



**Fa(s)ta Morgana? Estimating
Delusive Liquidity by Analysing the
High Frequency Market
Microstructure of Cross Listed
European Stocks**

P.A. Bonouvrie

FINANCIAL ENGINEERING & MANAGEMENT

31-10-14

EXAMINATION COMMITTEE

B. Roorda – University of Twente

R. Joosten – University of Twente

Preface

Nearly two years ago I started the master Financial Engineering & Management at the University of Twente. The master combined my interest in financial markets with a strong analytical basis which allowed me to understand the current financial industry much better than before. I quickly found out I wanted to graduate in the centre of the financial industry; the financial markets.

XXX

Additionally I thank my two supervisors from the University of Twente. Berend Roorda for the technical input which particularly convinced me to try to understand and explain my research much better than I did initially. Reinoud Joosten for the final remarks particular highlighting limitations of some chosen methods. However, I thank both mostly for their assistance during the part of the master before my graduation.

Finally I am grateful for the support of all people involved during my time in Enschede. If you are actually reading this preface you most likely are one of these people, so thanks! In particular I thank my family for allowing me to complete all the crazy things I loved to do, Huize Slappe Tuba for listening to all the nonsense I love to voice, and my best mates during my entire studies to which I owe many.

Management Summary

With the introduction of the Markets in Financial Instruments Directive (MiFID) in 2007 regulators aimed to ensure that investors were able to get the best possible execution of their orders by, amongst others, increasing competition in the market place. Due to this framework new trading facilities arose, leading to investors being able to buy or sell the same product at different venues. However, with fragmented trading across multiple exchanges it remains unknown at which of the exchanges an eventual order will arrive. We argue that as a result of this market makers send the same order to multiple exchanges to increase execution probabilities and therefore profits. This implies that we may overestimate the volume in the market if we combine the quoted volume available at each exchange as some of the orders will be withdrawn when the complementary order at another exchange gets filled. The volume arising from multiplication of such orders is what we call delusive liquidity. XXX

Contents

Preface	i
Management Summary	ii
1 Introduction.....	1
1.1 Background.....	1
1.2 Fragmentation	1
1.3 How Markets Work.....	2
1.4 Relevance to Thesis_Company_XXX	7
1.5 Research Goals.....	7
1.6 Contribution to Literature	8
1.7 Structure.....	8
2 Theoretical Framework.....	10
2.1 Black-Scholes-Merton Hedging Implications.....	10
2.2 Smart Order Routing.....	13
2.3 Tick Sizes.....	14
3 Data Selection & Sample Analysis	16
3.1 Three Measures of Delusive Liquidity.....	16
3.1.1 Intra Exchange Orders	16
3.1.2 Inter Exchange Top Down	16
3.1.3 Inter Exchange Bottom Up	16
3.2 Defining Variables at Delusive Liquidity	16
3.3 Defining Time Interval.....	17
3.3.1 Finding Minimum Time of Interval	17
3.3.2 Finding Maximum Delay of Interval	17
3.3.3 Calibrating Time Data	18
3.4 Intra Exchange	18
3.4.1 CHI-X Time Offset.....	20
3.4.2 XAMS Time Offset	20
3.5 Inter Exchange – Top Down Estimation.....	20
3.6 Inter Exchange – Bottom Up Identification.....	21
4 Analysis.....	23
4.1 Inter Exchange	23
4.1.1 ING Group.....	23
4.1.2 Aegon.....	23
4.2 Top Down Estimation of Delusive Orders.....	23
4.3 Tracking Orders	24

5	Conclusion & Discussion.....	25
5.1	Conclusion	25
5.2	Future research.....	25
5.3	Relevance Literature	26
5.4	Relevance Trading	26
	References.....	27
	Appendix.....	28
A	Glossary	28
B	Regression Plots.....	29

1 Introduction

1.1 Background

In financial markets liquidity is regarded to be of major importance. However, while we know several qualitative definitions of liquidity, a quantitative one remains hard to find. With trading speeds having increased considerably over the previous years the moment you start measuring liquidity it has already disappeared seems increasingly true. As Lehalle & Laruelle (2013, p.1) state “some simple qualitative definitions of liquidity exist ... [such as] an asset is liquid if it is easy to buy and sell it.” We also know several qualitative proxies of liquidity, each of them with limitations; the bid-ask spread (explained later) which puts little emphasis on available volumes, or round-trip costs which measure the immediate costs of buying or selling a security. When finding the costs for a range of volumes one can draw a curve that illustrates trading costs.

Additionally, when available order volumes are smaller than the volumes to be traded the trade price of the security shifts. We call this the market impact of a trade. The earlier definitions of liquidity ignore the concept of synchronization; if at near time moments a larger buy order is offset by a larger sell order the market impact is little. This offsetting appears when the initial order is not in line with the market consensus so that the opposing order will push the price back and therefore the market impact is only of temporary effect as prices return to pre-trade price levels.

One of the possible determinants of liquidity is the microstructure of the market. This is an area of the market that is also of great interest to market makers and algorithmic traders. O’Hara (2001, p.3) defines market microstructure as “the process and outcomes of exchanging assets under a specific set of rules. While much of economics abstracts from the mechanics of trading, microstructure theory focuses on how specific trading mechanisms affect the price formation process.”

Due to recent developments such as the introduction of dark pools (hidden liquidity) and high-frequency trading some securities may appear liquid one moment, while it vaporises the next. Additionally, fragmentation impacted the markets; average volumes per trade decreased but the total number of trades and orders increased. We aim to increase our understanding of liquidity in the current market landscape. We do not aim to quantify liquidity as a whole, but the analysis of market microstructure might be helpful in collecting early signals of liquid/illiquid environments.

1.2 Fragmentation

Over the past decades regulators have aimed to increase competition in the market place, with the implementation of Regulation National Market System (RegNMS) in the USA in 2005 and its European Union analogue Markets in Financial Instruments Directive (MiFID) in 2007 as the main regulations. Regulators emphasise the importance of ensuring market participants being able to get the best possible execution of their order, by amongst others enforcing increase pre- and post-trade transparency. They aim to achieve these aspects by encouraging increased competition in the market place, increased competition on individual orders, and increased consumer protection in the investment service industry.

As a result of these encouragements new market exchanges arose and competition among them increased. The same assets currently trade on multiple exchanges at the same time. These exchanges compete for clients' orders in for example transaction costs and execution speed. Possible effects of fragmentation will be discussed in future chapters, but by logical reasoning we can assume transaction costs for investors decreased due to competition on fees and minimum tick sizes. Lower regulatory minimum tick sizes will decrease transaction costs as minimum spreads will decrease, and therefore trading liquidity tends to move to the lowest tick size environment as shown during a 2009 tick size war in Figure 1.1.1 (Lehalle & Laruelle 2013). The spikes show the markets share of an exchange shortly after the exchange lowered the minimum tick size to a lower level than the competitors. After considering the impact of competition on tick sizes, exchanges agreed to standardise tick sizes across

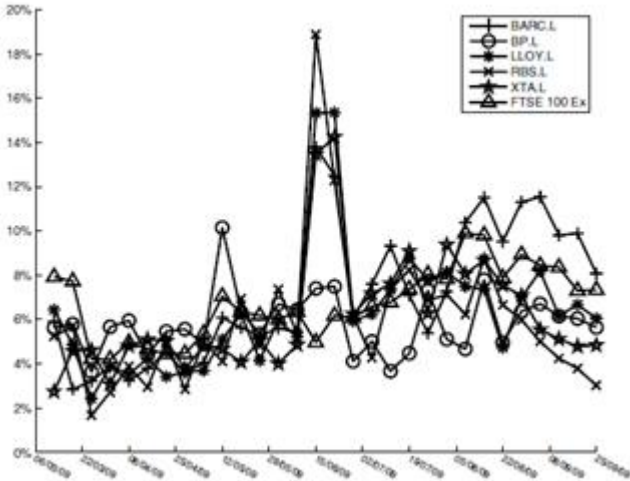


Figure 1.1.1: 2009 Tick Size War (Lehalle & Laruelle 2013).

exchanges for the same securities. While tick sizes of identical securities are harmonised across exchanges, different securities still trade on different tick sizes often related to the traded price of the security. With identical securities trading across exchanges, according to Van Kervel (2012, p.2) "... the order flow and liquidity of these [exchanges] become strongly interrelated." We aim to understand this order flow and the effects of the order flow at liquidity provision across different exchanges.

1.3 How Markets Work

In order to understand the general business model of market makers it is helpful to have a basic understanding of how financial markets work nowadays. Markets changed considerably over the previous 50 years through technological innovations. Before 1970 potential buyers and sellers of a security were matched at the floor of an exchange, also referred to as floor trading. You either needed a physical presence or you needed to be in contact with someone at the floor to be able to buy or sell a financial product. With the introduction of computer programs this evolved to traders placing their orders through the screen, while currently an increasing number of trades execute without direct human decision making through algorithms.

While these innovations have had numerous effects on trading, the basic rules that define who is entitled to the trade remain the same; an exchange prioritises the queue of interested buyers and sellers initially based on price, and if prices are equal the exchange prioritises the queue on speed; whoever placed the price first is leading the queue.

One of the main functions of exchanges is to establish a place where buyers are matched with sellers; a market. Exchanges profit mainly from transaction costs for each trade and from companies listing their shares at the particular exchange, and therefore exchanges aim to improve liquidity at their exchange.

Market makers supply this liquidity by sending quotes to the exchange, and according to Menkveld (2013) market makers appear crucial to liquidity improvements. A quote consist of four parts; the price for which market makers are willing to sell (ask price), the sell volume, the price for which they are willing to buy (bid price) and the buy volume. The difference between the ask and the bid price is called the bid-ask-spread, and market makers profit from investors *crossing the spread*; if a market maker is able to sell at this ask and buy at this bid it will generate a profit of $(ask - bid) * volume\ traded = spread * volume\ traded$.

Menkveld (2013) argues originally market makers largely competed on price rather than speed. As it is slow and therefore expensive to calculate prices for every security in the market, market makers usually have agreements with exchanges. This agreement includes quoting obligations for market makers for which they might get trade size protection or a reduction in fees in return. With the fragmentation of exchanges however, competition on both speed and prices increased. Not only do market makers quote at multiple exchanges to increase the likelihood of getting the actual order, they also should quote the price faster than other market participants at each exchange to increase the likelihood of getting the actual trade. As a result of this the high frequency market maker, or modern market maker emerged from the traditional market maker that just competed on price.

Foucault *et al.* (2005) explain the interaction between aggressive and passive traders. They distinguish passive traders which place *limit orders* and aggressive traders which place *market orders*. In general a limit order supplies liquidity for future traders and as Foucault *et al.* (2005, p.2) states “with a limit order, a trader can improve the execution price relative to the market order price, but the execution is neither immediate, nor certain”.

At a vanilla limit order traders supply a price and a quantity, the limit order then enters the order book at the back of the queue at the specified price. The limit order will not be executed until another investor is willing to trade for that price or better. Therefore, when sending limit orders traders add to volume quoted at that price level hereby supplying liquidity for future traders. We refer to participants sending limit orders as *liquidity providers*. There is one exception in which a limit order can get direct execution and therefore does not provide liquidity, namely when the traders buy limit order arrives when the pre specified buy price equals the best available ask price and vice versa. This will be illustrated with Example 1.3.1 later.

Foucault *et al.* (2005, p.2) define a market order as an “order [that] guarantees an immediate execution at the best price available upon the order arrival.” The market order therefore implies an investor’s demand for direct execution while taken the current quoted prices as given. At a market order traders supply a quantity only and the order will be executed directly at the best price available in the order book by being matched to an earlier placed limit order. Clearly, at both types traders specify whether they aim to sell or to buy. The market order’s direct execution decreases the volume quoted at the traded price by taking out volume equal to the trade size, so we refer to participants sending market orders as *liquidity takers*.

Consequently, the optimal order is then composed of the trade-off between the costs of time delay – waiting costs – when placing a limit order and the costs of the current market price – no price improvement - at an immediate market order. They observe an increase in passive traders results in more competitive limit orders with patient traders aiming to reduce waiting times, resulting in narrow spreads and increased market resilience with the resilience being the deviation of spreads from competitive levels due to liquidity demand shocks. An increase in aggressive traders implies the opposite. They predict a positive correlation between trading frequency and spreads as quoted volumes

decrease when more aggressive traders are present. Market makers can be seen as liquidity providers, so in line with the observations of Menkveld (2013), Foucault *et al.* (2005) agree on the importance of market makers related to liquidity improvements.

As noted earlier market makers aim to profit from buying and selling the same security and hereby collecting the spread. Foucault & Menkveld (2008) state the spread is compromised by three things; order handling costs (including fees paid to exchanges), the costs of being adversely selected on the bid or ask side of the quote and thirdly the premium market makers require for price risk on nonzero positions. This price risk occurs when just one side of their quote is executed and they are thus exposed to price changes in the underlying, which we call market risk.

Furthermore, O'Hara (2014, p.3) argues that “[a]t very fast speeds, only the microstructure matters” as exchange matching engines are the central focus of high frequency traders that optimise against market design. She states the goal of high frequency traders is generally to be the first in queue to trade and argues the degree of being able to achieve this goal depends mainly on the understanding of the microstructure of the market. This requires optimising the strategy against the matching engine of particular exchanges and therefore introduces concepts such as co-locations (a physical presence within the exchange) and fast connections lines. Additionally, she emphasises the increase in complexity of trading in today’s environment even for non-high-frequency traders. This follows from the ability of high frequency trading algorithms to observe deterministic trading patterns such as trading based on rules similar to minimising TWAP (time weighted average price) or VWAP (volume weighted average price) resulting in these strategies becoming less effective.

On the other hand, exchanges profit from creating attractive microstructures for high frequency traders to attract volume and liquidity, introducing a variety of order types, rebate schemes and access schemes. Names and characteristics of order types differ across exchanges. Most of them are designed for specific needs, such as iceberg orders which are limit orders allowing traders to present only parts of their order size in the market. The order then immediately refills to a pre-specified amount whenever the order gets (partial) execution. These orders are used by traders who aim to sell or buy large sizes without being ‘vulnerable’ by showing the entire size at once. Nearly all of these designed forms of orders are limit orders, but when observing the order book the nature of the orders remain unknown; we can solely observe aggregated volumes at each price level.

One order type of interest to this report is the *fill-or-kill* (also called *immediate-or-cancel*). Fill-or-kill orders are limit orders that allow participants to specify a price and size but will remove the order if it does not get an immediate fill (Garland *et al.* 2014). Their execution is similar to market orders, but at market orders you only specify a volume and fill up to whatever price for the full volume. The fill-or-kill order type will solely get an immediate fill when you specify the buy or sell level at the exact top level. Example 1.3.1 illustrates the orders discussed.

Example 1.3.1 - The current top level of the order book lists 10 lots on the ask side with ask price €10.00 and 10 lots on the bid side with bid price €9.00. The start table below shows the base situation, the marked cells in each of the discussed cases mark the changes from the base. We show two depth levels in the tables. The situation in the start table is as follows, if we are interested in buying the product we can buy up to 10 lots for price €10.00, if we are interested in buying 15 lots we then have to pay €10.50 for the final 5 lots *ceteris paribus*. Similarly, if we are interested in selling the product we can sell up to 10 lots for €9.00, but if we aim to sell 15 lots we can only sell the final 5 lots for €8.50. The tables shown below for each of the following cases show the change from the starting situation for that case with buy orders of size 1 and sell orders of size 2.

Start				
Level	Bid Size	Bid Price	Ask Price	Ask Size
1	10	€9.00	€10.00	10
2	5	€8.50	€10.50	5

- i. An arrived buy market order will be executed directly at €10.00 and an arrived sell market order at €9.00 respectively. The available volume at the top level will decrease by the order size.

i				
Level	Bid Size	Bid Price	Ask Price	Ask Size
1	8	€9.00	€10.00	9
2	5	€8.50	€10.50	5

- ii. An arrived sell limit order with specified price €10.00 will join the ask side at the end of the queue, an arrived buy limit order with specified price €9.00 will join the bid side at **the end of the queue**. When considering limit orders we reason the other way around. The two orders do not have direct execution as currently there is nobody in the market willing to buy your order for €10.00, or sell you the product for €9.00. You therefore have to wait before another trader enters who is willing to buy or sell for that price. However, as your limit orders were send in later than the orders visible in the market you will be in queue behind these orders. The available volume at the top level increases by the order size. As the specified prices are equal to the best available prices we still denote this order at the top level, however as the orders arrived later they will only get executed when the 10 lots in front of them are taken out of the market by either trades or by order removal. The 10 lots have what we call *queue priority*.

ii				
Level	Bid Size	Bid Price	Ask Price	Ask Size
1	10	€9.00	€10.00	10
1	1	€9.00	€10.00	2
2	5	€8.50	€10.50	5

- iii. An arrived sell limit order at €9.90 will improve the current ask price and therefore be on top of a new queue at €9.90. An arrived buy limit order at €9.10 will improve the current bid price and therefore be on top of a new queue at €9.10. This is often referred to as *jumping the queue*. Both orders then create a new best price with a queue size is equal to the order size until new limit orders join at that level.

iii				
Level	Bid Size	Bid Price	Ask Price	Ask Size
1	1	€9.10	€9.90	2
2	10	€9.00	€10.00	10
3	5	€8.50	€10.50	5

- iv. An arrived buy limit order at €10.00 or an arrived sell limit order at €9.00 however both get direct execution, hereby decreasing the available volume and thus taking liquidity similar to case i. The difference with a market order lies in the fact that if just before arrival the market changed to €10.10 - €8.90, then the limit order will not be executed directly since the current best price is worse than specified, whereas the market order will still be executed immediately at €10.10 and €8.90. In the latter case a *fill-or-kill* limit order will be withdrawn directly,

while a vanilla limit order will start a new queue at the specified price similar to case iii. When sending fill-or-kill limit orders with the characteristics of ii. the order will directly be cancelled as there will still be 10 lots in queue before this order. We will discuss this practice further in Chapter 3.

iv

Level	Bid Size	Bid Price	Ask Price	Ask Size
1	8	€9.00	€10.00	9
2	5	€8.50	€10.50	5

Some exchanges specialise into designing market microstructures that specifically limit the involvement of high frequency traders. Exchange IEX delays each order by a randomised 10 or 15 microseconds and exchange Aequitas only permits retail and institutional traders to take liquidity(O'Hara 2014).

With fragmented trading across multiple exchanges it still remains unknown at which of these exchanges the eventual trade will take place. As a result of this uncertainty, high frequency market makers present orders for the same underlying at multiple exchanges simultaneously. However, this spreading of orders increases market risk not only because they send out more orders at once - all exposed to price changes- , but also because market makers likely quote for more aggregated volume and therefore might trade a higher volume than desired. The spreading of orders is required to increase the likelihood of capturing the entire trade size if it arrives at a certain exchange. According to Lehalle & Laruelle (2013, p.63) these “strategies often imply that as soon as one of those orders is executed, all the remaining ones are immediately cancelled.”

The aggregated quoted volumes across exchanges therefore might overestimate the actual total volume available as a single trade may decrease the aggregated quoted volume by more than the trade size; in other words the aggregated quote volume contains *delusive liquidity*. This information however continues to be largely invisible in order books; due to anonymous quoting it remains unknown which party supplies the order. Additionally we expect different methods of fast limit order withdrawals that lead to an overestimation of visible liquidity at each price level if the volume signals are received in between the insertion and withdrawal of the orders. We aim to understand such presence of delusive liquidity, with the fast limit order withdrawals and equal orders sent to multiple exchanges as main focus areas.

We limit the scope of this report to the analysis of two products for just two trading facilities, XAMS and CHI-X. Strictly speaking XAMS is an exchange while CHI-X is a multilateral trading facility (MTF). The main difference between both is that the owner of the exchange will not trade on its own platform, where the owners of a MTF are able to trade on their own platform. We will refer to both as exchanges in the remainder of the report.

XXX

Even though the two exchanges combine to a market share of about 85%, we should keep in mind that by not considering all exchanges we may not get an overall overview of delusive liquidity. If a market participant sends orders to another exchange than one of the considered exchanges we will be unable to notice this delusive liquidity. Furthermore, equal sized orders arriving nearly simultaneously at three or more venues automatically have a higher explanatory value simply because the probability of multiple equal sized orders randomly arriving simultaneously at multiple exchanges decreases with an increasing number of considered exchanges. Given that rationale we believe the eventual outcome of

this analysis will underestimate the degree of delusive liquidity as we are most likely to neglect equal order pairs if they were partly sent to a non-considered exchanged.

1.4 Relevance to Thesis_Company_XXX

We will be analysing the liquidity effects in stocks even though the core business of Thesis_Company_XXX lies in the trading of derivatives, (equity) options in particular. In order to understand the relevance of this report to the company we explain one of the core principles applied in finding the price of financial derivatives; risk-neutral pricing. The foundation of many option pricing methods lies in a model developed by in 1973; the Black-Scholes-Merton model. Before deriving their the authors define a set of ideal market conditions as follows (Black & Scholes 1973);

1. The short-term interest rate (r) is known and constant trough time.
2. The stock price follows a random walk in continuous time with a variance rate proportional to the square of the stock price. Thus the distribution of possible stock prices at the end of any finite time interval is log-normal. The variance rate of the return on the stock is constant.
3. The stock pays no dividends or other distributions (q).
4. The option is 'European,' that is it can only be exercised at maturity ($T-t$).
5. There are no transaction costs in buying or selling the option.
6. It is possible to borrow any fraction of the price of a security to buy it or to hold it, at the short term interest rate.
7. There are no penalties to short selling. A seller who does not own a security will simply accept the price of a security from a buyer, and will agree to settle with the buyer at some future date by paying him an amount equal to the price of the security at that date.

They state that when these assumptions hold, the value of the option will depend on time, the price of the stock and all other variables that are to be taken as constant. While some of these assumptions do not hold in reality, the ideas that emerged from this model still apply.

The main idea is as follows; investors can create a hedged portfolio consisting of positions in the option and the underlying. If the hedge is maintained continuously then the return of this portfolio is independent of a change in the stock price and the return becomes certain. By non-arbitrage principles the expected return of this hedge portfolio should then be equal to the risk-free rate. Furthermore, when taking the short rebalancing hedge interval dt , the costs of hedging converges to the option price in the limit as $dt \rightarrow 0$.

We will explain the multiple ways you receive positions in the underlying (deltas) from trading options in Chapter 2 by elaborating on the Black-Scholes-Merton model. In practice, maintaining a continuous hedge is nearly impossible, but when reducing the costs of hedging by improved understanding of the underlying you are able to improve your theoretical option prices (Kat 2001). Additional insight regarding the available volumes in the underlying will also improve abilities to hedge perfectly for the required volume. Finally, considering the business model of market makers or high frequency traders, concepts resulting from the analysis of the stocks might very well be applicable to other asset classes as well.

1.5 Research Goals

Liquidity is of major importance to financial markets. Yet it seems a paradox that we know several qualitative definitions of liquidity, but still have not found an overall quantitative one. In order words, liquidity is fairly easy to describe but hard to measure. With the fragmentation of the market, liquidity has become harder to observe as the same security can be listed in multiple exchanges. Additionally,

with increased algorithmic trading a security may appear liquid at one moment while the liquidity disappears the next.

One of the possible determinants of liquidity is the microstructure of the market, i.e. how is the top of the order book composed. Within the microstructure we will aim to find measures of delusive liquidity. We do not aim to quantify liquidity as a whole, but the analysis of market microstructure aspects might be helpful in collecting early signals of liquid/illiquid environments. The following questions should be answered in order to reach these goals;

1. How do we identify delusive liquidity?

As explained earlier, we assume that liquidity providers send the same orders to multiple exchanges in order to increase execution probabilities (and therefore profits). This implies that we may overestimate the available volume in the market as some of the orders may be withdrawn when the complementary order at another exchange gets filled; delusive liquidity. We will suggest a methodology to estimate the relation between transactions at one exchange and quoted volumes at the same side of the order book in the same security at other exchanges. The degree of correlation between the orders and one side of the order book and transactions in the moment just before at another exchange should give us some idea of the degree of overestimation at available volumes. Secondly, we aim to give insight into practices that focus on the quick withdrawal of limit orders shortly after the order was inserted.

2. How much delusive liquidity is present at certain exchanges?

We now apply the developed methodology to two stocks at two exchanges, which will result in estimating delusive liquidity four times in total. We will aim to express the degree of delusive liquidity, and we aim to explain the influence of such delusive liquidity on daily trading.

3. How can we incorporate the findings into trading strategies?

Several trading strategies depend on the volume available at preferred price levels. Being able to better estimate the actual available volume should be beneficial to the performance of such strategies.

1.6 Contribution to Literature

O'Hara (2014, p.27) highlights several issues in the current field of research related to market microstructure suggesting the need for fundamental changes in this field as a result from the changed market characteristics in the past few years. With the increased quantities of data from several markets and trading venues she states these “pose challenges for even basic analyses of microstructure data. In the high frequency era, we need some new tools in the microstructure tool box”.

Fortunately we have access to data of much higher frequency than observed in currently published literature. XXX

This allows us to analyse the ‘jungle’ of the sub second order books, resulting in developing an initial understanding of the high speed market activity and resulting in blueprints for the new tools for analysing market microstructure O’Hara (2014) calls for.

1.7 Structure

The structure of the thesis is as follows;

Chapter 2 – We give a short review of related literature and explain the relevance of managing the underlying for an option trading company by showing how changes in several parameters of the Black-Scholes framework affect your positions in the underlying (the delta).

Chapter 3 – XXX

Chapter 4 – XXX

Chapter 5 – XXX

2 Theoretical Framework

Over the past twenty years several researchers analysed the microstructure of the market, such as Hendershott & Riordan (2009), Angel (1997), Menkveld (2013), Foucault & Menkveld (2008) and O’Hara *et al.* (2013). However, little research has been published analysing the microstructure of a market at microsecond intervalsXXX. Notably, an April 2014 working paper by O’Hara (2014, p.3) highlights the same calling for “a new research agenda, one that recognises that the learning models we used in the past are now deficient and that the empirical methods we traditionally employed may no longer be appropriate.” She believes that in modern markets the underlying order is now the unit of information rather than the trade itself. With large orders being split in child orders – and therefore these child orders being not independent – the trades resulting from the child orders will not be independent as well. Furthermore, while in the past a large buy order would have been executed as one or possibly several market orders, sophisticated algorithms now allow the order to be split into multiple limit orders therefore now showing up in the data as many small sell orders. The importance of this difference is illustrated by Menkveld & Yueshen (2013) who show that even though the U.S. Securities and Exchange Commission (SEC) identified the causal factor of the 2010 Flash Crash to be a large sell market order, the execution of this order involved large numbers of non-direct execution limit sell orders providing liquidity rather than taking it.

We will discuss some of the concepts offered by the earlier research and examine whether we can apply them into the higher frequency data we are able to access. But we start with a brief introduction into option theory to explain the role of the underlying in options trading. Thesis_Company_XXX is mainly trading options, and by understanding the basics of the option theory we show how an understanding of stocks is related to the Thesis_Company_XXX options trading strategies.

2.1 Black-Scholes-Merton Hedging Implications

Options are financial derivatives with pre-specified characteristics. A *call* option gives the holder the right buy the underlying asset for a certain price at a certain time. A *put* option gives the holder the right to sell the underlying asset for a certain price at a certain time. When buying options you acquire the right, not the obligation, to exercise the option contract against the defined characteristics. Logically, this right has a value and therefore you pay a premium for the contract. This option premium depends on several parameters which we explain below.

The payoff of a call at time t is $\max\{S_t - K, 0\}$, the payoff of a put at time t is $\max\{K - S_t, 0\}$, where K is the predefined price (strike), and S_t is the value of the underlying. The $S_t - K$ for calls respectively $K - S_t$ for puts is what we call the intrinsic value of the option. When the intrinsic value is positive we call the option *in-the-money* (ITM), when the intrinsic value is negative we call the option *out-of-the-money* (OTM) and when the intrinsic value is zero we call the option *at-the-money* (ATM) as illustrated in Table 2.1.1

Table 2.1.1: Intrinsic Value Options

Option	Call	Put
In the money (ITM)	$S_t > K$	$S_t < K$
At the money (ATM)	$S_t = K$	$S_t = K$
Out of the money (OTM)	$S_t < K$	$S_t > K$

As stated in Section 1.4 when following the assumptions under the Black-Scholes-Merton model, the price of a derivative is a function of the stochastic variables underlying the derivative and time. The model of stock price behaviour used by Black-Scholes-Merton assumes that percentage returns in a

short period of time are normally distributed, assuming the following stock price process for a non-dividend paying stock;

$$dS = \mu S dt + \sigma S dz \quad (2.1)$$

Suppose that the price f of an equity option is a function of S and t , as earlier stated. Then from Ito's Lemma we know the process followed by f is

$$df = \left(\frac{\partial f}{\partial S} \mu S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial f}{\partial S} \sigma S dz \quad (2.2)$$

While understanding of the derivation of these formulas is of low relevance to this report, knowing the intuition behind it benefits in interpreting how derivatives traders 'receive' positions in the underlying while trading the derivative. We see both the equity option and the stock price itself are affected by the same source of uncertainty; the Wiener process dz . This illustrates that whenever a portfolio is established with a position in the stock and an appropriate position in the option, the gain or loss in the position in the stock will be offset by the gain or loss in the position in the option. If the portfolio consists of -1 derivatives and $+\frac{\partial f}{\partial S}$ shares we can eliminate the source of uncertainty dz . This portfolio is therefore riskless for an instantaneously short period in time and should yield the risk-free rate, which yields the differential equation (Hull 2009);

$$\frac{\partial f}{\partial t} + rS \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf \quad (2.3)$$

The Black-Scholes-Merton formula is a solution to this equation for the prices of a European option;

$$call = S_0 e^{(q-r)T} N(d_1) - K e^{-rT} N(d_2) \quad (2.4)$$

$$put = K e^{-rT} N(-d_2) - S_0 e^{(q-r)T} N(-d_1) \quad (2.5)$$

With strike K , risk free rate r , cost of carry q , time to maturity T , and the cumulative probability distribution for a standardised normal distribution $N(X)$. $N(d_2)$ can be interpreted as the risk-adjusted probability that the option will be exercised, with

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(q + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}$$

$$d_2 = d_1 - \sigma\sqrt{T}$$

In order to retain the riskless portfolio the position needs to be rebalanced continuously. The ratio required to obtain the riskless portfolio for one unit of a derivative is called the delta (Δ). We identified this ratio earlier as $\Delta = \frac{\partial f}{\partial S}$, and thus know the delta is the first derivative of the option price with respect to the underlying, or in other words the delta is the slope of curve of the option price to the underlying price. In case of the equity options, for each position in the option you receive Δ shares in the underlying when retaining a hedged portfolio. When taking the first derivative of the options prices as given in (2.4) and (2.5), we can see that

$$\Delta(call) = e^{(q-r)T} N(d_1) > 0 \quad (2.6.1)$$

And

$$\Delta(put) = -e^{(q-r)T} N(-d_1) < 0 \quad (2.6.2)$$

We now know that long calls generate long deltas while long puts generate short deltas. Furthermore we know stocks have $\Delta = 1$. The hedge portfolio should be delta neutral, i.e. $\Delta(\text{portfolio}) = 0$ and therefore depending on the positions in the options we adjust the position in the stocks to bring the portfolio delta back to zero (Hull 2009).

The delta of an option changes when any of the pricing parameters change, requiring additional hedging steps to achieve or maintain delta neutrality, most notably;

- I. The first derivative of the delta with respect to the underlying rate; called the gamma (Γ). Gamma informs investors about the sensitivity of the delta with respect to changes in the underlying price. The underlying itself has zero gamma, options carry positive gamma with the gamma of put and call options with equal characteristics being equal as well. In general Gamma is the highest for at the money options; the area in which the delta moves a lot (Haug 2003a).

$$\Gamma = \frac{\partial \Delta}{\partial S} = \frac{\partial^2 f}{\partial S^2} = \frac{N'(d_1)e^{(q-r)T}}{S_0\sigma\sqrt{T}} \quad (2.7)$$

- II. The rate of change of the delta for a change in volatility ($D\text{delta}D\text{vol}$), which is mathematically the same as the sensitivity of volatility (Vega) to the underlying also called Vanna. An increased volatility indicates the underlying moves more, which increases the probability of an option moving from in the money to out of the money or vice versa. Therefore the absolute delta of in the money options decreases, whereas the delta of out of the money options should increase at an increase in volatility (Haug 2003b).

$$D\text{delta}D\text{vol}(\text{call}) = \frac{\partial \Delta}{\partial \sigma} = \frac{\partial c}{\partial S \partial \sigma} = \frac{-e^{(q-r)T}d_2}{\sigma} N'(d_1) \quad (2.8.1)$$

$$D\text{delta}D\text{vol}(\text{put}) = \frac{\partial \Delta}{\partial \sigma} = \frac{\partial p}{\partial S \partial \sigma} = \frac{e^{(q-r)T}d_2}{\sigma} N'(d_1) \quad (2.8.2)$$

- III. The sensitivity of the delta to changes in time, also called Delta Bleed, which gives an indication of what happens to the delta when we move closer to maturity. We can understand the intuition behind time effects on delta in similar reasoning as at the volatility effects on delta. When the time to maturity – the time left until the option expires – decreases, there will be less time for the underlying to move away from its current value. Thus in the money options will be less likely to end up out of the money and vice versa. Therefore the delta of in the money options is higher when nearing expiration, while the delta of out of the money options is lower when time to maturity is smaller. As we typically look at decreasing time to maturities we express the derivative as minus the partial derivative (Haug 2003b).

$$\text{Delta Bleed} = -\frac{\partial \Delta}{\partial t} = -\frac{\partial f}{\partial S \partial t} \quad (2.9)$$

Figure 2.1.1 shows how changes in the parameters discussed affect the delta of options, and therefore affect the required position in the underlying in order to retain the riskless portfolio. The figure shows asset (gamma) and time (delta bleed) effects on delta. For a non-dividend paying stock the absolute delta will always be between 0 and 1. However if the cost of carry is larger than the interest rate, or if the interest rate is negative, the multiplier $e^{(q-r)T}$ will be larger than one and as shown in Equations 2.6.1 and 2.6.2 this can result in deltas larger than 1. The latter is an addition to the original BSM method, and will rarely happen in reality unless options are far in the money with a large remaining time to expiration.

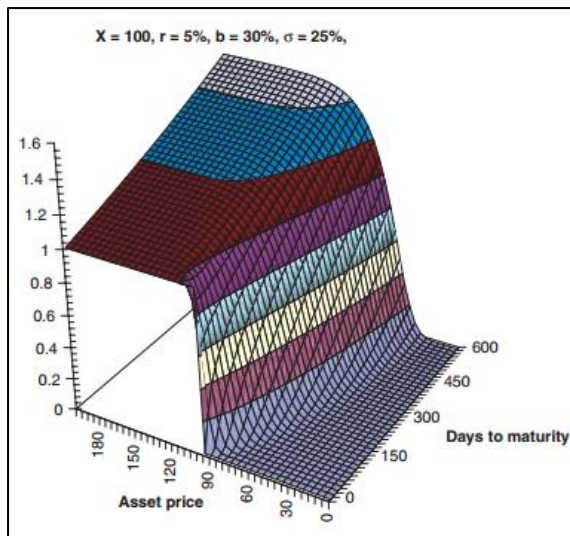


Figure 2.1.1: Sensitivity Spot Delta (Haug 2003a).

As stated in Section 1.4 the total hedging cost should amount to the option price. Now that we developed an understanding as to how option parameters affect positions in the underlying rate, we know that by improving understanding of the underlying rate we improve estimations of total hedging costs and therefore improve the option pricing.

2.2 Smart Order Routing

The increased emphasis on trading costs after the introduction of the financial market regulations MiFID and RegNMS mentioned earlier accelerated technological improvements at order routing, best execution and market transparency. MiFID defines this best execution as the need for investment firms to “take all reasonable steps to obtain .. the best possible result .. taking into account price, costs, speed, likelihood of execution and settlement, size, nature or any other consideration relevant to the execution of the order” (Gomber & Gsell 2006, p.6).

The basis of these technological improvements lies in the Direct Market Access (DMA), an execution concept offered by brokers where orders are forwarded directly to the market without intervention of the broker which could be offered at lower costs. In order to improve execution Smart Order Routing (SOR) systems were developed. SOR defines the best exchange out of a list of exchanges by taking into account different attributes of each exchange such as price, costs, market impact, liquidity, execution probability and speed. SOR concepts require immediate access to visible order books across the exchanges. The systems incorporate real time information mixed with historical data of the venues to be able to find the best execution at any moment in time. Gomber and Gsell (2006) find that sophisticated SOR systems include functionality for liquidity detection; sending so called immediate-or-cancel orders in order to find hidden liquidity. Additionally these concepts focus on splitting orders to minimise market impact.

Foucault and Menkveld (2008, p.120) stress the importance of brokers adapting Smart Order Routing systems for liquidity in the market. They state that in fragmented markets, orders might execute at a price worse than the best quoted price at another platform – so called *trade-throughs* – as brokers “may give up an improvement in execution price to economise on monitoring costs and the time required for splitting orders”. In general it is assumed that trade-throughs reduce liquidity provision, although they acknowledge there is little empirical evidence on the effects of trade-throughs. This assumption follows from the late limit order trader’s ability to jump ahead of the queue by placing a better limit order at the other venue when trade-throughs possibly exist, therefore reducing the

incentives for deeper order books as fill probabilities decrease further. The intuition of the latter is shown in Example 2.2.1.

Example 2.2.1 – At a less liquid exchange X the top level is quoted as €11.00 - €9.00, while the top level at exchange Y is quoted €10.50 - €9.50. This indicates the best available price is at exchange Y. If an order is executed at exchange X the order is not executed at the best available price in the market, which is called a *trade-through*. If a late limit order trader sends in a sell limit order of €10.50 or a buy limit order of €9.50 at exchange Y it will join the back of the queue. However, if the late limit order trader sends in the same at exchange X it will create a new queue at €10.50 respectively €9.50, thereby being on top of that queue. As it remains unknown at which exchange the eventual trade will take place, everyone quoting later in queue at exchange Y is now disadvantaged by imbalance caused through exchange X. If such situations occur regularly, incentives to join the queue at exchange Y diminish.

Additionally, Foucault & Menkveld (2008, p.120) define smart order routers as brokers who “automate the routing decision to obtain the best execution price” and consequently non-smart routers as brokers who “ignore quotes in the entrant market and thereby generate trade-throughs”. They relate an increase of smart brokers to an increase of liquidity as their presence increases fill probabilities for liquidity providers in the market. The observed correlation between these brokers and liquidity has several interesting implications. If exchanges protect liquidity providers against trade-throughs, liquidity supply increases and at a higher liquidity supply the benefit for using such automated routing increases. The opposite is also true, indicating a reinforcing or self-sustaining system.

However, Hendershott *et al.* (2011, p.2) argue that a positive correlation between algorithmic trading and liquidity should not always be taken as given. Liquidity providers increase trading options to others, but if smart routers result in liquidity takers being better equipped in finding the in-the-money trade, then the costs of providing such trade options should increase and therefore spreads should widen to compensate. They state that “AT could lead to an unproductive arms race, where liquidity suppliers and liquidity demanders both ... try to take advantage of the other side, with measure liquidity the unintended victim”. Additionally, competition between algorithms may be less vigorous than the competition between humans as a result from high development costs of algorithms.

2.3 Tick Sizes

According to Angel (1997) larger tick sizes encourage market making in the security as the tick size represents the minimum profit of buying the bid and selling the ask. Optimal tick sizes for a given firm depend on a trade-off between incentives that large tick sizes provide to liquidity providers and limit order trades – market makers collect a higher bid-ask spread at higher ticks sizes and it becomes harder for others to improve the best limit order as the next step size is larger – and the higher transaction costs that a larger tick size imposes on investors. Optimal tick sizes are nonzero according to Angel, as this simplifies traders information sets, reduces bargaining time and the potential for errors. Furthermore, nonzero tick sizes incentivise liquidity, introduces price and time priority and puts a floor on bid-ask spreads which results in market makers providing additional liquidity. Additionally, placing an order reveals information to the market, on larger tick sizes you have more protection when placing limit orders as it is harder for other market participants to gain queue priority by setting a better price. Tick sizes are optimal in only a certain price range, and as we see different tick size regimes across countries optimal price ranges will differ across countries as well. Although the regimes differ, resulting relative tick sizes are similar.

O’Hara *et al.* (2013) acknowledge that tick sizes can affect multiple aspects of liquidity as well as the interaction of different types of traders in the market. These aspects of liquidity include spreads, depths, depth resiliency, cancellation, and execution rates of limit orders. With smaller tick sizes the ability to set your price in front of another is enhanced, which results in reduced incentives to post limit orders (buy or sell orders at a specified price or better). For stocks with a low trade volume this

might result in illiquidity if fewer orders are posted. Furthermore, reduced tick sizes decreases profits for market makers which then reduces analyst coverage on less active stocks.

Foucault & Menkveld (2008) state that trading does not drift entirely to the lowest costs market. When queues at the lowest costs get large, execution probability declines. One might then route the order to a more expensive market so that one bypasses time priority at the current market and therefore increases execution probability. Additionally, liquidity takers are expected to split their orders across exchanges when similar orders are available at each of these exchanges, therefore incentivising liquidity providers to send orders to all exchanges as long as the costs are not too high.

3 Data Selection & Sample Analysis

In this chapter we discuss the applied methods to measure the degree of delusive liquidity at the top level of the order book. We discuss the methodology by showing the application of the methods at an initial sample set. We choose to show the methods directly alongside explaining it, as we believe understanding the nature of the dataset is essential for the logic behind both the assumptions made and the developed procedures.XXX. Up until now this has been an iterative process; on our growing understanding of the microstructure our understanding of the tools required to analyse this microstructure grew, we aim to show parts of the process as well.

We start by defining variables and explaining notation of the symbols, after which we continue by setting the time intervals that determine the decay effect of an individual trade. XXX

We picked these two exchanges as they constitute of most of the trading in the ING stock, with the public trading market share of XAMS being around 70% and the market share of CHI-X around 15%.

3.1 Three Measures of Delusive Liquidity

We will introduce the three suggested measures for delusive liquidity in this section, after which we will demonstrate the development and application of these measures in the remaining parts of this chapter. We elaborate on the terminology used in this section later in the chapter. XXX This three step procedure should be applied for each exchange for each underlying. As we analyse two exchanges we thus apply the procedure twice for each underlying, once to estimate the effects of trading at Exchange A on order book updates at Exchange B and one to estimate the reverse effect; trading at Exchange B on order book updates at Exchange A. These effects should not necessarily be equal, differences in exchange characteristics (technology, order types) attract different categories of traders and market shares of exchanges might impact its effects on the other exchanges.

3.1.1 Intra Exchange Orders

XXX

3.1.2 Inter Exchange Top Down

XXX

3.1.3 Inter Exchange Bottom Up

XXX

3.2 Defining Variables at Delusive Liquidity

We now continue by defining the variables and time windows used for each of the three methods.

We are interested in the change (at the side of the order book that is not related to a preceding trade at the same exchange. If a trade executes the same exchange we expect the volume quoted at the accompanying side of the order book to decrease with an amount equal to the trade size, and thus $dAsk$ or $dBid$ should be 0. A buy is traded on the ask side, and a sell is traded on the bid side. The definitions for changes at one side of the order book follow from the structure of the dataset which is similar a hypothetical replay of five market ticks in Table 3.2.1.

Table 3.2.1: Market Replay

Tick	Time	Ask Size	Ask Price	Bid Price	Bid Size	Flag	Trade Price	Trade Size	Feed
1	2014-08-09 09:00:00.000	14500	10.19	10.185	5700				XAMS- ING

2	2014-08-09 09:00:00.000	14750	10.19	10.185	5700				XAMS- ING
3	2014-08-09 09:00:00.001	13000	10.19	10.185	5700				XAMS- ING
4	2014-08-09 09:00:00.001					1	10.19	300	XAMS- ING
5	2014-08-09 09:00:00.002	12700	10.19	10.185	5700				XAMS- ING

When we observe two succeeding order book updates the change is equal to the difference between both. When we observe trade ticks between two order book updates the change on that side of the order book is equal to the difference between the current and the previous order book updates plus the sum of the trade volume in between both order book updates. At perfect data this should always be equal to zero, but at some occasions there may be a combination of changes at the order book due to an insertion or withdrawal of a limit order and a volume change resulting from the actual trade.

$$dASK_{i,e}^t = ASK_{i,e}^t - ASK_{i,e}^{t^-} + \sum_{t^-}^t SizeBuy_{i,e}^{t^-} \quad (3.1)$$

$$dBID_{i,e}^t = BID_{i,e}^t - BID_{i,e}^{t^-} + \sum_{t^-}^t SizeSell_{i,e}^{t^-} \quad (3.2)$$

Where t refers to the timestamp the current tick, t^- to the previous tick, i specifies the current tick index, i^- the index of the previous tick and e indicates the symbol of the particular exchange. Both i and e will be recorded for information purposes. Based on the market replay in Table 3.2.1 we can show the calculation of the changes;

- i. At tick two 250 volume is added to the end of the queue at the best ask price of 10.19. $dAsk = 250$. Tick two is an order book update tick.
- ii. At tick three 1750 volume is removed from the queue at the best ask price of 10.90. $dAsk = -1750$. Tick three is an order book update tick.
- iii. Tick four signals 300 volume was bought at 10.19. Tick four is a trade tick.
- iv. Tick five lists the new top level after the trade. The ask volume decreased by 300 which is equal to the trade size in tick four, thus $dAsk = 0$. Tick five is an order book update tick.

The delusive liquidity hypothesis states that market makers send the same order to multiple exchanges to maximise fill probabilities, and then retreat parts of the order at the other exchanges if one of the orders gets (partial) execution. When observing the changes at each side of the order book we can test whether the quoted volume decreases as a result of trades at another exchange. This implies we eventually only select the order book updates with a decrease in quoted volume.

XXX

3.3 Defining Time Interval

3.3.1 Finding Minimum Time of Interval

XXX

3.3.2 Finding Maximum Delay of Interval

XXX

3.3.3 Calibrating Time Data

Easley *et al.* (2012) discuss the concept of order flow toxicity; with trade information being channelled from multiple sources the risk of discrepancies in the actual order arrival process increases. With a larger emphasis on smaller time increments and different distances – and therefore different latencies until signal arrival – the time data may not be in the correct order especially when combining data from different exchanges. If the information flow from XAMS exhibits a different delay than the information flow from CHI-X, then our analysis might be completely wrong. Such differences could result in order book updates which in reality occur after trades at the other exchange now show up before and vice versa.

When analysing within the same exchange the probability of order flow toxicity is low as the data include exchange time stamps rather than arrival time stamps. This indicates that the intra exchange information flow will be in the correct order unless the exchange itself sends incorrect data. In the sample set we did not find any indications of wrong intra-exchange data.

3.4 Intra Exchange

MiFID states that each order should be sent with the intention to trade; an order sent without this intention would thus breach European market's regulation. XXX

We see four possible causes of a removal of the full order;

- i. External signal; a delusive order at another exchange is filled for the full size, and therefore 100% of that order is removed from the current exchange.
- ii. Information; updated information results in a response of removal of the order, not specifically related to a trade in the same security.
- iii. Algorithmic contradiction; some errors / conflicting rules within algorithms might result in the algorithm sending the order and then quickly cancelling the order again.
- iv. Quote stuffing; orders send with the intention to immediately cancel

Based on the time difference between two legs in an order pair we can exclude possible causes of the pair. Clearly, the causes can be ordered on expected time difference; a change relating to an external signal should be the slowest as the signal needs to travel further. Subsequent to the external signal an information update should cause a higher time difference than a pair originating from a quote stuffing strategy, simply because it has to respond to the information while at quote stuffing often the removal signal is sent at a predefined time difference after the first leg regardless of any signals.

Hendershott & Riordan (2009) show that in 2007 already a large number of limit orders are cancelled within two seconds.

XXX

CONFIDENTIAL

Figure 3.4.1:

XXX

CONFIDENTIAL

Figure 3.4.2:

CONFIDENTIAL

CONFIDENTIAL

CONFIDENTIAL

3.4.1 CHI-X Time Offset

XXX

3.4.2 XAMS Time Offset

XXX

3.5 Inter Exchange – Top Down Estimation

XXX

Example 3.5.1 – Say a simple series of buy trades at the first exchange and a series of changes at the ASK side of the order book at the second exchange occur as shown in Table 3.5.1.

Table 3.5.1: Series of buy trades

Tick Index	Time (ms)	Trade Exchange 1	dASK Exchange 2
1	1	500	
2	1.51	100	
3	3		-600
4	3.01		-800
5	6	200	
6	6.51		-500
7	15		-100

Now assume we set the interval minimum 1.5ms and the interval maximum 5ms. This indicates that both the -600 change and the -800 change (indices 3 and 4) might result from the 500 and 100 trades as these trades are within the [1.5ms-5ms] interval before the order book change. Additionally, the -500 change might result from both the 100 trade and the 200 trade (indices 2 and 5) as they fall within the defined interval. This demonstrates that the 500 trade would be included in two regression

equations (the equations for indices 3 and 4), the 100 trade would be included in three regression equations (the equations for indices 3,4, and 6), while the only 'perfect' chronology can be found for the 200 trade which would just be included once for the -100 change. Indices 3 and 4 furthermore show the difficulty of assigning order book updates to trades, as strong assumptions should be made to determine whether the 500 and 100 trade belong to index 3 or index 4 or both (partly).

XXX

XXX

Equal Trade Weight Assumption – If a trade possibly affects multiple order book updates we assume influence on each of the order book updates is equal, regardless of distance in time. We assign a weight of $1/x$ to each trade tick, where x is the number of order book decreases that fall within the trade tick's time interval of influence.

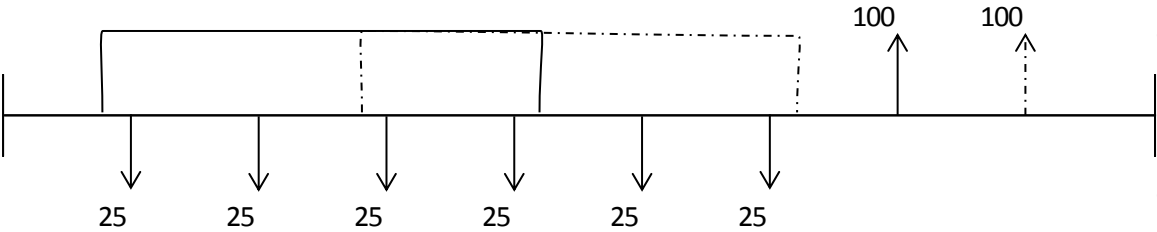


Figure 3.5.1 : Trade Weights Timeline.

XXX

XXX

XXX

XXX

XXX

3.6 Inter Exchange – Bottom Up Identification

XXX

Even though the queue remains unknown we differentiate several characteristics of each update which allow us to track an order through the order queue.

- 1) An order arrives at the end of the queue
- 2) At order arrival we know the volume in front of the order
- 3) The order can only be (partly) filled if it is on top of the queue, which will be the case if the existing volume at point of entry is depleted either through trades or through participants removing their orders.

The actual composition of the queue remains unknown and therefore we cannot be certain whether a decrease in top level volume unrelated to a trade relates to volume in front of or behind the tracked

order in queue. This leads to another assumptions we decide to introduce, the equal size – equal order assumption.

Equal Size Equal Order Assumption – If the size of an order removal is equal to the size of a placed order later in time than the to be tracked order but before the order removal, we assume the placed order is the same order as the removed order, and therefore we know the order removal should be behind the tracked order in queue.

For example, if we observe an increase in the top level order volume of x and shortly after we see an equal decrease in the top level order volume unrelated to a trade we assume this is the same order being removed again.

The validity of this assumption partly depends on the queuing priority rules of the exchanges. At some exchanges any change in order volume results in forfeiture of time priority and therefore results in little incentive to remove anything else but the full order size. At XAMS and CHI-X however, any increase in order size will result in forfeiture of time priority while a decrease in the order volume will not have an effect on the remaining order size. This implies there could be some incentive in removing parts of the order volume, although there would be few reasons from an investment perspective to decide to decrease size without changing price levels. Additionally, this assumption will result in conservative estimates – we now overestimate the volume in front of us rather than underestimate – which should increase reliability of the results.

Finally, if we observe a price change at the order level we stop tracking the order. We just observe top level data and since a price change indicates either the previous top level is now the second depth level or the top level is empty, we lose track of the changes in the queue.

XXXX

4 Analysis

XXX

4.1 Inter Exchange

XXX

4.1.1 ING Group

XXX

CONFIDENTIAL

Figure 4.1.1.:

XXX

4.1.2 Aegon

XXX

CONFIDENTIAL

XXX

4.2 Top Down Estimation of Delusive Orders

XXX

XXX

4.3 Tracking Orders

XXX

5 Conclusion & Discussion

5.1 Conclusion

We started this report with the observation that as a result of fragmentation high frequency market makers might send the same quotes in the same product to multiple exchanges in order to increase fill probabilities and therefore profits. Following from this spreading of quotes we expected that this would lead to delusive liquidity in order books; when the quotes get a fill at one exchange the market makers would remove their order from the other exchanges as they already traded their preferred size.

XXX

XX

XX

XXX

Finally, while we confirmed the existence of delusive liquidity we do not observe large numbers of delusive liquidity for the two products at the two exchanges. The infrastructure of Europe's markets might explain the low visible delusive liquidity partly. While new venues arose after the introduction of MiFID the trading landscape in Europe does not appear to have changed considerably. The traditional exchange for each country often still has at least 60% of the market share in that particular country. For instance the market share of Euronext Paris in France is 63.93%, the market share of London Stock Exchange in England is 62.25% and the market share of Deutsche Börse in Germany is 61.26%. On the contrary, the New York Stock Exchange has a market share of only 32.45%. We therefore argue markets in Europe are still clearly less fragmented than in the USA, resulting in a lower incentive for spreading orders as the probability of the order arriving at the top exchanges in Europe is higher than for the top exchanges in USA. Future research might test this potential explanation by application a similar methodology for estimating the degree of delusive liquidity at USA listed products.

5.2 Future research

XXX

Future research could also expand the analysis to multiple exchanges rather than just twoXXX This could result in stronger conclusions, but each added exchange will increase the possible order combinations send to the exchanges and thus complicate the analysis quickly.

Furthermore, even though we expect 'the speed game' is of higher importance at stocks than at options since less factors affect price discovery in stocks resulting in more participants aiming for the same prices, it would be interesting to expand the methodology to analyse options series. However, option traders in general adjust pricing parameters for a range of options strikes rather than for each option strike individually. This could be one of multiple reasons that results in the trader quoting the next strike rather than then the current one, and therefore we should adapt the methodology so that it succeeds in combining a range of options strikes considering them to be possibly quoted by the same.

Finally, one of the major determinants of the market microstructure is the regulatory minimum tick size for each stock. We discussed this concept in Chapter 2, with changing tick sizes incentives for market makers to quote the stock change as well. Financial products with different tick sizes may

therefore attract different investor bases, something we may have already seen signs of when comparing AGN to ING. O'Hara *et al.* (2013) offer a framework for finding comparable underlyings within a certain relative tick size. By applying the Fama-Fench industry models they show how to subdivide underlyings based on industry and market capitalization. This allows for partly accounting for differences in investors' base or appetite in the sampled stocks.

5.3 Relevance Literature

We contribute to the existing literature in several ways. At first we show that analysis of the market microstructure should be done at much higher frequencies than used currently. XXX.

Secondly we introduce ideas for new methodology when analysing modern microstructure datasets, we observe signs of new strategies that arose since the increase in fragmentation and improvements in technology which before were mainly hypothetical strategies. While the limitations of the proposed measures are clearly visible, hopefully the suggestions will be used as building blocks for future research.

Thirdly we demonstrate early results which highlight that even though we can trade the same underlying at different exchanges, difference in for example technology of these exchanges effects trading in the underlying itself. Finally, as far as we know of, the results found by applying the bottom up measure of tracking the order through the queue is the first scientific description for existence of delusive liquidity.

5.4 Relevance Trading

XXX

References

- Angel, J.J., 1997. Tick Size, Share Prices, and Stock Splits. *The Journal of Finance*, 52, 655-681
- Black, F., Scholes, M., 1973. The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy*, 637-654
- Easley, D., de Prado, M.M.L., O'Hara, M., 2012. Flow Toxicity and Liquidity in a High-Frequency World. *Review of Financial Studies*
- Egginton, J., Van Ness, B.F., Van Ness, R.A., 2012. Quote stuffing. *Manuscript, University of Mississippi*, 25
- Foucault, T., Kadan, O., Kandel, E., 2005. Limit Order Book as a Market for Liquidity. *Review of Financial Studies*, 18, 1171-1217
- Foucault, T., Menkveld, A.J., 2008. Competition for Order Flow and Smart Order Routing Systems. *The Journal of Finance*, 63, 119-158
- Garland, N., Vlasevich, M., Cheng, P., 2014. Traders' Guide to Global Equity Markets. ConvergEX Group
- Gomber, P., Gsell, M., 2006. Catching Up With Technology-The Impact of Regulatory Changes on ECNs/MTFs and the Trading Venue Landscape in Europe. *Competition & Reg. Network Indus.*, 7, 535-540
- Haug, E., 2003a. Know Your Weapon, Part 1. *Wilmott Magazine*, May, 49-57
- Haug, E., 2003b. Know Your Weapon, Part 2. *Wilmott Magazine*, August, 52-56
- Hendershott, T., Jones, C.M., Menkveld, A.J., 2011. Does Algorithmic Trading Improve Liquidity? *The Journal of Finance*, 66, 1-33
- Hendershott, T., Riordan, R., 2009. Algorithmic Trading and Information. *Manuscript, University of California, Berkeley*
- Hull, J., 2009. *Options, Futures and Other Derivatives* (8 ed.). New Jersey: Pearson education.
- Kat, H.M., 2001. *Structured Equity Derivatives*. England: John Wiley & Sons
- Lehalle, C.-A., Laruelle, S., 2013. *Market Microstructure in Practice*. World Scientific Publishing Co.
- Menkveld, A.J., 2013. High Frequency Trading and The New Market Makers. *Journal of Financial Markets*, 16, 712-740
- Menkveld, A.J., Yueshen, B.Z., 2013. Anatomy of The Flash Crash. *Available at SSRN 2243520*
- O'Hara, M., 2001. Overview: Market Structure Issues in Market Liquidity. *Bank for International Settlements Information, Switzerland*
- O'Hara, M., Saar, G., Zhong, Z., 2013. Relative Tick Size and the Trading Environment. *Working paper, Cornell University*
- O'Hara, M., 2014. High Frequency Market Microstructure. *Working paper, University of Warwick*
- Van Kervel, V., 2012. Liquidity: What You See is What You Get? *Available at SSRN 2021988*

Appendix

A Glossary

Bid-Ask Spread – Difference between the best price at the bid and the best price at the ask

Crossing the spread – Buying at the ask and selling at the bid, thereby paying the bid-ask spread

Fill-and-Kill order – A limit order that will be removed immediately if it does not get a immediate fill on arrival.

Limit Order – An order which supplies liquidity for future traders . With a limit order, a trader can improve the execution price relative to the market order price, but the execution is neither immediate, nor certain.

Liquidity Provider – A market participant who sends limit orders

Liquidity Taker – A market participant who sends market orders

Market Order – An order that guarantees an immediate execution at the best price available upon the order arrival

Quote – A quote is a combination of a bid offer and an ask offer showing the price at which a product can be bought or sold.

Trade-through – A violation of price priority in fragmented markets; an order that executes at a price worse than the best quote price.

B Regression Plots

XXX