction markets.

[2] Empirical research regarding call auction is provided by Mendelson (1982) or Ho, Schwartz, and Whitcomb (1985). They report that empirical findings seem to support, that if large uncertainty about fundamental data exists, or if a market failure is possible, periodic trading is more valuable. Moreover, several authors report that differences between periodic and continuous trading affects asset returns, see for instance Amihud and Mendelson (1991), Stoll and Whaley (1990), Forster and George (1996) or Amihud, Mendelson, and Lauterbach (1997).

to market maker or specialists[3], acts as a counterparty to every investor who places an order to the market. The institution is responsible at least for one security and is required to provide offers to trade, both for the buyer and for the seller side.[4]

Quote Driven Market Markets with designated market makers are called quote driven markets or dealer markets. Market makers are required to make trading possible all the time. They own an amount of cash and an inventory of the security they are responsible to. Thus they can always offer stocks to or buy stocks from the market. The bid price is the price at which the market maker is prepared to buy and the ask price is referred to as the price at which she is prepared to sell. The bid price is set slightly lower than the actual trading price. The ask price is set somewhat higher than this reference price. The difference between the bid and the ask price is called the bid-ask spread or sometimes simply the spread. Market makers can be responsible for more than only one security. In return, a security may also be assigned to more than one market makers.[5]

Order Driven Market In contrast to a quote driven market, an order driven market has no responsible intermediary institution for any traded stock. The supply and demand is only provided by the investor's submitted orders. Orders are entered into the trading systems of the various participants and routed directly to a central order book. Execution of the orders takes place in keeping with the principle of price-time priority.[6] Taking a buy order as an example, the booked orders with the highest price will be processed first and afterwards, in the event of price parity, according to the time of their arrival in the central order book.[7]

[3]The term market maker or specialist depends on the market which is considered. In this thesis these terms are treated as synonyms.

[4]In contrast to a market maker, the field of functions of a financial intermediary is wider defined as that of a market maker. Hence the two terms cannot be used as synonyms.

[5]Many financial markets operate with designated market makers, for instance the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), the NASDAQ Stock Exchange or in Europe the London Stock Exchange (LSE) or also parts of the Swiss Stock Exchange (SWX).

[6]For the principle of the time price priority see section 2.1.1.

[7]In this thesis, the Swiss Stock Exchange (SWX) is organized as an order driven market.

Order Forms

Order forms determine the restriction possibilities associated with the submission of an order to the market. Diverse forms of restriction exists. With the rules about the order placement of the market the restriction possibilities of an order are also defined. In general, an order can have restrictions or have none. Therefore an order can be placed as a market order or as a limit order. The two mentioned order forms beneath are the most common ones and are described in a very general way. Different markets can have different order forms or order forms with additional features.

Market Order A market order has no restrictions. It represents an order form to sell or to buy an asset at the best price currently available. The aim of such an order is to be executed as fast as possible - targeted immediacy. A market order without a given expiration date usually persists one trading day at the market and will be canceled after the closing bell in the evening.

Limit Order A limit order is an order form where a certain price limit is added by the investor. Such a limit order will be executed if the price reaches the limit. Therefore the main focus of such an order is not on an immediate execution rather on a specific price execution. A limit order is valid until a given expiration date. In the case that no such particular information is given at the order, the limit order usually expires in one trading day as well. In the case that the price never reaches the given limit the order becomes worthless after expiration and will be canceled by the market system.[8] Typically a limit order can be canceled anytime without causing any additional costs.

Protocols

Tick Size The minimum tick size is the smallest price change that a price process can move, either up or down. More precisely, it is the next attainable price. At US markets the minimum tick size is one sixteens of a US dollar.

[8] The mentioned forms are the most common ones. There exist a huge amount of other forms like fill or kill, hidden size or accept order for example. For a rigorous overview of the different order forms consult for instance, http://www.interactivebrokers.com/en/trading/orderTypesMatrix.php.

European markets trade in a decimal numeric system. The smallest possible price change is 0.01 units. Moreover the higher the price level of the asset the greater is the minimum tick size, in both US and European markets.[9] There are price ranges with assigned tick sizes. Each security is assigned to a specific category of price range. Once the price of a given security trades within a new price range, the corresponding price step takes immediate effect.[10]

Price-Time Priority One of the important protocols of exchange markets is the price-time priority rule. Before matching takes place at the market, the orders of the order book are arranged by price and time of entry. This procedure is used, regardless of which matching rule will be applied. In the first priority orders are arranged by the best to the worst price (price priority). This means for a buy order from the highest to the lowest offer. Ask orders are ranked conversely. Additionally market orders have priority over limit orders. Given equal price priority, orders are ranked within the same prices according to the time of entry, starting with the oldest one (time priority). Those orders that have been on the book the longest will be taken into account first.

Stop Trading Trading-opening will be delayed stopped or trading will be stopped if the proximate price diverge significantly from the reference price. These divergence-limits are specified by the stock exchange.[11]

Other Protocols There are a plenty of other protocols which can be considered: opening and closing time, opening-, reopening- and closing-restrictions and

[9] Not every asset is subjected to that rule. Some securities have stable minimum tick sizes independent of the price level. Additional information can be found on www.swx.ch

[10] As reported in Chung and Chuwonganant (n.d.), Ahn, Cao, and Choe (1996) showed that a reduction of the tick size at the Amex results in a decline of the bid-ask spread. Similar results found Bacidore (1997), D. Porter and Weaver (1997), Ahn, Cao, and Choe (1998) and Griffiths, Smith, Turnbull, and White (1998) for a purely order driven market, namely the Toronto Stock Exchange. Bollen and Whaley (1998) and Goldstein and Kavajecz (2000) confirmed similar results for the NYSE. Harris (1994) and Harris (1997) concluded that only in a market with a price-time priority rule are these findings are significant. He conjectured that only when market makers compete with public traders the effects on the bid-ask spreads are significant.

[11] Stop trading limits are set for safety reasons. It ought to help to avoid sharp retreats or crashes of prices.

so on.[12] This thesis is focused on liquidity aspects. The just mentioned aspects affect liquidity and are subjects of academic and practical researches.

2.1.2 Limit Order Book and Matching Rules

Limit Order Book The limit order book contains all submitted limited orders separated by the bid and the ask side. The orders are ranked according to the price-time priority rule. On the bid side, the best order with the highest price is on the top of the book, followed by the second best and so on. On the ask side, the order with the lowest price is on the top.

Table 2.1: Limit Order Book

The Table 2.1 displays a part of a limit order book. On the left hand side of the table the offers of the buy side are collected. On the right hand side the supply side are summarized. The limit orders are sorted according to the price-time priority, such that the corresponding best offer appears at the top of the list. Within same prices the orders are sorted according to their submitting time.

	Bid Side			Ask Side	
Time	Volume	Price	Price	Volume	Time
9:40	500	100.00	103.00	950	9:39
9:44	100	100.00	106.00	100	9:42
9:08	250	98.00	107.00	110	9:12
9:48	800	97.00			

The limit order book can be publicly available. But in general the book is not fully observable, except for some special investors. However, in some markets the first few orders close to the best quotes are visible. Other markets sell that information to the investor community.[13]

Matching Rules Matching rules describe how the orders submitted to the market are executed. Basically, there are three sets of rules depending on the current trading phase. These are the pre-opening-, opening- respectively re-opening- and ordinary trading phase. With focus on the matching mechanism

[12] See for a detailed overview www.swx.ch.

[13] The SWX and in particular the Virt-X department of the SWX publishes a set of orders around the best quoted offers.

we mainly concentrate on the rules of the ordinary trading phase.
Principally, each incoming order can initiate a sequence of trades at different price levels, provided that the incoming order has equal or even better conditions than the counterpart in the limit order book. The basic mechanism is that an incoming (limit) order is matched against available limit orders in the book as long as a restriction on one side stops the execution. For instance, an incoming limit order, with a price limit better than the best two offers in the book and a volume larger than the sum of the best two offers, is first executed against the best offer. The order is executed to the price of the limit order in the book and effects therefore a price movement in the corresponding direction.[14] Afterwards, the remaining part of the incoming order is executed to the next best offer in the limit order book, which again changes the price analogously. This process continues until one restriction stops the executions. Given that the incoming order cannot be fully executed because of restriction reasons, the remaining part is registered in the limit order book as an (partly) unexecuted order. Albeit the execution process depends on the restrictions of the order, the mechanism remains for all matching processes similar.[15]
The Swiss market, or the the part of the Swiss market which is considered in this thesis, is endowed as an auction market, without an intermediary, i.e. is an order driven market. The market has a limit order book with the above described matching rules and protocols.

2.2 Liquidity Concepts

In a financial context several risk forms are known. Related to financial markets mainly three risk forms can be distinguished. These are credit risk, market risk and operational risk.
Credit risk, on the one hand, describes the default possibilities of a debtor or a group of debtors, which are somehow connected to each other. More general, this is the risk that a counter party to a transaction will fail to perform according to the terms and conditions of the contract. On the other hand market risk is

[14]For safety reasons, each matching is, before execution, automatically monitored by the exchange to avoid large price movements. A maximal price change is prescribed by the surveillance authority.
[15]For more details consider www.swx.ch.

the risk that the price of a security may decline due to different reasons. These reasons can be for instance, unexpected movements in interest rates, changes in macro-economic data, volatility changes, and so on. Finally, risk forms covered by the term operational risk are risks of the kind like losses due to system breakdowns, employee fraud or misconduct, errors in models or natural or man-made catastrophes. It may also include the risk of loss due to the incomplete or incorrect documentation of trades.[16]

Additionally to these three risk forms, liquidity can be considered as a risk factor as well. It is a risk because the investor does not know whether the asset can be sold in future or not. In this sense liquidity can be understood as the ease with which an investor can convert an investment to cash and vice versa without suffering negative impact on either capital or return.[17] However, an investor bearing those risks has to be compensated in a certain form. Some of these compensations are explicit and can be observed directly. Some of them are implicit and have to be identified and calculated through different measurement approaches. Compensation for liquidity risk can be reflected in transaction costs, market impact, waiting costs or search costs. This form of compensation is also referred to as liquidity premium.

From this point of view this part of the chapter addresses the concept of asset liquidity. In the first part, we showed the complex nature to define liquidity, and provide a definition for our purposes. In a subsequent section we introduce the different dimensions asset liquidity can be understood with and measured by. Afterwards, we focus on some measures in more detail and close this chapter with a presentation of an alternative concept to measure liquidity costs. This chapter serves basically to build a knowledge fundament to comprehend the alternative liquidity measure introduced in the following chapter.

2.2.1 Definitions of Liquidity

The definition of asset liquidity has to take several aspects into account. An essential characteristic of a liquid market is that there are enough investors to buy and sell at all times, so that trading is always possible and can be warranted.

[16] An overview about risk in financial markets can be found, for instance, in Jorion (2005), among many others.

[17] For a more exact definition of liquidity or liquidity risk consider section 2.2.1.

Additionally a further aspect of such a liquid market is that large quantities of assets can be traded without causing large price effects. The time how long an order takes to be executed is another element of a liquid market.
Already a descriptive specification of a liquid market reveals the complex nature of liquidity considerations. Black (1971) for instance, defined a liquid market as one, in which "...bid and ask prices are always quoted, their spreads are small enough, and small trades can be immediately executed with minimal price effects...". A similar definition stated by Harris (1990b): "...a market is liquid if traders can quickly buy or sell large numbers of shares when they want and at low cost...". A more general approach to define liquidity or a liquid market have also been provided by O'Hara (2004): "...a liquid market is one in which buyers and sellers can trade into and out of positions quickly and without having large price effects...". Although the point of view of all these definitions is slightly different the main message rests the same. In a liquid market orders will be executed quickly, without a large influence on the current price and at low costs. In all of these definitions the order of the magnitude of the terms quickly, immediately, small, large or at low cost are not determined in more detail. Moreover, it is difficult to define these terms closer. Particulary on the one hand they are related to each other and to other components. On the other hand they depend on the preferences of the investor. For instance, a large price effect depends on the point of view of the investor and can therefore not be objectively measured. This emphasizes the complex nature of this concept. In this thesis a general definition of the term liquidity will be used.

Definition 2.2.1. *Market liquidity and asset liquidity refer to the ability to buy and sell assets quickly, independent of the quantity and with minimal influence on the current market price.*

The terms quickly, quantity and minimal influence on prices are not defined in more numerical detail since the relation between one of this terms is more important. What is important in the definition is that these three dimensions are covered with it. In comparing different assets, it is inevitable to fix two of the three dimensions and use the remaining one as a comparable figure. For instance, holding the price impact and the quantity constant, the order which is executed faster is referred to be more liquid.

2.2.2 Liquidity Dimensions and Liquidity Measures

As the definition of liquidity suggests liquidity is a complex concept of several dimensions. This multi-dimensionality was first intuitively described by Black (1971). According to today's understanding, Kyle (1985) identified three dimensions in the work of Black. He denoted the amount of stocks that can be traded at a given price as depth. The ability to trade across assets without affecting the price he named as breadth and the dynamic how quickly the price returns to its pre-traded price level have been considered as resiliency. As time proceeds, additional dimensions of liquidity have been identified. However, the terminology of the dimensions of liquidity is not always used in the same way. Sometimes, different authors use the same term with slightly different interpretations or even different definitions. According to Wyss (2004), in this thesis the dimensions depth, immediacy, resiliency and tightness are distinguished.

Depth The quantity of shares demanded at the best bid price is referred to as the bid depth. The volume of shares provided to sell at the ask side is named as the ask depth. The sum of the bid depth and the ask depth is referred to as depth.[18] So,

$$D_t = q_t^{ask} + q_t^{bid} \tag{2.1}$$

where q_t^{ask} refers to the best ask volume and q_t^{bid} refers to the best bid volume in a given time interval t. Figure 2.1 shows depth schematically. On the horizontal axis the volume is specified. Starting from the middle of the axis, on the right hand side the best quoted ask volume is depicted, denoted as the ask depth, and on the left hand side the best quoted bid volume is illustrated, named as bid depth.

In general the volumes on the best quoted volumes, rather than the overall volumes in a given time interval, are considered to calculate depth. According to Csavas-Szilard Erhart (2005) a reason for this is that this volume is the largest order that a market can absorb without evoking a price change. Therefore depth measures the volume of immediacy provided at the best quotes. Hence, Kyle (1985) characterized depth from the opposite point of view as "...the size of an order flow innovation required to change prices a given amount". How-

[18] Some authors use the term quantity depth, market depth or volume depth.

ever, as many authors already pointed out, the measurement concept of market depth is one of the most common approaches to measure liquidity. Studies like Goldstein and Kavajecz (2000), Chordia, Roll, and Subrahmanyam (2001) or Huberman and Halka (2001) used this concept to measure liquidity.[19] Amongst others, Chordia et al. (2001) modified this figure to an average depth measure by dividing depth by a factor of two.[20]

Immediacy/Trading time According to Harris (1990a) immediacy can be defined as the speed at which trades of particular sizes can be executed at a given cost. Immediacy can be interpreted and therefore be measured in several ways. According to Gouriéroux, Jasiak, and Le Fol (1999) immediacy is defined as the waiting time between two subsequent trades. Alternatively, some authors define the inverse of the waiting time or the number of trades per unit of time as immediacy. In contrast to this, immediacy can also be understood as the time until an order has been completely executed at a prevailing price. In this case an order can be executed in several smaller portions than the desired order volume. Then the time between the arrival of the order at the market place until the order has been completely executed, i.e. the last portion of the order has been carried out, is referred to as the trading time or immediacy.

Resiliency The dynamic how prices react to new information or to different order volumes is known as resiliency. From this point of view resiliency can be interpreted as the price-volume elasticity of a given asset.[21] In Figure 2.1 resiliency is presented as the dashed line connected between the aggregate volumes to the corresponding price.

Tightness The difference between the best demand quote, the bid or bid price and the best supply quote, the ask or ask price, is referred to as tightness. According to Demsetz (1968), tightness or the bid-ask spread can be interpreted as transaction cost. No transaction cost means that instantaneous buying and

[19] Brockman and Chung (2001), Lee, Mucklow, and Ready (1993) or Van Ness, R.A., and Pruitt (2000), among others are further examples.

[20] Additional variations of calculating depth can be found in, e.g. Wyss (2004).

[21] Kyle (1985) extend the interpretation of resiliency by considering the speed with which prices tend to converge towards the prevailing price level as resiliency.

selling of an asset for exactly the same price is possible. In reality this is not possible. An asset can only be bought for at least the best ask price or sold for maximum the best bid price. For example, a security offered at a price under the best bid price will be bought immediately and the order quits the limit order book.[22] In this case the seller suffers a loss of the difference of the best bid price and the current sold price. This cost can be interpreted as the price the investor has to pay for an immediate execution.[23]

This consideration reveals arguments why the bid-ask spread may be a good measure for liquidity. The more investors are in the market, the higher is the probability, that the current quoted best offers will be out- or underbid, respectively. This means that a higher bid price or a lower ask price than the currently available best price have been submitted to the market. As calculated above, in that sense the investor has to pay a smaller extra charge to execute an order. The transaction costs become smaller. Therefore a smaller bid-ask spread indicates a more liquid market. The absolute bid-ask spread (s_t) for a given time interval t is given as

$$s_t = p_t^{ask} - p_t^{bid} \tag{2.2}$$

where p_t^{ask} is the best quoted ask price and p_t^{bid} is the best quoted bid price. The absolute spread has the disadvantage that they depend on the absolute price level of the considered asset. The higher the price level the larger the bid-ask spread. This measure cannot be used to compare different assets with each other. In contrast to the absolute spread, the relative bid-ask spread is set in relation to the price level. This price level can be the artificially calculated mid price M_t. The relative spread s_t^{rel} is

$$s_t^{rel} = \frac{p_t^{ask} - p_t^{bid}}{M_t}. \tag{2.3}$$

where the mid price M_t is defined as $\frac{1}{2}(p_t^{ask} + p_t^{bid})$.[24]

---

[22] Assumed that the quantities agree with each other.

[23] From the context it becomes clear that liquidity is a symmetric problem. It affects seller and buyer in the same way, except of that on the bid side the costs occur as a loss and on the ask side they occur as an extra charge. For the sake of simplicity only the case of a sell will be considered.

[24] Numerous authors use and extend the concept of the bid-ask spread. Hamao and Hasbrouck

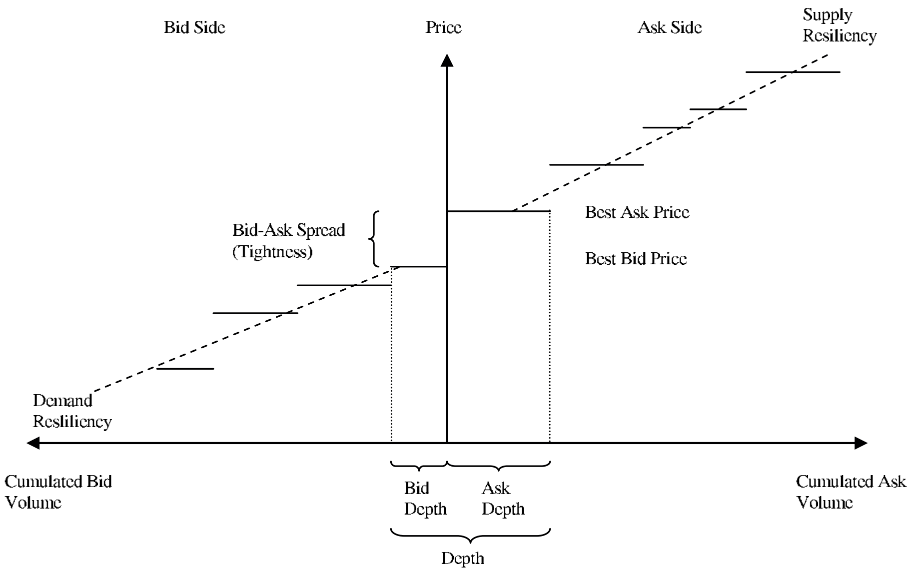


Figure 2.1: Liquidity Dimensions The Figure 2.1 shows the four different liquidity dimensions illustrated with a limit order book for a snap-shot. The horizontal axis represents the cumulative volumes of both the bid side on the left hand side and the ask side on the right hand side respectively. The vertical axis shows the quoted prices.

Figure 2.1 depicts an illustration of the four above mentioned liquidity dimensions with a limit order book in a snap-shot. On the horizontal axis the cumulative volumes of the orders have been drawn. On the left hand side the bid orders are collected and on the right hand side the ask orders are shown. The vertical axis indicates the quoted prices of the orders. The highest vertical line on the bid side is the best bid quote. The lowest line on the ask side is the best ask quote. The difference between the two best quotes corresponds to the bid-ask spread or tightness. The length of the vertical lines represents the quantity of the appropriate order. According to the corresponding best orders this are the bid depth or the ask depth, respectively. Additionally, on the bottom of the illustration the depth is marked. Resiliency is pictured as a dashed line.

(1995) for example, take the logarithm of the absolute spread, while M. Fleming and Remonola (1999) replace the mid price of the relative spread with the last traded price of the previous time interval. Hasbrouck and Seppi (2001), for instance, used the relative spread of the natural logarithm of the prices and in order to get a distribution function closer to the normal distribution Wyss (2004) applies the logarithm to the relative spread.

The line of the resiliency can also be interpreted as the demand and the supply curve. Figure 2.1 is a snapshot of the limit order book in a given time frame since incoming new orders change the picture of the book continuously.

2.2.3 Liquidity Risk

The bid-ask spread given in equation (2.2) is based on prices. No quantities are involved in the calculation. But the cost an investor has to pay depends basically on the amount of the asset which she wants to sell or to buy.
Now, if the amount of assets q the investor intends to trade does not exceed the amount of assets at the best quote, the order will be fully executed at the best quoted price level. The arising transaction costs, or we refer to as liquidity costs $L_t(q)^{best}$, are therefore defined as the difference between the revenue at the mid prices and the revenue at the best quote, or formally

Definition 2.2.2. *Liquidity costs $L_t(q)^{best}$ are defined as*

$$L_t(q)^{best} := q_t \cdot p_t^{bid} - q_t \cdot M_t. \tag{2.4}$$

where q_t refers to the best offered quantity, p_t^{bid} is the best offer on the bid side and M_t corresponds to the mid price between the best ask and the best bid price.

Exceeds the amount of shares (q) the investor aims to sell the amount of shares quoted at the bid price (q^{bid}), the remaining part of the order will be executed at the next lower price. The transaction costs increase once again. The cost to absorb all orders in the order book at a given time interval (t) is referred to as liquidation cost. We define the liquidation cost as

Definition 2.2.3. *The liquidation cost $L_t(q)$ of a liquidation of the position q at time t is the difference of the liquidation price $LP_t(q)$ and the mid price M_t.*

$$L_t(q) := LP_t(q) - q_t \cdot M_t = \int_0^q p_t(s)ds - q_t \cdot M_t \tag{2.5}$$

Figure 2.2 illustrates this liquidity and liquidation cost. At this time the horizontal axis represents the prices. The vertical axis shows the cumulated order volumes. The order volumes are cumulated because a buyer who is prepared to pay the current best price for an asset is also prepared to pay less than the best

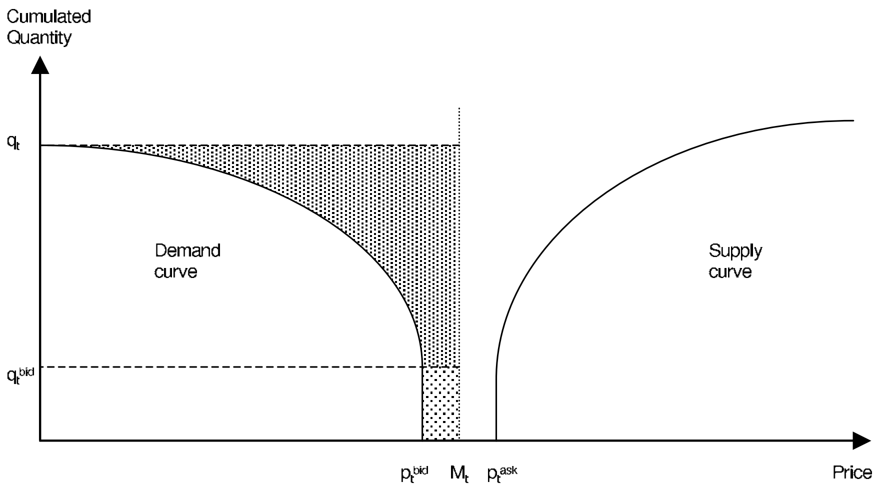


Figure 2.2: Liquidity and Liquidation The Figure 2.2 illustrates the liquidity cost (bright shadowed area) and liquidation cost (dark and the bright shadowed area) according to equation (2.4) and equation (2.5) respectively. The horizontal axis represents the prices while the vertical axis shows the cumulative volumes.

price. In that sense, the smaller the price the more investors are willing to trade. This is depicted with the concave volume-price function $p_t(q)$ on the demand side.[25] The bright shaded area in figure 2.2 shows the liquidity cost $L_t(q)^{best}$. Together with the dark shaded area, the liquidation cost $L_t(q)$ is represented. Since the liquidation cost $L_t(q)$ as stated in definition 2.2.3 is influenced by the volume-price function $p_t(q)$ and the mid price M_t. These functions are not deterministic, so liquidation cost is stochastic and represents a risk.

[25] If there are very few investors, i.e. the asset seems to be very illiquid, the volume-price function can also be of a convex form.

Chapter 3

Econometric Models

Auto Regressive Moving Average (ARMA) processes belong to discrete stochastic difference equation models and, according to Enders (2004), form the basis of time series analysis. The model consists of two parts, an autoregressive (AR) component and a moving average (MA) element. Albeit, the model adresses to several problems in economic time series, it features constant volatility. Many economic time series have neither constant means nor possess constant volatility. Generalized autoregressive conditional heteroscedastic (GARCH) processes address to this problem and offer possibilities to vary volatility over time. Both the ARMA and the GARCH model are commonly used in applied work.

This chapter provides a brief introduction to ARMA and GARCH processes. The chapter is basically oriented at McNeil, Frey, and Embrechts (2005), Enders (2004) and Leippold (2004). Starting with an overview about basic premises and assumptions in the first section, we introduce the ARMA model in a general way. Following on this, we show some model selection approaches and how to verify the model specifications. In the next section we extend the constant volatility models by introducing stochastic volatility processes, the set of GARCH models. Finally, in the last section we combine the ARMA models with GARCH volatility structures and we give a short approach for interpreting the parameters. By reason of the existence of substantial academic literature in the field of time series analysis, this section is kept very general.

3.1 Autoregressive Moving Average Process

A stochastic process $(X_t)_{t\in\mathbb{Z}}$ is a collection of random variables, or a random function, indexed by an integer t. It is a family of random variables defined on some probability space $(\Omega, \mathfrak{F}_t, \mathbb{P})$, where Ω is the space of events that describes all possible states of the world, $\mathbb{P}$ is the associated global objective probability of each event and $(\mathfrak{F}_t)_{t\geq 0}$ is the augmented natural filtration up to time t.

Moments of Time Series The first two moments of a time series model for a single risk factor $(X_t)_{t\in\mathbb{Z}}$ is defined as

$$\begin{aligned} \mu(t) &= \mathbb{E}[X_t] \qquad \forall t \in \mathbb{Z} \\ \gamma(t,s) &= \mathbb{E}[(X_t - \mu(t))(X_s - \mu(s))] \qquad \forall t, s \in \mathbb{Z}, \end{aligned}$$

where $\mu(t)$ describes the mean function and $\gamma(t,s)$ represents the autocovariance function. Note that it is that $\gamma(t,s) = \gamma(s,t)$ for all t, s and for $\gamma(t,t) = var(X_t)$.

Stationarity A common assumption in many time series techniques is that the data are stationary. A time series is called stationary if some properties are independent of time, i.e. that the process has the property that the mean, variance and autocorrelation structure do not change over time.[1] This means that a shift in the time origin does not affect these properties of the process.

Definition 3.1.1 (Strictly Stationarity)**.** *A time series $(X_t)_{t\in\mathbb{Z}}$ is said to be strictly stationary if*

$$(X_t, \dots, X_{t_n}) \stackrel{\text{dist.}}{=} (X_{t_{1+k}}, \dots, X_{t_{n+k}})$$

for all $t_1, \dots, t_n, k \in \mathbb{Z}$ and $n \in \mathbb{N}$.

Definition 3.1.2 (Covariance Stationarity)**.** *A time series $(X_t)_{t\in\mathbb{Z}}$ is covariance stationary if the first two moments are finite*

$$\begin{aligned} \mu(t) &= \mu, \qquad t \in \mathbb{Z} \\ \gamma(t,s) &= \gamma(t+k, s+k), \qquad t, s, k \in \mathbb{Z} \end{aligned}$$

[1] We treat the terms time series and process as synonyms.

and $\gamma(t, s)$ does not depend on t or s, but only on k.

In academic literature, a strictly stationary time series is also referred to as strong stationary and a covariance stationary time series is referred to as weakly stationary, second-order stationary, or wide-sense stationary.[2]
Stationarity is an elementary concept in time series analysis and particularly in time series analysis of autoregressive processes. There have been substantial approaches developed to test for stationarity. Most of the tests are hypothesis tests, where the null hypothesis states that stationarity is present in a time series against an alternative. Influential studies in this field include Dickey and Fuller (1979), Said and Dickey (1984), P. Phillips and Perron (1988) and Kwiatkowski, Phillips, Schmidt, and Shin (1992), among others.

White Noise Process Empirical data, in particular economic and financial empirical data, can be described by several artificial models. A part of the changes in the data can be explained by a corresponding model. The remainder is uncertain. A variety of different time series models introduce uncertainty with a so-called white noise process.

Definition 3.1.3 (White Noise)**.** *The time series $(X_t)_{t\in\mathbb{Z}}$ is a white noise process $(\varepsilon_t)_{t\in\mathbb{Z}}$ if it is covariance stationary without serial correlation,*

$$\varepsilon_t \sim i.i.d. WN(0, \sigma_t^2)$$

where i.i.d. means independent and identically distributed and WN represents a distribution function with mean zero and variance σ_t^2.

The conditional as well as the unconditional mean of a white noise process $(\varepsilon_t)_{t\in\mathbb{Z}}$ (hereafter denoted by ε_t) is zero. The process is conditional homoscedastic and is not restricted on Gaussian distributions. The distribution function WN can represent several forms of distribution functions.[3]
Several models are constructed with a white noise component which brings the

[2]For a detailed and augmented commented descriptions consider for instance, McNeil et al. (2005) or J. Cochrane (1997).
[3]In contrast to a white noise process, a strict white noise process has finite but not constant variance σ_t^2.

uncertainty in the model. The uncertainty is sometimes also called innovation or shocks.
In financial time series, the current state of a given situation often is influenced by the antecedent state or even by some of the last states. From a mathematical point of view this effect can be established in a process by a weighted factor of the last states. Since the process is related to itself, such processes are referred to as autoregressive processes (AR). An additional component in financial time series constitutes the effect of previous shocks or innovations. This element is constructed by a weighted average of previous innovations and hence referred to as the moving average part (MA) of the process.

Autoregressive Moving Average (ARMA) Process An ARMA processes is a combination of an autoregressive process (AR) and a moving average process (MA). It is a covariance stationary process where the innovation is generated by a white noise process.

Definition 3.1.4 (ARMA Process). *Let ε_t be a white noise. The process $(X_t)_{t\in\mathbb{Z}}$ is a zero-mean ARMA(p,q) process if it is covariance stationary and satisfy:*

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \qquad \forall t \in \mathbb{Z}$$

where (X_t) represents the ARMA process. ϕ_p is said to be the autoregressive parameter of the order p and θ_q is the moving average parameter of the order q.

The stationary behavior of the process is determined by the driving white noise. If the white noise is i.i.d. or strictly stationary, then the ARMA process exhibits the same properties.
This kind of time series can be divided in a deterministic and a stochastic part

$$X_t = m + \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \ldots + \alpha_i \varepsilon_{t-i} = m + \sum_{i=0}^{\infty} \alpha_i \varepsilon_{t-i}, \tag{3.1}$$

if the coefficients α_i satisfy that

$$\sum_{i=0}^{\infty} \mid \alpha_i \mid < \infty. \tag{3.2}$$

The coefficients α_i are the impulse response coefficients and the function $\{\alpha_i, i \geq 0\}$ is referred to as the transfer function or the impulse response function. This technical condition in equation (3.2) guarantees that $\mathbb{E}[|X_t|]$ exists and consequently that the infinite sum in equation (3.1) converges such that both $\sum_{i=0}^{\infty} |\alpha_i||\varepsilon_{t-i}|$ and $\sum_{i=0}^{\infty} \alpha_t \varepsilon_{t-i}$ are finite with probability one. This restriction is called the Wold theorem and ARMA processes restricted in this way are referred to as causal ARMA processes.[4]

First Order Autoregressive Process Consider for instance the example of the first-order autoregressive process

$$X_t = \rho X_{t-1} + \varepsilon_t.$$

By calculating successively recursive we obtain

$$\begin{aligned} X_t &= \rho^2 X_{t-2} + \rho\varepsilon_{t-1} + \varepsilon_t = \ldots \\ &= \rho^n X_{t-n} + \sum_{j=1}^{n-1} \rho^j \varepsilon_t. \end{aligned} \tag{3.3}$$

This process is causal if and only if $|\rho| < 1$.[5] This ensures that $\lim_{n\to\infty} \rho^j = 0$ such that the process in equation (3.3) can be represented by a MA(∞) process

$$X_t = \sum_{j=1}^{\infty} \rho^j \varepsilon_t. \tag{3.4}$$

The MA(∞) representation of the AR(1) process simplifies the computation of the moments like the mean and the variance. In addition this representation

[4] For more details, proofs and consequences consider for instance J. Cochrane (1997) or McNeil et al. (2005).

[5] The condition $|\rho| < 1$ causes that an AR(1) process can be represented by a MA(∞) process. Thus, this condition is known as an invertibility condition. A general ARMA(p,q) processes is causal if the two polynomials of the model parameters in the complex plane given by

$$\begin{aligned} \widetilde{\phi}(z) &= 1 - \phi_1 z - \ldots - \phi_p z^p \\ \widetilde{\rho}(z) &= 1 - \rho_1 z - \ldots - \rho_q z^q \end{aligned}$$

have no common roots and $\widetilde{\phi}(z)$ have no roots in the unit circle $|z| \leq 1$.

shows the impact of the marginal effects of previous shocks. The impact of the innovation occurring in time $t-j$, affects the today's value with ρ^j. Accordingly, the factor ρ^j is referred to as the multiplier or filter. In Figure 3.1 several functions of multiplier for different values for ρ^j are illustrated. Since the multiplier function is exponential the effect of previous shocks decreases disproportionately.

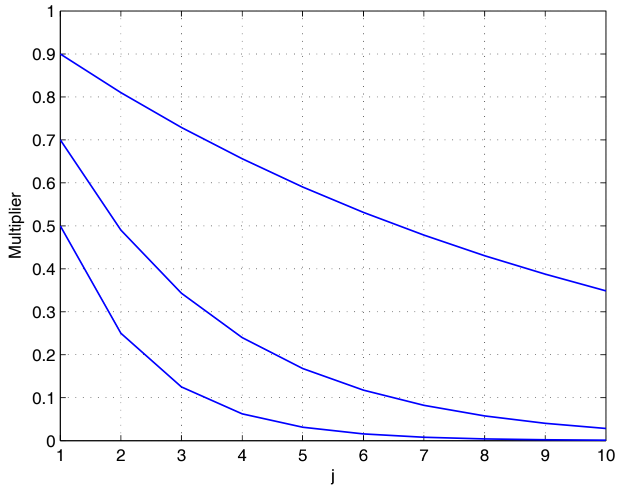


Figure 3.1: Multiplier Effect Figure 3.1 illustrates the the multiplier effect for different ρ. The smaller ρ is the faster converges it to zero. The picture shows the graph of $\rho = 0.9, 0.7$ and 0.5.

Analysis in the Time Domain Time series analysis usually starts with an examination of the set of serial correlations. In order to get a first impression about the order of a process, empirical estimates of serial correlations from real data are compared with their theoretical analogues. Studies of serial correlations is sometimes referred to as the analysis in the time domain.

Autocorrelation According to definition 3.1.2, the covariance between X_t and X_s depends only on the difference of $h = |t-s|$. This temporal separation (h) is known as lag. The autocovariance function (γ) can be rewritten with the lag function (h) as $\gamma(h) := \gamma(h, 0), \forall h \in \mathbb{Z}$. In addition, $h = 0$ implies that $\gamma(0) = var(X_t)$.

Definition 3.1.5 (Autocorrelation Function (ACF)). *The autocorrelation function (ACF) of a covariance-stationary time series* $(X_t)_{t\in\mathbb{Z}}$ *is defined as*

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{\gamma(h)}{var(X_t)}.$$

Accordingly, the ACF function with lag h describes the correlation between two values X_t and X_{t-h}, where the lag represents the shift of the time series to itself. Beside the autocorrelation function, a further analyzing tool to detect the order of an ARMA process is the partial autocorrelation function (PACF). In contrast to the ACF the PACF measures only the correlation between two fixed values, namely the value of the time series and the value at a given lag length, while the values between the two values are excluded. Consequently, it measures the linear relation between X_t and X_{t-h} by excluding the influence of the intermediate variables.

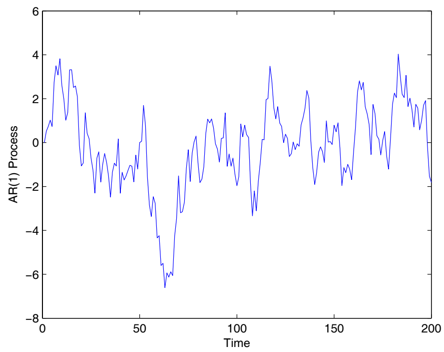
Definition 3.1.6 (Partial Autocorrelation Function PACF). *The partial autocorrelation at lag h is defined as the last component* ψ_h *in the matrix equation given by*

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_h \end{pmatrix} = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{h-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{h-2} \\ \vdots & \vdots & \cdots & \vdots \\ \gamma_{h-1} & \gamma_{h-2} & \cdots & \gamma_0 \end{pmatrix}^{-1} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_h \end{pmatrix}$$

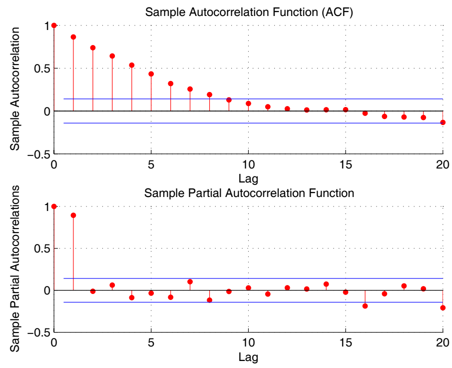
In order to determine the order of an ARMA process, ACF and PACF calculations with different lag length have to be considered. A graphical illustration of the values at different lag length is a correlogram. The x-axis refers to different lag length while the y-axis represents the corresponding ACF and PACF values respectively. The illustration contains a 95% significance band for the null hypothesis of a strict white noise process. Correlation estimates outside these bounds are considered as evidence against the null hypothesis that data are strict white noise; i.e. the data feature a kind of structural property.[6] A cut off at lag p in the partial autocorrelation indicates for a pure autoregressive behavior of order p, while a cut off at lag q is interpreted as a diagnostic for a moving

[6]For a more detailed description about correlograms, boundaries of the null hypothesis and why higher lagged correlations does not reject the null hypothesis consider McNeil et al. (2005).

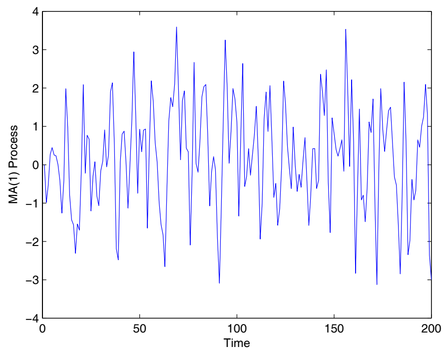
average process of order q. In Figure 3.2 and 3.3 several artificially generated processes with the corresponding correlograms are presented. In Figure 3.2 (a) a first order autoregressive process with $p = 0.9$ are simulated while Figure 3.2 (b) displays the function of a moving average process with $q = 0.9$. Figure 3.3 shows a autoregressive moving average (ARMA) process with $p = 0.9$ and $q = 0.9$. Simulation is done with 200 time steps.



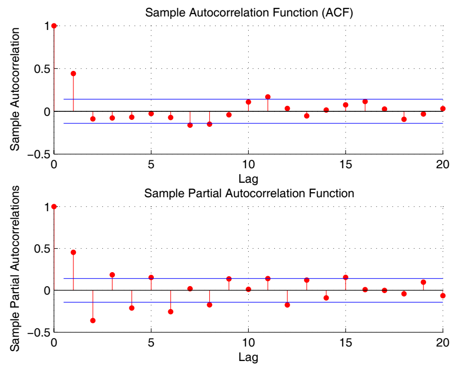
(a) AR(1) Process



(b) ACF and PCF

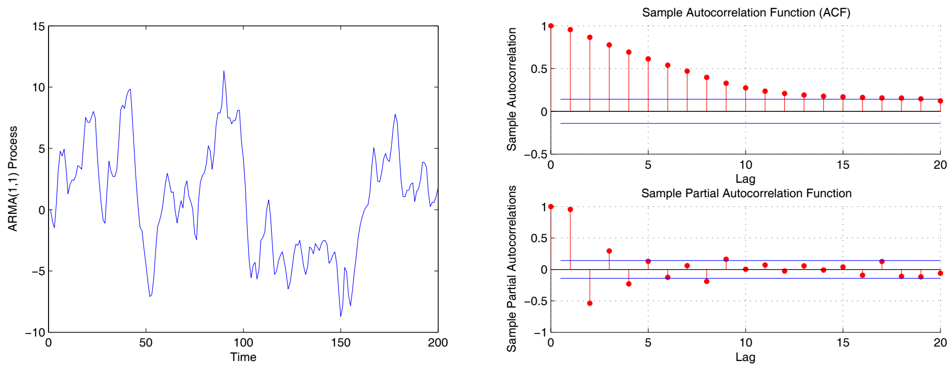


(c) MA(1) Process



(d) ACF and PCF

Figure 3.2: Auto Regressive and Moving Average Processes The Figure 3.2 displays an autoregressive and a moving average processes with their corresponding correlograms. The Figure (a) illustrates a pure AR(1) process while in Figure (c) a pure MA(1) process is shown. The first order AR process simulated with $p = 0.9$ and the MA process features $q = 0.9$. Simulation is done for 200 time steps.



(a) ARMA(1,1) Process (b) ACF and PCF

Figure 3.3: Auto Regressive Moving Average Process The Figure 3.3 shows a simulated ARMA(1,1) process with 200 time steps. The coefficients of the process are $p = 0.9$ and $q = 0.9$.

Beside visual investigations, formal analyses are typically used to determine the order of an ARMA process. There are several forms of analytical tests to assign the order of a process, such as hypothesis tests or the finite prediction error (FPE) criterion.[7] Nevertheless, these methods prefer mostly over-parameterized models. Albeit additional lags reduce the sum of squares of residuals, more parameters cause a reduction in forecasting performance of the fitted model. Additionally, the higher the order of the model is, the more imprecise the parameter estimation and the more difficult the parameter estimation process. The two most popular approaches accounting for that fact are the Akaike Information criterion (AIC) proposed by Akaike (1974) and the Schwarz-Bayesian information criterion (BIC) provided by Schwarz (1978). These model selection criteria methods can be viewed as measures of goodness-of-fit since they include a penalty function for every additional parameter estimated. The AIC is defined as

$$AIC = 2k - \log(L), \tag{3.5}$$

where k are the number of parameters and L is the likelihood function from the

[7] For more details consider for instance Enders (2004).

parameter estimation. The BIC is defined as

$$BIC = -2\log(L) + k\log(n), \tag{3.6}$$

where n is the number of observations.[8] The parameter combination minimizing the information criterion ascertains the order of the parameter set. For increasing sample sizes the BIC selects a more appropriate parameter set. Moreover, the BIC is asymptotically consistent while the AIC is biased toward selecting an over-parameterized model. In general, if the two approaches indicate the same number of orders, the set of parameters can be assumed to be most appropriate for the model.[9]

Alternatively, the analytical test of Ljung and Box (1978) can be used to determine the order of a process. Different from the AIC and the BIC approaches this method makes use of the autocorrelation function of the process. However, instead of testing randomness at each distinct lag, it tests the overall randomness based on a number of lags. More formally the Ljung-Box test statistic is given as

$$Q = n(n+2)\sum_{j=1}^{h}\frac{\rho^2(j)}{n-j}, \tag{3.7}$$

where n is the sample size, $\rho(h)$ is the autocorrelation function at lag j, and h is the number of lags being tested. The Ljung-Box test statistic is a Q-statistic testing the null hypothesis $\mathcal{H}_0$ of randomness of the data to the alternative that it is not. The hypothesis of randomness is rejected if the Ljung-Box test statistic exceeds the critical value of $\mathcal{X}^2_{h;1-\alpha}$ distributed function at a given significance level α with h degrees of freedom. In this case at least one value of the autocorrelation function is statistically significant different from zero. The Ljung-Box test is commonly used to reveal model misspecification and is therefore applied to the residuals of an estimated model.

[8] In time series literature, different authors report the AIC and the BIC in various ways. Nevertheless, the selection approach stays the same.

[9] However, since the BIC chooses the parsimonious model an additional test for the residuals is necessary. Since the AIC prefers the over-parameterized model, the t-statistic of all estimated coefficients has to be statistically significant at the given significance level.

3.2 Conditional Heteroscedastic Models

The innovation process modeled in ARMA processes have basically two drawbacks. On the one hand, volatility is held constant as time proceeds. On the other hand, the distribution function of the innovations captures no extreme events, i.e. the distribution function is not leptokurtosic. Generally, empirical findings report that financial time series typically feature leptokurtosis and have stochastic volatility. Moreover volatilities are not only stochastic but also exhibit forms of clustering effects, where periods of high volatilities are accompanied by high volatilities and periods of low volatilities are followed by low volatilities.
In contrast to models with constant volatility processes with stochastic volatility respectively, conditional heteroscedasticity induces leptokurtosis. Autoregressive Conditional Heteroscedastic (ARCH) models address to this attribute and offer possibilities to modeling conditional variance. With this kind of model at least three characteristics of financial time series can be replicated. First, it allows to simultaneously model the mean and the variance. Second, implicated by the conditional variance, the process is able to generate volatility clustering. And last, the unconditional distribution exhibits fat tails such that the probability of extreme events are higher than implicated by a normal distributed function.
The first model offering these features is introduced by Engle (1982). In this model the conditional variance is modeled by an AR(p) process using the squares of the estimated residuals.

Definition 3.2.1 (Autoregressive Conditional Heteroscedastic Model (ARCH)). *Let $(Z_t)_{t\in\mathbb{Z}}$ be a strict white noise with N(0,1). A process $(X_t)_{t\in\mathbb{Z}}$ is an autoregressive conditional heteroscedastic ARCH(p) process of order p if it is strictly stationary and it holds that*

$$X_t = \sigma_t Z_t, \tag{3.8}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2, \tag{3.9}$$

where $\alpha_0 > 0$ and $\alpha_i \geq 0$ for all $i = 1, \ldots, p$.

In order to avoid that the conditional variance becomes negative, the coefficients α_0 and α_i have to be restricted such that both are strictly positive. Additionally,

the restriction of $\sum_{i=1}^{p} \alpha_i \leq 1$ ensures covariance stationary of the process.[10]
The model is conditional heteroscedastic, since the conditional volatility changes continually, and it is autoregressive, since it depends on several previous states of the process. Because the volatility structure is based on an AR(p) model it features also similar characteristics. The autoregressive dependency of previous states causes that large values for the volatility tend to follow large values and small values tend to follow small values. This persistence produces volatility clustering effects.
Bollerslev (1986) extends the model of Engle (1982) by modeling the volatility structure not only by an AR(p) process but by an ARMA(p,q) process. The model is formulated as

Definition 3.2.2 (Generalized Autoregressive Heteroscedastic Model (GARCH)). *Let $(Z_t)_{t\in\mathbb{Z}}$ be a strict white noise with N(0,1). A process $(X_t)_{t\in\mathbb{Z}}$ is an generalized autoregressive conditional heteroscedastic GARCH(p,q) process of order p and q if it is strictly stationary and it holds that*

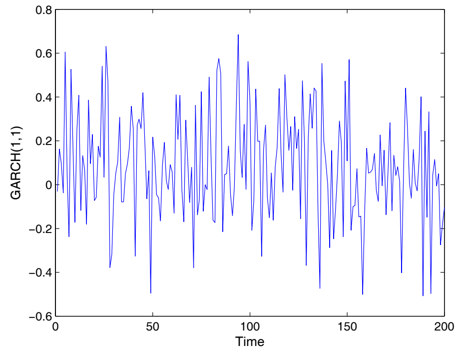
$$X_t = \sigma_t Z_t, \tag{3.10}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \tag{3.11}$$

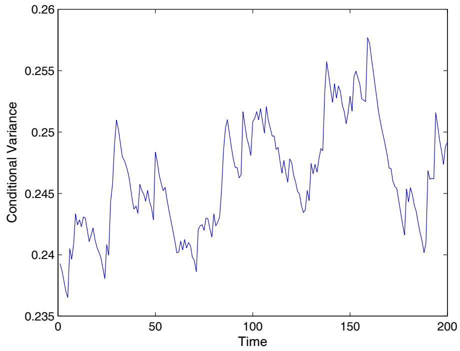
where $\alpha_0 > 0$ and $\alpha_i \geq 0$ for all $i = 1, \ldots, p$ and $\beta_i \geq 0$ for all $i = 1, \ldots, q$.

In order to ensure covariance stationary of the process, it is necessary to restrict α_i and β_i to $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} < 1$.[11]
This model is the general form of the previous ARCH model by meaning that the squared volatility σ_t^2 is additionally allowed to depend on previous squared volatilities. While the ARCH model effects persistence in the volatility structure the additional component of the GARCH model causes furthermore shocks. Overall, this model allows a more appropriate mapping of financial time series, not least because it is possible to enhance the probability of events at the tails of the distribution.

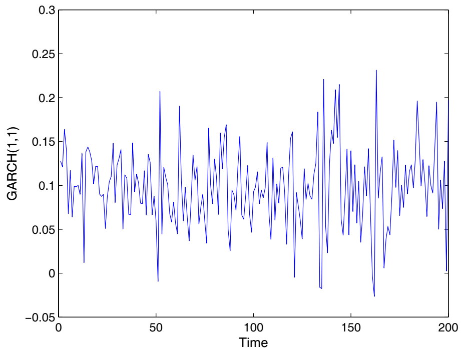
[10] For a detailed derivation of the mathematical conditions consider for instance, Hamilton (1994) or McNeil et al. (2005).
[11] See for instance Leippold (2004).



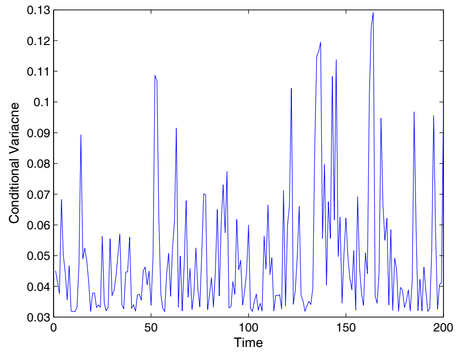
(a) GARCH(1,1) Process High Persistence



(b) Conditional Variance High Persistence



(c) GARCH(1,1) Process Low Persistence

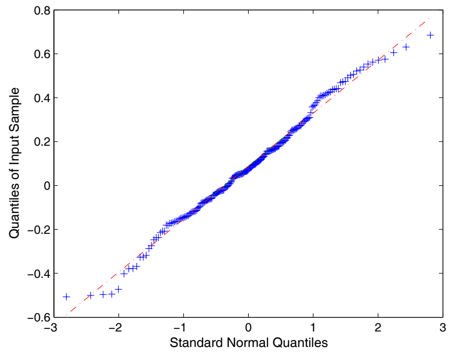


(d) Conditional Variance Low Persistence

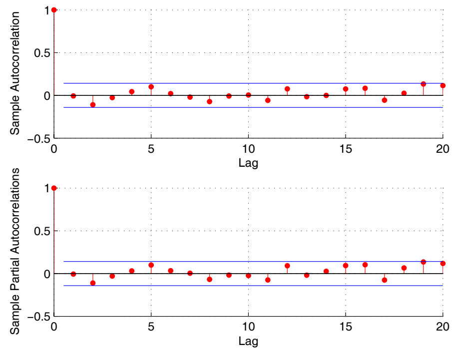
Figure 3.4: Generalized Auto Regressive Moving Average Process Figures 3.4 (a) and (c) show simulated GARCH(1,1) processes each with 200 time steps. Figures 3.4 (b) and (d) depict the corresponding conditional variances. The coefficients of the process in Figures 3.4 (a) and (b) are $p = 0.975$ and $q = 0.01$ while the constant is $\alpha_0 = 0.001$. The coefficients of the second simulation are $p = 0.01$, $q = 0.9$ and $\alpha_0 = 0.001$.

Figure 3.4 shows two trajectories of simulated GARCH(1,1) processes with a graph of the corresponding conditional variances. The simulation displayed in Figure 3.4 (a) is carried out with a parameter set of $p = 0.975$, $q = 0.01$ and $\alpha_0 = 0.001$ while the parameter set of the process in Figure 3.4 (c) is given by $p = 0.01$, $q = 0.9$ and $\alpha_0 = 0.001$. The conditional variance in Figure 3.4 (b) highlights high persistence in volatility while Figure 3.4 (d) emphasize high

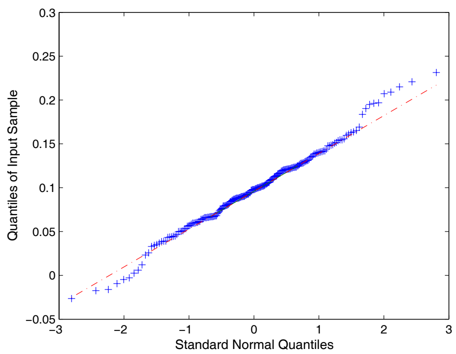
sequences of volatility shocks.
Figure 3.5 depicts QQ-plots, ACF and PCF estimations of the corresponding processes. Albeit the driving innovation process of the model is strict white noise, the simulated time series has fatter tails than a white noise implies, which is graphically shown in the QQ-plots.



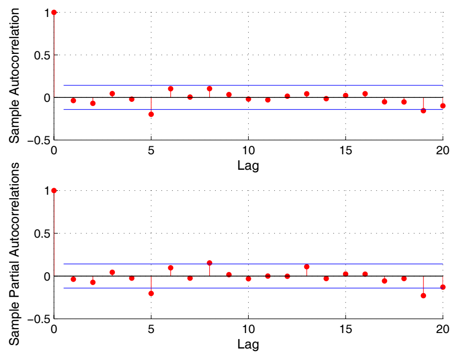
(a) QQ-Plot GARCH(1,1) Process



(b) ACF and PCF



(c) QQ-Plot GARCH(1,1) Process



(d) ACF and PCF

Figure 3.5: QQ-Plot GARCH(1,1) Process and ACF and PCF Plots
Figures 3.5 (a) and (c) display the corresponding QQ-plots of the GARCH(1,1) process from Figure 3.4. Figures 3.5 (b) and (d) shows the ACF and PCF function of the process.

ARCH respectively GARCH processes describe solely the variance structure of a process. However, the underlying process can be formulated in different ways. Basically, replacing the innovation process of an ARMA process with a GARCH volatility structure delivers a family of ARMA models with GARCH errors that combines the features of both model classes.

Definition 3.2.3 (ARMA processes with GARCH variance structure)**.** *Let $(Z_t)_{t\in\mathbb{Z}}$ be a strict white noise with N(0,1). A process $(X_t)_{t\in\mathbb{Z}}$ is an ARMA(p,q) process of autoregressive order p and moving average order q with GARCH(a,b) volatility structure of persistence a and shocks b if it is covariance stationary and it holds that*

$$\begin{aligned} X_t &= \mu_t + \sigma_t Z_t, & (3.12)\\ \mu_t &= \mu + \sum_{i=1}^{p} \phi_i (X_{t-i} - \mu) + \sum_{j=1}^{q} \theta_j (X_{t-j} - \mu_{t-j}), & (3.13)\\ \sigma_t^2 &= \alpha_0 + \sum_{k=1}^{a} \alpha_k (X_{t-k} - \mu_{t-k})^2 + \sum_{l=1}^{b} \beta_l \sigma_{t-l}^2, & (3.14) \end{aligned}$$

where $\alpha_0 > 0$ and $\alpha_k \geq 0$ for all $k = 1, \ldots, a$ and $\beta_l \geq 0$ for all $l = 1, \ldots, b$ and $\sum_{k=1}^{a} \alpha_k + \sum_{l=1}^{b} \beta_k < 1$. In addition Z_t is independent of $(X_s)_{s<t}$ for all t.

The coefficients of the ARMA process extended with stochastic volatility are also subjected to the restrictions given in the previous section for general ARMA(p,q) processes.[12]
This family of processes allows to map time series respective to financial time series in a more appropriate way than with constant volatilities. It is possible to describe the underlying process with a theoretical economic model while the stochastic volatility accounts for inconsistency in the innovation. With the conditional variance the structure of shocks and its persistence can be visualized. This extension may contribute additional information to the underlying theory and improve the explanatory power of the economic model.

Analysis in the Time Domain The approach to determine the orders of the parameters is similar to the method for ordinary ARMA models. In general, the first step is an analysis of correlograms. Following on that first impression analytical approaches like for instance the AIC information criterion or the BIC information criterion can be applied to ascertain the orders of the process.

[12] For more details consider for instance McNeil et al. (2005).

Having found a set of orders and after parameter estimation, a test for model misspecification, like the Ljung-Box test can be applied.

This chapter is by no means complete. Basically, it serves to catch an overview and intuition about the models applied in later chapters. From this point of view the chapter is kept in a very general way. However, in recent literature in particular, the GARCH approach has been extended in several forms. GARCH models with leverage effects are constructed, a threshold is implied or exponential GARCH, EGARCH structures are developed for instance. In order to gain advanced insights in time series analysis and in particular the different parameter estimation methods consider for instance McNeil et al. (2005), Enders (2004) or Hamilton (1994) among many others.

Chapter 4

Order Book Dynamics and Asset Liquidity

This chapter deals with the dynamical behavior of the limit order book and its importance to asset liquidity. In general, measurement approaches to detect liquidity omit this source of influence. However, measurement concepts based on a given situation in a specific time interval may not reflect the entire liquidity situation adequately. Starting from this point of view, we introduce in this chapter an alternative measurement concept based on the dynamical behavior of the limit order book.
The chapter is structured in three sections. In the first section, we highlight the dynamic of the limit order book. Starting with the order flow around the best quotes we broaden our considerations to the entire limit order book and the time dependent changes of it. In the following section we introduce a time dependent measurement concept. The chapter closes with the third section, where we specify the model, the parameters and their economic interpretations.

4.1 Limit Order Book Dynamics

Different definitions and various measurement concepts for liquidity have been developed and several dimensions of liquidity are identified. For all these dimensions a vast amount of different measurement approaches and concepts have been found and discovered. But the previous considerations and treatments of the liquidity question have been based on a static point of view. Static in the sense that the measures capture a specific situation in time. The dynamical behavior of the measurement approaches is in general not taken into account.

Order Flow around the Best Quotes From an auction theoretic perspective the dynamic of the limit order book and in particular of the bid-ask spread is a direct consequence of the order flow. Different orders submitted to the market, are settled against or collected in the limit order book. Hence, this order submission process generates the dynamical behavior of the book. Diverse order types shape the face of the limit order book in various ways, whereas two superior attributes of an order can be identified. While a limit order adjacent to the best quotes has little impact on the bid-ask spread, an order at or within the bid-ask spread has a larger effect. Beside the order price, the order size influences the dynamic of the limit order book additionally. Large orders, large compared with the currently available quoted volume, have a larger impact than small orders. In this context, order types can be differentiated according to their aggressiveness. An order with a larger order size than the currently available best offer placed within the bid-ask spread or at the best quote is referred to as an aggressive order. Orders placed deeper in the limit order book are said to be less aggressive.[1] Since in a more aggressive environment aggressive orders can be executed to more favorable prices, order aggressiveness is a possible indication for liquidity.
Moreover, the order placement procedure or the order flow is first, more frequent around the best quotes. Second, the order flow induces a mean reverting structural character of liquidity measured by the bid-ask spread or measures based on the bid-ask mechanism. This effect is well documented for instance by Coppejans, Domowitz, and Madhavan (2004), Wyss (2004), Handa and Schwartz (1996) or Biais et al. (1995), among many others. To illustrate this structural nature we adopt the arguments presented in Biais et al. (1995). Trades occur

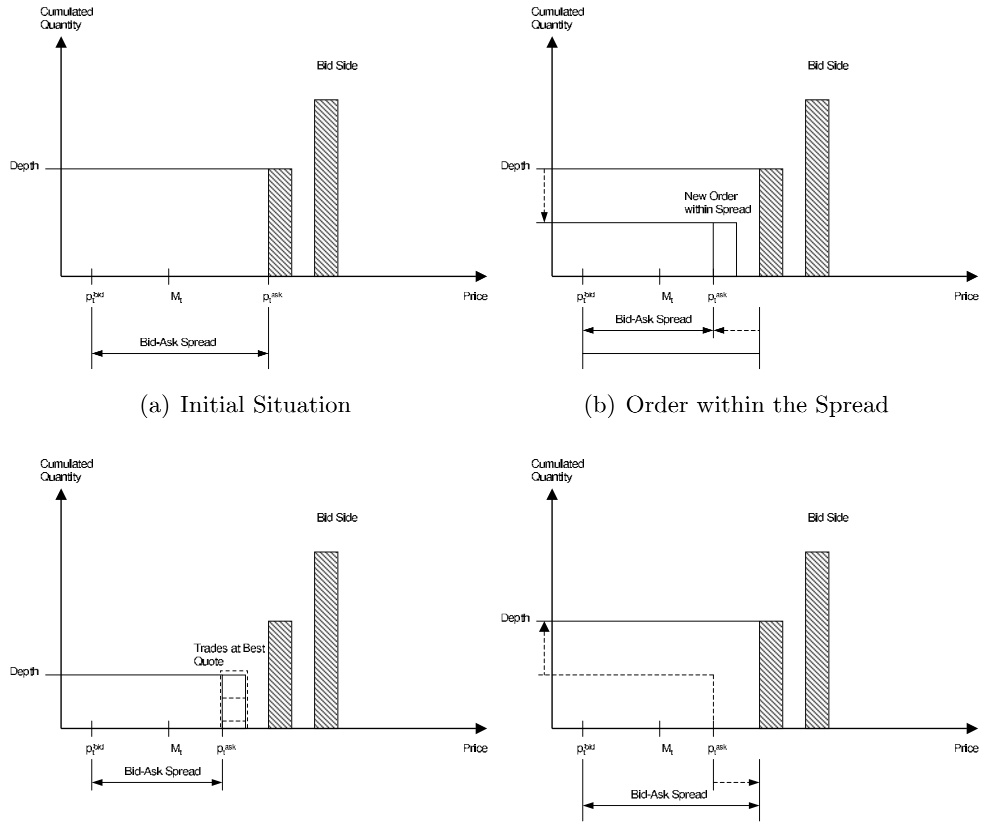
[1]Harris and Hasbrouck (1996) or Biais, Hillion, and Spatt (1995), among others, separate the order placement according to its aggressiveness. Biais et al. (1995), for instance, differentiate the orders according to several events. They show evidence, amongst other things, that orders placed around the best quotes are more frequent than a placement deeper in the limit order book. In particular, they show that order placement decreases monotonically by moving away from the quotes.

relatively more often when the bid-ask spread is tight. In contrast when the bid-ask spread is wide, orders are placed more frequently within the bid-ask spread. Furthermore, new orders are placed within the spread when depth around the spread is large and are placed at the quotes when depth is thin.[2] Therefore the mean reverting mechanism works as follows. Consider a starting situation where depth is relatively high and the bid-ask spread is wide. Since the new order is placed within the bid-ask spread, the new placed order becomes the best quoted offer. Because the bid-ask spread is generally calculated as the difference between the two best quoted offers, the new arrived order tightens the spread. Moreover, for the reason that the older best quoted offer is an aggregation of several unexecuted or uncanceled offers at this quote, the volume at the new best offer is in general smaller. Hence, depth is reduced around the bid-ask spread and spread is tighter. Now, the limit order book has thin depth and tight spread. In this situation, agents hit the quote and trading takes place at the best offer. The last trade clearing the best offer widens the spread and the limit order book appears as in the initial position.[3]
Figure 4.1 illustrates this mechanism of the ask side schematically. In each Figure, the horizontal axis represents the prices and the vertical axis the cumulative volumes, respectively, all the offers prepared to trade at this particular price. p_t^{bid} and p_t^{ask} are the corresponding best quoted offers while M_t represents the arithmetic mean between the bid-ask spread. A bar illustrates the cumulative volume of the orders at the corresponding price level. Figure 4.1 (a) displays the initial situation of the limit order book. Depth is relatively high and the bid-ask spread is wide. Figure 4.1 (b) shows an arriving new order which tightens the bid-ask spread and lowers the depth. The absorption of the order is shown in Figure 4.1 (c) while the rebounding is illustrated in Figure 4.1 (d).

From an economic point of view a couple of hypotheses can be assumed to explain this behavior. Orders are collected in the limit order book according to the price-time priority rule. Since limit orders with equal price restrictions are sorted according to their submission time, new limit orders restricted to

[2]See for instance the empirical results documented by Biais et al. (1995) or our findings reported in section 5.

[3]This mechanism is not constricted to one side of the book. Moreover, this mechanism appears on both sides, may be at the same time, and may effect price movements.



(a) Initial Situation

(b) Order within the Spread

(c) Trades Hitting the Quote

(d) Rebounding of the Spread

Figure 4.1: Mean Reverting Mechanism The Figure 4.1 displays the order flow mean reverting mechanism for the ask side schematically. In each Figure, the horizontal axis represents the prices and the vertical axis the cumulative volumes, respectively, all offers prepared to trade at the given price. p_t^{bid} and p_t^{ask} are the corresponding best quoted offers while M_t represents the arithmetic mean between the bid-ask spread. A bar illustrates the cumulative volume of the orders at the corresponding price level. Figure 4.2 (a) shows a given initial position of the limit order book with assumed relatively high depth and wide bid-ask spread. An order placement within the bid-ask spread is illustrated in Figure 4.2 (b). The Figure 4.2 (c) shows trading at the quotes until the offers at the quotes are cleared. Figure 4.2 (d) displays the reverting of the bid-ask spread to its previous location.

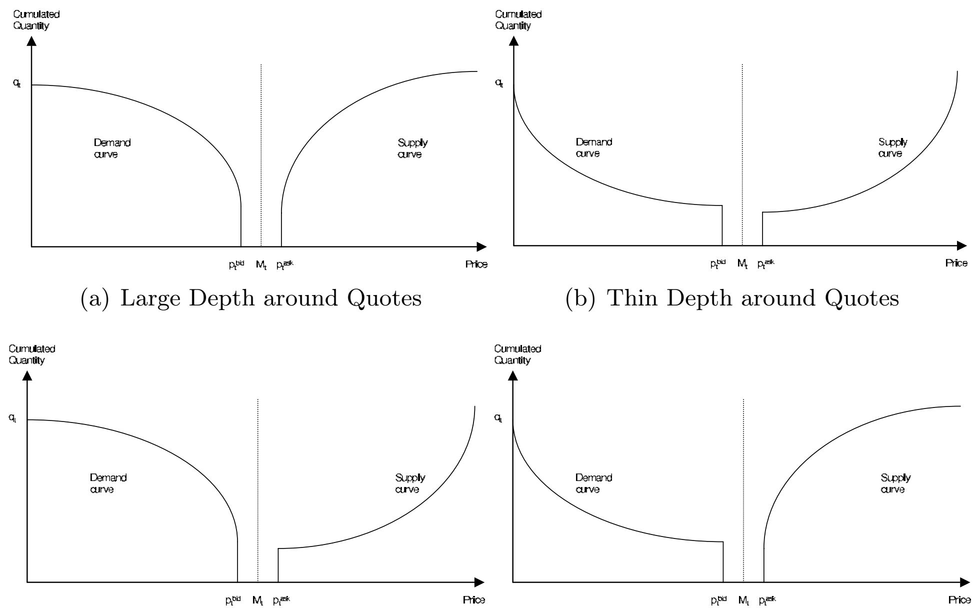
the same price are unlikely to be executed, in particular when depth is large at this quote. Such an order opens additional risks to the investor. First, it enhance the waiting time until execution. Associated with transaction costs, Lo, MacKinlay, and Zhang (2002) show evidence that the cost of submitting limit orders increases with the expected time of execution. Second, it bears the risk not to be executed at all. Therefore the investor is better off by undercutting the best offer. As Biais et al. (1995) point out, the difference to undercut the current quote could reflect the price to obtain time priority. Moreover, Harris (1994) notes that there is first mover advantage in supplying liquidity if market enforces time precedence, since it protects traders who exposes their quotations or limit orders to the market. The reason for this fact is that, by revealing their orders, these traders disclose their information to all market participants and take the risk that other investors may act based on that information to their disadvantages. Another, more practical oriented argument is suggested by Biais et al. (1995).[4] Consider for instance a noninformed agent splitting her order to reduce market impact. She first buys a given limited amount of the desired asset. Then, she waits hoping that additional liquidity will be provided on the corresponding side of the order book. Finally, the agent submits the next order to the market and hits the quotes again. This means that the order first hits the best quote in the book, ends up in execution, consuming liquidity and widening the bid-ask spread. Then liquidity will be supplied and tightened the bid-ask spread.

Order Flow beyond the Best Quotes Albeit the order flow is mainly concentrated around the bid-ask spread, the quotes beyond the best quotes provide additional information to the value finding process of a security. Pascual and Veredas (2006) evaluate the informational content of the limit order book in order to the explanatory power of long run volatility. By separating liquidity-driven from information-driven volatility using a state space co-integrated model for the bid and ask quote, they find at least that the book beyond the best quotes contribute additional explanatory power to the best quotes. Furthermore, Cao, Hansch, and Wang (2004) for instance show empirical evidence that order flow beyond the best quotes adds approximately 30% of the information to short-term future returns. Moreover they show that traders use the available information

[4] Albeit Biais et al. (1995) proposed this argument to explain the time interval between to similar orders, this point can be used in a broader sense to explain investors behavior.

on the state of the book to submit orders strategically to the market.

For trading activity beyond the best quotes we offer a possible explanation. An investor or a financial intermediary acting for an investor has a distinct view about the value of the considered security. This view may arise from strategic trading considerations. Or, the investor owns private information. By submitting an order to the market the investor discloses his information and contributes explanatory power to the best quotes, i.e. to the informational content of the limit order book.



(a) Large Depth around Quotes

(b) Thin Depth around Quotes

(c) Demand Pressure

(d) Supply Pressure

Figure 4.2: Limit Order Book Dynamic The Figure 4.2 shows several situations of the limit order book schematically. Figure 4.2 (a) illustrates large depth around the best quoted offers while Figure 4.2 (b) represents low depth on both sides. Figure 4.2 (c) and (d) displays asymmetric depth between the ask and the bid side.

A possibility to bring out the dynamical behavior of the limit order book is to illustrate the cumulative volumes with the corresponding prices. A graphical interpretation of this relation is schematically shown in Figure 4.2. Each figure is a frozen in time window of the limit order book. The horizontal axis corresponds to the prices while the vertical axis displays the cumulative volumes. In each Figure, the bid offers are represented on the left hand side and on the right hand side the ask offers are shown. The corresponding best quotes (highest bid price p_t^{bid} respective to the lowest ask price p_t^{ask}) are given in the middle of the graphic and represents at the same time the bid-ask spread. M_t is the arithmetic mean of the bid-ask spread.

Time Dependent Limit Order Book However, as time proceeds the limit order book affected by the order flow appears in one of these four situations. In each time step the shape of the limit order book looks different. Already the mean-reverting mechanism around the best quotes shapes the slope of the book continuously. In order to get an impression about this dynamical behavior of the limit order book, consider Figure 4.3. The x-axis shows the prices. The y-axis displays the time where the time grid is one hour. The third dimension, the z-axis, illustrates the cumulative volumes at the corresponding prices. Considering the illustration from the right front side, the shape of the limit order book, as presented in one of the pictures in Figure 4.2 appears for a given snap-shot in time. The shaded valley between the best quotes, shows the time dependent changes of the mean reverting bid-ask spread process.

4.2 Time-Dependent Liquidity Measure

Liquidity has multiple dimensions which consist of key elements of volume, time and transaction cost. According to B. R. Porter (2003) an ideal measure of liquidity should therefore incorporate elements of depth, breadth and resiliency. In general we share this opinion and propose the construction of a new liquidity measure concept. As mentioned above, the bid-ask spread features a mean reverting behavior. This mean reverting behavior is driven by the order flow and particularly by the traders of the market. However, amongst other things this mean reverting characteristic serves as a basis of our measurement concept.

Consider a market situation consisting of one buyer and one seller. Both hand in an offer either to buy or to sell the asset. When the offers matches, trade trade immediately takes place to the first offered price. Otherwise the offers are collected in the limit order book and the agents disclose the information

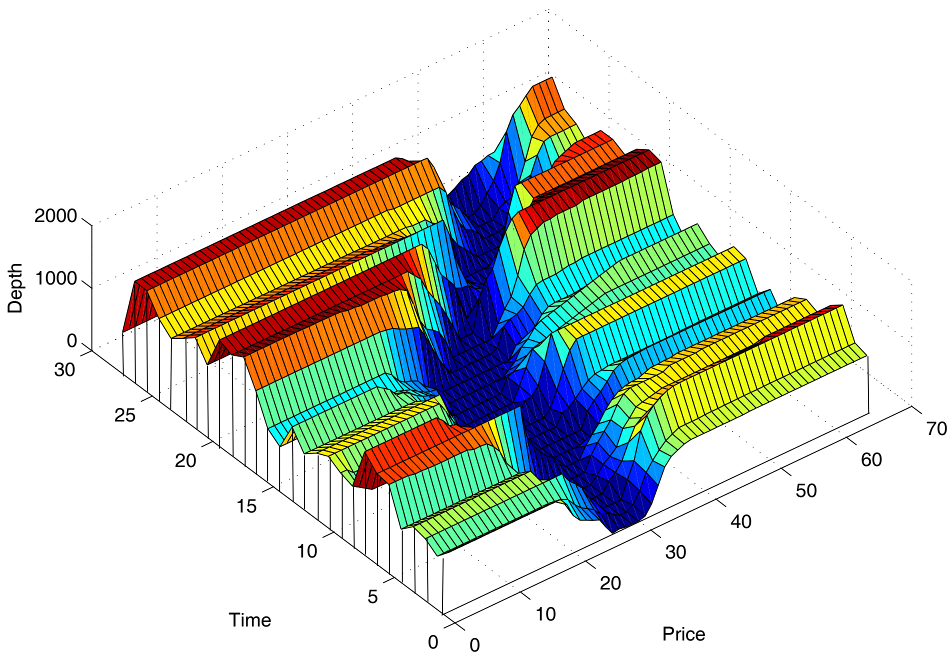


Figure 4.3: Three Dimensional Limit Order Book The Figure 4.3 shows the limit order book as time precedes. On the x-axis the prices are illustrated. On the y-axis the time is shown where the time grid is one hour. The z-axis displays the cumulative volumes at the corresponding price level. From the right front side, the shape of the limit order book appears for a given snap-shot.

connected with the offer. In order to attain matching, the agents can adjust their offers as time proceeds. However, if there are more than two actors in the market, the agents standing outside observe the limit order book. Based on the new acquired information they can submit a competitive order.[5] The new agents entering the market either tighten the bid-ask spread by submitting a competitive offer, or widen the bid-ask spread by hitting the quote.[6] From an

[5]In a later section we show that agents tend to use previous order characteristics as benchmark for their own order. We find that similar endowed orders tend to follow each other. Biais et al. (1995) report results alike to our findings.

[6]See the mechanism explained in more detail in section 4.1.

auction theoretic and competitive point of view, the more agents in the market the faster new orders are submitted. Which in turn can be reflected by more movements, respectively a faster converging of the bid-ask spread to its mean. Moreover the more market participants in the market the tighter the mean of the spread. However, a tighter bid-ask spread reflects a more liquid market. From this perspective, we suggest that the converging rate of the mean reverting process is a liquidity measure.

However, as pointed out above, the limit order book owns informational content even beyond the best quotes. From this point of view, we do not rely on a measurement approach solely based on the bid-ask spread. Moreover we construct a measure based on the bid-ask spread mechanism but extended by some additional dimensions. In order to incorporate the dimension of resilience and depth of the book, we propose to measure the limit order book for a given time interval by a density measure approach. The density measure D_t is the difference between the natural logarithm of the density at the ask and the bid side. In order to compute the density at each side the sum of the cumulated volumes $q_{(i,t)}$ weighted with the difference between the mid price P_t^m and the corresponding price $p_{t,t}$ is calculated.

$$D_t = \log\left[\left(\sum_{a=1}^{A} \frac{q_{a,t}^{ask}}{p_{a,t}^{ask} - P_t^m}\right)^{-1}\right] - \log\left[\left(\sum_{b=1}^{B} \frac{q_{b,t}^{bid}}{P_t^m - q_{b,t}^{bid}}\right)^{-1}\right], \tag{4.1}$$

where the integers a respectively b represents the corresponding price level of the offers. Since all the available volumes on each side are take into account, the measure is constructed with respect to depth and to the informational content beyond the quotes of the book. Additionally, since the cumulated volumes are weighted with the difference between the mid price P_t^m and the corresponding price level $p_{i,t}$ also resiliency is considered. The limit order book contains all not executed orders. From this fact $p_{a,t}^{ask} > P_t^M$, respectively $p_{a,t}^{bid} < P_t^M$ is true because otherwise orders immediately lead to execution. Then orders are traded and disappear from the limit order book.
However, since this measure is an extended version of the bid-ask spread it behaves like the bid-ask spread and tends to converge to a mean. Therefore, we

propose this measure as time series to the estimation of the converging rate, i.e. the dynamical liquidity measure.

4.3 Model Approach

In order to obtain the converging rate of the density measure D_t, we apply a first order autoregressive model X_t extended by a first order generalized conditional heteroscedastic volatility structure to capture the time series generated by D_t. The model is given by

$$\begin{aligned} X_t &= \mu_t + \sigma_t Z_t, & (4.2)\\ \mu_t &= \mu + \phi_1(X_{t-i} - \mu), & (4.3)\\ \sigma_t^2 &= \alpha_0 + \alpha_1(X_{t-1} - \mu_{t-1})^2 + \beta_1\sigma_{t-1}^2, & (4.4) \end{aligned}$$

where ϕ_1 is a member of $[-1, 1]$, $\alpha_0 > 0$ and $\alpha_1 \geq 0$ and $\beta_1 \geq 0$ and $\alpha_1 + \beta_1 < 1$. In addition Z_t is independent of $(X_s)_{s<t}$ for all t.
Since the autoregressive parameter is bounded by $|\phi_1| < 1$ the process is stationary.[7] Stationarity is equivalent to mean reversion. From this point of view a first order autoregressive model can be interpreted as a mean reversion model. However, a faster reverting to a mean is reflected by a smaller dependency of the previous state of the process. The coefficient ϕ_1 describes how strong a previous state is present in the current state. A value of ϕ_1 close to zero shows a small presence of the previous state and indicates therefore a fast converging rate while a value of ϕ_1 closer to one constitutes higher dependency on the previous state and represents thus a lower speed of reversion. Hence ϕ_1 can be considered as an inverse converging rate.

Asset Liquidity Measure Liquidity can be measured for several time intervals. In each snapshot of the limit order book all different measurement approaches can be applied and the current liquidity situation can be determined. However, as mentioned in the previous section our focus is not on a measure acquired for a given time interval. Moreover, we propose that the changes be-

[7] Otherwise the process is a random walk, see for instance Dickey and Fuller (1979) among many others.

tween the sequence of time intervals reflect asset liquidity. We suggest, based on the mean reverting behavior of the bid-ask spread, liquidity can be measured by the converging rate of the mean reverting process. From an economic point of view matching of the orders take place faster and to more favorable transaction costs when bid-ask spread is small. The mean of the bid-ask spread, respectively of the density measure, is zero. Thus, the faster the process converges to this mean, the more liquid the considered asset. According to this, we suppose that a value of ϕ_1 close to zero denotes a liquid asset and conversely a value of $|\phi_1|$ close to one describes an illiquid asset.
The GARCH volatility structure reveals additional characteristics of the time series. In contrast to a model with constant volatility the extended model with stochastic volatility is able to highlight the arrival of sudden extraordinary events. Such events are referred to as shocks. The β_1 parameter illustrates the frequency of sudden shocks, while the α_1 parameter describes the persistence of them. To both applies that the larger the value the higher the frequency or the longer the persistence of the shocks. From an economic point of view a shock in the time series equals either a sudden deficit or a sudden surplus of orders. In general such events are associated with large price movements or long waiting times till execution. In that sense, most liquidity is provided when α_1 and β_1 tend to zero.[8]
According to those arguments the most liquid market environment is given by tending all parameters to zero. Hence, we propose that

$$\lambda = \frac{1}{2}\left(|\phi_1| + \alpha_1 + \beta_1\right), \tag{4.5}$$

where λ is the dynamic liquidity measure. Since the parameters are bounded by $|\phi_1| < 1$ and $\alpha_1 + \beta_1 < 1$, while $\alpha_1 > 0$ and $\beta_1 > 0$ due to positive volatility, the dynamic liquidity measure λ is bounded by $0 < \lambda < 1$. Based on the construction this liquidity measure allows for an estimate of the liquidity attribute of a particular asset. Moreover, since the measure is independently of dimensions different assets can be compared with each other and a liquidity ranking can be established.

[8]Since α_0 only measures the scale of the volatility we omit this parameter in the liquidity measure.

Chapter 5

Data and Methodology

The purpose of measuring asset liquidity is to provide additional information to investors to advance asset management and in particular to improve asset selection and active trading. In this chapter we present an analysis of the entire limit order book in the context of asset liquidity. Several authors recently evidenced that, compared with a truncated limit order book, an entire limit order book contributes additional information to market development, market behavior and price discovery process. In accordance to these findings we provide an alternative measurement approach to capture the movements of the entire limit order book, the density approach. Additionally we introduce a new measurement concept by defining the dynamic of measurement changes as the liquidity measure. Instead of calculating an average measure, such as for instance the bid-ask spread, we observe the changes of such a measure as time proceeds. The basis of this concept builds on the density measure, while the liquidity measure is based on the movement of this measure.

The chapter is structured in five sections. The first section introduces the market microstructure of the considered market and presents the available data set. We show the reconstruction of the limit order book. The next section deals with the dynamic of the limit order book. In the subsequent section we concentrate on the shape of the book. We analyze the liquidity costs, the shape of the book according to Naes and Skjeltorp (2006) and in particular the order placement procedure according to Biais et al. (1995). We prove that there are no random effects in the data even if we refine the order placement procedure increasingly

and give several economic interpretations of this behavior. This section serves to a theoretical and empirical background for the subsequent section. The next section introduces the density measure and provides an extensive technical analysis of the time series. In addition we apply the same analyses as in the previous section to connect the measure with economical interpretations. The section closes providing the parameter estimations of time series which gather the changes of the density measure as time proceeds. In the following section, the asset liquidity based on the findings in the previous section are represented. In the last section, the liquidity premium based on the liquidity ranking from the previous section according to Fama and French (1992) is computed.
Calculations are performed with Matlab. However, to verify the results, all calculations and estimations are additionally computed with S-Plus, R, Excel and SPSS. The results in this thesis can be verified except for rounding errors or errors caused by the converging of the numerical optimizer of the corresponding software.

5.1 Order Book Data

The market At the Swiss Exchange SWX several kinds of securities are traded. The exchange's assets spectrum runs from equities and bonds to derivative instruments or exchange traded funds (ETF). The market microstructure for the individual assets is different. It exist mainly in two market places, the Swiss Exchange SWX and a subgroup of it, the Virt-X. Swiss equities are traded on both market places. In terms of market capitalization, small and middle capitalized equities are assigned to the Swiss Exchange SWX. Large capitalized equities, or so-called blue chips, are traded at Virt-X.[1]
The Virt-X is organized as a double auction market. During a regular trading day, which starts at 9.00 o'clock and ends at 17.20 o'clock, equity trading takes place continuously and fully automated. An investor can choose between on order book trading and off order book trading. Orders submitted to the on order book trading are routed directly to the trading system and are forwarded to the central order book (limit order book). Passing a control and security system, the

[1]Beside Swiss Blue Chips also derivative Instruments, Pan-European equities and ETF's are traded at Virt-X.

order will be executed according to the matching rules.[2] Alternatively, orders assigned to the off order book trading are executed between two market participants. These trades will also be reported at Virt-X. Thus the market and the order execution process is fully computerized and no designated market maker exists. This part of the market is organized as a pure order driven market.
The exchange supports various order forms. Orders submitted to on order book trading can feature forms of restriction. Orders restricted only to quantity are market orders. Orders containing additional restrictions are referred to as limit orders.[3]
The minimum tick size is ruled by the exchange and depends on the current price level. In general, the minimum tick size increases as price moves up. Once the price reaches a next higher price range, the corresponding minimum tick size is valid immediately. The minimum tick size can even change within a trading day. All these different kinds of offers are collected and stored in the limit order book. The book follows the price-time priority rule.[4]

Data Description We consider all orders submitted to the Swiss Stock Exchange (SWX) between January 2 and December 31, 2002. Excluding days the exchange is closed, these are 251 business days. The Swiss Stock Exchange provides order histories for every single Swiss blue chip.[5] Data of 23 of 26 Swiss blue chips are available for the purpose of this thesis. The dataset consists of total 152,488,698 records including records in Tables of trade prices, orders, canceled orders and other information about trading and trading activities.[6] The Tables comprise fields about transaction time and date when the order has been sub-

[2]Orders are executed according to the price-time priority rule. An order is first arranged to limit price. Within the same limit price orders are sorted according to the submission time. For more details consider chapter 2 and www.swx.ch.

[3]Orders may have various forms of restrictions. At Virt-X, five different forms for trading purposes are available. Hidden size orders, with a visible order size and an additional hidden order size, accept orders, where the order can be partially executed canceling the remaining part, fill-or-kill orders, which can only be executed fully, conditional orders with respect to a price condition and unreleased orders entered together with a time condition. For more detail see chapter 2 and www.virt-x.ch.

[4]See for more details chapter 2.

[5]A part of the data are publicly available. For scientific research purposes, the SWX hands out the remaining part of the dataset.

[6]The database accommodates data which are not relevant for the purpose of this thesis. However, no confidential information about trades, trading activities or investors is available.

mitted to the exchange. It contains fields about buy or sell indication, original order size and the order price. The order form is assignable as well. Expiration date, delete time and reason are also available. In addition, some order identification numbers are inserted. Only orders submitted to the SWX during regular trading periods are considered.[7]

Order Flow Based on the available data we perform an analysis of the order flow. For instance, regarding the order flow of Novartis, 1,789,433 orders have been submitted during the considered period. 1,645,246 of them, or 91.94%, are limit orders and 144,187 or 8.05% are market orders. We find similar dimensions between limit orders and market orders for all other securities. 50% of the limit orders concerning Novartis are buy-orders. The other half are sell-orders. Of all limit orders submitted, 44.8% are not executed at all, 32.7% are involved in at least one trading process and 27.4% are fully matched in one trade.[8] 46.5% of the limit orders were deleted and 4.5% expired. 40.3% of the market orders are buy orders. 15.2% of all market orders were fully executed in one trade, 78.8% were involved in at least one trade. The remainder have been canceled. None of the market orders expired. Order sizes and revenues varies more than the order forms. The average order size of the ask side for Novartis comprises 4940 shares and the average order size on the bid side amounts 4788 shares. Basic trading and market data for all considered assets during the period of 2002 are reported in Table 5.1. The Table shows data concerning stock price development, market capitalization, outstanding shares, average trading volume and average order sizes. The column percentage changes of stock prices documents the market environment for the year 2002. In general, markets strongly declined in this period.

[7]Continuous trading takes place in the state of trading between 9.00 a.m. and 17.20 p.m. From 6.00 a.m. until 9.00 and 5.30 p.m. until 10.00 p.m. is the pre-opening state. During this period orders can be submitted or modified but no trades are executed. The auction pre-opening takes place from 5.20 p.m. to 5.30 p.m. No securities can be traded in this period but derivative instruments can. For a more detailed description consider www.swx.ch.

[8]Exceeds the order size the currently available trade size, the order will be split in several smaller portions. For more details see the matching rules described in section 2.1.2 or www.swx.ch.

Table 5.1: Trading and Market Data for assets of the Swiss Market Index (SMI)

For the 251 trading days in the period of January 2 and December 31, 2002, for 23 of 26 securities including in the Swiss Market Index (SMI) at that time, the Table 5.1 reports data of market and trading activities. The closing price refers to December 31, 2002. The percentage change is calculated on the relative difference between the closing price on January 1 and December 31, 2002, while the year high and low are ascertained during the trading days in the year. Market capitalization and the number of outstanding shares are computed on December 31, 2002. Average trading volume and average order size is based on data of an hourly limit order book.

	Closing Price	Year High	Year Low	Change %	Market Capitalization Overall	Market Capitalization Free Float	Outstanding Shares	Average Volume	Average Order Size Bid	Average Order Size Ask
ABB	3.93	68.88	1.69	−75.85	4716.0	4374.6	1113.129771	1373200	9038	9778
Adecco	54.20	113.11	34.45	−38.30	10124.4	6731.7	124.201107	123860	1322	1320
Baloise	55.00	155.21	46.22	−64.05	3041.9	3041.9	55.307273	21987	819	875
Richemont	25.80	40.88	19.87	−16.77	13467.6	13467.6	522.000000	194130	3284	3286
Ciba	96.40	127.87	91.47	−7.55	6363.2	6363.2	66.008299	37699	960	917
Clariant	22.10	40.11	21.73	−29.62	3391.0	2989.2	135.257919	76309	2269	2313
Credit Suisse	30.00	74.09	19.97	−57.73	35666.4	35666.4	1188.880000	939790	3856	4086
Givaudan	620.00	681.32	502.51	24.00	5409.9	4874.9	7.862742	5968	229	218
Holzim	251.00	393.13	197.00	−33.72	7287.2	6907.5	27.519920	14366	297	292
Kudelski	18.75	104.18	16.45	−57.82	875.2	634.6	33.845333	43482	1132	891
Nestle	293.00	395.80	273.47	−16.76	116019.0	116019.0	395.969283	153150	935	2001
Novartis	50.45	68.88	50.12	−14.35	142478.4	133345.5	2643.121903	1097700	4788	4940
Roche	96.35	133.76	89.93	−17.30	67691.9	67691.9	702.562532	514430	2487	2532
Swiss Re	167.00	170.88	71.18	−45.69	29193.4	29193.4	321.867696	180000	1409	1486
SGS	416.00	527.66	258.44	−53.22	3254.0	2488.3	5.981490	2848	662	153
Sulzer	255.00	373.26	158.09	−27.97	683.9	683.9	3.637766	2142	185	182
Swatch	23.40	42.42	20.68	−27.22	3207.0	1251.4	53.478632	34583	2172	2094
Swisscom	460.00	516.81	364.99	−12.93	26514.4	9881.9	24.673908	16348	345	298
Swisslife	715.00	738.48	61.36	16.30	2336.0	1871.1	17.365197	15104	285	274
Syngenta	86.00	108.98	72.60	−5.85	9010.8	8119.6	101.431605	48288	1209	1187
UBS	83.80	84.35	50.43	−18.24	84307.8	84307.8	1254.580357	599940	4008	2702
Unaxis	179.00	209.58	64.28	−47.86	1218.2	871.9	9.425946	6834	389	401
Zurich	389.50	420.07	44.21	−57.99	18560.4	18560.4	143.879070	132710	1395	1150

Reconstruction of the Limit Order Book The data received from the SWX allows for an estimate of the limit order book. The dataset provides information to every single limit order that has been submitted to the exchange. Each order record specifies, among other things, investor generated data as the date and time of submission, the type of order, which characterizes the order as a buy or sell intention, the order original size and the limit price. Exchange created data as expiration date, delete reason and delete time are also included in the same record. Delete reason denotes if an order has been fully executed or not. Not fully matched orders appear several times in the dataset with adjusted order volumes. Data are collected in separate tables for every single asset. The time resolution in this database is 0.01 seconds. This dataset enables to reproduce the limit order book for several time intervals.

We reconstruct the limit order book similar to the approach proposed by Kavajecz (1999) and Naes and Skjeltorp (2006) and in particular according to the matching rules of the SWX. The reconstruction is carried out in the following steps. In a first step the bid and ask limit orders are separated and saved in individual files. Then, a filter for a desired time interval runs over the data. The bid limit orders respective to the ask limit orders, which belong to the corresponding time interval, are clustered and stored in a multi-dimensional vector. These multi-dimensional vectors contain all orders separated according to the chosen time interval. In a third step, data which can be matched within the chosen time interval are identified. Matching rules are applied to these identified records and the matched orders are eliminated from the dataset. Within the time interval data are sorted in descending order, respectively ascending order such that the corresponding best quote appears on the top. It results two multi-dimensional vectors with the limit order book for the desired time interval for both the bid side and the ask side, respectively.
To verify the correctness of the reconstructed limit order book, the mid price of the best ask quote and the best bid quote are calculated. We compare this time series of the mid prices with the mean of the real traded prices for several time intervals. In Table 5.2 the Pearson's correlation coefficients of selected securities for different time intervals are reported. The correlation coefficient for the 5-minute time interval, for Adecco for instance is 0.999654, for 30-minutes 0.999946 and for 60-minutes 0.999976. We find similar results for other time

series. Since the correlation's coefficients are all close to 1, this shows evidence of the value of the reconstructed limit order book.

Table 5.2: Correlation of the Limit Order Book

The Table reports Pearson's correlations between traded prices received from our data set and artificially computed mid prices from the reconstructed limit order book. Time interval for the snapshots of the limit order book are 5, 30 and 60 minutes. Calculation is based on 2008 observations for the 60 minutes, 4016 for the 30 minutes and 8032 for the 5 minutes time intervals for each security.

	5 Minutes	30 Minutes	60 Minutes
Adecco	0.999654	0.999946	0.999976
Credit Suisse	0.999994	0.999997	0.999999
Novartis	0.999994	0.999999	0.999999
Richemont	0.999873	0.999889	0.999904
Nestle	0.999913	0.999947	0.999977
Syngenta	0.999243	0.999353	0.999655

After reconstruction of the limit order book, we define a feasible time interval. An appropriate time interval is important for a representation of the dynamic of the limit order book most accurately. A too short time grid runs the risk of containing less or even no limit orders. Larger time intervals may not reflect the dynamic behavior of the limit order book appropriately. Ahn, Bae, and Chan (2001), for instance, address this problem and conclude a 15-minute time interval to be most applicable for their dataset. Kavajecz (1999) takes 30-minute time intervals and Ekinci (2005) investigates several time intervals beginning with 1-seconds over 2-seconds, 5-seconds and 15 seconds to 1-minute, 5-minute, 15-minute and 30-minute time intervals. Pascual and Veredas (2006) take intervals ranging from 30-minutes to 2-hours. However, the academic theory does not provide any suggestions to the right time interval. We calculate reconstructed limit order books of 5-minute, 10-minute, 15-minute, 30-minute and 1 hour time intervals. The main part of investigation is based on an hourly data set.

Having reconstructed the limit order book on an hourly basis, the time horizon for the analysis has to be determined. While the data set of Biais et al. (1995) comprise 19 trading days, data of Naes and Skjeltorp (2006) range from February 1999 to June 2001 and consists of 597 trading days. Data in the study of Jain and Joh (1988) runs from 1979 to 1983, comprising 1263 trading days and the data set of Gerety and Mulherin (1992) ranges beginning of 1933 and ends in 1988. Our data set in this thesis ranges from the 1st of January until the end of December 2002 and covers a one year period. Consequently our analysis horizon is based on one year.

5.2 Dynamic of the Limit Order Book

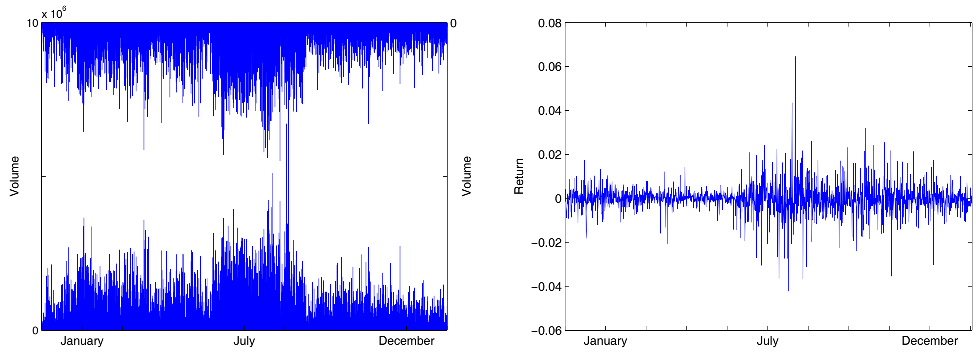
As Gerety and Mulherin (1992) pointed out, theoretical, experimental and empirical science often analyzes different kinds of structural characteristics. The aim is to recognize and comprehend regularities, normalities and recurrences in patterns, systems or structures. This section serves to an introductive overview about the trading activities and the limit order book. The main focus is on the visualization of the dynamic of the limit order book. Following Chordia et al. (2001) we distinguish between trading activity measures as volumes and liquidity measures as depth or bid-ask spread based figures. We start with illustrations based on data of an aggregate level and refine data increasingly. The end of the section closes with analyzes of interdependency between different aspects of the limit order book with each other or with other variables.

Volume Dispersion A vast amount of academic literature and economic investigations address to the relations between volume and return, respectively, volumes and volatility in different ways. He, Velu, and Chen (2004) or Hasbrouck and Seppi (2001) for instance, model the relation through interaction of several independent components of each factor. Campbell, Grossmann, and Wang (1993) consider serial correlation structures.[9]

We are not concerned with detailed analyzes regarding the relation of daily volumes and returns at this stage of investigations. However, to get a first impression about trading activities, we provide a graphic of the offered volumes and returns. In Figure 5.1 the bid and ask volumes and the natural logarithm of the returns of Novartis for one year are illustrated. The measurement interval is one hour. On the left hand side in Figure 5.1 (a) the sums of bid and ask volumes are depicted. Each line represents the sum of the demanded respectively

[9]And many other authors as Jain and Joh (1988), Niemeyer and Sandås (1993) or Hedvall (1994).

supplied volume of the limit order book in the measurement interval. On the top of the picture the ask volumes with inverted ordinate are illustrated. The corresponding ordinate is on the right side. On the bottom of the picture the bid volumes are drawn with corresponding ordinate on the left side. The white part between the two volume processes can be interpreted as the volume pressure process. On the right hand side in Figure 5.1 (b) the natural logarithms of the returns are presented. From a structural point of view, clusters on each time series can be detected. Serial correlation of both volumes and returns are well documented in academic literature.

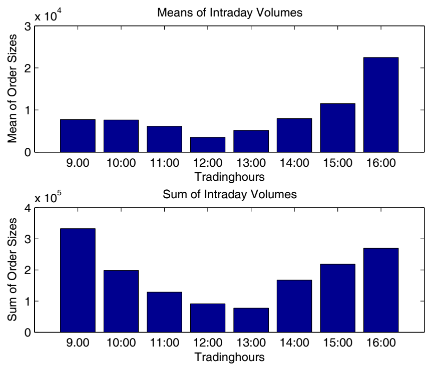


(a) Intraday Volume Dispersion Bid Side (b) Log-Returns

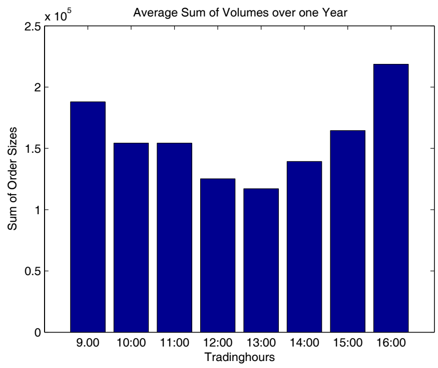
Figure 5.1: Bid and Ask Volumes and Log-Returns of Novartis For the period between January 2 and December 31, 2002, the Figure 5.1 (a) illustrates the intraday volumes for Novartis on hourly basis. The Figure (b) shows time series of natural logarithms of the returns calculated on hourly time intervals. Calculations are based on 2008 observations.

Figure 5.1 (a) depicts higher volumes in the first half of the year as in the second half of the year. In particular clusters of higher volumes can be found around January and between the time of May and June. The chart of the log-returns in Figure 5.1 (b) shows a contrary picture. The returns oscillate strongest in the second half of the year, especially around June and July.

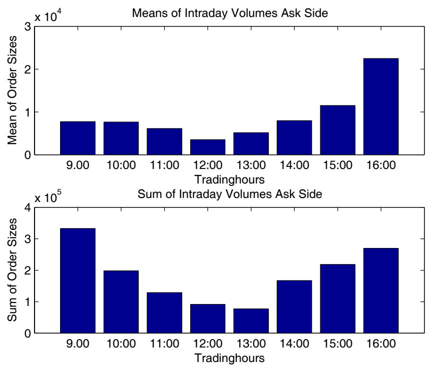
In contrast to this the analysis of the intraday volume variabilities reveals a more systematic behavior. Jain and Joh (1988), among others report a U-shaped pat-



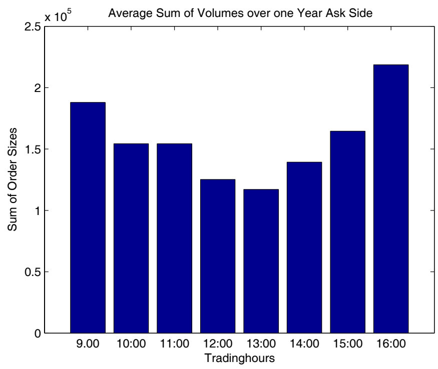
(a) Intraday Volume Dispersion Bid Side



(b) Volume Dispersion One Year Bid Side



(c) Intraday Volume Dispersion Ask Side



(d) Volume Dispersion One Year Ask Side

Figure 5.2: Intraday Volume Dispersion The Figure 5.2 displays the intraday volume dispersion separately for both, the bid side (upper Figure) and the ask side (lower Figure). The pictures on the left hand side give an idea about calculations for a given day, while the illustrations on the right side show the average of a year.

tern of intraday trading volumes.[10] Additionally, they consider and compare every single day in the week separately and discover that in average, trading activity is small at the beginning and the end of the week, and is large around Wednesday. Gerety and Mulherin (1992) investigate data from the New York Stock Exchange (NYSE) from 1933-1988 and confirm a U-shape in intraday trading volumes also for long time horizons. They find that most activities oc-

[10]See also Harris (1986).

cur at the beginning and the end of a trading session, and argue that much of that clustering is due to fear of bearing risk in overnight positions. Biais et al. (1995) provide a more extensive analysis of trading activities for the Paris Stock Exchange. They additionally differentiate between orders according to their aggressiveness and trades according to their size. In general, both exhibit the intraday U-shape pattern. The study reports that small sized trades occur in the morning, when depth is low while large sized trades take place more in the late afternoon, when depth is large. Biais et al. (1995) offer four different interpretations of these effects. The first interpretation states that small trades in the morning contribute to price discovery while large trades tend to arise after prices are already discovered. Second, they assert that fund managers, who intend to trade a large trade, are evaluated with respect to the closing price. Third, the high frequency of small orders in the morning are caused by financial intermediaries who execute orders which they received from their customers early in the morning. Fourth, the intraday U-shape patterns of the volumes arise from strategic order splitting. We differ from these analyzes since we consider the offered volumes for the bid and the ask side separately. Figure 5.2 illustrates the situation for Novartis. Figure 5.2 (b) depicts the average sum of offered volumes of one year on the bid side of a business day. Figure 5.2 (d) illustrates the same for the ask side. Observations take place in hourly time intervals and calculations are based on the entire limit order book in the corresponding time interval. We find a U-shape with lowest value around noon, for both the bid and the ask side thus also for the Swiss market. Besides our results and the findings for the US market, Niemeyer and Sandås (1993) present similar patterns for Sweden and Hedvall (1994) shows the intraday U-shape in volumes also for Finnland.

On the left hand side of Figure 5.2 in subfigure (a) and (c) additionally the volume dispersion of a single day is plotted. In each of these cases the illustrations on the top pictures are the mean of the volume in the corresponding trading hour. The lower illustrations depict the sum of all order volumes in this trading hour. The U-shape appears here as well. All considered securities exhibit this kind of U-shape of intraday volume dispersion.

5.3 Shape of the Limit Order Book

In order to get a deeper insight into the structure and the order flow process and in particular of the trading activity, we refine the grid of the data once again. Instead of considering the cumulative volumes, we now disjoint the data additionally with respect to their prices submitted. In the limit order book all orders submitted to the market place are collected according to their intent, price and submission time.[11] The book is separated into a bid- and an ask-side and orders are assigned to the corresponding side. Within the bid respectively the ask side, the orders are sorted first according to the submitted price and second, within the same price, according to the submitted time.[12]
A graphical interpretation of the limit order book is presented in Figure 5.3. Here, the time aspect within the same prices is omitted and orders submitted to same prices are summarized. In this graphic we assume that an investor intending to buy (sell) for the best quoted price is also prepared to buy (sell) at the next lower (higher) price. As a result we aggregate the orders with increasing depth of the book.[13] This diagram allows a view to depth at a number of price levels.[14]
A majority of analyses are mainly based on the best quoted prices, such as the bid-ask spread for instance. Recent academic literature and scientific economic research discovers that quotes beyond the best quotes accommodate additional information. Cao et al. (2004) evidence that quotes beyond the best quotes have informative properties. They show that firstly, the order book beyond the best quote contains about 30 % of information of the true value of the underlying asset. Second, the imbalance information between demand and supply as expressed in the entire book contributes additional explanatory power to future returns.[15] And third, trading and submission strategies are affected by the avail-

[11]For a more detailed description consult also chapter 2.
[12]This is called the price-time priority rule. See for more details chapter 2.
[13]From a microeconomic perspective the presented graphic represents the demand and supply functions.
[14]This kind of illustration of the limit order book is also used by Glosten (1994) or Naes and Skjeltorp (2006). In contrast to our pictures Naes and Skjeltorp (2006) visualize the shape of the book on a relative basis. Instead of using the absolute aggregated volumes on the ordinate they calculate the percentage of the accumulated share volume. The horizontal axis shows the ticks away from the best quotes.
[15]They find that in a regression framework the imbalance information increases the adjusted

able additional information of the limit order book. Harris and Panchapagesan (2005) confirm these findings and extend research by analyzing trader's behavior regarding different actions a trader can take. They find strong evidence that traders use the additional information provided by the entire limit order book.[16] Moreover, to show that traders rely on their order book, they measure order book information with trading options and future price changes.[17] As a result, they find strong evidence that the limit order book is informative about future price changes and that order properties such as duration, price relative to the market, and order size have information content. Pascual and Veredas (2006) evaluate the informational content of the limit order book by its explanatory power in long run volatility.[18] By separating liquidity driven volatility from informative driven volatility they report that changes in immediacy costs, for trades of different sizes, indicates posterior fluctuations in long run volatility.

In Figure 5.3 (a) to (d) hourly snapshots of several situations of the limit order book are depicted. With focus on the shape of the book, Figure 5.3 (a) illustrates a high concentration of orders around the bid-ask spread on both sides. This entails a steep shape of the limit order book. In contrast to this, in Figure 5.3 (b) orders are less dispersed around the bid-ask spread, which results in a more gentle shape of the book. From an auction theoretic perspective Figure 5.3 (a) is considered as more liquid situation than Figure 5.3 (b). To illustrate this, consider the example of a buy (sell) order, price limited on the best quoted ask (bid) price. The order size is equal to the corresponding offered size in Figure 5.3 (a). Since order size corresponds to the offered size in 5.3 (a), the order will be fully executed at this quoted price. In situation 5.3 (b) due to lack of enough volume at this price the order can not be fully executed and has to be split in several portions. The split part either is executed at an adjacent price level or the agent waits until volume is provided again.[19] In the first case, transaction costs

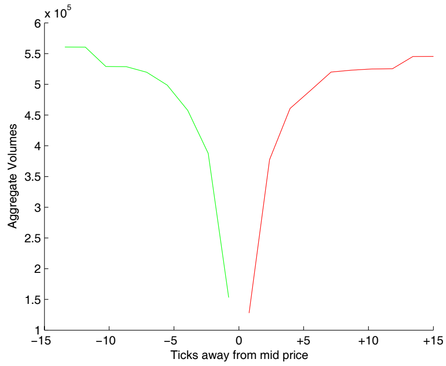
R^2 by 11-18% in comparison to the results using information based only on the best quotes.

[16] They characterize information according to their dependency of the limit order book and investigate traders behavior to the information available only from the entire book.

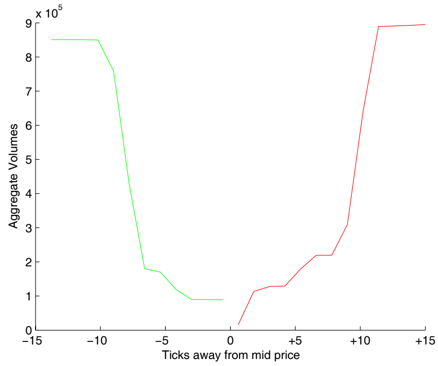
[17] They use two option pricing models to estimate the value of limit orders and analyze the value pattern of these artificially constructed options.

[18] Similar results are provided by Ahn et al. (2001) and Coppejans et al. (2004).

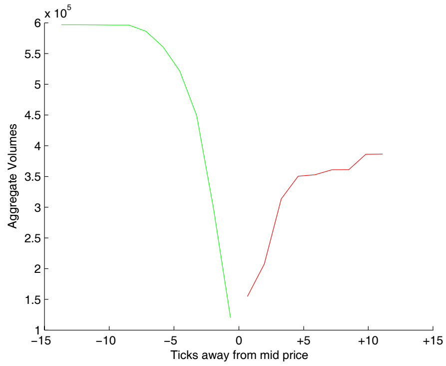
[19] Execution on an adjacent price level is only possible if the order price limit complies with the adjacent price, the order has been submitted as market order or no other restrictions according to order splitting have been done.



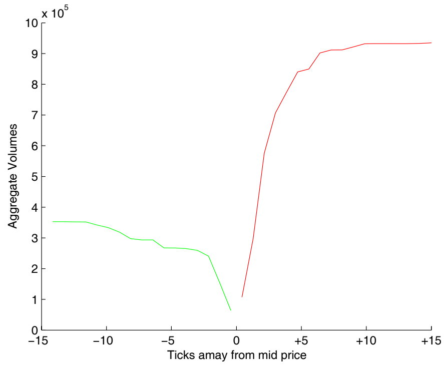
(a) Liquid Limit Order Book



(b) Illiquid Limit Order Book



(c) Demand Pressure in Limit Order Book



(d) Supply Pressure in Limit Order Book

Figure 5.3: The Limit Order Book The Figure 5.3 illustrates hourly based snapshots of several order book situations of Novartis. In each picture the left curve represents the bid volumes aggregated with increasing tick levels away from the mid price, while a tick level corresponds to 0.05 units of currency. The right curve shows the same for ask volumes. The gap between the two curves represents the bid-ask spread with the corresponding quoted quantities. From an auction theoretic perspective Figure (a) is considered as more liquid situation than Figure (b). Figure (c) represents a buy pressure while Figure (d) shows a sell pressure.

measured by the price impact are higher and in the second case, the waiting time until the order is executed is longer. Both accord with the definition of a less liquid market environment.[20] In this context, Figures 5.3 (c) and (d) represent

[20]See definition of liquidity in section 2.2.3.

a demand pressure and a supply pressure in the order book, respectively

Liquidity Costs With focus on the price impact which arises of different order sizes we analyze the shape of the limit order book regarding to its informational contribution to the price impact. There has been substantial academic literature in the field of transaction- and liquidity cost. With focus on price impact several authors including Chan and Lakanishok (1993), Chan and Lakanishok (1995) and Kalay, Sade, and Wohl (2004) constitute that buy orders have larger price impacts than sell orders. Similar results are reported by Naes and Skjeltorp (2006). Moreover, Chiyachantana, Jain, Jiang, and Wood (2004) find that the underlying market condition is a major determinant of price impact, in particular of the asymmetry between the price impact of sell and buy orders. They report that in bullish markets, institutional purchases have bigger price impacts than sells, and in bearish markets sell orders have larger price impacts.
Stock markets declined in the year 2002, in particular the Swiss stock market. Table 5.1 reports the situation. Except Givaudan all other considered companies lost substantial market value. In 2002 the market was bearish. According to the mentioned literature in such a market environment sell orders affect prices more than buy orders.
In order to measure the price impact we calculate the liquidation cost based on the limit order book of the bid and the ask side separately. Since the data is not continuous but discrete, we use instead of the definition 2.2.3 the discrete form of

$$L_s(q) = \sum_{\tau=1}^{R} q_{(\tau,s)} p_{(\tau,s)} - q \cdot P_s, \tag{5.1}$$

where τ denotes the tick level away from the mid price. $q_{(\tau,s)}$ is the aggregated volume at the corresponding tick level and $p_{(i,s)}$ represents the price at the corresponding tick level. $\sum_{\tau=1}^{R} q_{(\tau,s)} = q$ represents the depth of the entire limit order book in the snapshot s. Instead of using the mid price M_s as presented in the theory section, we use the current trading price P_s in the corresponding time interval.[21]
To calculate the liquidation cost we assume that an order submitted to the

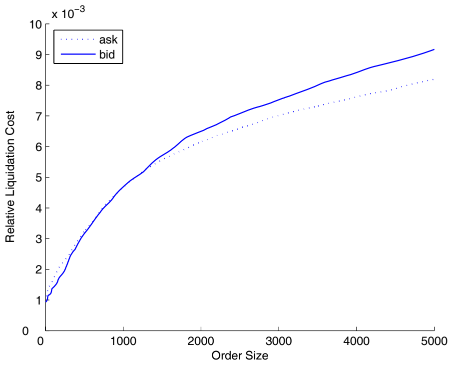
[21] Mid prices are highly correlated with trading prices. See section 5.1. For that reason this change does not affect the result, but facilitate calculation.

market will be executed at the available prices in the corresponding time interval of the limit order book. In particular, if the order size exceeds the size of the current best quoted offer, the order will be split and the offered size will be cleared. The remaining, not executed part of the order will be matched against the next best quoted price. Additionally, exceeds the order size the currently available size of the limit order book in time (s), the remaining part of the order is transferred in the proximate limit order book in time ($s + 1$). The matching process is carried out until the order is fully cleared. A fully executed order may affect several price levels and consequently, may be split several times. Hence liquidation costs increase due to execution at proximate price levels. This reflects the price impact effected through the order size.

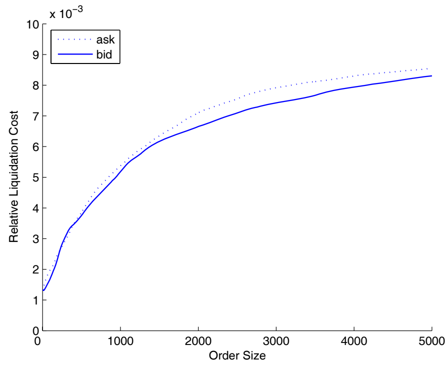
Order splitting is associated with additional costs. Financial institutions charge additional fees for each split order. Additionally, in the case that an order cannot be fully matched in one time interval, even more costs arise. Since these costs depend on the financial institution and in particular on the conditions of the contract between the investor and the institution, we exclude this kind of cost in this calculation. Because only the liquidation cost in form of an immediate price impact is calculated, it is herein after referred to as pure liquidation cost.

The calculations are performed for an order size starting at 1 share and ending with 10'000 shares. The estimation for the average cost is based on a time period of one year. The liquidation-costs in equation (5.1) are absolute values and reflect the effective costs in Swiss Francs. In order to compare the liquidity costs we calculate the relative costs with respect to the corresponding revenues. In Figure 5.4 the liquidation cost structure with respect to different order sizes are illustrated. The dashed line represents the liquidation costs on the bid side, the solid line on the ask side. In Figure 5.4 (a) and (c) the liquidation costs of the bid side are higher. In contrast, in Figure 5.4 (b) and (d) liquidation cost on the ask side are higher.

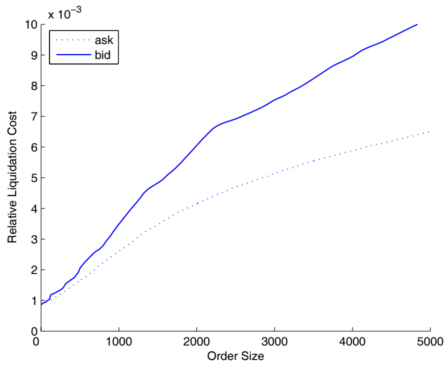
The bid side absorbs possible sell orders. In order to measure the price impact of a sell order, the bid side is relevant. To decide which price impact is greater, we apply a right-sided T-test to the liquidation costs. For 56,52% of all considered assets, or 12 out of 23, the null hypothesis that the mean of the bid side is greater than the mean of the ask side is accepted. Of the rejected null hypotheses 5 of 23 or 21,74% even do not accept the inversely arranged T-test, that the mean



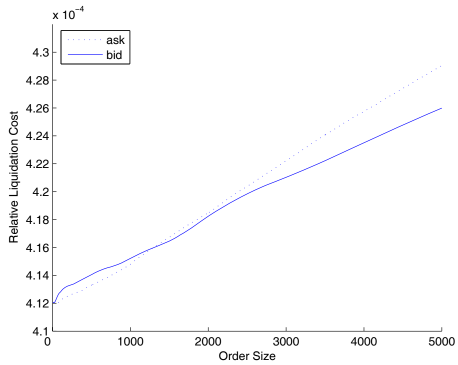
(a) SGS



(b) Sulzer



(c) Swisslife



(d) Novartis

Figure 5.4: Pure Liquidation Cost The liquidation costs relative to the price are graphically illustrated in Figure 5.4. The costs are calculated for the price impact function of the liquidity cost given in equation (5.1), where the order size increases starting of 1 share to 10,000 shares. The solid line represents the relative costs on the bid side, while the relative costs on the ask side are shown by the dashed line.

of the ask side is greater than the mean of the bid side. The remainder, again 21,74%, accepts the alternative T-test and supports a higher price impact of a buy order. Based on these statistical results we can not confirm the findings of Chiyachantana et al. (2004) that sell orders have larger price impact. In contrast, we can neither decline nor accept the hypothesis that in a bearish market environment sell orders have large price impacts.

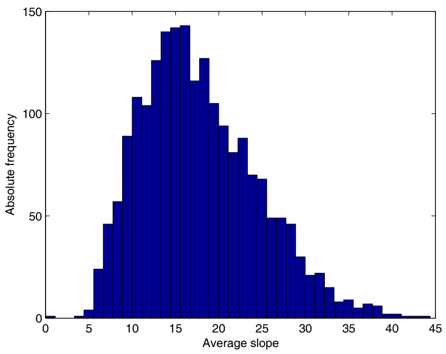
However, in several empirical papers the shape of the limit order book is investigated. Glosten (1994), for instance, uses similar representations of the shape of the limit order book in his investigations. In his model, he explains the slopes of the limit order book shapes with homogeneous liquidity suppliers. Goldstein and Kavajecz (2004) provide evidence that there exists a negative relation between the shape of the limit order book and volatility during extreme market situations. Naes and Skjeltorp (2006) address to discover price formation in a limit order market. Their main findings are that volume, volatility and the volume-volatility relation are negatively related to the order book slope. In addition, they show evidence that there exists a negative relation between order book slope and analysts earning forecasts. They interpret these findings as an implication of the order book slope as a proxy for the beliefs of investors about the asset value. The analysis of the slope of the order book also delivers insights in different aspects of the order flow and hence in the dynamic of the limit order book. Based on this, we follow Naes and Skjeltorp (2006) and partially reconstruct their analysis.

Measuring the Slope of the Order Book The slope of the limit order book is basically the elasticity measure $\partial q / \partial p$, which expresses how quantity (q) supplied in the limit order book changes as a function of the price (p). Considering Figure 5.3, the slope changes along the price axis. To obtain an average measure, Naes and Skjeltorp (2006) propose to first average the slopes along the corresponding offer side of the limit order book for the given snapshot. Obtaining the average slope for the bid and the ask side separately, in a second step the average of a trading day is calculated for each side. In order to get the average of the slope of the entire limit order book, the daily average of the bid and the ask side is computed. Following Naes and Skjeltorp (2006) we calculate the slope $S^{x,i}_{s,t}$ for the corresponding offer side x, for each snapshot $s \in [1, \ldots, 8]$, security i and date t as

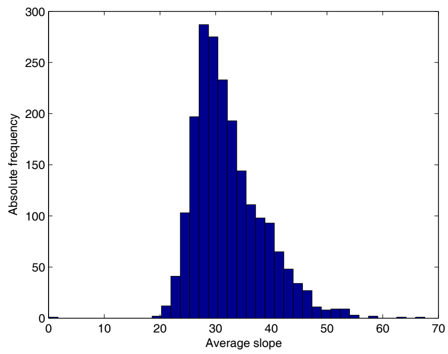
$$S^{x,i}_{s,t} = \frac{1}{N_x} \left\{ \frac{\nu^x_1}{p^x_1 / p^x_0 - 1} + \sum_{\tau=1}^{N_x} \frac{\nu^x_{\tau+1} / \nu^x_\tau - 1}{p^x_{\tau+1} / p^x_\tau - 1} \right\}, \tag{5.2}$$

where N_x are the total number of offered prices containing orders at the corresponding book side. Let τ index the tick level, where the bid-ask midpoint

level is represented by $\tau = 0$, and $\tau = 1$ denotes the best quoted price with a positive accumulated share volume.[22] Let the mid price denoted by p_0^x and let ν_τ^x represent the natural logarithm of accumulated total offered shares at each tick level τ. The first term in the brackets of equation (5.2) measures the slope from the bid-ask mid price to the best quoted price level. The second term sums the local slopes along the remaining price levels of the limit order book.[23]
Having calculated the mean of the slope for a trading day separately we average the daily slope of the bid and the ask side arithmetically. To scale the slope measure for parameter estimates of the regression analysis we divide the measure by 100.



(a) Average slope



(b) Average slope ± 10 ticks

Figure 5.5: Histogram of Slope Estimates The Figure shows the absolute frequency of the daily averaged slopes of the limit order book of Novartis. In Figure (a) estimations are based on the entire limit order book. In Figure (b) frequency distribution is counted on a truncated limit order book limited to ± 10 ticks away from the mid price.

In Figure 5.5 histograms of the absolute slopes of two different subsets of the limit order book for Novartis are presented. Figure 5.5 (a) represents the frequency of

[22]For each security traded at the SWX, the SWX rules the tick size depending on the currently traded absolute price level. See for more details chapter 2. We adopt this tick size for every single considered asset.

[23]Naes and Skjeltorp (2006) pointed out that the first and the second term inside the brackets are not measured in the same units since no volume at the bid-ask mid price is available in order to calculate the elasticity at this price level. However, they verify their choice of the measure by calculating alternative approaches proposed by Kalay et al. (2004) or Kim, Lee, and Morck (2004).

the slopes calculated on the entire limit order book. In Figure 5.5 (b) the slopes are estimated on a truncated limit order book, where depth is limited to ± 10 ticks away from the mid price.
The mean of the histogram of the slope based on the entire limit order book is 17.79 (median 16.79), while the mean of the frequency of the slope based on the truncated book is 32.35 (median 30.97). According to Naes and Skjeltorp (2006), we find similar outcomes that the slope increases the less of the depth of the book is involved. Based on this, we assume that in general average volumes are larger the closer they have been submitted to the best quotes. Considering order sizes equally, this means that the frequency of order placement around the best quotes is higher. In fact, Biais et al. (1995) show evidence that order placement decreases monotonically as one moves away from the quotes.
From an economical perspective the shape of the limit order book can disclose information about the beliefs of investors about the real value of the traded security. A wider book indicates that investors disagree more about the real value of the asset. A steep book implies that investors have similar value perceptions.

Table 5.3: Correlation between slope measures and liquidity measures
The Table reports Pearson's correlations of the results of 23 assets for slopes with several liquidity and trading activity measures. The first estimation is based on the entire limit order book, while the latter two are based on a truncated book bounded to 15 respectively 10 ticks away from mid price. Estimations take place on a daily basis which involves 2008 observations.

	Slope (Full order book)	Slope (±15 ticks)	Slope (± 10 ticks)
Number of Trades	−0.356	−0.224	−0.187
Trading Volume	0.060	0.305	0.348
Market Capitalization	0.381	0.555	0.579
Spread	−0.358	−0.464	−0.471
Slope (Full order book)	1.000		
Slope (+/- 15 ticks)	0.871	1.000	
Slope (+/- 10 ticks)	0.782	0.935	1.000

The proceeding analysis explores how the slope of the limit order book is related to typical measures for liquidity or trading activity. Typical measures include the number of trades, executed within the considered snapshot, trading volume, the current market capitalization and the bid-ask spread. We estimate first the slope for every considered asset and compare then the results with the liquidity and trading activity measures. In Table 5.3 the results of the Pearson's correlation between these measures for different restricted limit order books are represented. As also reported by Naes and Skjeltorp (2006), we find that market capitalization is positively and the bid-ask spread is negatively correlated to the book slope. The full order book slope has a positive correlation to market capitalization of 0.381 and is negatively correlated with -0.358 to the bid-ask spread. Correlation increases the more the limit order book is restricted. The book slope, where the slope is estimated by a book bounded by $\pm$ 10 ticks away from the mid price has a positive correlation with market capitalization of 0.579 and a negative correlation of -0.471 with the bid-ask spread. These findings indicate that in general larger capitalized assets are more liquid in terms of smaller spreads and steeper order book slopes.[24] In contrast to Naes and Skjeltorp (2006) we find a negative correlation between the order book slope and the number of executed trades. The full order book slope has a negative correlation of -0.356 with the number of executed trades. But the negative correlation tends to be smaller the more the book is bounded. The restricted limit order book slope is correlated with -0.187 to the number of trades. Trading volume however behaves conversely. For the entire book we find a correlation of 0.060. The correlation to the bounded book is 0.348. This indicates that trading is more frequent closely around the bid-ask spread. These findings are supported by the outcomes reported to Figure 5.5 and by Biais et al. (1995).

Naes and Skjeltorp (2006) proceed with analyzing the volume-volatility, price volatility and slope volume relation in different ways. However, our main focus is on the relation between the slope of the limit order book and the price process, respective of the price volatility. In this sense, we mainly concentrate on the regression framework involving these parameters. In order to estimate this relation R. D. Huang and Masulis (2003) as well as Naes and Skjeltorp (2006) follow the two step regression approach provided by C. M. Jones, Kaul, and Lipson (1994). Following C. M. Jones et al. (1994) we measure in a first step

[24]Naes and Skjeltorp (2006) arrive to a similar conclusion.

the return volatility for each security i by

$$R_{i,s} = \sum_{j=1}^{12} \beta_{i,j} R_{i,s-j} + \hat{\varepsilon}_{i,s}, \tag{5.3}$$

where $R_{i,s}$ is the return on security, i in the corresponding snapshot s. Short term movements enter conditional expected returns by the 12 lagged return regressor. The residual $\hat{\varepsilon}_{i,s}$ is the estimate of the unexpected return on security i in the time interval s. The absolute value $\mid \hat{\varepsilon}_{i,s} \mid$ represents the volatility. The second step according to C. M. Jones et al. (1994) consists of estimating the relative effects of explanatory variables to the response variable by running for each security i the regression

$$\begin{aligned} \mid \hat{\varepsilon}_{i,s} \mid &= \beta_1 N_{i,s} + \beta_2 OV_{i,s} + \beta_3 MCAP_{i,s} + \beta_4 SPREAD_{i,s} \\ &+ \beta_5 SLOPE_{i,s} + \sum_{j=1}^{12} \rho_{i,j} \mid \hat{\varepsilon}_{i,s-j} \mid + \eta_{i,s}, \end{aligned} \tag{5.4}$$

where N reflects the number of trades executed in the corresponding time interval s, OV is an abbreviation of the average order book volume, $MCAP$ is the average of the market capital in the snapshot, $SPREAD$ is the relative bid-ask spread and $SLOPE$ stands for the slope measure computed according to equation (5.2). The term $\rho_{i,j}$ weights the volatility persistence across 12 lags.[25] In addition to the order book slope, several variables to control for liquidity are included in the regression.

Less liquid assets generally have smaller order book volumes. A large offer cannot be absorbed by the best quotes and affects therefore a price impact. In general price impacts comprise an additional source of risk. Since agents require liquidity premium for additional risk the bid-ask spread is higher for less liquid assets. From this microeconomic theoretical perspective a positive relation between the order book slope and the volatility is expected.

In Table 5.4 the regression results for Novartis for four different models are

[25] Naes and Skjeltorp (2006) include a constant term $\alpha_{i,s}$ in some regression frameworks. We estimate different regressions including and excluding the constant parameter and conclude, that estimation is improved by omitting the constant parameter.

Table 5.4: Price volatility and slope regression

The Table reports the estimation results for Novartis for the entire limit order book of four different versions of the regression

$$\begin{aligned}\mid \hat{\varepsilon}_{i,s} \mid &= \beta_1 N_{i,s} + \beta_2 OV_{i,s} + \beta_3 MCAP_{i,s} + \beta_4 SPREAD_{i,s} \\ &+ \beta_5 SLOPE_{i,s} + \sum_{j=1}^{12} \rho_{i,j} \mid \hat{\varepsilon}_{i,s} \mid + \eta_{i,s},\end{aligned}$$

The regressand $\mid \hat{\varepsilon}_{i,s} \mid$ is the absolute value of the unexpected return on security i in the snapshot s obtained by equation (5.3). N is the number of executed trades within the considered snapshot, OV represents the average order book volume, $MCAP$ the market capitalization, $SPREAD$ the bid-ask spread and $SLOPE$ the slope of the entire limit order book computed according to equation (5.2). For each model in the first row the parameter estimations are presented while in the second row the corresponding t-values of the parameter estimations are shown. In model 2 and 3, negative respectively positive related parameters are omitted. In model 4 we test only the effect of the limit order book slope. The parameter estimation results for the volatility persistence are excluded. F-tests denoted by ** are significant at a 1 percent level. Estimation is based on 2008 observations.

	Model 1		Model 2		Model 3		Model 4	
Variables	$\hat{\beta}$	t-Value	$\hat{\beta}$	t-Value	$\hat{\beta}$	t-Value	$\hat{\beta}$	t-Value
N	0.000	−1.32	0.034	6.90				
OV	0.081	2.49	0.000	0.11				
Market Cap	−0.061	−1.39			0.005	8.91		
Spread	−0.108	−1.40			0.083	6.64		
Slope	−0.065	−12.06	−0.065	−12.48	−0.066	−12.69	−0.076	−15.64
Adj. R^2	0.290		0.289		0.289		0.258	
F-test	819^{**}		817^{**}		815^{**}		699^{**}	
DW	2.001		2.004		2.006		1.989	

summarized. The estimations are based on the entire limit order book. Model 1 includes all parameters, while the model 2 and 3 excludes highly correlated variables of the estimation. To extract the effect of the slope to the volatility the last model is set up only of the slope parameter.

As expected from microeconomic considerations, the slope variable causes a negative effect to the price volatility. The results are in the same region and in particular, highly significant across all model specifications. The parameter estimation in model 4 for the slope $\hat{\beta}_5$ is -0.065 with a corresponding t-value of -15.64, which indicates significance even on a 1 percent significance level. That

means that volatility decreases the steeper the order book slope is. These findings are identical to that of Naes and Skjeltorp (2006). Moreover, we confirm the outcomes by Naes and Skjeltorp (2006) and Biais et al. (1995) of the positive relation between the order book volume and volatility. As pointed out by Biais et al. (1995) this shows that more trades are executed when the order book is thick. The estimation results of the market capitalization as well as of the bid-ask spread alters signs by changing the model.
The adjusted R^2 is 0.290 for the model 1 and declines by omitting control variables. Since in model 4 only the slope parameter is estimated the adjusted R^2 indicates that 0.258 of the proportion of the response variation is explained by the slope measure. In order to show the goodness of fit, the results of the F-tests show evidence that the adjusted R^2's are significant even on a 1 percent significance level. The Durbin Watson (DW) values indicate no autocorrelation in the residuals.[26]

Order Flow Persistence So far, we have concentrated on the analysis of the limit order book dynamic generated by orders aggregated on each price level. In order to explore how order flows affect order book dynamic, we now turn to the analysis of the order placement procedure. We follow Biais et al. (1995) and consider all orders according to their restrictions regarding price and quantity and in particular according to the submission time. The orders are differentiated between the sell and the buy side and within the corresponding side according to their aggressiveness. According to Biais et al. (1995) on each side, seven categories of orders, corresponding to decreasing degrees of aggressiveness, are defined.[27] The first three categories of orders are limit orders price limited at or higher than the currently available best quoted price. The most aggressive order is "large buy", an order limited on a larger quantity and a higher price than the currently available best offer. The second category is "market buy", an order to buy a larger quantity than the quantity offered at the best quote and limited to the best quoted price. Orders assigned to the category of "small buy"

[26]Naes and Skjeltorp (2006) provide deeper analysis regarding the explanatory components of the limit order book slope. Since our work does not rely on the book slope we finish analysis at this point.
[27]Biais et al. (1995) differentiate between trades and orders. We consider all orders before they are executed and call them consequently all orders.

are restricted in quantity lower than the offered quantity, while the price is set as high as the currently best quoted price. The next two categories are price limited within the bid-ask spread. The order size of orders of the category "large buy within" are larger or equal to the best offer, while the quantity of an order from the category "buy within" is smaller than the best quoted size. The price limits of the last two categories are ascertained below the price of the best offer. The category "large bid below" consists of orders with quantity limited to larger than the best offer. Orders which belong to the category "small bid below" are quantity limited below the best quoted offer. In the Table 5.5, the frequencies of the 14 categories of orders conditional on the previously submitted order for Novartis are reported.[28] Each row corresponds to a type[29] of submitted order in time $t-1$, while each column represents a type of submitted order in time t. The Table is a transition probability matrix, where each row is a probability vector summing up to one. The Table 5.5 documents that the probability that a given type of order occurs after the same type of order has just occurred is larger than that a different type of order occurs. Thus, for example large orders tend to follow large orders or small orders tend to follow small orders.[30] These outcomes also find Biais et al. (1995), whereas compared to our results the probabilities of the persistence of the same order type is smaller.[31]
Several empirical analyzes examine the behavior of limit order submission process and discover that order flows are serially correlated. Additional work has been done in this direction by Danielsson and Payne (2001) for foreign exchange markets, Choi and Lee (2000) for the Korea stock exchange and Yeo (2006) for the NYSE. Furthermore, Parlour (1998) provides a theoretical framework to explain the behavior of the order placement process. She shows that within this model framework, even with absence of asymmetric information and with random arrivals of different trader types, non-random patterns in order flow are observable. Limit order submission is not random and order flow has a non-zero

[28] Albeit the presented estimations are computed for Novartis, the calculations for all other considered assets lead to similar outcomes.
[29] In this case the order type refers to as the category of aggressiveness.
[30] In order to facilitate overview, the largest numbers in each row are in bold type.
[31] Since the results appear in a diagonal matrix Biais et al. (1995) call this occurrence the diagonal effect. However, since the probabilities attribute a characteristic of the order placement procedure, we refer it to as the persistence in the order placement procedure or the persistence in order flow.

first order serial correlation. In particular, the state of the limit order book affects an agent's order submission strategy.[32]

[32]However, the theoretical model of Parlour (1998) suggests a negative serial correlation, while empirical investigations show evidence of a positive serial correlation in the order flow.

Table 5.5: Frequency of orders conditional on the previous type of submitted order

The Table 5.5 reports the order frequency, conditional on the previous type of submitted order for Novartis. The dataset comprises 1,645,246 records gathered between January 2 and December 31 in 2002. On the bid and on the ask side orders are sorted in seven categories. On each side the seven categories are arranged in descending order according to their aggressiveness. The most aggressive order is "large buy" or "large sell", respectively. Each row corresponds to a type of submitted order in time $t-1$, while each column represents a type of submitted order in time t. The Table is a transition probability matrix, where each row is a probability vector summing up to one.

t-1	Large Buy	Market Buy	Small Buy	Large Bid Within	Bid Within	Large Bid Below	Bid Below	Large Sell	Market Sell	Small Sell	Large Ask Within	Ask Within	Large Ask Above	Ask Above
Large Buy	43.46	1.23	35.95	0.83	0.66	2.85	2.44	0.05	0.02	0.61	0.04	0.19	6.62	5.04
Market Buy	8.82	18.39	16.04	7.04	6.46	8.65	7.19	0.14	0.30	3.01	0.97	3.68	11.16	8.15
Small Buy	6.62	0.41	64.70	0.35	2.97	1.61	7.27	0.08	0.06	1.23	0.05	1.42	1.54	11.67
New Bid Within	5.05	4.34	8.88	22.84	8.53	12.61	10.23	0.27	0.87	6.75	1.07	3.52	9.24	5.79
New Bid At	0.67	0.89	16.21	1.69	26.49	2.38	19.32	0.25	0.80	7.76	0.31	7.44	2.05	13.73
New Bid Below	1.54	0.90	5.14	1.56	1.55	46.36	27.48	2.11	0.34	6.80	0.27	0.93	2.94	2.08
Cancel Bid	0.21	0.15	6.53	0.30	3.14	7.80	64.40	1.51	0.32	8.69	0.07	1.96	0.50	4.41
Large Sell	0.05	0.01	0.46	0.06	0.26	4.68	10.21	41.85	1.43	33.98	0.45	1.23	1.36	3.97
Market Sell	0.15	0.29	2.86	1.37	5.84	3.79	11.79	7.83	15.83	15.95	2.25	11.28	5.84	14.93
Small Sell	0.07	0.05	0.99	0.18	1.09	1.68	7.77	4.48	0.36	73.39	0.15	2.46	1.61	5.73
New Ask Within	0.56	1.74	4.49	2.62	3.23	7.18	4.08	4.71	3.35	10.72	20.64	9.71	13.87	13.10
New Ask At	0.16	0.47	7.61	0.67	6.43	0.76	12.53	0.88	1.13	13.19	0.72	28.53	3.44	23.47
New Ask Below	3.99	0.75	3.23	0.94	0.42	10.50	0.88	0.58	0.28	4.49	0.39	1.15	40.33	32.07
Cancel Ask	0.52	0.12	7.17	0.10	1.47	0.20	5.81	0.40	0.16	3.66	0.09	2.19	6.92	71.17

Biais et al. (1995) summarize three alternative economic hypotheses explaining this observation. The first hypothesis states that serial correlation arises due to strategic order splitting. An order is split to reduce market impact, and thus to lessen transaction cost. Another point of view for order splitting is the informational content an order owns. Large orders may signal additional positive private information. This influences order flow in two ways. First, the investor owning such information, sometimes also referred to as insider, repeatedly buys and respectively sells the asset until his private information is revealed in the price. Second, investors observing this trading behavior imitate the insider. Easley and O'Hara (1987) for instance, show that order clustering arises if agents own private information. This informational aspect of order splitting outlines the content of the second hypothesis. The third hypothesis assumes that agents react similarly to the same events, since they own the same economical or technological knowledge. However, for proceeding analyzes what is important is that the order placement is serially correlated.

To get more consolidated findings in the order placement process, we proceed following Biais et al. (1995) and extend the computation of the frequency of orders conditional on the previous type of orders by the analysis of order frequency conditional on the state of the book. The state of the book is differentiated by means of the size of the bid-ask spread and the size of the depth. The threshold of assigning the state is defined by the median of each of the time series. A state with depth at the best quotes below the median of the depth in the considered period is referred to as "small depth." Similarly, if the bid-ask spread of a state is below the median of the bid-ask spread, the state is assigned to the group "small spread." In the Table 5.6 the results of these estimations are represented. Orders are categorized according to the order aggressiveness scheme. With focus on movements around the bid-ask spread, orders are subsumed in groups above, within, at and below the bid-ask spread.

In general, we find that independently what the current state of the book is orders are more frequent within and in particular at the bid-ask spread than below for the buy side and above for the sell side. These findings emphasize the outcomes presented earlier of the frequency of the slopes around the bid-ask spread of a truncated limit order book compared with the entire book. Trading activity basically takes place closely adjacent of the bid-ask spread. Nevertheless, 30% of the orders submitted to the market are orders subjected to less aggressiveness. Additionally, we find that buy orders aimed to be immediately executed are relatively more frequent when spread is tight, while buy orders within the bid-ask spread are relatively more frequent when spread is large. Moreover, sell orders intended with immediate execution are more frequently placed than buy offers. One possible explanation for this effect is that stock markets substantially declined in the year 2002. Sell pressure is always large and the risk to suffer price loss is high. Strategic investors try to sell as fast as possible in order to avoid larger losses. For this reason, order placement of sell orders below the best quotes is preferred.

With respect to the depth of the limit order book we find that orders within the spread occur more often when depth is large. Orders submitted at the quotes are more often when depth is low. This result is in line with the findings of Biais et al. (1995). From an auction theoretic point of view an argument for this market behavior is that when depth is low, a new order limited to the best quote is more likely to be executed than when depth is high. In contrast, when depth is high, an order placed within the quotes are more likely to be executed than at submitted at the quotes. In terms of economic theory to undercut the best quote is the price to obtain time priority.

Table 5.6: Frequency of orders given the state of the limit order book
The Table 5.6 shows the frequency of submitted orders conditional on the previous state of the limit order book. The estimation is based on observations for Novartis between January 2 and December 31 in 2002, which comprises 1,645,246 submitted orders. Each row is a probability vector summing up to one. Orders are categorized according to the order aggressiveness scheme. The categories are simplified by subsuming orders above, within, at and below the bid-ask spread each in one group. The bid-ask spread and depth is considered to be large, if it is larger than its time series median.

	Bid side				Ask Side			
	Buy	Within	At	Below	Sell	Within	At	Above
				Panel A: Large Spread				
Small Depth	13.14	11.11	10.51	14.22	24.48	5.20	6.95	14.40
Large Depth	9.96	9.84	3.81	23.72	29.77	4.70	1.03	17.17
				Panel B: Small Spread				
Small Depth	16.75	13.70	6.25	16.27	19.02	2.43	5.65	19.93
Large Depth	16.28	15.66	1.28	19.85	17.93	2.87	1.03	25.09

Summary In the last two sections we illustrated how data of the limit order book is structured. In the first section we showed evidence that u-shaped intraday volume patterns can also be found for the Swiss market and in particular, this patterns also exists by considering the entire limit order book. Afterwards, in the subsequent section, we illustrated different forms of shape of the limit order book and connected it to auction theory. Measured on the effect of the shape to unexpected volatility we supported the thesis that the entire limit order book contains additional information compared to a truncated data set. With the approach of Naes and Skjeltorp (2006), and later on with the approach of Biais et al. (1995), we disclosed that orders are relatively frequently placed closely around the bid-ask spread and, additionally, that the order flow process is serially correlated.

By continuously refining the perspective of the data respective to the order placement procedure we collected all relevant components to extend the analysis and to build the fundament of our measurement concept. Basically two arguments found in the previous section are important. One aspect is that large depth at the quotes induces more relative frequent order placement at the spread, while a thin book implies order submission within or above the spread. From an auction theoretic perspective this behavior implies mean reversion in market liquidity, as measured, for instance, by the bid-ask spread.[33] The other aspect states that serial autocorrelation is detectable in several elements of the limit order book. With this basis we explore in the next section the limit order book as time proceeds and introduce a new measurement concept for liquidity.

[33]See for instance Handa and Schwartz (1996) or Biais et al. (1995).

5.4 Density Measure Approach

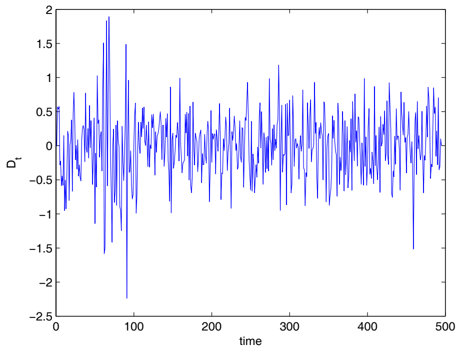
In this section we analyze the dynamic behavior of the limit order book as time proceeds. In the previous section the time aspect played a subsidiary role. We present a density measure approach, which includes and extends findings of Biais et al. (1995) and Naes and Skjeltorp (2006), and many other authors, and in particular the results performed in the previous section. Albeit the analyzes are performed for every considered asset, the results of only two securities are presented in more detail. The outcomes of the remaining assets are mentioned in short overviews or where useful are summarized in tables. The approach is represented by means of the shares of Novartis and Swisslife.
The section is structured as follows. First we start with a short description of the density model approach. Fundamental analyzes of the time series follow, i.e. descriptive statistics, tests of the distribution and stationarity. In order to highlight the density measure from an economical point of view we apply the concept provided by Naes and Skjeltorp (2006) to our time series. Next, analyzes in the time domain follows. Based on these results we continue with model selection and verification. The section closes by providing an appropriate time series model to capture the dynamic of the density measure.

Model Description According to the description in section 2.2.3 the density measure D_t is calculated as the difference between the natural logarithm of the density on the ask side and the bid side. The density on each side is computed as the sum of the volumes $q_{(i,t)}$ weighted with the difference between the mid price P_t^m and the corresponding quoted price $p_{i,t}$. Formally this is

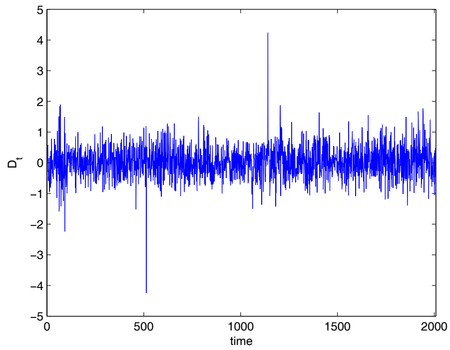
$$D_t = \log\left[\left(\sum_{a=1}^{A}\frac{q_{a,t}^{ask}}{p_{a,t}^{ask}-P_t^m}\right)^{-1}\right] - \log\left[\left(\sum_{b=1}^{B}\frac{q_{b,t}^{bid}}{P_t^m-q_{b,t}^{bid}}\right)^{-1}\right]. \tag{5.5}$$

The integers a, respectively b, represent the position of the offer in the current limit order book on the ask and bid sides. The sum $\sum_i^I q_{i,t}^{ask}$ and $\sum_h^H q_{h,t}^{bid}$ are the sums of the offered volumes in the time interval t.[34]
Due to the construction of the density measure a negative value of D_t reflects

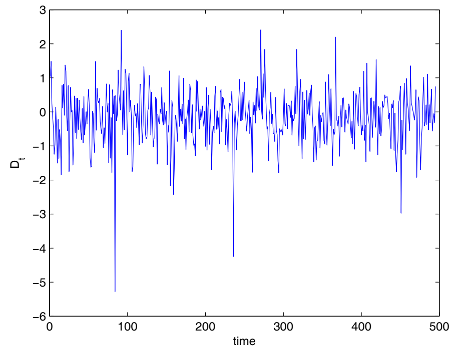
[34]Remember the time interval $t = t_{i+1} - t_i$ is defined as one hour.



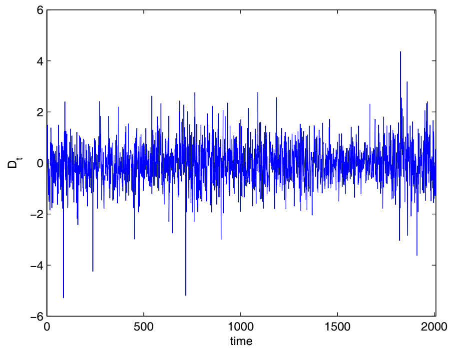
(a) Novartis first Quarter



(b) Novartis one Year



(c) Swisslife first Quarter



(d) Swisslife one Year

Figure 5.6: Order Book Dynamic Based on Density Model The Figure 5.6 illustrates the limit order book dynamic measured with the density measurement approach for Novartis (a) and (b) and for Swisslife (c) and (d). The Figures on the left hand side are restricted to the first quarter of the year and contains 496 observations. The illustrations on the right hand side shows the dynamic for the entire period and comprises 2008 observations.

a supply pressure. A positive value of D_t indicates an excess demand. The smaller the absolute value of D_t the more balanced market situation. Figure 5.6 illustrates the time series of the order book dynamic measured with the density model approach. The sub-Figures 5.6 (a) and (b) show the dynamical behavior of the limit order book of Novartis for the first quarter and for one year. In Figures 5.6 (c) and (d) the corresponding time series for Swisslife are presented.

Since the density is measured in relation to the mid price, the measure takes account of the higher importance of orders placed close to the bid-ask spread. As Biais et al. (1995) find, order placement frequently occurs not only within the bid-ask spread but also at and beyond the spread and influences order placement behavior.[35] In contrast to alternative liquidity measures, for instance the bid-ask spread of our measurement approach includes the order placement process explicitly. Additionally, the density measure incorporates the mechanism of the slope, the dynamic of the entire limit order book and the aspect of mean reversion.[36]

Table 5.7: Descriptive Statistic of Novartis Order Book Dynamic

For the period between January 1 and December 31, 2002, the Table 5.7 reports the descriptive statistic for the limit order book dynamic measured with the density measure according to equation 5.5.

	1^{st} Quarter	2^{nd} Quarter	3^{rd} Quarter	4^{th} Quarter	1 Year
Min	−2.2383	−4.2418	−1.4938	−1.3917	−4.2418
1^{st} Quantile	−0.3014	−0.2970	−0.3071	−0.2957	−0.3021
Mean	−0.0033	0.0149	0.0347	0.0653	0.0280
Median	0.0313	0.0120	0.0313	0.0441	0.0285
3^{rd} Quantile	0.2976	0.2953	0.3345	0.4008	0.3317
Max	1.8921	1.4950	4.234	1.7672	4.2349
Variance	0.2368	0.2488	0.2717	0.2780	0.2591
Std.Dev	0.4866	0.4988	0.5212	0.5273	0.5090
S.E. Mean	0.0218	0.0223	0.0228	0.0236	0.0113
Skweness	−0.1111	−0.9797	1.0357	0.1962	0.0990
Kurtosis	1.6416	10.0860	7.7304	0.1697	4.8989
Total Number	495	496	520	496	2008

Descriptive Statistical Analysis A descriptive statistical analysis delivers first insights of the form and behavior of the process. In Table 5.7 the results of the descriptive statistic to the time series of the order book dynamic of Novartis are collected. The minimum is at -4.2418 and the maximum is at 4.2349 for the one year observation of the process. The lowest value is reached in the

[35]We confirm these findings also for our data set in the previous section.
[36]See a detailed presentation of the measure in chapter 3.

second quarter, whereas the highest value is in the third quarter. Since the process is represented in relative values, this shows that the density in the second quarter is on the ask side more than twice as high as on the bid side. This implicates a large supply pressure for that quarter. The annual mean of the process is 0.0280 and volatility ranges from 0.2368 to 0.2780 with an annual mean of 0.2591. Considering the distribution behavior, the values of the skewness and kurtosis already indicate that the order book dynamic measured with the density model may not be normally distributed. This indication is also supported by the quantiles, in particular for the fourth quarter.[37]

Table 5.8: Descriptive Statistic of Swisslife Order Book Dynamic
For the period between January 1 and December 31, 2002, the Table 5.7 reports the descriptive statistic for the limit order book dynamic measured with the density measure according to equation 5.5.

	1^{st} Quarter	2^{nd} Quarter	3^{rd} Quarter	4^{th} Quarter	1 Year
Min	−5.2811	−5.1915	−2.0397	−3.6221	−5.2811
1^{st} Quantile	−0.5748	−0.4937	−0.5599	−0.4424	−0.5201
Mean	−0.1601	0.0081	−0.0353	0.0372	−0.0374
Median	−0.1315	0.0119	−0.0511	0.0414	−0.0369
3^{rd} Quantile	0.3102	0.4869	0.4279	0.4858	0.4307
Max	2.4112	2.7578	2.7702	4.3643	4.3643
Variance	0.6324	0.7520	0.5444	0.6760	0.6542
Std.Dev	0.7952	0.8672	0.7378	0.8222	0.8088
S.E. Mean	0.0357	0.0389	0.0323	0.0369	0.0180
Skweness	−0.8074	−0.3146	0.1659	0.2453	−0.1765
Kurtosis	4.7691	2.9800	0.3055	2.8772	3.0035
Total Number	495	496	520	496	2008

Table 5.8 summarizes the results of the descriptive statistic of Swisslife. Compared with Novartis, the minimum is with a value of -5.2811 larger. The maximum is similar. The annual mean is negative with -0.0374 and it exhibits the lowest value in the first quarter with -0.1601. Volatility ranges from 0.5444 to 0.6760 with an annual mean of 0.6542. Skewness and kurtosis as well as the values of the quantiles indicate that the data may be normal distributed, in particular for a one year observation.

[37] Analytical test for this assumption follows.

Data Distribution In order to get more insight about the distribution of the data, we perform graphical and analytical distribution tests. Figure 5.7 illustrates the empirical cumulative function (CDF) and the quantile-quantile (QQ) plot of both, Novartis and Swisslife. The sample period is one year. The pictures on the left hand side, Figure 5.7 (a) and (c) are the empirical cumulative distribution functions of the corresponding asset. The dashed lines in these pictures represent the cumulative distribution function of a normal distribution. A closer observation of the plots supports the indications by the values for the skewness and kurtosis. It shows that the empirical cumulative distribution functions have heavier tails than that of the normal distribution, at least for Novartis. Additionally, it is noticeable that the empirical cumulative distribution function of Novartis is heavier tailed than that of Swisslife.

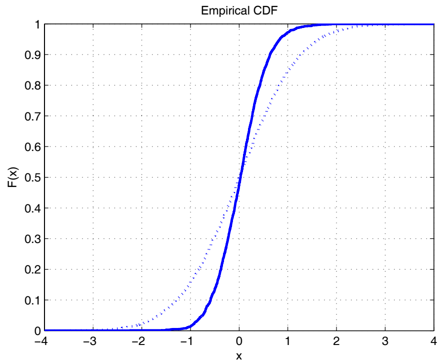
These findings can also be confirmed by the quantile-quantile plots on the right hand side in Figure 5.7, picture (b) and (d) respectively. The dashed line represents the quantiles of a normal distribution. In contrast to the CDF plots, this analysis indicates that the distribution of Novartis is closer to a normal distribution than that of Swisslife.

In addition to the graphical interpretations, analytical investigations should ensure the distribution function of the processes more accurately. First, a test for normality is applied. Miscellaneous test forms for this topic are available. Jarque and Bera (1987), for instance provide a test for normality, the so-called Jarque-Bera-Test. It is a goodness-of-fit measure of departure from normality, based on the sample skewness and kurtosis. It is a test against the null hypothesis that the data are from a normal distribution.

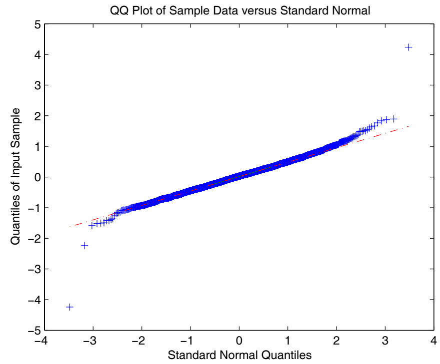
An alternative test for normality provides Kolmogorov and Smirnov. The so-called K-S test compares the empirical cumulative distribution function with the cumulative distribution function specified by the null hypothesis. This test can be applied on different distribution functions. For a test of normality the null hypothesis is the cumulative normal distribution function, according to the dashed lines in Figure 5.7 (a) and (c), respectively.[38]

Lilliefors (1967) extends the test of Kolomogorov and Smirov. This test uses also the null hypothesis that data arises from a normally distributed population, but the null hypothesis does not specify the mean and the variance of the normal

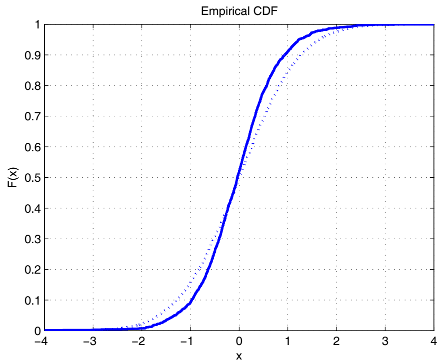
[38] A detailed description of the test provides for instance, Boes, Graybill, and Mood (1974).



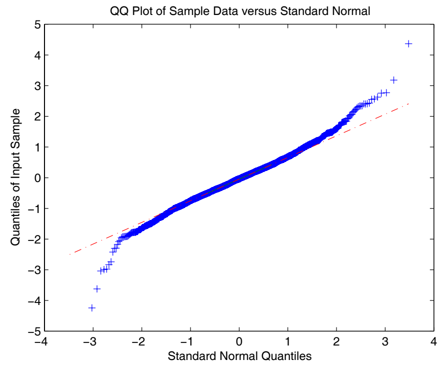
(a) CDF Plot of Novartis



(b) QQ Plot of Novartis



(c) CDF Plot of Swisslife



(d) QQ Plot of Swisslife

Figure 5.7: CDF and QQ Plots of Novartis and Swisslife For the 251 trading days with 2008 observations in the period between January 2 and December 31, 2002, the Figure 5.7 shows the graphical results for tests of normal distribution. The Figures on the left hand side exhibit the empirical cumulative distribution function (solid line) compared with the cumulative distribution function arose from a normal distribution (dashed line). On the right hand side, the empirical quantiles (solid line) are compared to the quantiles of a normal distribution (dashed line).

distribution. Recent literature propose in place of the test of Lilliefors to use the test provided by Shapiro and Wilk (1965). In particular, this test is better for data sets with a large number of observations than the test of Lilliefors.[39]

[39]See for instance Royston (1982).

In Table 5.9 and 5.10 the results of the Jarque-Bera and Kolmogorov-Smirnov tests for Novartis and Swisslife are summarized. All tests are performed at a 5% significance level. The Jarque-Bera critical value at this significance level is 5.9914 for both, all quarters and the one year time series. The corresponding test value against this critical value is presented in the JB-stat row. A larger test value than the critical value leads to a rejection of the null hypothesis. These results are presented in the $\mathcal{H}_0$-column. The null hypothesis $\mathcal{H}_0 = 0$ represents that the data comes from a normal distributed population. A rejection of the null hypothesis $\mathcal{H}_0 = 1$ implies that the data from the time series does not support the assumption of a normal distributed population. The p-values represent the corresponding probability levels for a rejection of the null hypothesis.

Table 5.9: Tests for Normal Distributed Population of Novartis
The Table 5.9 reports tests for analytical test for a normal distributed population as the Jarque-Bera test and the Kolmogorov-Smirnov test of Novartis for several periods. The null hypothesis $\mathcal{H}_0$ states that data comes from a normal distributed population. The critical value for the Jarque-Bera test for a 5% significance level is 5.9914, while the corresponding critical values for the Kolmogorv-Smirnov test are reported in the CV column.

	Jarque-Bera Test			K-S Test				
	$\mathcal{H}_0$	JB-stat	p-Value	$\mathcal{H}_0$	KS-stat	p-Value	CV	N
1^{st} Quarter	1	53.43	0.0000	1	0.1977	0.0000	0.0607	494
2^{nd} Quarter	1	2140.30	0.0000	1	0.1973	0.0000	0.0606	495
3^{rd} Quarter	1	1344.70	0.0000	1	0.1866	0.0000	0.0592	519
4^{th} Quarter	0	3.75	0.1536	1	0.1863	0.0000	0.0606	495
1 Year	1	1987.20	0.0000	1	0.1812	0.0000	0.0302	2008

The critical value for the Kolmogorv-Smirnov test for a significance level of 5% is shown in the CV column. The test values are summarized in the KS-stat row. A larger test value than the critical value leads again to a rejection of the null hypothesis $\mathcal{H}_0$. In the case of Novartis the Jarque-Bera test proposes not to reject all of the null hypothesis. The Kolmogorov-Smirnov test supports a rejection in every single case. Since the two tests deliver contradictory results at least in one case, we additionally apply the Shapiro-Wilk test for further analyses. These test results endorse the outcomes of the KS test. In the case of Swisslife no such contradictory statements exist. All test results show a clear rejection of the null hypothesis.

Table 5.10: Tests for Normal Distributed Population of Swisslife

The Table 5.10 reports tests for analytical test for a normal distributed population as the Jarque-Bera test and the Kolmogorov-Smirnov test of Swisslife for several periods. The null hypothesis $\mathcal{H}_0$ states that data comes from a normal distributed population. The critical value for the Jarque-Bera's test for a 5% significance level is 5.9914, while the corresponding critical values for the Kolmogorv-Smirnov test are reported in the CV column.

	Jarque-Bera Test			K-S Test				
	$\mathcal{H}_0$	JB-stat	p-Value	$\mathcal{H}_0$	KS-stat	p-Value	CV	N
1^{st} Quarter	1	53.43	0.0000	1	0.1977	0.0000	0.0607	494
2^{nd} Quarter	1	2140.30	0.0000	1	0.1973	0.0000	0.0606	495
3^{rd} Quarter	1	1344.70	0.0000	1	0.1866	0.0000	0.0592	519
4^{th} Quarter	0	3.75	0.1536	1	0.1863	0.0000	0.0606	495
1 Year	1	1987.20	0.0000	1	0.1812	0.0000	0.0302	2008

Considering all assets, 78.26% of the Jarque-Bera tests reject the null hypothesis $\mathcal{H}_0$ of a normal distributed population. The Kolmogorov-Smirnov test rejects 93.43% of all null hypotheses and the Shapiro-Wilk test rejects 100.00% of all null hypotheses.[40]Since the sample size accords to the Kolmogorov-Smirnov test, and moreover to the Shapiro-Wilk test, we assume from now on, that the sample data does not come from a normal distributed population.

Stationarity In order to get a first impression of how the order book dynamics evolves as time progress, the evolution of the order book dynamic are plotted. In Figure 5.6 (a) and (c) these plots are presented. The pictures indicate that the processes tend to converge to a mean. The data sample has to be tested for stationarity.[41] Analytical approaches to test for stationarity are provided by Dickey and Fuller (1979), P. Phillips and Perron (1988) or Kwiatkowski et al.

[40] In contrast to these results the Lillifors test rejects 53.91% of all null hypotheses. These findings seem to support the argument that the Lillifors test is not applicable for a large number of observations. For this reason the detailed results of the Lilliefors test are omitted in the thesis.

[41] With stationary process the weak form of stationarity is meant, i.e. the variance is not time dependent and the covariance between u_t and $u_{t+\tau}$ depends only on τ. Consider for a short description of stationarity section 3. A more detailed overview can be found for instance in Hamilton (1994) or McNeil et al. (2005).

(1992).[42]

The test of Dickey and Fuller (1979) tests whether a unit root is present in a time series sample. A unit root is not stationary. The null hypothesis $\mathcal{H}_0$ of the Dickey-Fuller test proves the unit root assumption in the time series sample. A rejection of the null hypothesis gives evidence that the time series sample is stationary.[43] Said and Dickey (1984) augment the basic unit root test to accommodate general autoregressive moving average (ARMA) models with unknown orders. Their test is referred to as the augmented Dickey-Fuller (ADF) test.

The test approach provided by P. Phillips and Perron (1988) computes for the null hypothesis $\mathcal{H}_0$ also that the time series sample has a unit root against a stationary alternative. The Phillips-Perron (PP) unit root tests differ from the ADF tests mainly in how they deal with serial correlation and heteroskedasticity in the errors. In particular, where the ADF tests use a parametric autoregression to approximate the ARMA structure of the errors in the test regression, the PP tests ignore any serial correlation.[44] In contrast to these two tests, the test approach provided by Kwiatkowski et al. (1992), usually abbreviated by KPSS, proves the null hypothesis $\mathcal{H}_0$ that the data sample is stationary against the alternative of a unit root.

The ADF test statistic with lag=0 for both Novartis and Swisslife are reported in Table 5.11 and 5.12, respectively. For the 1^{st} quarter of Novartis, for instance, the ADF test statistic is -0.9235 with a t-value of -20.57 and has a corresponding p-value of 3.388×10^{-39}. Hence, the null hypothesis is not rejected at a $3.79 \times 10^{-35}\%$ level. Since analysis is based at a 5% significance level the null hypothesis in this case is rejected. A rejection of the null hypothesis at a 5% significance level is denoted with $\mathcal{H}_0 = 1$. Table 5.11 reports for every single case a rejection of the null hypothesis. The results obtained by the ADF-test are also confirmed by the PP and KPSS tests.[45]

Similar results are found for Swisslife. Here as well, all null hypotheses are

[42] There are more than these mentioned approaches. These three tests are the most common ones, see Enders (2004).

[43] Consider section 3 or Hamilton (1994), Enders (2004), or McNeil et al. (2005) among many others.

[44] See for a more detailed description Enders (2004).

[45] The results are available on request.

Table 5.11: Unit Root Test for Novartis

For the period between January 2 and December 31, 2002, the Table 5.11 reports for several subperiods the test results of the augmented Dickey-Fuller (ADF) test for unit root and stationarity, respectively. The ADF test is performed for a lag level of zero, with no intercept nor a time trend. The results of the Durbin-Watson test are documented in the DW column.

	$\mathcal{H}_0$	ADF	S.E.	t-value	p-value	adj. R^2	DW	N
1^{st} Quarter	1	-0.9235	0.0449	-20.57	$3.388e{-}39$	0.4618	1.9831	494
2^{nd} Quarter	1	-0.9535	0.0450	-21.20	$1.422e{-}39$	0.4765	1.9925	495
3^{rd} Quarter	1	-0.8704	0.0436	-19.97	$7.911e{-}39$	0.4350	2.0194	519
4^{th} Quarter	1	-0.9596	0.0450	-21.34	$1.228e{-}39$	0.4797	2.0023	495
1 Year	1	-0.9265	0.0223	-41.60	$6.416e{-}60$	0.4632	1.9990	2008

rejected at a 5% significance level, indicated by $\mathcal{H}_0 = 1$ in Table 5.12. These findings are also confirmed by the PP and KPSS tests.[46]

Table 5.12: Unit Root Test for Swisslife

For the period between January 2 and December 31, 2002, the Table 5.12 reports for several subperiods the test results of the augmented Dickey-Fuller (ADF) test for unit root and stationarity, respectively. The ADF test is performed for a lag level of zero, with no intercept nor a time trend. The results of the Durbin-Watson test are documented in the DW column.

	$\mathcal{H}_0$	ADF	S.E.	t-value	p-value	adj. R^2	DW	N
1^{st} Quarter	1	-0.8953	0.0448	-20.00	$9.059e{-}39$	0.4480	2.0085	494
2^{nd} Quarter	1	-0.9678	0.0450	-21.51	$1.040e{-}39$	0.4838	1.9983	495
3^{rd} Quarter	1	-0.9263	0.0437	-21.21	$1.142e{-}39$	0.4649	2.0088	519
4^{th} Quarter	1	-0.8977	0.0448	-20.05	$8.079e{-}39$	0.4489	1.9999	495
1 Year	1	-0.9234	0.0223	-41.47	$3.566e{-}60$	0.4616	2.0036	2008

An important issue for the implementation of the ADF test is the specification of the lag length. If it is too small then the remaining serial correlation in the errors will bias the test. The efficiency of the test will suffer if the length of the lags are too large. The Durbin-Watson (DW) test, provided by Durbin and Watson (1950) and Durbin and Watson (1951) proves a correlation of the errors. A DW-value between 1.5 and 2.5 shows evidence that the residuals are not correlated for the given lag length. The DW values computed for the two examples

[46] Additional results are available on request.

support an efficient setting of the ADF test.[47]
Since a constant parameter as well as a trend in the time series cannot be observed directly from the plots, we perform additionally ADF tests and PP tests with constants, trends and constants combined with trends. Like in all other cases, the results are not affected. The null hypotheses of all considered assets are rejected even on a 0.01% level.

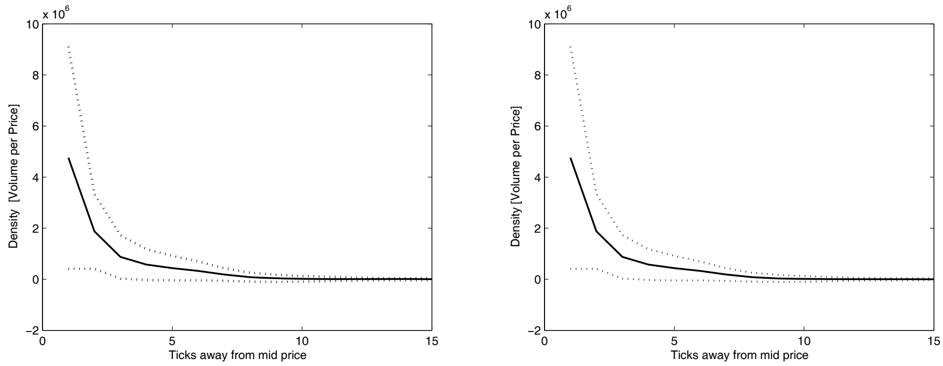
We perform all the tests with different setting to all considered assets. The ADF and PP tests for stationarity reject all null hypotheses even on a 0.01% significance level, such that additional tests like the KPSS tests are omitted. Based on that statistical analysis and findings, for the subsequent part of the analyses it can be assumed that the dynamic of the limit order book is a stationary process, without trend or constant and with no normal distributed innovations.

Density Measure, Return and Volatility Relation So far, we have performed an analysis of basic characteristics of the time series generated by the density measure approach. Before we proceed to analysis regarding the time domain we investigate the relation of the density measure to return and volatility. In order to examine the relation to volatility we apply the same approach as already used for the slope measure in a previous section.

In Figure 5.8 we illustrate the density against the depth of the limit order book. The depth level is determined by the ticks away from the mid price. The density is computed for each depth level separately and collected for the whole sample period. For each depth level we calculate the median of the density and its corresponding standard error. In Figure 5.8 the solid line represents the median of the density on each tick level, while the dashed lines above and below the solid line are the corresponding standard errors. The Figure 5.8 reveals that depth decreases disproportionately by increasing tick level. In addition it shows that the variance of the density adjacent to bid-ask spread is larger, than more ticks

[47] Nevertheless tests with different lag lengths have been performed. The lag lengths thereby have been determined and implemented according to Schwert (1989) or Ng and Perron (1995). Tests for different lag lengths show similar results. The null hypotheses are rejected for all tests.

away.[48] Due to the construction of the density measure this means that order sizes close to the bid-ask spread varies more than far away of the bid-ask spread. In combination with the analysis of the order flow we find that orders are placed more frequently close to the bid-ask spread, order sizes are generally larger and in addition vary more around the bid-ask spread than deeper in the book.



(a) Bid Density (b) Ask Density

Figure 5.8: Book Density as a Function of Depth The Figure 5.8 illustrates the limit order book density as a function of the depth of the book, while the depth is indexed by the ticks away from the mid quote. On the left hand side, the density of the bid side is shown, while on the right hand side the density of the ask side is illustrated. Estimations are based on 1,645,246 observations for Novartis.

From an economical point of view this behavior may have several reasons. As already mentioned earlier, agents may disagree about the real value of an asset. This is reflected in the slope of the limit order book. Looking at this interpretation in more detail, we assume that larger disagreement in real value of the asset is reflected by a small slope of the book. In contrast small nuances in the

[48] Analytical tests support this statement. We apply a right-sided t-test between the time series generated by the density at the quotes and 10 ticks deeper in the book. We perform the test for each side separately. The null hypothesis states that the means do not differ. The null hypothesis for the bid side is rejected by a test value of 40.244 compared to the critical value of 0.266 for a 0.01 % significance level. A similar result is found for the ask side. The test value amounts 20.975, while the critical value for the same significance level is 0.299. In order to affirm these results we additionally calculate an analysis of the variances. The F-value is for the bid side 1619.44 and for the ask side 439.95. Both induce a rejection of the null hypothesis on a 0.01 % significance level.

belief of the real value are reflected by the occurrence of small variance of the order sizes around the bid-ask spread.
The Table 5.13 shows how the density measure is related to other liquidity and trading activity measures. It represents the Pearson's correlations between averages of corresponding measures of 23 assets. We find positive correlations between the numbers of executed orders respective of trades and the density measure. As already expected from previous findings the correlation is stronger the more the book is restricted. A similar picture is shown for trading volumes. Market capitalization is positively correlated to the density measure while the bid-ask spread has a negative relation. This indicates that larger firms generally feature higher density in the limit order book and exhibit smaller bid-ask spreads. Thus considered economically, larger firms are commonly more liquid than smaller firms.

Table 5.13: Correlation between density and liquidity measures
The Table 5.3 reports Pearson's correlations of the results of 23 assets for density with several liquidity and trading activity measures. Calculations are based first on the entire limit order book and afterwards on a book truncated to 15 or 10 ticks away from the best quote. Calculation is based on 2008 observations for every security.

	Slope (Full order book)	Slope (±15 ticks)	Slope (± 10 ticks)
Number of Trades	0.035	0.176	0.247
Trading Volume	0.176	0.334	0.301
Market Capitalization	0.207	0.220	0.070
Spread	−0.150	−0.781	−0.961
Slope (Full order book)	1.000		
Slope (+/- 15 ticks)	0.767	1.000	
Slope (+/- 10 ticks)	0.466	0.615	1.000

To examine the relation between the density measure and price volatility we apply the analogous analytical investigation already performed to the slope of the limit order book. Table 5.14 shows the results of the regression between the density measure and several combinations of parameters. We find that the parameter estimations for the density are negative and weak significant. The estimations for the density across all parameter combinations are robust even

Table 5.14: Price Volatility and Density Regression

The Table reports the estimation results for Novartis for the entire limit order book of four different versions of the regression.

$$\begin{aligned} \mid \hat{\varepsilon}_{i,s} \mid &= \beta_1 N_{i,s} + \beta_2 OV_{i,s} + \beta_3 MCAP_{i,s} + \beta_4 SPREAD_{i,s} \\ &+ \beta_5 DENSITY_{i,s} + \sum_{j=1}^{12} \rho_{i,j} \mid \hat{\varepsilon}_{i,s} \mid + \eta_{i,s}. \end{aligned}$$

The regressand $\mid \hat{\varepsilon}_{i,s} \mid$ is the absolute value of the unexpected return on security i in the snapshot s obtained by equation (5.3). N is the number of executed trades within the considered snapshot, OV represents the average order book volume, $MCAP$ the market capitalization, $SPREAD$ the bid-ask spread and $DENSITY$ the density of the entire limit order book. For each model in the first row the parameter estimations are represented, while in the second row the corresponding t-values of the parameter estimations are shown. In model 2 and 3, negative and positive parameters are omitted. In model 4 we test only for the effect of the density measure. The parameter estimation results for the volatility persistence are excluded. Parameter estimations denoted by * are not significant at a 1 percent level. Estimation is based on 2008 observations.

	Model 1		Model 2		Model 3		Model 4	
Variables	$\widehat{\beta}$	t-Value	$\widehat{\beta}$	t-Value	$\widehat{\beta}$	t-Value	$\widehat{\beta}$	t-Value
N	0.000	−2.038	0.000	−0.322				
OV	0.159	4.727	0.053	10.298				
Market Cap	−0.139	−3.060			0.070	14.054		
Spread	−0.214	−2.677			0.159	13.593		
Density	−0.002*	−0.385	−0.002*	−0.322	−0.003*	−0.493	−0.004*	−0.698
Adj. R^2	0.212		0.198		0.208		0.126	
F-test	543		498		528		291	
DW	2.033		2.026		2.036		1.999	

if all other explanatory variables are removed. This suggests that the density measure captures weak liquidity effects on volatility. The adjusted R^2 shows that approximately 12% of the variation in the response variable can be explained by the density measure. An F-value of 291 shows evidence of the significance of the adjusted R^2 on a 0.01% significance level.

In order to examine the relation between the density measure and the return we perform a linear regression analysis. The first analysis simply regresses the density measure to the return. We find that the Durbin-Watson value indicates autocorrelation in the residuals. We re-estimate the regression with a correction

Table 5.15: Price Return and Density Regression

The Table 5.15 shows the results of the regression analysis between the density and the return. The estimation referred to as CO regress is corrected for autocorrelation in the error terms according to the approach provided by D. Cochrane and Orcutt (1949). The regression results labeled by MA regress is the estimation approach proposed by C. M. Jones et al. (1994) with a 12 lagged moving average component of the return as a correction term. Estimations are based on 2008 observations for Novartis. Parameter estimations and F-tests are significant on a 0.01% significance level.

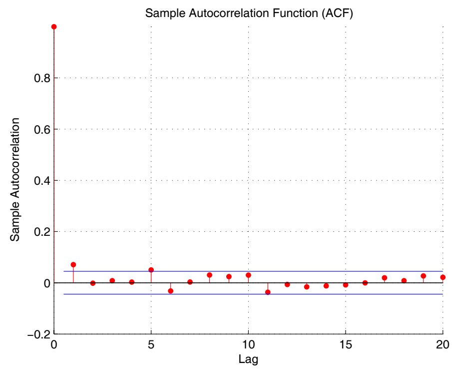
	$\hat{\beta}$	t-value	Adj. R^2	F-value	DW
CO regress	0.110	4.956	0.011	24.340	1.995
MA regress	0.035	4.247	0.028	58.950	2.028

for autocorrelation in the residuals with the approach provided by D. Cochrane and Orcutt (1949). The results are reported in Table 5.15 in the CO regress row. The regression shows a positive and significant relationship between the density measure and the return. Albeit the F-value evidences high significance of the R^2, only 1.1% of the variation in the return can be explained by the density measure. The alternative regression analysis, the Jones regression, shows similar results. This regression is based on the approach provided by C. M. Jones et al. (1994). The approach includes a moving average component of the return as an additional explanatory variable in order to smooth outlier effects.[49] While this second approach increases the explanatory power regarding variation, it remains still on a low level of 2.8%.

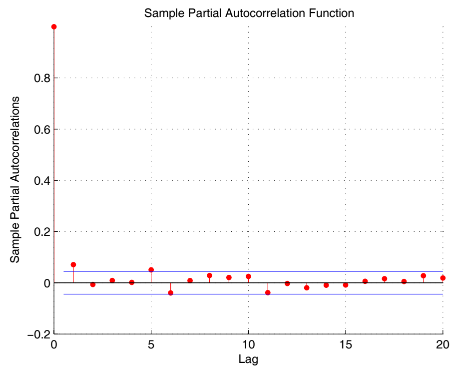
Time Domain Analysis Stationary processes with and without normally distributed errors can be captured and reproduced by several concepts and different models. A review of the serial correlation should give closer insight to the properties of time series generated by the density measure approach. In Figure 5.9 the correlograms of Novartis and Swisslife for the period of one year are illustrated. Correlations can be detected in both the sample autocorrelation (ACF) and the sample partial autocorrelation (PCF), respectively. The dependency characteristics are in favor of models of stationary univariate autoregressive moving average

[49]It is basically the same approach as presented in equation 5.4 with a 12 lagged return regressor.

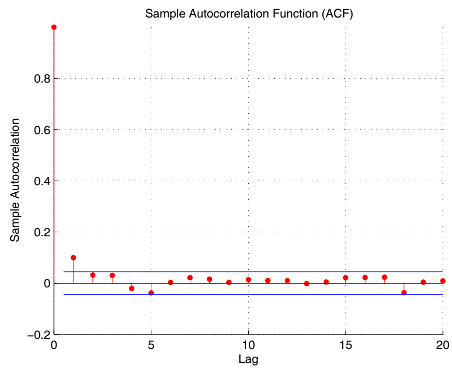
(ARMA) classes.[50]



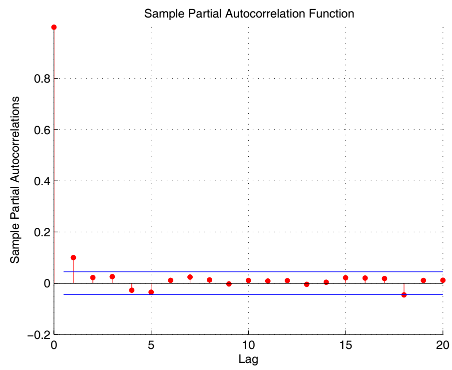
(a) ACF of Novartis



(b) PCF of Novartis



(c) ACF of Swisslife



(d) PCF of Swisslife

Figure 5.9: ACF and PCF Plots The Figure 5.9 illustrates the empirical autocorrelation function (ACF) and the partial autocorrelation function (PCF) for the density measure approach for Novartis and Swisslife. The estimations are based on 2008 observations.

Model Selection Criteria and Parameter Estimation A widely used approach to ascertain the orders to an ARMA process is to analyze the correlograms

[50]For a short description about premises, properties an characteristics of ARMA models consider section 3. A more detailed overview is provided by Greene (2000) or Hamilton (1994), among many others.

of the data sample.[51] The last significant lag of the sample partial autocorrelation function (PCF) determines the order of the autoregressive (AR) part. In order to assign an order to the moving average (MA) part of the model, the sample autocorrelation (ACF) function has to be considered. The last significant lag represents the order of the MA part.

A standard 95% significance level is depicted with a horizontal solid line in Figure 5.9. The range in between the lines marks all not significant lag parameters. Significant parameters are illustrated beyond the solid lines. According to this interpretation the limit order book dynamic of Novartis for a one year period is reflected best with an ARMA(1,1) process.[52] For the case of Swisslife we find similar results. The correlogram analysis suggests again an ARMA(1,1) process.

Alternatively to the just mentioned approach, a statistical model selection criteria may be used. The two most common information criterion are the Akaike information criterion (AIC) proposed by Akaike (1974) and Schwarz-Bayesian information criterion (BIC), provided by Schwarz (1978). Both approaches use a certain combination of the number of parameters of the considered model and the corresponding result of the log-likelihood function to rule a decision.[53] According to Wagenmakers, Grünwald, and Steyers (2006) what is important to model selection is to keep the model as general as possible. As Zivot and Wang (2003), among others, state the AIC criterion asymptotically overestimates the order of the model, whereas the BIC estimate the order more consistently. But for finite samples the BIC shares no particular advantage over the AIC.

We perform AIC as well as BIC calculations and choose the model according to the minimizing of the information criterion rule. Since the BIC criterion suggests in more cases a more simple model we rely the analysis mainly on the BIC criterion. In Table 5.16 the results of the Schwarz-Bayesian information criterion for different ARMA models estimated for Novartis for a period of one year are reported. Similar results are presented in Table 5.17 for Swisslife.[54]

In both cases an increase of order enhances the value of the information criterion.

[51] A good introduction to estimation and forecasting methods see for instance Box, Jenkins, and Reinsel (1994), Granger and Newbold (1986), Mills (1990), Enders (2004).

[52] The 5th lag is graphically hard to assign. Analytically it is in between the no-significance band.

[53] For more information consider section 3, or Zivot and Wang (2003), for instance.

[54] The results for every quarter are available on request.

The lower the order, the lower the BIC such that only models with a maximal order of 1 have to be taken into account.[55] According to the minimization of the information criterion rule, the Schwarz-Bayesian information criterion proposes an ARMA(1,0) process for both, Novartis and Swisslife, respectively. This analytical investigation differs of the conclusions resulting from the correlogram analysis.

Table 5.16: BIC values for different ARMA models

The Table 5.16 represents the Schwarz-Baysian information criterion (BIC) for different ARMA(p,q) models. In the rows the p-values of the model increases. In the columns the q-values are heightened. The ARMA model is estimated with t-distributed innovations and estimations are based on 2008 observations of Novartis.

p\q	0	1	2	3	4
0	0	2895.6	2902.9	2910.4	2918.0
1	2895.4	2902.9	2909.6	2917.2	2923.7
2	2902.9	2909.6	2917.2	2920.0	2925.1
3	2910.4	2917.2	2919.4	2931.8	2928.5
4	2918.0	2923.9	2925.2	2936.0	2942.1

Since we have proven that the process does not have normal distributed errors, estimations represented in Table 5.16 and 5.17 are based on t-distributed innovations. The BIC value for an ARMA(1,0) process with gaussian distributed errors for the same period for Novartis is 2996.5 and that for Swisslife is 4853.3. Comparing these findings with the corresponding results in Table 5.16 and 5.17 shows evidence for a better fit by t-distributed innovations.
The BIC model selection criterion suggests in 100% of the cases an ARMA(1,0), ARMA(0,1) or an ARMA(1,1) model. In 91.30% of all considered time series either an ARMA(1,0) or an ARMA(0,1) model is selected. The ARMA(1,0) model is chosen in 56.62%. But in cases where the ARMA(0,1) model is preferred to an ARMA(1,0) model, the difference between the information criterion is small. For instance for Zurich's third quarter the BIC value for the ARMA(0,1) is 976.38. The BIC value for the ARMA(1,0) model is 976.48. According to the next higher BIC value the ARMA(1,1) model is suggested with a BIC value of 982.25. In order to be able to compare the dynamic behavior of the limit

[55] This is in contrast to the Akaike information criterion. The results vary between several higher ordered models, such that no exact model can be assigned to.

Table 5.17: BIC values for different ARMA models

The Table 5.17 represents the Schwarz-Baysian information criterion (BIC) for different ARMA(p,q) models. In the rows the p-values of the model increases. In the columns the q-values are heightened. The ARMA model is estimated with t-distributed innovations and estimations are based on 2008 observations of Swisslife.

p\q	0	1	2	3	4
0	0	4729.0	4732.8	4739.8	4746.6
1	4727.7	4730.9	4738.3	4745.9	4753.0
2	4732.2	4738.3	4740.6	4753.5	4756.7
3	4739.3	4745.9	4748.1	4750.8	4758.2
4	4746.1	4753.1	4756.7	4758.2	4765.1

order book of all considered securities with each other, the ARMA(1,0) model is defined for all processes. It is clear that in cases the model does not fit best, some accuracy losses have to be taken into account. But in the majority of the cases where the ARMA(1,0) model is not the first choice, it is close to the first choice and therefore the bias is small. In addition, the analysis of the correlograms supports several times an ARMA(1,0) model.

Table 5.18 and 5.19 represent the parameter estimations of the ARMA(1,0) model based on t-distributed innovations. The corresponding degrees of freedom of the t-distribution are reported in the DoF row. Except for the autoregressive parameter of the fourth quarter of Novartis all parameter estimations are statistically significant at a 5% significance level.[56]

Verification of the Model Selection To ensure that the model fit is adequate, we perform an analysis of the residuals. This analysis checks whether independency and heteroscedasticity are still present in the innovations. The analysis is mainly based on the correlograms, the DW test statistics, the Ljung-Box test, Engle's test of heteroscedasticity as well as the Goldfeld Quandt test. Figure 5.10 depicts the residuals of the estimated ARMA(1,0) model and its corresponding correlograms. In order to gain insight into the independency of the residuals, consider the correlograms in Figure 5.10 on the right hand side. No significant autocorrelation can be detected, neither in the ACF in the up-

[56] Statistical significant parameters at a 5% significance level are marked with a single asterisk. Estimations significant at a 10% significance level are highlighted with a double asterisk.

Table 5.18: Parameter Estimation for ARMA(1,0) Model

The Figure 5.18 represents the parameter estimations for ARMA(1,0) models for Novartis. The values in brackets below the parameter estimations are the corresponding t-values. The σ^2 represents the volatility of the model, DoF specifies the degrees of freedom for the t-distributed innovations, N is the number of observations for the different periods and DW is the corresponding Durbin-Watson value. T-values labeled with an asterisk indicate a statistical significance on a 10% significance level while parameter estimations without an asterisk are significant on a 5% significance level.

	p	σ^2	DoF	N	DW
1^{st} Quarter	0.094	0.234	7.452	494	2.017
	(2.385)	(12.113)			
2^{nd} Quarter	0.064	0.231	7.980	495	2.022
	(1.7404)	(12.882)			
3^{rd} Quarter	0.102	0.254	9.228	519	1.960
	(2.462)	(13.321)			
4^{th} Quarter	0.040	0.282	24.153	495	1.999
	$(1.000)^*$	(14.435)			
1 Year	0.074	0.251	9.114	2008	2.000
	(3.770)	(26.638)			

per Figure nor in the PCF picture in the lower Figure. These results are also supported by the Durbin-Watson values in Table 5.18 and 5.19. Since the DW statistic is in the range of 1.5 and 2.5, this indicates a negative but low, not significant autocorrelation. This negative autocorrelation can also be observed by the oscillation of the values around zero in the ACF and PCF pictures in Figure 5.10 (b) and (d), respectively.

The portmanteau test of Ljung and Box (1978) is commonly used to reveal model misspecification of a time series model. If no significant autocorrelation is found in the residuals from the fitted model, the model is declared to pass the test. The advantage of the Ljung-Box test is that it tests for different lags at the same time. The null hypothesis $\mathcal{H}_0$ is that the model fit is adequate to the alternative of the existence of autocorrelation. The test is based on Q-statistic. In Figure 5.11 the p-values for the Ljung-Box test for 200 lags are illustrated. The hypothesis test is performed on a 5% significance level, which is depicted with a solid line in the Figure 5.11. Each time the graph of the p-Value's hits below the 5% significance level the null hypothesis is rejected for the corresponding lag length. This is

Table 5.19: Parameter Estimation for ARMA(1,0) Model

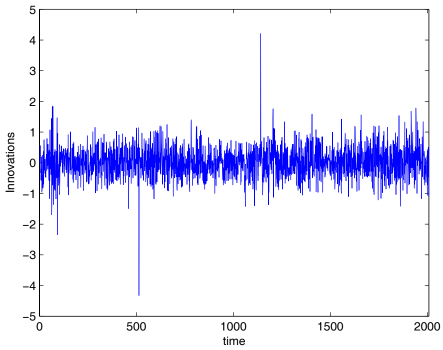
The Figure 5.19 represents the parameter estimations for ARMA(1,0) models for Swisslife. The values in brackets below the parameter estimations are the corresponding t-values. The σ^2 represents the volatility of the model, DoF specifies the degrees of freedom for the t-distributed innovations, N is the number of observations for the different periods and DW is the corresponding Durbin-Watson value. T-values labeled with an asterisk indicate a statistical significance on a 10% significance level while parameter estimations without an asterisk are significant on a 5% significance level.

	p	σ^2	DoF	N	DW
1^{st} Quarter	0.102	0.633	5.366	494	1.996
	(2.599)	(10.422)			
2^{nd} Quarter	0.060	0.754	5.427	495	2.055
	$(1.378)^*$	(9.574)			
3^{rd} Quarter	0.085	0.542	20.746	519	2.020
	(1.861)	(14.680)			
4^{th} Quarter	0.093	0.671	5.208	495	1.984
	(2.386)	(9.584)			
1 Year	0.089	0.646	6.117	2008	2.028
	(4.304)	(22.106)			

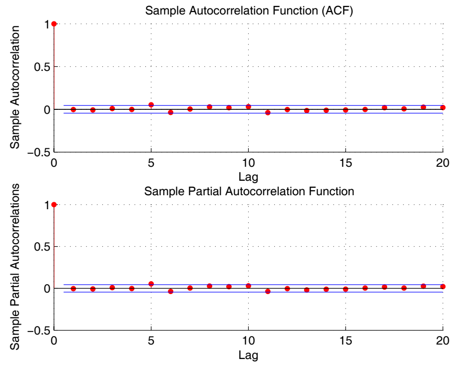
the case, several times for the first quarter, the third quarter and the one year period.

Although the Durbin-Watson test as well as the graphical analysis by correlogram shows evidence of no autocorrelation in the residuals, the Ljung-Box test does not support these results in every case. In particular an increasing lag length entails significant autocorrelation. Similar results are also found for all considered assets.

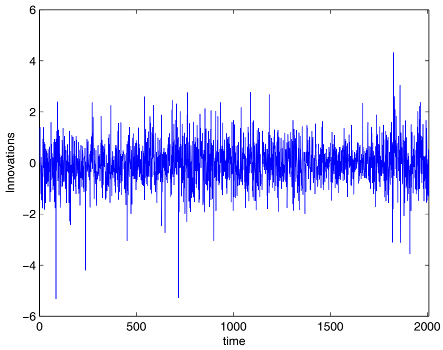
Beside tests for autocorrelation we perform tests of the existence of heteroscedasticity in the residuals. The test for heteroscedasticity is provided by Goldfeld and Quandt (1965). The null hypothesis $\mathcal{H}_0$ assumes a homoscedastic time series. In Table 5.20 the results of the Goldfeld-Quandt test for Novartis are summarized. The Goldfeld-Quandt test statistic (GQ stat) is compared with a critical value (CV). The critical value is calculated on a 5% significance level. A larger test statistic than the critical value leads to an acceptance of the null hypothesis, which is designated with $\mathcal{H}_0 = 1$. The corresponding p-value shows the significance level at which the null hypothesis is rejected. In three of five



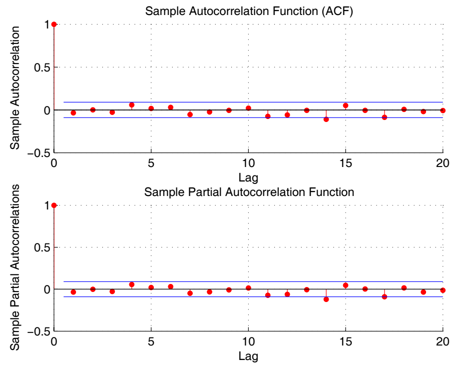
(a) Residuals of Novartis



(b) ACF and PCF of the Residuals



(c) Residuals of Swisslife



(d) ACF and PCF of the Residuals

Figure 5.10: Residual Analysis of ARMA(1,0) The Figure 5.10 illustrates the innovations, figure (a) and (c) and the corresponding empirical autorcorrelation function (ACF) and the partial autocorrelation function (PCF) of both, Novartis (a) and (b) and Swisslife (c) and (d). Estimtations are represented on 2008 observations.

cases the assumption of homoscedasticity is rejected and heteroscedasticity can be assumed.

An additional test provided by Engle (1982) investigates for the presence of autoregressive conditional heteroscedasticity (ARCH) effects in the time series. The null hypothesis $\mathcal{H}_0$ states that no ARCH effects are detectable in the residuals. A rejection of the null hypothesis is marked with $\mathcal{H}_0 = 1$. In Table 5.20, the results for this test are shown. ARCH effects can be detected in three of five

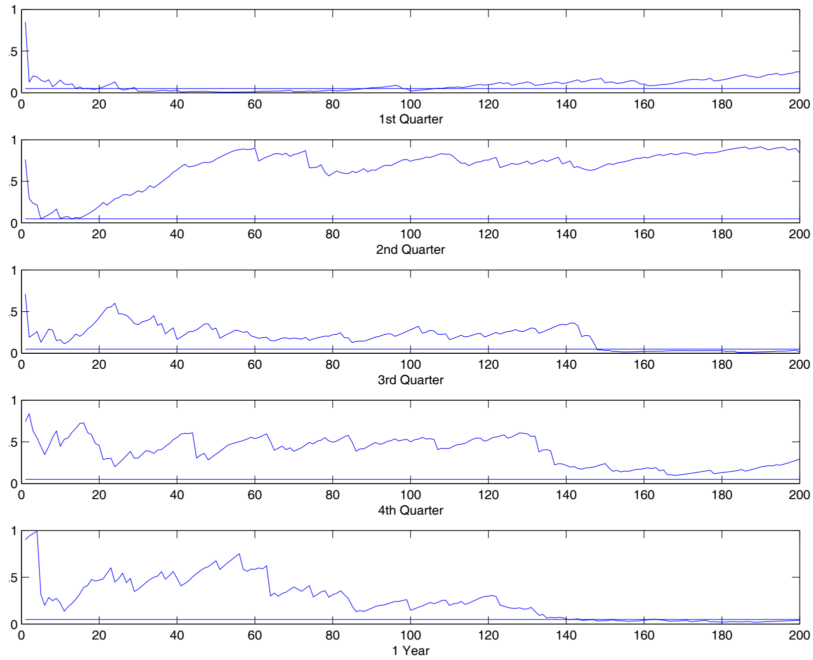


Figure 5.11: Ljung-Box for ARMA(1,0) The Figure 5.11 shows the Ljung-Box test for the ARMA(1,0) model for different periods. The null hypothesis $\mathcal{H}_0$ is that the model fit is adequate to the alternative of the existence of autocorrelation. The hypothesis test is performed on a 5% significance level which is illustrated with a solid line on the bottom of each picture. The p-values for a rejection of the null hypothesis is on the ordinate while the corresponding lag length is on the x-axis.

cases.

Although residuals are for the most part not correlated, based on the results of the ARCH test, they are not stochastically independent. Hence, the innovation process of the ARMA(1,0) model features volatility-clustering.

In order to gain a more appropriate model based on the results of the analysis of the residuals, the ARMA model has to be extended with at least heteroscedastic volatility. This can be done by adding a generalized autoregressive conditional heteroscedastic (GARCH) component to the ARMA model.[57]

[57]For further information about ARMA and ARMA-GARCH models, consider Hamilton

Table 5.20: Engle's ARCH Test and Goldfeld-Quandt Test

The Table 5.20 reports the results for Engle's test for ARCH effects and Goldfeld-Quandt test for heteroscedasticity. The null hypothesis $\mathcal{H}_0$ states that there is no heteroscedasticity in the time series. An acceptance of the null hypothesis on a 5% significance level is assigned by $\mathcal{H}_0 = 1$. The critical value (CV) for the ARCH test is 3.8414 for all periods. The critical value (CV) for the Goldfeld-Quandt test is reported in the CV column. The test values stands in the ARCH stat column respectively in the GQ stat column. The corresponding p-values are in the next column.

	$\mathcal{H}_0$	ARCH stat	p-value	$\mathcal{H}_0$	GQ stat	CV	p-value	N
1^{st} Quarter	1	28.879	0.000	1	1.433	1.234	0.002	494
2^{nd} Quarter	0	3.196	0.073	1	1.925	1.233	0.000	495
3^{rd} Quarter	0	0.013	0.908	1	1.366	1.227	0.006	519
4^{th} Quarter	1	13.375	0.000	0	0.765	1.233	0.981	495
1 Year	1	11.156	0.000	0	0.870	1.109	0.985	2008

ARMA-GARCH Model As pointed out above, the analysis of the residuals of the ARMA(1,0) model suggests an extension of the ARMA(1,0) model. The extension should include the found properties of the residual analysis. First, it should satisfy the heteroscedastic behavior. Second, with respect to the non gaussian distribution, the volatility model should have other distribution possibilities. To augment the ARMA model with these characteristics the generalized autoregressive conditional heteroscedastic volatility (GARCH) model provided by Bollerslev (1986), which is a generalized model of the model proposed by Engle (1982), is taken into account.[58]

Model Selection Criteria and Parameter Estimation The parameter selection criteria for finding an accurate number of parameters for a GARCH variance model is similar to the approach already applied to the ARMA model in earlier sections. Beside the analysis of the correlograms, the Akaike information criterion (AIC) and the Schwarz-Baysian information criterion (BIC) are used to determine the order of the model. Albeit these approaches have been developed to define the order of ARMA models, they obtained also practical relevance for

(1994), Greene (2000), Enders (2004) or McNeil et al. (2005).

[58]Robert F. Engle and Clive W.J. Granger have been honored with the Nobel prize for that model in the year 2003.

GARCH modeling.[59]
Hence, the analysis proceeds as follows. First the ARMA setting is fixed. Then various GARCH settings are added to the given ARMA model and the corresponding information criteria are computed. The GARCH setting, which minimizes the information criteria is selected. Finally, the GARCH setting is fixed and different ARMA settings are tested and compared with each other.[60] The combination with the lowest information criteria represents the most accurate model setting.
In order to gain a proper analysis we perform four different tests. On the one hand the BIC criteria is used. On the other hand, the AIC criteria is performed. Each of the approaches are estimated with both, gaussian distributed and t-distributed innovations, respectively.[61] In the Table 5.21 the results of the BIC calculation for Novartis for the one year period are reported. The estimations are based on an ARMA(1,0) model without a constant element, which is combined with several GARCH components. The two GARCH parameters α and β vary between one and four.

Table 5.21: BIC values for ARMA(1,0)-GARCH(α,β)
The Table 5.21 reports the Schwarz-Baysian information criterion (BIC) for ARMA(1,0) models with different GARCH extensions for Novartis. The α-GARCH component increases with increasing rows and the β-GARCH component increases with additional column. On the left hand side, estimations are based on t-distributed innovations. On the right hand side, estimations are performed for gaussian distributed errors. The estimation period is one year with 2008 observations.

	T-distributed Errors				Gaussian Errors			
$\alpha\backslash\beta$	1	2	3	4	1	2	3	4
1	2846.4	2848.4	2850.4	2852.3	2919.6	2921.6	2920.9	2922.9
2	2847.5	2849.5	2851.5	2853.5	2918.3	2920.2	2922.2	2924.2
3	2845.9	2847.9	2849.9	2851.9	2920.3	2922.2	2924.2	2926.2
4	2846.9	2849.0	2851.0	2853.0	2921.3	2923.2	2925.2	2927.2

The reported results constitute the lowest BIC value for the GARCH volatility

[59] The use of these information criterion is disputed. See for instance the comment of Bollerslev, Engle, and Nelson (1994). Because of the lack of a useful selection criteria these approaches will be used nevertheless.
[60] This last step serves for a confirmation of the selected model.
[61] The estimation results for the AIC approach are available on request.

extension with $\alpha = 1$ and $\beta = 1$. The BIC value for the GARCH(1,1) model is 2874.4 with t-distributed innovations and 2942.0 for the GARCH(1,1) model with gaussian errors. Higher orders in one of the two parameters α or β effect an increase of the BIC value. This is independent of which parameter is enhanced. Consequently the highest BIC value is represented by the GARCH(4,4) model. Additionally, with higher orders the statistical significance of the estimated parameters decrease. Parameters of the order higher than 4 lead to estimations below the 5% significance level. For that reason and that higher orders cause an increase of the BIC value, estimates are made only until fourth orders.

Additionally to the underlying ARMA(1,0) model without a constant parameter, we perform BIC estimations with a model including a constant. The BIC value for the ARMA(1,0)-GARCH(1,1) with a constant and t-distributed innovations is 2878.9 , for the ARMA(1,0)-GARCH(2,1) model is 2886.0 and for the ARMA(1,0)-GARCH(1,2) model is 2886.5. Comparing these findings with the model setting without a constant, an upward shift in the information criterion is revealed. For every parameter setting a shift from the same size exists. Since these BIC values are larger than the BIC values of the same model without a constant, we omit the constant parameter.

The analysis delivers additional insights regarding the distribution of the errors. Comparing the results of the t-distributed innovations with the gaussian distributed errors reported in Table 5.21, one can discover lower BIC values for t-distributed errors. For instance the BIC value of a GARCH(1,1) specification with t-distributed innovation is 2874.4. For the same specification but with gaussian errors the value is 2942.0. This is true for every single specifications and for both test methods. This kind of econometric analysis confirms a GARCH(1,1) model with t-distributed innovations to additionally improve the model fit.

In order to justify the extension with modeled volatility, the BIC values of the ARMA models are compared with their extended version, the ARMA-GARCH models. The BIC value of the ARMA(1,1) process of Novartis without additional modeled volatility as presented in Table 5.16 is 2895.4. The corresponding extended ARMA(1,1)-GARCH(1,1) model has an BIC value of 2874.4.

Based on the minimization of the information criteria, the introduction of the GARCH component enhance the model fit. Both, the AIC and the BIC estimations of the extended model are lower than the information criterion of the

Table 5.22: BIC values for ARMA(p,q)-GARCH(1,1) models
The Table 5.22 represents the estimation results for the Schwarz-Baysian information criterion (BIC) for different ARMA models while the GARCH extension is held constant with $\alpha = 1$ and $\beta = 1$. Estimations are based on t-distributed innovations and for 2008 observations.

p\q	0	1	2	3	4
0	0	2874.6	2881.7	2889.3	2896.6
1	2874.4	2881.7	2889.3	2896.5	2901.5
2	2881.7	2889.0	2893.6	2903.8	2905.2
3	2889.3	2896.5	2903.8	2905.5	2912.6
4	2896.8	2901.8	2905.2	2909.9	2916.9

ARMA model. According to these outcomes, the time series of the density model can be fitted the most appropriately by an ARMA-GARCH model with t-distributed innovations.

Finally, in order to ensure that the best fit of the density dynamic is done by an ARMA(1,0)-GARCH(1,1) model, we perform additional estimations of some alternative models. In Table 5.22 the BIC estimations of different ARMA models with GARCH modeled volatility are summarized. The lowest BIC value with 2874.4 is presented by the ARMA(1,0)-GARCH(1,1) model, which confirms the preceding computations. Comparing of the results of all other models in Table 5.22 with the outcomes in Table 5.16 it shows that all BIC values are below the BIC estimations of the ARMA models with constant volatility. This underlines an improvement of the introduction of heteroscedastic variance.
The analysis based on the method of minimizing the Akaike information criterion or the Schwarz-Baysian information criterion suggests a GARCH(1,1) model for an extension of the basic ARMA model. This is not only the case for Novartis or Swisslife. In fact, in 99.13% of all considered assets and periods a GARCH(1,1) extension improves the model fit.
In Table 5.23 the parameter estimations of a ARMA(1,0)-GARCH(1,1) model are represented. The Table is arranged according to the considered periods starting with the first quarter. The parameter p is the autoregressive parameter of the ARMA component. The remaining parameters k, α and β belong to the GARCH element. The values in the brackets are the t-statistics calculated from

Table 5.23: Parameter Estimation for ARMA(1,0)-GARCH(1,1) model

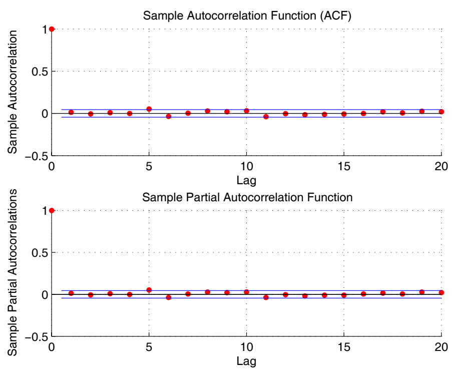
The Figure 5.23 represents the parameter estimations for ARMA(1,0)-GARCH(1,1) models for Novartis. The values in brackets below the parameter estimations are the corresponding t-values. The p represents the autoregressive parameter, k the moving average parameter, α and β the GARCH parameters, DoF specifies the degrees of freedom for the t-distributed innovations, N is the number of observations for the different periods and DW is the corresponding Durbin-Watson value. T-values labeled with an asterisk indicate a statistical significance on a 5% significance level while parameter estimations without an asterisk are significant on a 1% significance level. With two asterisk labeled values are significant at least on a 10% significance level.

	p	k	α	β	DoF	N	DW
1^{st}Quarter	0.072	0.019	0.835	0.079	16.087	494	1.978
	(1.591)	(0.162)*	(11.094)	(2.498)			
2^{nd}Quarter	0.040	0.176	0.125	0.108	9.081	495	1.977
	(0.797)*	(1.834)	(0.294)**	(1.697)			
3^{rd}Quarter	0.079	0.033	0.784	0.088	8.952	519	1.909
	(1.685)	(1.374)*	(6.438)	(1.931)			
4^{th}Quarter	0.015	0.244	0.000	0.131	27.524	495	1.948
	(0.296)**	(2.01)	(0.000)	(1.789)			
One Year	0.058	0.130	0.337	0.146	11.112	2008	1.968
	(2.384)	(4.581)	(2.717)	(4.643)			

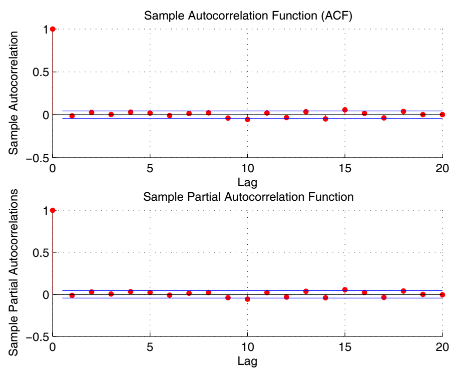
the corresponding standard errors and degree of freedom of the corresponding parameter. In addition to these outcomes the Durbin-Watson (DW) value and the calculated degree of freedom (DoF) of the innovations are reported. Statistically significant parameters on a 5% significance level are marked with a single asterisk. Parameter estimations signed with a double asterisk are statistically significant at a 10% significance level. The calculation for the statistical significance is based on the standard error combined with the corresponding degree of freedom as well as the chosen significance level.

Verification of the Model Selection In order to ensure the quality of the model fit we perform again an analysis of the residuals by means of an example of Novartis and Swisslife.[62] Again, as in the previous section the analysis is based on exploring correlograms and analytically on Engle's ARCH-test for conditional heteroscedasticity and the Goldfeld-Quandt test for heteroscedasticity and the Ljung-Box test.

In Figure 5.12 the correlograms of Novartis and Swisslife are presented. No correlation is significant at a 5% significance level. Neither in the sample autocorrelation nor in the sample partial autocorrelation is a significant lag constituted. This is also true for all other considered assets.



(a) ACF and PCF of Novartis



(b) ACF and PCF of Swisslife

Figure 5.12: Residual Correlograms of ARMA(1,0)-GARCH(1,1) Model The Figure 5.12 shows the correlograms of the errors of the ARMA(1,0)-GARCH-(1,1) model for Novartis, figure (a)and Swisslife, figure (b), respectively. Estimations are based on 2008 observations.

The test for heteroscedasticity is rejected in less cases than in the more simple case of the ARMA model. Albeit, extending the ARMA model with a GARCH component reduces GARCH effects in the innovations, not all cases of heteroscedasticity can be eliminated. For instance for the first quarter Swisslife documents a rejection of the null hypothesis that no GARCH effects can be assumed. We find similar results for all other considered assets. With the

[62]Whereas partially the corresponding analysis of Swisslife is available on request.

Goldfeld-Quandt test we find similar outcomes. Heteroscedasticity is reduced but not eliminated.[63]

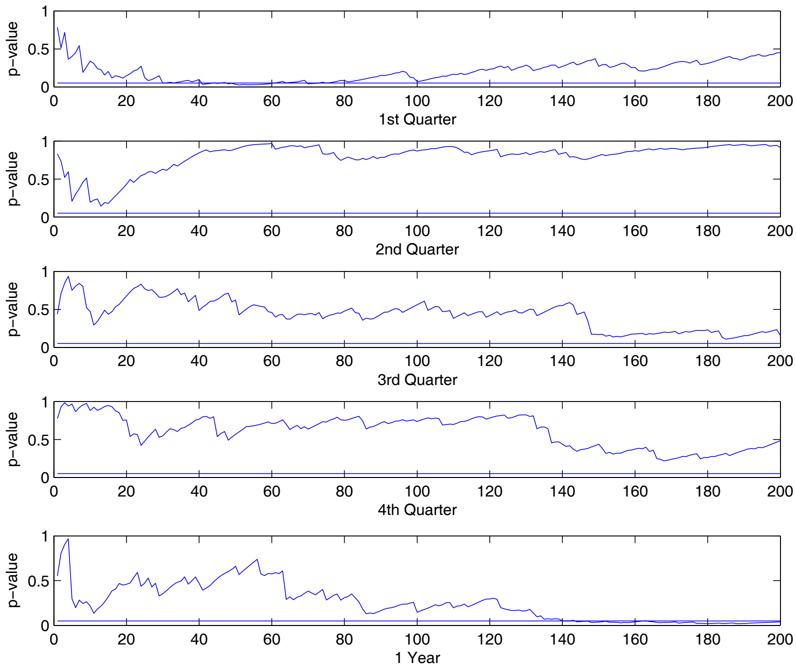


Figure 5.13: Ljung-Box for ARMA(1,0)-GARCH(1,1) The Figure 5.13 shows the Ljung-Box test for the ARMA(1,0)-GARCH(1,1) model for different periods. The null hypothesis $\mathcal{H}_0$ is that the model fit is adequate to the alternative of the existence of autocorrelation. The hypothesis test is performed on a 5% significance level which is illustrated with a solid line on the bottom of each picture. The p-values for a rejection of the null hypothesis is on the ordinate while the corresponding lag length is on the x-axis.

In Figure 5.13 the p-values for 200 lags of the Ljung-Box test for Novartis are illustrated. The significance level is 5% and is implemented with a solid line in the pictures. The Ljung-Box test implies an improvement of the model specification with GARCH elements for every considered period. In particular the model fit of the one year time series is enhanced. Comparing the one year Ljung-Box test

[63]The results for the ARCH-test and the Goldfeld-Quandt test for both, Novartis and Swisslife are available on request.

results from Table 5.11 with the corresponding test results from Table 5.13 it can be seen that the p-values are more rare under the significance level. An upwards shift of the p-values can be observed for all considered assets and periods.
Comparing all test results of the ARMA-GARCH model with the ARMA model we can conclude that the ARMA-GARCH model reflects more appropriately the properties of the time series of the limit order book. All results show evidence for this fact, at least the lower AIC and BIC values for the new model specification. All results are at least as good as the results of the ARMA model.
In Table 5.24 the ARMA(1,0)-GARCH(1,1) parameter estimations of all considered assets are represented. The presented period is one year. The Table is arranged in the same way as Table 5.23. The remaining estimations for the different quarters are available on request.

Table 5.24: Parameter Estimation all Assets
The Figure 5.24 presents the parameter estimations for ARMA(1,0)-GARCH(1,1) models for all considered assets. The values in brackets below the parameter estimations are the corresponding t-values. The p represents the autoregressive parameter. The constant of the GARCH process is represented by k, while α and β are the GARCH parameters. DoF specifies the degrees of freedom for the t-distributed innovations and DW is the corresponding Durbin-Watson value. T-values labeled with an asterisk indicate a statistical significance on a 5% significance level while parameter estimations without an asterisk are significant on a 1% significance level. With two asterisk labeled values are significant at least on a 10% significance level. Estimations are based on 2008 observations.

	p	k	α	β	DoF	DW
ABB	0.0031^{**}	0.3103	0.1198^{**}	0.0753	16.45	1.9918
	(0.130)	(2.665)	(0.392)	(2.439)		
Adecco	0.0503^{*}	0.0603^{*}	0.7807	0.0427	10.17	1.9522
	(2.174)	(1.706)	(6.956)	(2.751)		
Baloise	0.1025	0.0784^{*}	0.8811	0.0309^{*}	8.18	1.9784
	(4.567)	(1.587)	(13.420)	(2.125)		
Richemont	0.0826	0.0021^{**}	0.9910	0.0054^{*}	8.16	2.0218
	(3.867)	(0.983)	(163.400)	(1.685)		
Ciba	0.1110	0.0434^{*}	0.8996	0.0304	8.97	2.0023
	(5.016)	(1.705)	(18.139)	(2.292)		
Clariant	0.0717	0.3729^{*}	0.3221^{**}	0.0321^{*}	9.24	2.0260
	(3.159)	(1.191)	(0.587)	(1.309)		
Credit Suisse	0.0334^{*}	0.2046^{**}	0.3308^{**}	0.0201^{**}	9.46	1.9779
	(1.494)	(0.812)	(0.412)	(0.986)		

Table 5.24: (continued)

	p	k	α	β	DoF	DW
Givaudan	0.0903	0.0137	0.9507	0.0291	10.29	1.9748
	(4.165)	(2.005)	(61.618)	(3.601)		
Holzim	0.0914	0.0110*	0.9653	0.0142*	10.83	1.9998
	(4.129)	(1.119)	(40.202)	(1.795)		
Kudelski	0.0591	0.6821**	0.0000	0.0198**	9.74	2.0099
	(2.655)	(0.767)	(0.000)	(0.831)		
Nestle	0.0280*	0.2753**	0.4022**	0.0226**	6.52	2.0385
	(1.293)	(0.965)	(0.664)	(0.918)		
Novartis	0.0579	0.1296	0.3369	0.1462	11.11	1.9678
	(2.384)	(4.581)	(2.717)	(4.643)		
Roche	0.0966	0.0031*	0.9737	0.0207	12.97	1.9756
	(4.347)	(1.459)	(111.000)	(3.156)		
Swiss Re	0.0445	0.0228*	0.9305	0.0146*	9.65	1.9854
	(2.011)	(1.015)	(15.165)	(1.276)		
SGS	0.0725	0.1029	0.8698	0.0407	4.64	2.0048
	(3.553)	(2.114)	(16.590)	(2.596)		
Sulzer	0.0842	0.2823*	0.6722	0.0496	5.91	2.0226
	(3.816)	(1.690)	(3.780)	(2.120)		
Swatch	0.1679	0.0015**	0.9886	0.0089	15.54	2.0284
	(7.826)	(0.712)	(150.300)	(2.108)		
Swisscom	0.0720	0.0878	0.8150	0.0592	6.26	2.0027
	(3.271)	(2.276)	(12.103)	(3.124)		
Swisslife	0.0853	0.0028*	0.9832	0.0124	6.81	2.0223
	(3.966)	(1.396)	(164.100)	(3.171)		
Syngenta	0.0835	0.1271	0.6033	0.0748	10.26	2.0470
	(3.524)	(2.150)	(3.652)	(2.723)		
UBS	0.3997	0.4005	0.0000	0.1349	8.08	2.2201
	(17.673)	(5.295)	(0.000)	(3.557)		
Unaxis	0.1076	0.0050*	0.9772	0.0179	7.15	1.9574
	(5.074)	(1.351)	(129.200)	(3.324)		
Zurich	0.0676	0.1267**	0.6251*	0.0188**	14.46	2.0132
	(2.996)	(0.625)	(1.077)	(0.935)		

5.5 Asset Liquidity

In regard to asset liquidity, we analyze and interpret the parameters of the ARMA-GARCH model in an economic and auction theoretic sense. We consider the ARMA- and the GARCH-part of the equation separately. We assume that the ARMA part represents the liquidity measure and the GARCH component contributes additional information about liquidity characteristics. The liquidity measure reflects the integral part of the dynamic of the process. It shows the kernel of the dynamic and the changes which liquidity experiences as time proceeds. In return, the GARCH component reveals additional effects of the process. It discloses information about the occurrence of liquidity shocks and its persistence.[64]

The dynamic of the limit order book captures the changes between the unbalance of investors prepared to sell and to buy. In general, a balanced market absorbs orders faster and to more favorable conditions than an unbalanced market. Hence, according to definition 2.2.1 of liquidity given in chapter 2.2.3, an unbalanced market is less liquid than a balanced market. In an unbalanced environment either the waiting time until execution is longer or the price impact of the order is larger. In either case the agent pays for the additional risk. In the first case, the investor loses at least the risk free interest rate and bears the additional risk of disadvantageous price changes. In the second case, the price impact of the order directly reflects the costs. However, regarding the dynamic behavior of the limit order book, a market which reaches a balanced market situation faster is more liquid. Thus, a liquid market is characterized by a high converging-rate, while a less liquid market has a low converging-rate. Consequently, the converging rate itself is the liquidity measure.[65]

---

[64] The basis of the interpretation of the parameter is described in chapter 4.

[65] To illustrate this, consider the order placement process. According to Biais et al. (1995) the aspect of the mean-reverting behavior of the bid-ask spread and the order placement process are closely connected to each other. Agents place limit orders when the bid-ask spread is large or the book is thin. Conversely, they hit the quote when spread is tight. This implies the mean-reverting mechanism of the bid-ask spread. Large depth at the best quotes induces order placement within the bid-ask spread. This reduces depth and the slope of the book as shown by Naes and Skjeltorp (2006) and confirmed by us in the previous section. New orders are more frequently placed at or even above the best quotes as found by Biais et al. (1995), until large depth is reached again. Thus, a thick book absorbs orders faster than a thin. Consequently the faster this favorable situation appears, the more liquid is the market.

A fully balanced market is represented by a zero difference of the orders on the bid and the ask side. A faster reverting to this mean is reflected by a smaller dependency of the previous state of the limit order book.[66] Consequently the integral part of the dynamic behavior is represented by how strong this previous state is present in the current state. This component is measured by the autoregressive part p of the ARMA(1,0)-GARCH(1,1) model. The parameter p is bounded by $|p| < 1$ by definition. A p-parameter close to $|1|$ indicates a high influence of the previous state on the current state. A p-value close to zero shows a smaller influence of the previous state on the current state and therefore a higher dynamic.

Additionally, an illiquid market is also characterized by a confrontation of a convex shape on the one side to a concave shape on the other side.[67] Since the construction of the measure by the density approach also takes the shape of the limit order book in account it reflects such kind of situations. We constitute as a liquidity shock a sudden arrival of such an imbalance. The β parameter of the GARCH component describes the occurrence of sudden events or shocks. How long such a shock persists is measured by the α parameter of the GARCH component. In order to avoid negative volatility, α and β are greater than zero by definition.[68] Additionally they are jointly bounded by $\alpha + \beta < 1$. The existence of many liquidity shocks is denoted by a high value of β. A high value of α represents a long persistence of the liquidity shock. Since these two parameters provide additional information to liquidity behavior they are subjected to the group of additional liquidity characteristics.

The k parameter of the GARCH component scales the volatility axis. A large value of k induces large absolute values of the axis. This parameter relies mainly on the absolute values of the underlying time series. A process with large absolute values arouses large differences and hence an estimation of a large k value.

---

[66] A faster reverting does not depend on higher trading activity. Chordia et al. (2001) show evidence that trading activity measures such as volume or the number of trades is low correlated with liquidity measures such as depth, effective or quoted spread.

[67] See section 4.

[68] For more details consider section 3 or Hamilton (1994), Greene (2000) or McNeil et al. (2005),

Asset Liquidity Measure An imbalance between the bid side and the ask side reports a price pressure in the corresponding direction. Besides, this imbalance means that on the respective side too few shares are offered near the best quoted price. Comparing with a balanced market situation, it is more difficult to trade in this state. Since trading is affected by an imbalance of the limit order book this imbalance can be used as a proxy of liquidity.[69] According to this the most balanced limit order book is reflected by a zero difference between the price densities.

Hence we assume that the faster the market converges to its equilibrium the more liquid it is. Since p reflects the dependency on the previous state, a small value of p indicates a faster converging rate.[70] We define the liquidity measure λ as

$$\lambda = \frac{1}{2}\left(|p| + \alpha + \beta\right). \tag{5.6}$$

Since the parameters in equation (5.6) are bounded by $|p| < 1$ and are jointly bounded by $\alpha + \beta < 1$, the impact of the converging rate is higher weighted than the remaining parameters. In addition the liquidity measure λ is bounded by $0 < \lambda < 1$.[71] A liquidity value λ close to zero represents a very high asset liquidity, while a value of λ close to one indicates a very low liquidity situation. In Table 5.25 the considered assets arranged according to their liquidity ranking λ are represented. A low liquidity measure λ indicates a more liquid asset. Beside the liquidity ranking λ, the corresponding weight in the Swiss Market Index (SMI) is reported.[72] The liquidity ranking shows that the 7 most liquid assets represent 65.67% of the market capitalization of the Swiss Market Index. From the 5 largest assets, 4 are ranked under the 7 most liquid securities.

Liquidity shocks and its persistence are summarized under additional liquidity characteristics. It provides additional information to the investor for a considered

[69] For more details, see section 4.
[70] For a detailed analysis of the ARMA parameters interpreted as converging rate consult for instance R.-R. Chen and Yang (1995).
[71] Since the parameter k measures only the scale of the volatility we omit this parameter in the liquidity measure.
[72] The weights are calculated based on market capitalization at 31^{th} of December 2002. For additional information about calculating the weights consult
http://www.swx.com/trading/products/indices/stock_indices/smi/smi_de.html.

Table 5.25: Liquidity Ranking

For the period between January 2 and December 31, 2002, the Table 5.25 reports the liquidity ranking in descending order of the considered securities of the SMI. The most liquid asset, Kudelski, is represented by the lowest value of the liquidity measure λ. The measure λ is calculated according to equation

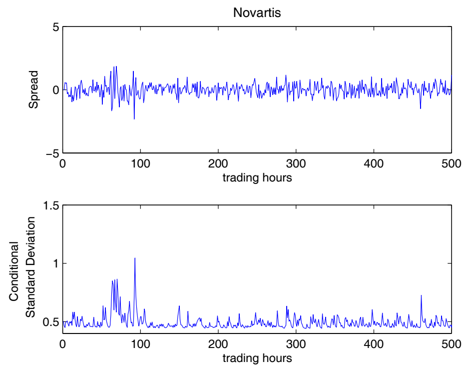
$$\lambda = \frac{1}{2}(|p| + \alpha + \beta),$$

Where the variables are the estimated parameters of an ARMA(p)-GARCH(α, β) process. For each asset the estimations are based on 2008 observations.

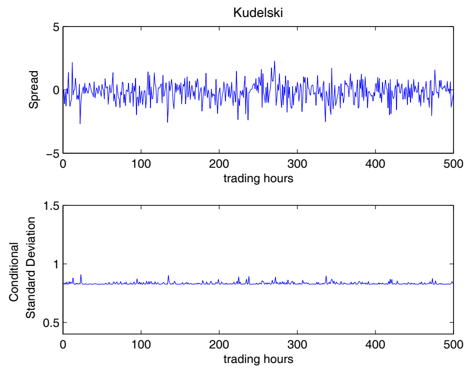
#	Name	λ	Weight in %	#	Name	λ	Weight in %
1	Kudelski	0.0395	0.11	13	SGS	0.4916	0.43
2	ABB	0.0991	0.76	14	Swiss Re	0.4948	5.08
3	Credit Suisse	0.1922	6.21	15	Baloise	0.5073	0.53
4	Clariant	0.2130	0.52	16	Ciba	0.5205	1.11
5	Nestle	0.2264	20.19	17	Givaudan	0.5351	0.85
6	UBS	0.2673	14.67	18	Holzim	0.5355	1.20
7	Novartis	0.2705	23.21	19	Richemont	0.5395	2.34
8	Zurich	0.3558	3.23	20	Swisslife	0.5405	0.33
9	Syngenta	0.3809	1.41	21	Roche	0.5455	11.78
10	Sulzer	0.4030	0.12	22	Unaxis	0.5514	0.15
11	Adecco	0.4368	1.17	23	Swatch	0.5827	0.22
12	Swisscom	0.4731	1.72				

investment horizon. As already mentioned a high α measures a long persistence of a shock. A high β indicates the existence of sequent shocks. In order to get an intuition of the mechanism, several situations are visualized in Figure 5.14. The Figure comprises four different sub-Figures from (a) to (d). The sub-Figure (a) depicts the time series of Novartis, sub-Figure (b) is Kudelski, (c) Givaudan and the last sub-Figure (d) is the time series of UBS. The upper pictures in each sub-Figure represent the limit order book dynamic. The lower pictures show the corresponding conditional standard deviations. The effect of α and β are identifiable in the conditional standard deviations. Compared to each other, in Figure (a) both values are high, whereas in Figure (b) both are low. Figures (c) and (d) represent high α and low β.

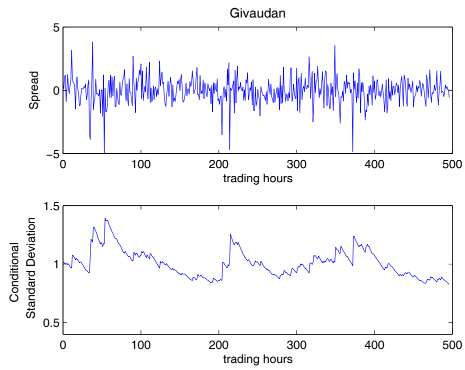
The chart of Figure 5.14 (c) exhibits less shocks. But the available shocks persist long in the limit order book dynamic. This effect is illustrated by a sudden



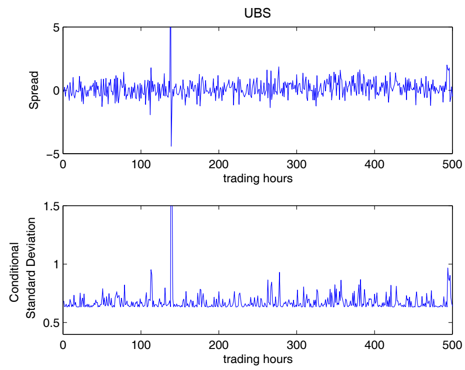
(a) $\alpha = 0.3369$ and $\beta = 0.1462$



(b) $\alpha = 0$ and $\beta = 0.0198$



(c) $\alpha = 0.9507$ and $\beta = 0.0291$



(d) $\alpha = 0$ and $\beta = 0.1349$

Figure 5.14: Additional Liquidity Characteristics The Figure 5.14 illustrates different aspects of the GARCH(α, β) effects. The upper pictures show the time series of the density measure, while the lower pictures exhibits the corresponding conditional standard deviations. A high α indicates a long persistence of the shocks, while a high β identifies the existence of frequently shocks.

increase followed by a slow decrease of the chart of the standard deviation. In contrast to this, Figure 5.14 (d) shows many sequent shocks, which declines instantaneously. A combination of high α and β is represented by the standard deviation of Figure 5.14 (a). The chart evolves disquietingly compared to the other pictures. To visualize these effects, it is necessary to plot the conditional standard deviation. These effects are not recognizable by only considering the charts of the limit order book dynamic.

Based on this model-framework and considered from an economic point of view, the two parameters are useful concerning the timing of submitting an order to the market. Knowing the behavior of the current period signals either to trade or to wait, depending on the corresponding preferences. For instance in times, where an order imbalance remains longer (situation (c) described in Figure 5.14), the waiting time until an order has been executed may be longer. This causes additional costs in the form of loss of interest on the invested capital, price changes of the corresponding asset or monitoring expenses.

In Table 5.26 a liquidity ranking of the industrial sectors where the considered assets belong is performed. The averages are calculated on a yearly basis. According to this analysis the two most liquid sectors are food and beverages and the bank sector. These two sectors make up a market capitalization of 42.19%.

Table 5.26: Liquidity Ranking according to Industrial Sectors

The Table 5.26 shows the liquidity ranking of the industrial sectors of the SMI. For each sector the average of the liquidity value of the firms from the corresponding sector is calculated. The market capitalization is computed on December 31, 2002.

Rank	Industrial Sector	Market Cap. in %	Number of Companies	Liquidity Measure λ
1	Food & Beverages	20.74	1	0.2264
2	Banks	21.45	3	0.2298
3	Industrial Goods & Services	2.67	1	0.2940
4	Chemicals	3.12	2	0.3714
5	Pharma	35.94	4	0.4080
6	Telecommunication	1.77	1	0.4731
7	Insurance	9.42	5	0.4746
8	Non Cyclical Goods & Services	0.87	1	0.5351
9	Construction	1.23	2	0.5355
10	Electronic & Elect. Equipment	0.16	2	0.5514
11	Cyclical Goods & Services	2.63	1	0.5611

At the end of this ranking reside sectors like cyclical goods and services, electronical equipments, constructions and non cyclical goods and services. These last four sectors make up a market capitalization of less than 5%. The largest industrial sector, measured at the market capitalization is the pharma sector with 35.94% of the SMI market capital. The pharma sector is located at the 5th place. The top 5 most liquid sectors make up a market capitalization of 83.92%.

Measuring liquidity depends on the time horizon considered. In shorter periods, estimations capture different events and effects in more detail. This influences the result in a variation of the liquidity measure. In Table 5.27 the estimation results of the liquidity measure λ for all four quarters and for all considered assets are reported. The Table is sorted in ascending alphabetic order. It reveals that this liquidity measure is affected by the chosen time horizon. Although the number of the measure changes, the classification remains the same. In particular, the 5 most liquid sectors determine around 75% of the market capitalization.

Table 5.27: Liquidity Measures for Quarters

For the period between January 2 and December 31, 2002, the Table reports the liquidity measures λ for the quarters. In addition, the estimation for the whole year is calculated. For security, calculations are based on 494 observations for the first quarter, 495 for the second quarter, 519 for the third quarter and 495 for the fourth quarter. The average as well as the standard error refers to the quarters.

	Quarters					Std.	
	1^{st}	2^{nd}	3^{rd}	4^{th}	Average	Error	1 Year
ABB	0.1659	0.0742	0.1793	0.4802	0.2249	0.1765	0.0991
Adecco	0.4477	0.4172	0.0215	0.4154	0.3254	0.2032	0.4368
Baloise	0.3966	0.5079	0.3000	0.4612	0.4164	0.0901	0.5073
Richemont	0.0159	0.0737	0.0343	0.1148	0.0597	0.0440	0.5395
Ciba	0.5067	0.5109	0.1832	0.3091	0.3775	0.1602	0.5205
Clariant	0.4763	0.2322	0.0480	0.1506	0.2268	0.1827	0.2130
Credit Suisse	0.1289	0.3785	0.1647	0.3001	0.2430	0.1166	0.1922
Givaudan	0.2598	0.4991	0.0475	0.1045	0.2277	0.2020	0.5351
Holzim	0.2680	0.0457	0.4958	0.4420	0.3129	0.2029	0.5355
Kudelski	0.4522	0.0107	0.4895	0.3869	0.3348	0.2203	0.0395
Nestle	0.4325	0.0270	0.2315	0.3692	0.2650	0.1795	0.2264
Novartis	0.4930	0.1361	0.4755	0.0730	0.2944	0.2209	0.2705
Roche	0.5224	0.5715	0.2568	0.5740	0.4812	0.1515	0.5455
Swiss Re	0.4800	0.2348	0.0051	0.0479	0.1919	0.2164	0.4948

Table 5.27: (continued)

	Quarters					Std.	
	1^{st}	2^{nd}	3^{rd}	4^{th}	Average	Error	1 Year
SGS	0.4053	0.0574	0.1601	0.1169	0.1849	0.1529	0.4916
Sulzer	0.0747	0.3597	0.5604	0.4575	0.3631	0.2090	0.4030
Swatch	0.3036	0.5318	0.1291	0.0873	0.2629	0.2023	0.5827
Swisscom	0.5456	0.2985	0.4866	0.3546	0.4213	0.1144	0.4731
Swisslife	0.5514	0.0429	0.5473	0.5329	0.4186	0.2506	0.5405
Syngenta	0.3953	0.1234	0.4393	0.3210	0.3198	0.1397	0.3809
UBS	0.2384	0.2491	0.3333	0.2125	0.2583	0.0523	0.2673
Unaxis	0.5165	0.5869	0.4569	0.4426	0.5007	0.0658	0.5514
Zurich	0.1671	0.4955	0.0459	0.0338	0.1856	0.2152	0.3558
LM Market	0.3584	0.2811	0.2648	0.2951	0.2999	0.0410	0.4001
LM SMI	0.3894	0.2297	0.2877	0.2581	0.2912	0.0696	0.3284

At the end of the table, two additional estimations are presented. The LM market is a simple average of all liquidity measures of all considered assets. The LM SMI is the average of all liquidity measures of all considered assets weighted according to the SMI weights. This calculation of the measures reveals that the liquidity of the market is more stable than of one single asset.

5.6 Liquidity Premium

A basic principle in asset pricing theory is that any form of risk is compensated in some form of additional return. The risk and return relation is a central role in several economic concepts and models. The basic idea is that the higher the risk, the higher is the probability to achieve a larger return. Less liquid assets bear the risk of, for instance, longer execution times, higher price impacts by execution or menace of order splitting, which again is connected with higher transaction costs.[73] From this point of view liquidity can be considered as a form of risk.[74] Since higher risk is compensated in the form of higher returns, liquidity risk is a pricing component. This priced liquidity risk component is referred to as the liquidity premium.

[73]See more detailed descriptions ahead.
[74]See also the definition of liquidity given in chapter 2.

Depending on which measure serves as a liquidity proxy, several studies investigate the relationship between liquidity premium and asset return. Amihud and Mendelson (1986), for instance, study the effects of the bid-ask spread on returns and find that higher yields are required on higher spread stocks, i.e. that market-observed expected return is an increasing and concave function of the spread. Later, they verify their findings by connecting additional risk forms with liquidity risk in Amihud and Mendelson (1989). Datar, Naik, and Radcliffe (1998) extend the work of Amihud and Mendelson (1986) by using the turnover rate as proxy for liquidity.[75] They evidence that liquidity plays a significant role in explaining the cross-sectional variation in stock returns. Brennan and Subrahmanyam (1996) find that there is a significant return premium associated with both the fixed and variable elements of the cost of transacting while Pereira and Zhang (n.d.) focus on the relationship between the liquidity premium and the volatility of liquidity. They examine the liquidity premium in a dynamic model with price impact and give an explanation of that liquidity premium decreasing with the volatility of liquidity. Acharya and Pedersen (2005) introduce a liquidity adjusted capital asset pricing model. They show that investors require return premium for illiquid securities and are willing to pay a premium for liquid assets when the whole market is illiquid, investors are willing to pay a premium for liquidity when market return is low and additionally, illiquid assets tend to have commonalities in liquidity with market liquidity, return sensitivity in market liquidity and liquidity sensitivity to market returns.[76]

Brennan et al. (1998) as well as B. R. Porter (2003) calculate, by means of the Fama-French three factor model provided by Fama and French (1992) the liquidity premium. In order to extract the liquidity component of the returns according to our liquidity hierarchy respective to liquidity measure, we follow Brennan et al. (1998) and B. R. Porter (2003) and apply the Fama-French three factor model to our dataset.

The capital asset pricing model (CAPM) derived independently by Sharpe (1964),

[75] The turnover rate is defined as the number of shares traded as a fraction of the number of outstanding shares.

[76] Work in this direction has also been done by Chordia, Roll, and Subrahmanyam (2000), Huberman and Halka (2001), Hasbrouck and Seppi (2001), Amihud (2002) or Pastor and Stambaugh (2003), among others.

Lintner (1965) and Mossin (1966) relates the expected return of an asset to the risk free rate and a weighted risk premium.[77] The risk premium is the difference between the expected market return and the risk free rate. The weight, referred to as β_i, serves as the asset (i) specific risk measure.[78] Later on, several authors extend the capital asset pricing model by relaxing assumptions or appending additional factors.[79] By combining the original market risk factor with two new developed risk factors, Fama and French (1992) present a model which separates the asset's expected return in several risk specific components. Beside market risk, they add size risk, measured as the difference between small minus big assets (SMB) and value risk, measured as the difference between assets with high book-to-market ratio minus assets with a low ratio (HML) as supplementary explanatory variables into the model. Since we are interested in the liquidity risk premium we additionally attach a risk factor measuring liquidity risk (LMI). Therefore the model appears as

$$r_i - r_{rf} = \beta_1(r_m - r_{rf}) + \beta_2\text{LMI} + \beta_3\text{HML} + \beta_4\text{SMB},$$

where r_f is the risk free interest rate, and β_i is a measure of the exposure an asset (i) has to market risk. The factor size risk (SMB) is the difference between

[77] Formally the capital asset pricing model (CAPM) is

$$\mathbb{E}[r_i] = r_f + \beta_i(\mathbb{E}[r_m] - r_f),$$

where $\mathbb{E}[r_i]$ is the expected return of the asset (i), r_f is the risk free interest rate and $\mathbb{E}[r_m]$ is the expected return of the market. For more details of the model and its assumptions, particularly a critical discussion consider for instance, Black (1972), Black, Jensen, and Scholes (1972), Fama and MacBeth (1973), Banz (1981) or Kothari, Shanken, and Sloan (1995), among many others.

[78] The parameter β_i is a measure of systematic risk. The systematic risk is the portion of volatility measured by the degree to which the assets return vary relative to those of the market return. Accordingly a β_i of zero indicates a risk less investment, while a β_i of 1 denotes average market risk. If β_i is larger than one the investment is riskier than an average market risk investment, if β_i is lower than one, the investment is less risk than an average market risk investment. Formally, β_i can be written as $\beta_i = \frac{\text{cov}(r_i, r_m)}{\sigma_m^2}$, where $\text{cov}(r_i, r_m)$ is the covariance between the return of the market r_m and the return of the asset r_i and σ_m^2 is the variance of the return of the market.

[79] Ross (1976), for instance, introduces the arbitrage pricing theory (APT), which is a linear function of various macro-economic factors or risk factors. It differs from the CAPM in that it is less restrictive in its assumptions. For instance, the APT assumes only arbitrage free markets instead of the existence of market equilibrium.

the average return of the returns of the smallest market capitalized assets and the average return of the returns of the largest market capitalized assets. The parameter s_i measures the level of exposures to size risk. Constructed in a fashion similar to SMB, HML is calculated as the difference between the average return of stocks with the highest book-to-market ratio and the average return of stocks with the lowest book-to-market ratio. The HML factor is referred to as the value risk. The parameter h_i is a measure of the level of exposure to value risk.[80]

In contrast to the intention of Fama and French (1992) our aim is not to improve the overall explanatory power of the capital asset pricing model by enhancing the adjusted R^2. From our liquidity ranking reported in Table 5.25 we pick the 7 most liquid and the 7 fewest liquid assets to construct the liquid-minus-illiquid (LMI) portfolio.[81] The performance of the LMI portfolio is calculated as the difference between the average return of the 7 most liquid and the 7 fewest liquid assets. From an economic perspective, the LMI is designed to measure the additional return investors have received by investing in illiquid assets. Hence, the β_2 measures the level of exposure to liquidity risk.

For the period between January 2 and December 31, 2002 the results of the Fama-French three factor model estimations are reported in Table 5.28. The calculations are based on daily notations which include 251 observations. To determine the SMB portfolio, market capitalization at December 31, 2001 is considered. In order to establish the HML portfolio, two parameters have to be computed separately. The market value is taken of the closing prices at December 31, 2001 while the book value of the asset is calculated of the corresponding annual report.[82]

[80] However, the size risk factor and the value risk factor are known as the Fama-French factors.

[81] Fama and French (1992) originally use a portion of 30% to create the portfolios. We follow them and take 30% of the most liquid assets to build the liquidity portfolio.

[82] Source of the Data is Reuters. The annual reports are publicly available on the corresponding homepages.

Table 5.28: Fama-French Three Factor Model

For the period between January 2 and December 31, 2002, the Table 5.28 reports the results of the Fama-French three factor analysis. The model is $r_i - r_{rf} = \beta_1(r_m - r_{rf}) + \beta_2\text{LMI} + \beta_3\text{HML} + \beta_4\text{SMB}$, where r_i is the return of security i, r_{rf} is the risk free rate, r_m is the return of the SMI. LMI is the liquidity portfolio, where the average return of the 7 most liquid assets are subtracted from the average return of the 7 fewest liquid assets. The HML portfolio is constructed based on the book-to-market ratio and SMB is selected according to the market capitalization of the entire assets at the end of December 2002. Calculation is based on 251 observations for each security. Values market with an asterisk are not significant on a 5% significance level. For the sake of simplicity not market values are significant on a 5% significance level.

Name	β_1	t-Value	β_2	t-Value	β_3	t-Value	β_4	t-Value	adj. R^2	F-test	DW
ABB	0.21	1.28	−2.03	−8.67	2.97	10.81	2.97	10.81	0.70	580.25	1.93
Kudelski	1.13	9.87	−1.28	−7.72	−1.11	−6.22	−1.11	−6.22	0.53	284.81	2.01
Clariant	1.05	13.56	−0.44	−3.92	0.19	1.64	0.19	1.64	0.60	379.02	2.44
Adecco	1.29	13.51	−0.33	−2.75	−0.46	−3.19	−0.46	−3.19	0.40	168.06	1.98
UBS	1.19	34.06	−0.21	−3.86	−0.23	−3.57	−0.23	−3.57	0.78	912.28	1.67
Credit Suisse	1.61	20.96	−0.19	−1.73	0.34	2.84	0.34	2.84	0.69	568.68	1.79
Nestle	0.72	21.97	−0.07	−1.48	−0.07	−1.12	−0.07	−1.12	0.60	376.12	2.00
Novartis	0.77	23.23	−0.07	−1.44	−0.30	−5.33	−0.30	−5.33	0.65	462.48	1.68
Ciba	0.66	16.11	−0.06*	−1.08	−0.01*	−0.20	−0.01*	−0.20	0.56	326.82	2.14
Unaxis	1.36	15.98	−0.03*	−0.27	0.03*	0.22	0.03*	0.22	0.51	262.42	2.07
Swisscom	0.45	10.10	0.07	1.29	−0.09	−1.32	−0.09	−1.32	0.30	106.05	2.15
Roche	0.97	28.00	0.17	2.96	0.02	0.31	0.02	0.31	0.69	550.45	2.10
Swatch	1.07	15.17	0.18	1.70	0.43	3.77	0.43	3.77	0.52	267.63	2.21
SGS	1.05	11.63	0.18*	1.40	0.01*	0.06	0.01*	0.06	0.42	183.20	2.09
Swiss Re	1.42	23.82	0.18	1.83	0.24	2.09	0.24	2.09	0.63	435.56	2.14
Zurich	1.92	16.52	0.28	1.62	0.53	3.09	0.53	3.09	0.52	269.07	1.92
Givaudan	0.38	9.52	0.32	5.66	0.12	2.07	0.12	2.07	0.24	81.45	2.20
Holzim	0.91	17.93	0.33	5.14	0.32	3.90	0.32	3.90	0.51	265.18	2.13
Syngenta	0.59	8.31	0.40	4.31	0.43	3.77	0.43	3.77	0.27	94.65	2.01
Baloise	1.57	23.78	0.49	5.48	0.78	6.64	0.78	6.64	0.65	459.55	2.38
CFR Richemont	1.20	21.04	0.62	9.25	0.67	9.64	0.67	9.64	0.62	406.70	1.98
Swisslife	1.93	19.32	0.62	4.72	0.21	1.26	0.21	1.26	0.55	303.28	1.79
Sulzer	1.01	12.20	0.87	7.18	0.77	5.97	0.77	5.97	0.50	252.58	2.43

The Table 5.28 is sorted in ascending order according to β_2. In general a negative value of β_2 describes a small liquidity premium while a positive value of the parameter displays a larger liquidity premium. The table 5.28 reports that assets tend to have liquidity premium according to the liquidity ranking proposed in Table 5.25. For instance, from Table 5.25 Kudelski is considered to be one of the most liquid assets. Since liquidity is a risk factor, these assets should have a small liquidity premium. Table 5.28 documents a large negative value for Kudelski and therefore a small liquidity premium. In contrast, according to Table 5.25 Swisslife is less liquid. In Table 5.28 a large positive value is attached to Swisslife which reports a large liquidity premium. In general, the assets tend to be sorted in Table 5.28 according to their liquidity ranking from Table 5.25. This analysis implies that illiquidity is a cost factor and the market tends to compensate for it.

Chapter 6

Conclusion/Discussion

In this thesis we have focused on the modeling of an asset liquidity measure based on the dynamic of the limit order book. Asset liquidity is an essential characteristic of a well working financial market. In fact, the absence of liquidity can influence the trading process considerably. The simple situation that an investor is not able to sell any given amount of assets at a given point of time can cause a difficult financial situation. From this point of view a proper liquidity measure is crucial.

We assumed in this work that the limit order book comprises all the relevant information to construct an appropriate measure for asset liquidity.[1] This assumption suggests that it is possible to construct a measure which reflects asset liquidity based on data only generated by the activities in the limit order book. Therefore our main focus was to research the setting, structure and activities in and around the limit order book to reflect its dynamic behavior most accurately. In academic literature as well as in practical uses there is a huge amount of measures and key figures based on the limit order book. But in contrast to existing measures we observed the changes of the limit order book as time proceeds, i.e. the dynamics of the book, and used these changes to construct a new liquidity measure.

We reflected the dynamic of the limit order book by means of the order disbalance

[1] In chapter 4 we showed evidence of this fact by means of the collection of different studies as Harris and Hasbrouck (1996), Coppejans et al. (2004), Pascual and Veredas (2006) or Cao et al. (2004) among many others.

between the bid and the ask side. Since a large disbalance causes greater price jumps and hence increasing transaction cost, large disbalances show illiquidity. In contrast small differences offer more advantageous trading situations. From an asset liquidity point of view a tighter difference between the bid and the ask side is preferable. The faster the limit order book turns to small differences between the bid and the ask side, the faster a trader can act at this market to favorable transaction costs. As time proceeds the pattern of the these differences shows a mean reverting behavior. Since the mean is zero by construction a faster converging to the mean shows a more liquid asset. Therefore we interpret the converging rate of the mean reverting process as a measure for liquidity. In order to reach this model and to explore the way leading to this model, we collected a set of sequent research questions to which we addressed to answer.

The first research question we had to answer is whether there is a possibility to display the limit order book and its movements. For this, we reconstructed the limit order book of 23 Swiss blue chips for the year 2002. We have collected in a first step each limit order book entry for every single considered asset and generated an overall dataset consisting of 152,488,698 records. This dataset included all relevant information for all submitted orders. In a second step we programmed a Matlab-Code which reconstructed all order arrivals, executions and cancellations. Since every market has its own market architecture and trading rules we used the setting and principles of the Swiss Stock Exchange for programming the code. For each asset the limit order book is separated in two sides; the bid and the ask side. The submitted orders are collected for each side individually. For an arbitrary time frame the Matlab-Code ranked orders according to the best offered price for each side separately. On the bid side, orders are classified according to the highest price and on the ask side they are sorted according to the lowest price. Within same prices orders are arranged according to the submitting time. This Matlab-Code allowed displaying the limit order book for different time frames.
In order to illustrate the movements of the limit order book, we applied a density measurement approach. The measure can be calculated for different time frames. In a given time frame, the difference of the sum of the volume per price of the ask side and the bid side are taken, while the volume per price is computed for all order levels in the current book. The advantage of this measure was that

it is symmetric according to the bid and the ask side, it is independent of the price level and it weights orders less which are further away from the best offer. The measure shows for each time frame if the book was in an excess supply or an excess demand situation. With this concept it was possible to illustrate each trading situation, changes and the movements of the limit order book.

The second research question was to discover characteristic structures in the limit order book and particularly in the time series of the density measure. The idea of this question was to find out if investors behave strategically and whether this can be recognized by observing the limit order book. For this reason we performed a top to bottom analysis with respect to the order placement process. We started with a daily grid of the data and refined our perspective continuously. The first analysis attended to the daily volume dispersion. Like Gerety and Mulherin (1992) for the New York Stock Exchange, Niemeyer and Sandås (1993) for the Swedish Market or Hedvall (1994) for Finland we found a similar U-shape pattern in daily volumes for the Swiss Market. Based on an hourly observation interval, most volume occurred at the beginning and the end of a trading session. Around noon, volumes usually declined and formed the daily U-shape pattern. In contrast to other investigations we additionally distinguished between the volumes on the bid and on the ask side. We found that within the same time interval, the bid and the ask volumes are equally scaled. Albeit mostly in common with findings of other researchers this was a first impression that the order placement process possess a certain structure which can be connected with a strategic behavior of the investors.
The first analysis with data on tick level was to illustrate the shape of the limit order book. We tightened the time grid to an hourly basis and aggregated the order volumes on the same submitted prices. We found that first, shape does change as time proceeds. Moreover we discovered that for different assets the shape of the limit order book does change at different speeds. This is a first indication that different assets have different levels of trading activities. Second, we found that the shape of the limit order book does also change beyond the best quotes. Cao et al. (2004) showed evidence that quotes beyond the best quotes have informative properties and Harris and Panchapagesan (2005) found strong evidence that traders use the additional information provided by the entire limit order book. To verify these findings, we performed a further analysis of the shape

of the book by investigating the slope of the shape. We applied the slope measure provided by Naes and Skjeltorp (2006). We found that the more the considered range of the limit order book has been tightened, the steeper the average slope of the shape. In addition we provided a correlation analysis between the spread and the slope and found that the spread is negatively correlated to the slope. Moreover the correlation is increasing the more the range for the slope calculation is tightened. This means that in general larger volumes at quotes are associated with a larger bid-ask spread. In chapter 4 we introduced theoretically that a limit order book shaped with a steep slope is, in general, associated with larger spreads and also that the converse of this is true. These empirical findings support the theoretical concept and support the idea of the mean reverting character of the bid-ask spread and hence the mean reverting behavior of the limit order book.
To find out more about the order flow influence on the limit order book we also explored order flow persistence in this context. According to Biais et al. (1995) we investigated the frequency of orders conditional on the previous type of submitted order and the frequency of orders given the state of the limit order book. To evaluate the order flow according to their type of orders we first had to sort the orders in different categories. We followed Biais et al. (1995) and ranked the orders according to their aggressiveness. The most aggressive order was an order close to or in between the bid-ask spread, while the least aggressive order was one far away from the current trading price. We found that the probability that an order of the same aggressiveness follows each other is significantly higher than that of any other aggressiveness class. Biais et al. (1995) call that event the diagonal-effect. Biais et al. (1995) provide three economic hypotheses to explain this observation of serial correlation in the order submission process. Serial correlation may arise due to strategic order splitting. Strategic order splitting reduce transaction costs. Moreover a large order in the market suggests that the investor owns private information. This can be observed by the other market participants and can therefore be imitated. This reason is also supported by the findings of Easley and O'Hara (1987). The third hypothesis assumes that agents react similarly on the same events, since they own the same economical and technical knowledge. The second investigation paid attention to the state of the book with respect to the depth and the bid-ask spread. The threshold of assigning the state is defined by the median of each time series. From this point

of view we distinguished four different combinations of states. From theoretical considerations and findings of diverse researchers orders hit the quote if spread is tight or depth is small. Conversely, orders are placed within the spread if spread is large or depth is large. We found that, in general, order placement behavior for the bid and the ask side was different for the year 2002. For the ask side, orders are placed more aggressively if spread is large and are even more aggressive if depth is additionally large. Our theoretical aspects are, for the most part, confirmed by the findings of the ask side. The bid side shows another picture. Compared with the ask side, the orders of the bid side are in general less aggressive. A possible explanation for this fact can be that stock markets strongly declined in the year 2002. In order to avoid large losses, investors tried to sell their positions fast. Hence investors acted more aggressive on the sell side of the market. However, what is relevant is that first we found a serial correlation for the order flow process and second, that the correlation mainly follows the theoretical considerations of the mean reverting behavior explained in more detail in the previous chapter.

Beside these mentioned analyses, we further made regressions of different parameters on return and on volatility. We only found a slightly negative connection between volatility and the slope of the limit order book. We tested several models where we used different parameter settings. Across all assets and all models we found a statistical significant weak negative relation. That means that volatility decreases the steeper the order book slope is. Additionally we investigated the price impact by calculating the liquidity cost according to our cost calculation presented in chapter 2. Along with several empirical findings we found that in a bearish market environment, sell orders have larger price impact than buy orders. Since we are mainly interested in the liquidity question, we ranked the assets according to their liquidity cost for an average order size. This ranking served later on to compare the results with the liquidity measure introduced in the following chapter.

We found that the limit order book contains relevant information for asset liquidity and that it is observable. We now turned to answer the third question. Is it possible to capture the time series of the movements with a dynamical model? In this context we first analyzed the characteristics of the time series generated with the density measure approach. The measure is constructed out of the sum of the price-weighted order size, while the weights are calculated as the difference between the corresponding price of the order and the current trading price for the considered time frame. Hence a positive value of the measure reflected an excess demand and a negative value indicated a supply pressure.

Then we turned to the statistical analysis of the time series for all considered assets. In a first step we examined the distribution of the data. We found that it is not normally distributed but t-distributed. Then, we confirmed that it is a stationary process by applying three different tests for stationarity.
Before we proceeded with analysis regarding the time domain, we investigated the relation to return and volatility. We plotted the density as a function of the depth. We found that volatility is statistically significant closer to the bid-ask spread rather than deeper in the book. As mentioned above, several other researches find similar results that trading activities do not only take place close to the bid-ask spread. Since we considered more than a truncated book, all this information is incorporated in this measurement concept.
The correlation matrix presented a negative correlation between the density measure and the bid-ask spread. This is in line with the findings of the investigation of the frequency of orders given the state of the limit order book. Moreover we find that correlation becomes stronger the more the limit order book is restricted to the range around the bid-ask spread.
Additionally we tested the relation of the density measure to return and volatility by applying the similar tests already used to the slope time series of the limit order book. We applied the same four models provided by Naes and Skjeltorp (2006) to the return and volatility time series. Regarding volatility, the regression resulted in a weak negative relation to the density measure. The R^2 showed that approximately 12% of the variation in the response variable can be explained by the density measure. For the regression with respect to the price returns, we applied two regressions. The first one was the same regression as for the volatility and the second was corrected for autocorrelation in the error terms. For both the result's explanation of the variation of the response variable is 1.1% and 2.8% respectively. From this point of view the additional tests of the density measure did not contribute explanatory power neither to price volatility nor to price returns. Since we are interested in capturing the time series with an appropriate model we continued with the analysis of the time domain. Since

the time series is stationary we applied tests for autoregressive moving average (ARMA) processes. First we plotted the sample autocorrelation and the sample partial autocorrelation. We found a first indication in these correlograms across all assets that the processes can be reflected with first order ARMA processes. From this first illustrated impression we turned to statistical analytical methods to assign the orders of the processes. In this context we applied the Akaike information criterion (AIC) and the Baysian information criterion (BIC) for different order combinations. Since we have previously shown that the time series is not normally distributed, tests are based with t-distributed innovations. The BIC model suggested, in all cases, a more general model selection with less number of parameters than the AIC approach. Wagenmakers et al. (2006) proposed to keep the models as general as possible by reason of a.o. better results for parameter estimation. Since the majority of the correlograms, and moreover the BIC model selection, suggested a first order autoregressive model we decided to select this model. To verify that the model fit is adequate, we performed an analysis of the residuals. For the analyses we used correlograms, the Durbing-Watson test statistics, the Ljung-Box test, Engle's test of heteroscedasticity as well as the Goldfeld-Quandt test. Although residuals have been not correlated for the most part, based on the ARCH test of Engle, we found that they have not been stochastically independent. Hence, the innovation process of the ARMA(1,0) model featured volatility clustering. Due to these results we decided to extend the model with an heteroscedastic part, the generalized autoregressive conditional heteoroscedastic (GARCH) component.
We extended the ARMA model with GARCH components of several parameter combinations and found that the ARMA(1,0)-GARCH(1,1) model fitted the best. Selecting this model combination, we additionally tested the distribution of the innovation of the process. Again, the Baysian information criterion proposed t-distributed errors. Moreover, all values of the BIC were smaller than the previous results of the model without the extension. By repeating all the tests for verification, the model fit we now concluded that the extended model delivered the most adequate results.

The last question remaining to answer is to test whether the measure of these dynamic changes serves as an appropriate measure for asset liquidity. The density measure captures the differences between investors prepared to sell and to buy. In general, a balanced market absorbs orders faster and to more favorable conditions than an unbalanced market. From this point of view, an unbalanced market is less liquid than a balanced market. Regarding the dynamic behavior of the limit order book, this means that a market reaching a balanced market situation faster is more liquid. In that sense a liquid market is characterized by a faster converging rate from an unbalanced in a balanced market situation. Since the autoregressive parameter measures the converging rate of the process, this figure serves as the dynamic liquidity measure. The two parameters of the GARCH component describe first the occurrence of sudden shocks and second the persistence of such shocks. Applying to our liquidity time series this means that this component measures the arrivals of sudden imbalances of orders and how long such an imbalance persists in the market. Hence it measures liquidity shocks.[2]

The dynamic liquidity measure is an additive combination of the converging rate of the ARMA(1,0) process and the GARCH(1,1) parameters. Since all parameters are bounded and standardized the measure is standardized as well by 0 and 1, while 1 reflects an illiquid and zero a liquid market. We applied the measure to all considered assets and ranked them according to their liquidity attribute. Our measure identified that Kudelski was the most liquid asset in the year 2002 while Swatch represented the least liquid asset. In the year 2002 the liquidity ranking shows that the 7 most liquid assets represents 65.67% of the market capitalization of the Swiss Market Index. Moreover from the 5 largest companies, 4 are ranked under the 7 most liquid securities. Naes and Skjeltorp (2006) evidenced that in general larger capitalized assets are more liquid than small capitalized. Our findings presented similar results except of Kudelski. According to our liquidity ranking, the most liquid asset has been Kudelski albeit it is one of the smallest capitalized asset in the Swiss Market Index. A possible explanation for this fact may be that Kudelski has been excluded from the Swiss Market Index in that year. Since several investors pursued a strategy involving imitating the Swiss Market Index, a drop out of an asset caused large portfolio regroupings from institutional and private investors. This in turn entailed large trading activities in the corresponding asset. Consequently asset liquidity increased.

[2]See for more details chapter 4 and chapter 5.

An additional feature of the dynamic liquidity measure is that it shows the liquidity behavior concerning liquidity shocks and the persistence of such shocks of an asset. We showed that Novartis experienced a similar number of liquidity shocks like UBS during that trading period. But in contrast to UBS, the liquidity shocks of Novartis persisted longer in the market. Givaudan for instance had less shocks but, compared to the sample, they persisted a long time. In the same period, we found no liquidity shocks in the data of Kudelski. Considered from an economic point of view, this information is useful for an investor concerning the timing of submitting an order to the market. Knowing the behavior of the current period opens the possibility to trade with an appropriate strategy.
We then extended our liquidity calculations to the industrial sectors of the Swiss Market Index. We found that the most liquid categories was the food and beverages and the bank-sector. These sectors covered more than 42% of market capitalization of the Index. Moreover the 5 most liquid sectors accounted for around 85% of market capitalization. Additionally we found that the most liquid assets were in general the strongest weighted components in the Swiss Market Index.
In order to evaluate the explanatory and measurement power of the model we applied the Fama-French three factor model. With the Fama-French three factor model we were able to calculate whether an asset has a liquidity premium or not. We found that liquid assets according to our liquidity ranking had a small liquidity premium. Assets which were harder to trade compensated this with a larger liquidity premium. The liquidity premium ranking generated by the Fama-French three factor model corresponded by the majority to our liquidity ranking generated by the dynamic liquidity measure. From this results we concluded that the model is an appropriate measurement approach for asset and market liquidity.

In this thesis we presented a new approach to measure asset liquidity. We provided a substantial analysis of the model and verified its measurement power. In addition we showed several aspects and uses of the concept. This model and the idea behind it open different opportunities for further researches. The first use of this concept can be to support an investor in its asset allocation and asset selection process. The model can contribute additional information to a decision making process. Second, from a trading point of view the model can be evalu-

ated for a trading strategy. For instance a short-long position strategy can be tested, where a portfolio consisting of liquid assets are sold against a portfolio composed of illiquid assets. Another application can be to use the figure for risk monitoring and control for risk management. As mentioned in the beginning of this thesis we said that one assumption in a lot of valuation models is that assets can be sold immediately and for no costs. From this point of view, a further research field can be to incorporate the liquidity aspect in valuation models.
Liquidity is a fundamental attribute of a well working financial market. In fact, a market without liquidity can hardly be seen as working at all. The simple situation that an investor is not able to sell any given amount of assets at a given point of time can cause financial distress even up to its insolvency. From this point of view, market liquidity can be seen as the life elixir of financial markets.

References

Acharya, V. V., & Pedersen, L. H. (2005). Asset pricing with liquidity risk. *Journal of Financial Economics*, *77*, 375-410.

Admati, A. R., & Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial Studies*, *1*, 3-40.

Ahn, H., Bae, K., & Chan, K. (2001). Limit order, depth, and volatility: Evidence from the stock exchange of hong kong. *The Journal of Finance*, *56*(2), 767-788.

Ahn, H., Cao, C., & Choe, H. (1996). Tick size, spread and volume. *Journal of Financial Intermediation*, *5*, 2-22.

Ahn, H., Cao, C., & Choe, H. (1998). Decimalization and competition among stock markets: evidence from the toronto stock exchange cross listed securities. *Journal of Financial Markets*, *1*, 51-87.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716-723.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31-56.

Amihud, Y., & Mendelson, H. (1980). Dealership market: Market-making with inventory. *Journal of Financial Economics*, *8*, 31-53.

Amihud, Y., & Mendelson, H. (1982). Asset price behavior in a dealership market. *Financial Analysts Journal*, *29*, 50-59.

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

Amihud, Y., & Mendelson, H. (1989). The effects of beta, bid-ask spread, residual risk and size on stock returns. *The Journal of Finance*, *44*(2), 479-486.

Amihud, Y., & Mendelson, H. (1991). Volatility, efficiency and trading: Evidence from the japanese stock market. *The Journal of Finance*, *46*(5), 1765-1789.

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

Bacidore, J. (1997). The impact of decimalization on market qualitiy: an empirical investigationofthetorontostock exchange. *Journal of Financial Intermediation*, *6*, 92-120.

Baker, M., & Stein, J. C. (2004, June). Market liquidity as a sentiment indicator. *Journal of Financial Markets*, *7*(3), 271-299.

Banz, R. (1981, March). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*, 3-18.

Biais, B., Hillion, P., & Spatt, C. (1995, Dec.). An empirical analysis of the limit order book and the order flow in the paris bourse. *The Journal of Finance*, *50*(5), 1655-1689.

Black, F. (1971). Towards a fully automated exchange, part i. *Financial Analysts Journal*, *27*, 29-34.

Black, F. (1972). Capital market equilibrium with restricted borrowing. *Journal of Business*, *45*, 444-455.

Black, F., Jensen, M. C., & Scholes, M. S. (1972). The capital asset pricing model: Some empirical tests. In M. C. Jensen (Ed.), *Studies in the theory of capital markets* (p. 79-121). New York: Praeger Publishers Inc.

Boes, D., Graybill, F., & Mood, A. (1974). *Introduction to the theory of statistics* (3rd ed.). New York: McGraw-Hill.

Bollen, N., & Whaley, R. (1998). Are teenies better? *Journal of Portfolio Manager*, *25*, 10-24.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*, 307-327.

Bollerslev, T., Engle, R., & Nelson, D. (1994). Arch models, in: Handbook of econometrics. In R. Engle & D. McFadden (Eds.), (Elsevier ed., Vol. 4, p. 2959-3038). Amsterdam: North Holland Press.

Box, G., Jenkins, G., & Reinsel, G. (1994). *Time series analysis: Forcasting and control* (3 ed.). England: Prentice-Hall, Inc.

Brennan, M., Chordia, T., & Subrahmanyam, A. (1998). Alternative factor specifications, security characteristics and the cross-section of expected stock returns. *Journal of Financial Economics*, *49*, 345-373.

Brennan, M., & Subrahmanyam, A. (1996). Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics*, *41*, 441-464.

Brockman, P., & Chung, D. (2001). Managerial timing and corporate liquidity: Evidence from actual share repurchases. *Journal of Financial Economics*, *41*, 417-448.

Campbell, J. Y., Grossmann, S. J., & Wang, J. (1993, November). Trading volume and serial correlation in stock returns. *The Quarterly Journal of Economics*, *108*(4), 905-939.

Cao, C., Hansch, O., & Wang, x. (2004). The informational content of an open limit order book. *SSRN eLibrary.*

Cetin, U., Jarrow, R., Protter, P., & Warachka, M. (2003). *Option pricing with liquidity risk.* (Singapore Management University, Office of Research)

Chan, L., & Lakanishok, J. (1993). Institutional trades and intra-day stock price behavior. *Journal of Financial Economics*, *33*, 173-199.

Chan, L., & Lakanishok, J. (1995). The behavior of stock prices around institutional trades. *The Journal of Finance*, *50*, 1147-1174.

Chen, A., Kim, H., & Kon, S. (1975). Cash demand, liquidation costs and capital market equilibrium under uncertainty. *Journal of Financial Economics*, *2*, 293-308.

Chen, R.-R., & Yang, T. (1995). The relevance of interst rate processing in pricing mortgage back securities. *Journal of Housing Research*, *6*(2), 315-332.

Chiyachantana, C., Jain, P., Jiang, C., & Wood, R. (2004). International evidence on institutional trading behavior and price impact. *The Journal of Finance*, *59*(2), 869-898.

Choi, W.-S., & Lee, S.-B. (2000). Price limits and limit order flow: Evidence from the korea stock exchange. *Review of Pacific Basin Financial Markets and Policies*, *3*(4), 451-474.

Chordia, T., Roll, R., & Subrahmanyam, A. (2000). Commonality in liquidity. *Journal of Financial Economics*, *56*, 3-28.

Chordia, T., Roll, R., & Subrahmanyam, A. (2001). Market liquidity and trading activity. *The Journal of Finance*, *56*(2), 501-530.

Chung, K., & Chuwonganant, C. (n.d.). *Tick size, order handling rules and trading costs.*

Cochrane, D., & Orcutt, G. (1949). Applications of least square regression to relationships containing autocorrelates error term. *Journal of Amarican Statistcal Association*, *44*, 32-61.

Cochrane, J. (1997). *Time series for macroeconomics and finance* (Tech. Rep.). University of Chicago: Graduate School of Business.

Constantinides, G. (1986). Capital market equilibrium with transactions costs. *Journal of Political Economy*, *94*, 842-862.

Copeland, T. E., & Galai, D. (1983, Dec.). Information effects on the bid-ask spread. *The Journal of Finance*, *38*(5), 1457-1469.

Coppejans, M., Domowitz, I., & Madhavan, A. (2004). Resiliency in an automated auction. *Working Paper.*

Csavas-Szilard Erhart, C. (2005). *Are hungarian financial markets liquid enough?* http://econpapers.repec.org/paper/mnbopaper/2005_2F44.htm.

Danielsson, J., & Payne, R. (2001). *Measuring and explaining liquidity on an electronic limit order book: Evidence from reuters d2000-2.* SSRN.

Datar, V., Naik, N., & Radcliffe, R. (1998, August). Liquidity and stock returns: An alternative test. *Journal of Financial Markets*, *1*(1), 203-219.

Demsetz, H. (1968). The cost of transacting. *Quarterly Journal of Economics*, *82*, 35-53.

Dickey, D., & Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unitroot. *Journal of the American Statistical Association*, *74*, 427-431.

Dunbar, N. (2001). *Inventing money: The story of long-term capital management and the legends behind it.* John Wiley & Sons.

Durbin, J., & Watson, G. (1950). Testing for serial correlation in least squares regression i. *Biometrika*, *37*, 409-428.

Durbin, J., & Watson, G. (1951). Testing for serial correlation in least squares regression ii. *Biometrika*, *38*, 159-178.

Easley, D., & O'Hara, M. (1987). Price, trade size and information in securities markets. *Journal of Financial Economics*, *19*, 69-90.

Eisfeldt, A. L. (2004). Endogenous liquidity in asset markets. *The Journal of*

Finance, *59*, 1-30.

Ekinci, C. (2005, October). *Limit order book reconstruction and beyond: an application to istanbul stockexchange* (Finance No. 0510025). available at http://ideas.repec.org/p/wpa/wuwpfi/0510025.html: EconWPA.

Enders, W. (2004). *Applied econometric time series* (1 ed.). 111 River Street, Hoboken, NJ 07030: John Wiley & Sons, Inc.

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, *50*, 987-1007.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, *47*(2), 427-465.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return and equilibrium: Empirical tests. *Journal of Political Economy*, *81*, 607-636.

Fleming, M., & Remonola, E. (1999). Price information and liqudity in the u.s. treasury market: The response to public information. *The Journal of Finance*, *54*, 1901-1915.

Fleming, M. J. (2003). Measuring treasury market liquidity. *Economic Policy Review*, *9*(3).

Forster, M., & George, T. (1996). Pricing effects and the nyse open and close: Evidence form internationally cross listed stocks. *Journal of Financial Intermediation*, *5*, 95-126.

Garbade, K. (1982). *Securities markets*. New York: McGraw-Hill.

Gerety, M., & Mulherin, J. (1992). Trading halts and market activity: an analysis of volume at the open andtheclosse. *The Journal of Finance*, *47*(5), 1765-1784.

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

Glosten, L., & Harris, L. (1988, May). Estimating the compononents of the bid-ask spread. *Journal of Financial Economics*, *21*(1), 123-142.

Glosten, L., & Milgrom, P. (1985, March). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, *14*(1), 71-100.

Goldfeld, S., & Quandt, R. (1965). Some tests for homoscedasticity. *Journal of American Statistical Association*, *60*, 539-547.

Goldstein, M., & Kavajecz, K. (2000). Eighths, sixteenths and market depth:

change in the tick size and liquidity provision on the nyse. *Journal of Financial Economics*, *56*, 125-149.

Goldstein, M., & Kavajecz, K. (2004). Trading strategies during circuit breakers and extreme market movements. *Journal of Financial Markets*, *7*, 301-333.

Gouriéroux, C., Jasiak, J., & Le Fol, G. (1999). Intra-day market activity. *Journal of Financial Markets*, *2*, 193-226.

Granger, C., & Newbold, P. (1986). *Forecasting economic time series* (2 ed.). San Diego: Academic Press.

Greene, W. (2000). *Econometric analysis* (4 ed.). Upper Saddle River, New Jersey 07458: Prentice-Hall, Inc.

Griffiths, M., Smith, B., Turnbull, D., & White, R. (1998). The role of tick size in upstairs trading and downstairs trading. *Journal of Financial Intermediation*, *7*, 393-417.

Grossman, S. J., & Miller, M. H. (1988, July). Liquidity and market structure. *The Journal of Finance*, *43*(3), 617-633.

Hamao, Y., & Hasbrouck, J. (1995). Securities trading in the absence of dealers: Trades and quotes on the tokyostockexchange. *The Review of Financial Studies*, *8*, 849-878.

Hamilton, J. (1994). *Time series analysis* (1 ed.). Chichester, West Sussex: Princeton University Press.

Handa, P., & Schwartz, R. (1996, Dec.). Limit order trading. *The Journal of Finance*, *51*(5), 1835-1861.

Harris, L. (1986). A transaction data study of weekly and intradaily patterns in stock returns. *Journal of Financial Economics*, *16*(1), 99-117.

Harris, L. (1990a). *Liquidity, trading rules and electronic trading systems* (Paper 91-8). Southern California: School of Business Adminisitration.

Harris, L. (1990b, Jun.). Statistical properties of the roll serial covariance bid-ask spread estimator. *The Journal of Finance*, *45*(2), 579-590.

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

Harris, L. (1997). *Decimalization: a review of the arguments and evidence.* (Working Paper)

Harris, L., & Hasbrouck, J. (1996, Jun.). Market vs. limit orders: The super dot evidence on order submission strategy. *The Journal of Financial and*

Quantitative Analysis, *31*(2), 213-231.

Harris, L., & Panchapagesan, V. (2005). The information content of the limit order book: evidence from nyse specialsttrading decisions. *Journal of Financial Markets*, *8*(1), 25-67.

Hasbrouck, J., & Seppi, D. (2001). Common factors in prices, order flows and liqudity. *Journal of Financial Economics*, *59*, 383-412.

He, X., Velu, R., & Chen, C. (2004). *Commonality, information and return/return volatility - volume relationship.*

Heaton, J., & Lucas, D. (1996). Evaluating the effects of incomplete markets on risk sharing and asset pricing. *Journal of Political Economy*, *104*, 443-487.

Hedvall, K. (1994). *Essays on the market microstructure of the helsinki stock exchange* (Tech. Rep. No. 56). Helsinki: Swedish School of Economics and Business Administration.

Ho, T., Schwartz, R., & Whitcomb, D. (1985). The trading decision and market clearing under transaction price uncertainty. *The Journal of Finance*, *40*, 21-40.

Ho, T., & Stoll, H. (1981). Optimal dealer pricing under transactions and return uncertainty. *The Journal of Finance*, *45*, 579-590.

Ho, T., & Stoll, H. (1983). The dynamics of dealer markets under competition. *The Journal of Finance*, *38*, 1053-1074.

Holmström, B., & Tirole, J. (2001, October). Lapm: A liquidity-based asset pricing model. *The Journal of Finance*, *LVI*(5), 1837-1867.

Huang, M. (2003). Liquidity shocks and equilibrium liquidity premia. *Journal of Economic Theory*(109), 104-129.

Huang, R. D., & Masulis, R. W. (2003, 3). Trading activity and stock price volatility: Evidence from the london stock exchange. *Journal of Empirical Finance*, *10*, 249-269,.

Huberman, G., & Halka, D. (2001, Spring). Systematic liquidity. *Journal of Financial Research*, *24*, 161-178.

Jain, P., & Joh, G. (1988). The dependency between hourly prices and trading volume. *Journal of Financial and Quantitative Analysis*, *23*(3), 269-283.

Jarque, C., & Bera, A. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, *55*(5), 163-172.

Jones, C. (2002). *A century of stock market liquidity and trading cost.* (Working

Paper)

Jones, C. M., Kaul, G., & Lipson, M. L. (1994). Transactions, volume, and volatility. *The Review of Financial Studies*, *7*(4), 631–651.

Jorion, P. (2005). *Financial risk manager handbook* (Vol. 3). Hoboken: John Wiley and Sons.

Kalay, A., Sade, O., & Wohl, A. (2004). Measuring stock illiquidity: an investigation of the demand and supply schedules at the tase. *Journal of Financial Economics*, *74*, 461-486.

Kavajecz, K. A. (1999, April). A specialist's quoted depth and the limit order book. *The Journal of Finance*, *54*(2), 747-771.

Keynes, J. M. (1936). *The general theory of employement, interest and money.* New York: Harcourt, Brace and Company.

Kim, J.-W., Lee, J., & Morck, R. (2004). *Heterogeneous investors and their changing demand and supply schedules for individual common stocks.* SSRN.

Kothari, S. P., Shanken, J., & Sloan, R. G. (1995, March). Another look at the cross-section of expected stock returns. *The Journal of Finance*, *50*(1), 185-224.

Kwiatkowski, D., Phillips, P., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of aunitroot. *Journal of Econometrics*, *54*, 159-178.

Kyle, A. S. (1985, November). Continuous auctions and insider trading. *Econometrica*, *53*(6), 1315-1336.

Lee, C., Mucklow, B., & Ready, M. (1993). Spreads, depths and the impact of earnings information: An intraday analysis. *Review of Financial Studies*, *6*, 345-374.

Leippold, M. (2004). *Ökonometrie der finanzmärkte.* (Vorlesungsskript, Universität Zürich)

Levy, H. (1978). Equilibrium in an imperfect market: A constraint on the number of securities in a portfolio. *American Economic Review*, *68*, 643-658.

Lilliefors, H. (1967). On the kolmogorov-smirnov test for normality with mean and variance unknown. *Journal of American Statistical Association*, *62*, 339-402.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, *47*, 13-37.

Ljung, G., & Box, G. (1978). On a measure of a lack of fit in time series models. *Biometrica*, *65*, 297-303.

Lo, W., Andrew, MacKinlay, C., & Zhang, J. (2002). Econometric models of limit-order executions. *Journal of Financial Econometrics*, *65*, 31-71.

Longstaff, F. (1995). Option pricing and the martingale restriction. *Review of Financial Studies*, *8*(4), 1091-1124.

MacKenzie, D. (2003). Long-term capital management and the sociology of arbitrage. *Economy and Science*, *32*, 349-380.

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

McNeil, A., Frey, R., & Embrechts, P. (2005). *Quantitative riks management - concepts, techniques, tools* (P. U. Press, Ed.). New Jersey: Princeton University Press.

Mendelson, H. (1982). Market behavior in a clearing house. *Econometrica*, *50*, 1505-1524.

Mills, T. (1990). *Time series techniques for economists.* New York: Cambridge University Press.

Milne, F., & Smith, C. (1980). Capital asset pricing with proportional transaction costs. *Journal of Financial and Quantitative Analysis*, *15*, 253-265.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, *35*, 768-783.

Muranaga, J., & Shimizu, T. (1997). Market microstructure and market liquidity. *Working Paper*.

Naes, R., & Skjeltorp, J. (2006, July). Order book characteristics and the volume-volatility relation: Empiricalevidencefromalimit order book. *Journal of Financial Markets*, *9*, 408-432.

Ng, S., & Perron, P. (1995). Unit root tests in arma models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association*, *90*, 268-281.

Niemeyer, J., & Sandås, P. (1993). An empirical analysis of the trading structure at the stockolm stock exchange. *Journal of Financial Management*, *3*(3/4), 63-101.

O'Hara, M. (1995). *Market microstructure theory* (C. MA, Ed.). Blackwell.

O'Hara, M. (2003). Liquidity and price discovery. *The Journal of Finance*, *58*(4), 1335-1354.

O'Hara, M. (2004). *Liquidity and financial market stability.* National Bank of Belgium.

Parlour, C. (1998, Winter). Price dynamics in limit order markets. *The Review of Financial Studies*, *11*(4), 789-816.

Pascual, R., & Veredas, D. (2006). *Does the open limit order book matters in explaining long run volatility?* (Presentation University Zurich)

Pastor, L., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, *111*, 642-685.

Pereira, J. P., & Zhang, H. H. (n.d.). The liquidity premium in a dynamic model with price impact.

Phillips, P., & Perron, P. (1988). Testing for unit roots in time series regression. *Biometrica*, *75*, 335-346.

Phillips, S., & Simth, C. (1980). Trading costs for listed options: The implications for market efficiency. *Journal of Financial Economics*, *8*, 179-200.

Porter, B. R. (2003). *Measuring market liquidity.* Available at SSRN: http://ssrn.com/abstract=439122 or DOI: 10.2139/ssrn.439122.

Porter, D., & Weaver, D. (1997). Dezimalization and market qualitiy. *Financial Management*, *26*, 5-26.

Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 341-360.

Royston, J. (1982). An extension of shapiro and wilk's w test for normality to large samples. *Applied Statistics*, *31*, 115-124.

Said, E., & Dickey, D. (1984). Testing for unit roots in agressive moving average models of unknown order. *Biometrica*, *71*, 599-607.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461-464.

Schwert, W. (1989). Test for unit roots: a monte carlo investigation. *Journal of Business and Economic Statistic*, *7*, 147-159.

Shapiro, S., & Wilk, M. (1965). An analysis of variance test for normalitiy (complete samples). *Biometrica*, *52*(3 and 4), 591-611.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*, 425-442.

Stoll, H. (1985). Alternative view of market making. In T. H. Y. Amihud & R. Schwartz (Eds.), (Market Making and the Changing Structure of the Securities Industry ed.). Lexington Health, Lexington MA.

Stoll, H., & Whaley, R. (1990). Stock market structure and volatility. *Review of Financial Studies*, *3*, 37-71.

Strebulaev, I. (n.d.). Liquidity and asset pricing: Evidence from the u.s. treasury securities market.

Van Ness, B., R.A., V. N., & Pruitt, S. (2000). The impact of reduction in tick increments in major u.s. markets on spreads, depth, and volatility. *Review of Quantitative Finance and Accounting*, *15*, 153-167.

Vayanos, D. (1998). Transactions costs and asset prices; a dynamic equilibrium model. *Review of Financial Studies*, *11*, 1-58.

Wagenmakers, E.-J., Grünwald, P., & Steyers, M. (2006). Accumulative prediction error and the selection of time series model. *Journal of Mathematical Psychology*, *50*, 149-166.

West, R., & Tinic, S. (1971). *The economics of the stock market.* New York: Praeger.

Wyss, R. von. (2004). *Measuring and predicting liquidity in the stock market.* Unpublished doctoral dissertation, Universitiy of St. Gallen.

Yeo, W. Y. (2006). *Persistence in limit order flow.* SSRN.

Zivot, E., & Wang, J. (2003). *Modeling financial time series with s-plus* (I. Corporation, Ed.). Berlin: Springer-Verlag.

Georg Pristas studied finance and economics at the University of Zurich, Switzerland. During his graduate studies he worked for McKinsey & Company as a student. After his graduation in 2003 he joined the Swiss Banking Institute at the University of Zurich as a research assistant and started his doctoral studies in finance. In 2007 he completed his thesis under the supervision of Prof. Dr. Markus Leippold, University of Zurich.