



Measuring the price impact from large trades and order flow: an empirical study in market microstructure on Oslo Stock Exchange.

---

Master Thesis , autumn 2011

**Authors: Skjalg Yngve Thomassen and Thomas Rasmussen**

**Supervisor: Jonas Andersson**

NHH - Norwegian School of Economics

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Neither the institution, the advisor, nor the sensors are - through the approval of this thesis - responsible for neither the theories and methods used, nor results and conclusions drawn in this work.

## **Abstract**

Theories of market microstructure suggest that large transactions can reveal information and hence impact prices. Extensive research finds support for such a price impact. However, we are not aware of any similar studies at Oslo Stock Exchange (OSE). Other studies have typically been conducted at hybrid markets, e.g. New York Stock Exchange, where there are specialists that facilitate trading. OSE, on the other hand, is a fully electronic limit order market, thus the price dynamics may be different. The implication of a price impact for a trader who plans to submit multiple orders in a stock is that the first trades affect the price of the later trades.

We analyze the temporary and permanent impact on security prices from large buy-initiated and sell-initiated transactions. We find that large trades are associated with significant price impacts 5 seconds and 10 minutes after the transactions for most of the stocks in the sample. There are significant intraday differences in the estimated price impacts. Furthermore, we study the aggregated difference between buy-initiated and sell-initiated turnover, i.e. order flow. We analyze a model where returns over 15 minute intervals are explained by the past order flows. Normalized order flow has a significant effect on returns, but it explains little of the variance.

## **Acknowledgements**

First and foremost we would like to thank our supervisor Professor Jonas Andersson for valuable guidance and for introducing us to market microstructure. Secondly we would like to thank Ingvild Resaland at Oslo Stock Exchange for providing the dataset which made this thesis possible. We would also like to express our gratitude towards Marte Enge and Eirik Egeland Nerheim for reading through the thesis and giving insightful comments.

# Table of Contents

Chapter 1: Introduction .....	1
1.1. Market microstructure .....	1
1.2. Oslo Stock Exchange.....	2
Chapter 2: Literature .....	4
2.1. Theory.....	4
2.1.1. Prices.....	4
2.1.2. Informed and uninformed traders .....	4
2.2. Empirical studies .....	8
Chapter 3: Framework and hypotheses .....	9
3.1. Modeling the price impact.....	10
3.1.1. Temporary and permanent impact.....	10
3.1.2. Trade size .....	10
3.1.3. Buy-initiated and sell-initiated transactions.....	11
3.1.4. Expected return .....	11
3.2. Research questions and hypotheses.....	12
3.2.1. Research question 1.....	12
3.2.2. Research question 2.....	14
Chapter 4: Data.....	16
4.1. Data sample .....	16
4.2. Data processing.....	17
Chapter 5: Empirical analysis .....	22
5.1. Research question 1 .....	22
5.1.1. Time dimension.....	22
5.1.2. Estimated models .....	23
5.1.3. Summary results.....	24
5.1.4. Temporary impact.....	25

5.1.5. Permanent impact.....	26
5.1.6. Economic significance .....	27
5.2. Research question 2.....	29
5.3. Robustness.....	30
Chapter 6: Conclusions and further research .....	31
References .....	33

## Appendices

Appendix A: Estimated coefficients research question 1.....	37
Appendix B: Likelihood ratio test.....	42
Appendix C: Output from the model fitting (Acergy) .....	43
Appendix D: Panel data analysis.....	45
Appendix E: Data processing in Perl. ....	46

## List of tables

Table 1 - Models and hypotheses for research question 1 .....	14
Table 2 - Model and hypotheses for research question 2.....	16
Table 3 - Descriptive statistics for the 112 stocks in the dataset.....	18
Table 4 - Decomposition of when trades occur.....	18
Table 5 - Transactions in our sample .....	20
Table 6 - Sample after removing small trades and trades between bid and ask quotes .....	21
Table 7 - Result summary .....	24
Table 8 - Results from the panel data regression model .....	29
Table 9 - Hypotheses for research question 2 .....	30

## List of figures

Figure 1 - Illustration of an order book before and after an order.....	3
Figure 2 - Intraday distribution of turnover in fifteen minute intervals.....	19
Figure 3 - The average deciles of the trade size in percentages of daily volume.....	21
Figure 4 - Expected temporary impact.....	25
Figure 5 - Expected permanent impact .....	26
Figure 6 - Root mean squared error .....	31

# Chapter 1: Introduction

## 1.1. Market microstructure

A basic assumption in financial economics is that individual buying or selling of a security does not change its price. One solid argument for this assumption is that trading a security cannot change the firm's underlying cash flows. Since investors presumably agree that the share price should reflect the discounted value of all future cash flow, we do not expect that the trading of a share can change the price permanently. E.g. a positive deviation from the true value, which is caused by a large buy transaction, is expected to be offset by others selling.

The theory of microstructure<sup>1</sup> on the other hand, argues that the more informed traders trade larger quantities than uninformed traders. Hence, there may be price impacts from larger trades, because the transactions contain information. For a trader who is paying more today, there is little comfort in an assumption stating that the stock price eventually will return to the true price. E.g. if a trader submit a series of buy orders, the first transactions may increase the prices for the future transactions.

To measure the price impact of transactions and order flows we use a high frequency dataset from Oslo Stock Exchange. These types of datasets grant the researcher with more possibilities than with daily data, but they also contain challenges from a data processing and econometric perspective. Processing the raw data is an extensive task that made it necessary to acquire certain skills in programming. To ease the effort for future scholars of market microstructure, we have enclosed a stylized example in Appendix E that shows how one can extract similar data.

In the first research question we study the price impact of large buy-initiated and sell-initiated transactions. Like other empirical studies we measure the price impact in two time dimensions, temporary and permanent. We define the temporary and permanent price impacts as five seconds and ten minutes returns respectively. For most of the stocks in the sample we find significant price impacts. There are also significant intraday differences in the estimated price impact. In the second research question we study a model of returns in 15 minute intervals explained by the past aggregate buy-initiated and sell-initiated transactions, i.e. the

---

<sup>1</sup> The term for studies on the trading mechanisms in the financial securities markets is *market microstructure* (Hasbrouck, 2007).

order flows. We find that a positive (negative) normalized order flow is associated with a positive (negative) return the next 15 minute interval. We also control for differences in intraday return and reject the null hypothesis that they are equal. However, the R-squared is low compared to similar studies performed with indices.

## 1.2. Oslo Stock Exchange

Oslo Stock Exchange (OSE) is a fully electronic limit order market located in Oslo, Norway. Relative to other exchanges has OSE an overweight of commodity industries, such as energy producing, oil-service, and aquaculture. The market capitalization March 2007 and March 2010 was NOK 1835 (USD 301) and NOK 1324 (USD 223) billions respectively<sup>2</sup>. The continuous trading session last from 09:00 Central European Time (CET) to 17:20<sup>3</sup> CET (henceforth, all hours are in CET). Before the continuous trading session begins there is an opening auction where traders can submit orders. Crossing of these orders are done at a point in time between 09:00 to 09:05 at a price that maximizes the nominal value traded. The opening auction starts at different times for each stock (the most liquid first). Since the start for the continuous trade session can vary both for securities and different days, we conveniently define the continuous trade session to start 09:05 in our analysis. During the continuous trade session orders are crossed automatically according to a strict price-time priority rule. The continuous trade session ends 17:20 and after there is a closing auction equal to the opening auction.

Limit orders have both a price and a quantity limit, usually limit means the price limit (henceforth, limit is price limit, unless otherwise stated). Most fully electronic limit order markets follow a similar strict price-time priority rule as OSE. This means that an unconditional<sup>4</sup> buy (sell) limit order is crossed if the limit is equal or higher (lower) than a previous submitted sell (buy) order (in this chapter we assume orders are unconditional). OSE allows traders to submit orders that “walk the book”, i.e. buy (sell) orders are crossed to the best available prices given the price and quantity limit of the order. Traders can also submit market orders that we can interpret as limit orders with an infinite limit. The order

---

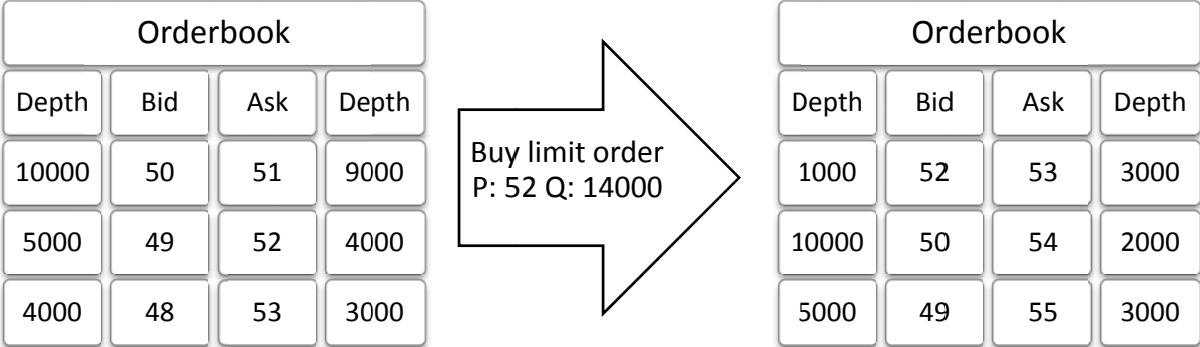
<sup>2</sup> <http://www.world-exchanges.org/statistics/monthly-reports> (accessed 29.11.2011)

<sup>3</sup> OSE extended trading hours by one hour the 1<sup>st</sup> of September 2008 (Oslo Børs ASA, 2008).

<sup>4</sup> Traders can also submit orders that are canceled if there are not enough offers to fill the order immediately (fill or kill). Alternatively the trader can submit orders that are canceled after it is filled with the maximum number of shares given the limit (fill and kill).

book consists of asking (ask) and bidding (bid) quotes, which are offers to sell and buy a given quantity of shares. A hypothetical order book is illustrated below.

**Figure 1 - Illustration of an order book before and after an order**



The quotes are usually referred to as levels, where the first level consists of the best ask price and bid price. The figure only shows the upper three levels. After a trader buy or sell all shares offered at a level, the lower levels are pushed up and thus the best quote change. We consider the buy order illustrated in Figure 1, and see that the quantity limit is larger than the available quantity at the first two levels. The remaining quantity enters the order book as best bid quote. This illustrates that one order can result in several transactions, in this case minimum three if the last part is filled. However, quantity offered at each level may consist of several orders and each crossing of two orders is recorded by the exchange as one transaction.

In this thesis we study the price impacts of large trades and some might suggest that one should consider the price impact from buy and sell orders. This is generally rejected in literature, because of the implication that traders can manipulate the share price by first submitting a limit order and afterwards submitting a cancel order. In this study we consider only the first level of the order book, and do not consider the volume offered. One might argue that a weighed price of the offers would better reflect that the traders of large quantities face different average prices than the traders of small quantities. However, the volume available in the order book is not always the true quantity available since traders can submit

orders with partially hidden volume<sup>5</sup>. Quoted prices, on the other hand, are accurate since the order book always shows a part of the hidden volume submitted by traders.

## **Chapter 2: Literature**

### **2.1. Theory**

#### **2.1.1. Prices**

In economics, prices are determined by the equilibrium between supply and demand. How the prices actually reach the equilibrium is metaphorically referred to as the “invisible hand” or the “black box” of trading. It is not obvious when and how this equilibrium actually occurs, neither does it seem that the general economic literature is concerned with this issue. Theories of market microstructure, on the other hand, provide possible answers by describing and analyzing the trading of assets under explicit rules, i.e. how the specific trading mechanism affects the price (O'Hara, 1995).

According to Hasbrouck (2007) there are no comprehensive and realistic models for limit order markets. The theory of market microstructure has traditionally been developed with regards to the traditional dealer market. Since we study a market without dealers, specialists or market makers one might argue that theories for dealer markets do not apply for limit order markets. However, the basic insights are relevant for the empirical analysis, and hence we review some of these theories.

#### **2.1.2. Informed and uninformed traders**

A common assumption in financial economics is that the security prices reflect all publicly available information. The oft-quoted paradox is that for the share prices to reflect all publicly available information someone has to analyze the securities, and why should anyone bother to do so when the prices already reflect all information? The theory of market microstructure analyzes some of the traders' game theoretical issues. One class of models is information-based that allow for the presence of individual traders with superior information, i.e. asymmetric information. Superior information may be private information that is not publicly

---

<sup>5</sup> Orders with partially hidden volume (iceberg order) have a maximum public volume that is shown in the order book and the rest of the volume is hidden. When the public volume is crossed a new equivalent part of the hidden volume is revealed, with a new time priority.



available, and the ability to interpret information better or faster than the other market participants.

An important question when analyzing the market participants is why uninformed traders willingly would transact with informed traders. One might compare it to a novice poker player entering a game against a world champion. It may be realistic that they would do so just for the mere entertainment, but not with the expectation to win and certainly not to participate in a series of games. A possible explanation was proposed by Bagehot<sup>6</sup>(1971), he argues that uninformed traders confuse trading gains with market gains. By attributing profits to trading skills rather than realizing that the market tends to move upwards, traders may perceive that they have an edge and trade more frequently than they should. According to Bagehot, traders are seduced by the random walk argument and believe that even an idea or hunch will give them a return over time. Furthermore, he assumes that there are another group of traders that in fact have superior information.

The market makers' role is to provide liquidity by transacting with anyone wishing to trade, this includes trading with both uninformed and informed traders. Given that the market makers have no private information about the true value of the firm, they will on average lose against the informed traders. For market makers to survive in the long run, the profits they make from uninformed traders must exceed the losses inflicted by informed traders. Since market makers typically are obliged to provide quotes at all times the only way to balance profits and losses is by setting the difference between bid and ask quotes large enough. The important notion for all models of auctions with market makers is that the presence of asymmetric information will result in a positive bid ask spread, even with a risk neutral market maker that makes zero profits (Glosten and Harris, 1988). Trading is a zero sum game where the informed traders have an expected positive gain and the market makers none, this implies that uninformed traders on average lose on their trading activities. Uninformed traders are also called noise traders in the market microstructure literature (see Black (1986) for a general discussion on noise).

Milgrom and Stokey (1982) show that private information is valueless given that all the participants have rational expectations and the initial allocation is Pareto optimal. The latter condition is true in reality at OSE, because each trading day starts with an opening auction. Since the largest possible nominal value is crossed in the opening auction it is reasonable to

---

<sup>6</sup> Bagehot is a pseudonym, the real Walter Bagehot died in 1877.

assume that the initial distribution of stocks 09:05 is Pareto optimal. Milgrom and Stockey (1982) argue that after an initial auction the only reason to trade would be an advantageous bet, i.e. private information that is not reflected in the share price. However, if everybody knows that other traders only trade given that they have private information, there would be no reason to buy or sell the stock. Because of this it is a general assumption that some investors trade for pure liquidity reasons, e.g. pension funds that needs to convert securities to cash to pay retirees.

Kyle (1985) propose a dynamic model for sequential auction equilibrium. He considers a market with multiple market makers, informed and noise traders (i.e. liquidity traders). Noise traders are assumed to have a distribution independent from the informed traders quantities at all times. In the model Kyle assume that there is one informed trader, who is profit maximizing and risk neutral, i.e. an intertemporal monopolist. This is a strict assumption and it seems reasonable that there can be more than one informed trader. Furthermore, he assumes that market makers earn zero profits on average and have no private information. Hence, price changes are always a consequence of the observed aggregated order flow. One of the key insights from the model is that the informed trader must consider the price impact of transactions on future prices to maximize his monopoly profits, i.e. divide their total demand into smaller trades. The oft-quoted Stealth trader hypothesis suggests that this would make the informed traders concentrate their trading in medium sizes, because of the cost associated with small trades (Barclay and Warner, 1993).

Due to the normality assumption for noise traders, Kyle's sequential auction model converges to a Brownian motion process for prices when the time between the sequential auctions goes towards zero. This model relies on the crucial assumption that the informed traders submit orders in such a way that the information is gradually reflected in the security prices, i.e. if there is more noise trading one period there is also more informed trading. However, if there are more informed traders there would clearly be incentives to trade before other informed traders increase (decrease) the price. It may also be restrictive to assume that noise traders are submitting pure random orders, regardless of order flow and time.

Admati and Pfleiderer (1988) consider a more realistic model, building on the framework proposed by Kyle (1985). The model differs from Kyle's by allowing for multiple informed traders and strategic behavior for noise traders. The rationale for the latter is that in reality we observe volumes that are typically larger early and late in the continuous trading sessions. The

informed traders (minimum one) are assumed to observe part of future public information one period ahead. Furthermore, there are two kinds of liquidity traders, nondiscretionary and discretionary. The latter have flexibility to split their orders over periods restricted to a quantity traded within a given time. Admati and Pfleiderer (1988) show that noise traders will buy relatively more in the period prior to their given time limit. However, the model lacks as the private information only is useful for one period.

The models suggest that market makers will change prices as a function of net order flow and that prices gradually will reflect the information. They both assume that there is no way to distinguish informed trading from uninformed. Another perspective is that the trade size of each transaction can reveal information (Easley and O'Hara, 1987). Easley and O'Hara argue that an informed investor would trade larger quantities at any price, given that he knows that the price do not reflect the true value. Since the uninformed do not have this quantity bias, one should expect relatively larger trades to contain more information than small.

Neither of the models (Kyle, 1985; Admati and Pfleiderer, 1988) considers the choice between limit and market order, i.e. traders could submit a limit order within the spread. This trade-off between waiting for a better price and the cost of trading now with certainty is defined as the implementation shortfall (Perold, 1988). Almgren and Chriss (2000) quantify this relationship in an efficient frontier for optimal execution strategy.

An important question is whether we should expect the same dynamics in a limit order market as predicted by these theories or not. The absence of market makers can conceptually be solved by interpreting limit orders as market maker quotes. Furthermore, it is reasonable to sustain the assumption that some traders are more informed than others and that some traders trade for liquidity reasons. The market maker's inclination to buy (short) stocks on his own account and short the stocks to satisfy take sell (buy) orders, may impact the results of studies of market impact. Hence, studies at e.g. the New York Stock Exchange may differ from our findings because of the organization of the market. Nevertheless, the most critical assumption for our thesis is that larger trades contain an information component. We claim that the informed traders' preference for large quantities still apply without a dealer, hence it is reasonable to test for a price impact from large trades.

There may also be traders that passively submit both bid and ask orders to profit of the spread, thus acting as temporary market makers. Hasbrouck and Schwartz (1988) divided traders into two groups; active and passive traders. Passive traders can avoid execution cost imposed by

the bid-ask spread by waiting for the contra side of the market to take their offers. Thus, they provide liquidity for other traders and may reduce the bid-ask spread. Active traders want immediate transactions, e.g. a trader that wants to buy will rather pay a price that is rounded up than waiting. By this definition active traders are the ones that affect the traded prices and initiate the transactions.

## **2.2. Empirical studies**

There is extensive research on price impacts of large trades. However, we are not aware of any studies performed with data from Oslo Stock Exchange (OSE). Holthausen, Leftwich and Mayers (1987) study the effects of large block transactions on the New York Stock Exchange (NYSE). Their results suggest that buy-initiated block transactions are associated with a permanent increase in the stock price. However, sell-initiated block transactions seem to have temporary effect on prices, but only weak evidence of permanent effects. In later work Holthausen, Leftwich and Mayers (1990) analyze how quickly prices reach a new equilibrium after large block transactions. They find that prices adjust within at most three trades after the block transaction. Another study at NYSE, conducted by Chakavarty (2001), analyzes which trade sizes that move prices. He finds that medium-size trades are associated with the largest cumulative price impact.

Chan and Lakonishok (1995) analyze sequences of trades (packages) that they interpret as one order. The study uses data of orders and trades submitted by investment management firms at the New York and American Stock Exchanges. These orders are in most cases submitted over several days. They find that the weighted average price impact is higher when orders are considered as a package and claim that it is naïve to consider one order or trade isolated. We argue that if one considers multiple orders as a sequence, the results might depend on the investment manager's reactions on the stocks return after the initial order. Hence, we believe it is reasonable to consider individual transactions.

Koski and Michaely (2000) study the information content of different sized transactions during periods with varying degree of asymmetric information. They find that large trades have the largest price impacts during periods when asymmetric information is at its highest. Furthermore, they find that the spread increase and depth decrease significantly after large trades, but not after small trades. Another interesting finding is that the effect of the trade size

is non-linear. Hasbrouck (1991) also find that the relationship for the permanent price impact is concave.

Order flow is a term used in many empirical studies related to return and variance, i.e. the aggregate of buy-initiated and sell-initiated transactions. Relative order flow (Blume, et al., 1989) is a measurement for the imbalance between the value of buy-initiated and sell-initiated transactions. A positive (negative) result from this calculation indicates a net buying (selling) pressure. An alternative measures for this imbalance is the normalized order flow (Lakonishok, Shleifer and Vishny, 1992). Blume, MacKinlay and Terker (1989) finds that the relative order flow has a positive and significant effect on returns in 15 and 30 minute intervals. Other studies have found similar conclusions and that various order flow measurements describe much of the variation in stock returns (e.g. Chordia and Subrahmanyam, 2004; Moberg, 2008; Dunne, Hau and Moore, 2010).

### **Chapter 3: Framework and hypotheses**

The common assumption from the theories we have described is that there are informed and uninformed traders. Traders can observe the transactions (but not other traders' identities) at the exchange and may be influenced by other market participants' trades. Hence, traders' reactions to large transactions may cause price impacts. Basically there are two views traders can have on anonymous transactions. The first view is that only the large trades contain information and can cause a price impact. The rationale for this view is that informed traders have a demand for larger quantities of shares, regardless of the price (Easley and O'Hara, 1987). In Research Question 1 (RQ1) we analyze this price impact from large trades. The second view is that informed or smarter traders split their orders (Barclay and Warner, 1993) and that we cannot distinguish between informed and uninformed trades. Then traders may analyze the aggregated order flows, and interpret an overweight of buy-orders (sell-orders) in a period as an indicator of a future price increase (decrease). This is analogous to the market makers' behavior in models of informed and uninformed trading (Kyle, 1985; Admati and Pfleiderer, 1988). In Research Question 2 (RQ2) we analyze the price impact from the aggregated order flow. The two views are complementary in the sense that traders can both take large trades and the order flow into account.

### 3.1. Modeling the price impact

#### 3.1.1. Temporary and permanent impact

The temporary price impact is a liquidity shock due to the trade which results in a short term disequilibrium. E.g. when a trader buys all stocks offered at the best ask price it may take some time before new orders arrive at this level. If the deviation persists, we define it as a permanent impact, i.e. information related.

Bertsimas and Lo (1998) propose a model where the temporary impact is the difference between the transaction price and the quote midpoint<sup>7</sup> ( $q_t$ ) and the permanent impact is the difference between the present and a future quote midpoint. Our measurement differs from Bertsimas and Lo (1998) since we measure both temporary and permanent impact as the difference in quote midpoints, for increased comparability. Hasbrouck (1991) estimate the return as the change in the quote midpoint<sup>8</sup>. Because we estimate models for multiple stocks a percentage measure is more suitable.

We define the temporary and permanent impact as five seconds and ten minutes percentage returns ( $r_t$ ):

$$r_t^{\text{Temp}} = \log\left(\frac{q_{t+5 \text{ sec}}}{q_t}\right) \quad r_t^{\text{Perm}} = \log\left(\frac{q_{t+10 \text{ min}}}{q_t}\right)$$

By estimating the mid quote returns we avoid autocorrelation caused by the bid-ask bounce. This phenomenon occurs when a trade at the ask (bid) price is followed by a trade at the bid (ask) price and hence the quoted price change regardless of changes in the bid and ask prices. The bid-ask bounce causes an expectation of the return series to be negatively autocorrelated.

#### 3.1.2. Trade size

There is no single definition of what constitutes a small or large transaction. In order to determine whether a trade is a small or large it is necessary with a benchmark that is comparable across securities. One alternative is to use the number of shares traded, e.g.

---

<sup>7</sup>  $q_t = \frac{(P_t^{\text{ASK}} + P_t^{\text{BID}})}{2}$

<sup>8</sup>  $r_t = q_t - q_{t-1} = \Delta q_t$

Chakravarty (2001). This definition lacks comparability because the value per share and total outstanding differs between stocks. One alternative would be to use the value of the transaction as measurement. However, this ignores that some stocks may have different levels of trading activity, which may affect the price impact. We find a more suitable measure to be percentage of the total daily traded volume ( $S_t$ ) as done in Kissell and Malamut (2005).

### **3.1.3. Buy-initiated and sell-initiated transactions**

The size of a transaction is a strict positive variable. For there to be any meaningful interpretation of the price impact we make a distinction between buy-initiated and sell-initiated transactions. Otherwise, one could argue that for every buyer there is a seller. Active traders demand immediate transactions and hence submit market orders. Presumably active traders have an urgent need for buying (selling) the stock that may indicate an expectation of positive (negative) short term return. If the trader have no expectations about the short term return he will know that submitting limit buy (sell) orders with a limit lower (higher) than the current ask (bid) are associated with lower average execution cost. We also realize that there are other reasons to submit market orders, such as an urgent hedging need. Passive traders, on the other hand, submit limit orders lower (higher) than the current ask (bid) quotes and the order may not be crossed. Passive traders may also submit orders on both bid and ask like a market maker, and profit from the spread. From this we have that an transaction at the current ask (bid) quote is buy-initiated (sell-initiated).

### **3.1.4. Expected return**

Incorporating expected return is a crucial part when modeling asset prices and returns. In our case the intervals are diminutive, thus adding a drift term to our model will most likely disturb more than it explains<sup>9</sup>. This measurement error comes before choosing the actual expected return, e.g. CAPM or another factor model. Hasbrouck (2007) analyse the removal of expectation and identify a negative bias, but a significant reduction in estimation error. We consequently believe that omitting the expected return and the dividend rate will give a more parsimonious model for describing the data generating process.

---

<sup>9</sup> Assuming 250 trading days, 8.5 hours trading sessions, evenly distributed returns, and 10 % annual returns give an expected 10 second (10 minute) return of 0.00001 % (0.00075%).

In order to model price impact of large trades we assume a drift term ( $\beta$ ) that is conditional on trade size (Almgren, Thum and Hauptmann, 2005). Since we are only measuring the returns and over such a small time span, the difference between the arithmetic and the more complex geometrical Brownian motion will be trivial (Almgren and Chriss, 2000).

## 3.2. Research questions and hypotheses

### 3.2.1. Research question 1

*Are large stock transactions followed by temporary and permanent price impacts, and are there intraday differences?*

We assume that the temporary price impact of a trade is liquidity related. Traders of large transactions may consume large parts of the available quantity offered, and hence it may take some time for other traders to submit new orders. However, the price is expected to return to its equilibrium after new orders arrive. For measuring the temporary impact ( $r_t^{\text{Temp}}$ ) from large trades we use a linear model.

Other empirical studies find that the impact is greatest for medium sized trades (e.g. Barclay and Warner, 1993; Chakravarty 2001). This could suggest that an information component is declining for some trade sizes. We use the square root of trade size ( $\sqrt{S_t}$ ) to model the permanent impact. The reason for this is that other empirical studies (e.g. Hasbrouck, 1991; Koski and Michaely, 2000) find a concave relationship for the permanent impact, i.e. increasing, but diminishing with trade size.

We measure the permanent impacts ( $r_t^{\text{Perm}}$ ) in ten minute intervals, and it is reasonable to assume that there are more factors influencing returns compared to the temporary impacts. Therefore, we take a more comprehensive approach by including two additional variables in our model. We include normalized order flow ( $v_t^{\text{Normalized}}$ ) to account for the omitted small trades. Lakonishok, et al. (1992) propose normalized order flows as a measurement of the imbalance between buy-initiated and sell-initiated trades. Positive normalized order flow means that there are an overweight of buy-initiated transactions, i.e. net buying pressure. Hence, we would expect that a positive (negative) normalized order flow is related to a positive (negative) return the next 10 minutes. The normalized order flow is calculated for the interval ten minutes before the transaction to one second before the transaction. The reason



for the one second lag is that including the same second mean that normalized order flow also contains the value of the transaction ( $T_t$ ).

We also include the lagged return ( $r_t^{\text{Perm lag}}$ ) in case there are momentum or mean reversal effects in returns.

$$r_t^{\text{Perm lag}} = \log\left(\frac{q_t}{q_{t-10 \text{ min}}}\right) \quad v_t^{\text{Normalized}} = \frac{\sum_{t-1 \text{ sec}}^{t-15 \text{ min}} T_t^{\text{Buy}} - \sum_{t-1 \text{ sec}}^{t-15 \text{ min}} T_t^{\text{Sell}}}{\sum_{t-1 \text{ sec}}^{t-15 \text{ min}} T_t^{\text{Buy}} + \sum_{t-1 \text{ sec}}^{t-15 \text{ min}} T_t^{\text{Sell}}}$$

Theory, e.g. Admati and Pfleiderer (1988), suggests that trading will be concentrated in certain periods of the day. Because informed traders are likely to trade when volume is high (Kyle, 1985; Admati and Pfleiderer, 1988), we will test for differences in the impact during intraday trading. Moberg (2008) find that the volume pattern is U-shaped on OSE, i.e. more volume traded at the start and at the end of a trading day. Other studies (e.g. Andersen and Bollerslev, 1997; Almgren, et al., 2005) find the same U-shaped pattern for the volatility. We identify the same volume characteristic in our sample and the result is shown in the data chapter. There may be several reasons for increased volume parts of the day, e.g. passive funds that trade at the end of the day because they track an index that is measured by closing prices. We test for intraday differences by including slope dummies for each hour of continuous trading for both the buyer and seller initiated trades, resulting in 16 dummies. The base case is trading from 0905 to 0930, then one slope dummy ( $D_j$ ) for each trading hour<sup>10</sup>. Sell-initiated transactions are modeled with a slope dummy ( $D_t^{\text{Sell}}$ ), i.e. the price impact for a buy-initiated (sell-initiated) transaction is  $\beta_1(\beta_1 + \beta_2)$  multiplied with the square root of trade size in percentage of daily traded volume.

---

<sup>10</sup>  $D_1(09:30 - 1030) = 1, D_2(10:30 - 1130) = 1, \dots, D_8(16:30 - 1720) = 1$  | other hours  $D_j = 0$

**Table 1 - Models and hypotheses for research question 1**

**Model 1, temporary impact**

$$r_t^{\text{Temp}} = \beta_1 S_t + \beta_2 S_t D_t^{\text{Sell}} + \sum_{j=1}^8 \delta_j D_j S_t + \sum_{j=1}^8 \phi_j D_j S_t D_t^{\text{Sell}} + \varepsilon_t$$

$$\mathcal{H}_0: \beta_1 \leq 0, \beta_2 \geq 0, [\beta_1 + \beta_2] \geq 0, \delta_1, \delta_2, \dots, \delta_8, \phi_1, \phi_2, \dots, \phi_8 = 0$$

$$\mathcal{H}_1: \beta_1 > 0, \beta_2 < 0, [\beta_1 + \beta_2] < 0, \delta_1, \delta_2, \dots, \delta_8, \phi_1, \phi_2, \dots, \phi_8 \neq 0$$

**Model 2, permanent impact**

$$r_t^{\text{Perm}} = \beta_1 \sqrt{S_t} + \beta_2 \sqrt{S_t} D_t^{\text{Sell}} + \beta_3 v_t^{\text{Normalized}} + \beta_4 r_t^{\text{Perm lag}} + \sum_{j=1}^8 \delta_j D_j \sqrt{S_t} + \sum_{j=1}^8 \phi_j D_j \sqrt{S_t} D_t^{\text{Sell}} + \varepsilon_t$$

$$\mathcal{H}_0: \beta_1 \leq 0, \beta_2 \geq 0, [\beta_1 + \beta_2] \geq 0, \beta_3 \leq 0, \beta_4 = 0, \delta_1, \delta_2, \dots, \delta_8, \phi_1, \phi_2, \dots, \phi_8 = 0$$

$$\mathcal{H}_1: \beta_1 > 0, \beta_2 < 0, [\beta_1 + \beta_2] < 0, \beta_3 > 0, \beta_4 \neq 0, \delta_1, \delta_2, \dots, \delta_8, \phi_1, \phi_2, \dots, \phi_8 \neq 0$$

1. Alternative hypothesis: Large buy-initiated (sell-initiated) trades are followed by a positive (negative) temporary price impact. The temporary price impact from large trades varies intraday.
2. Alternative hypothesis: Large buy-initiated (sell-initiated) trades are followed by a positive (negative) permanent price impact. The lagged return coefficient is different from zero. Normalized order flow coefficient is positive. The permanent price impact from large trades varies intraday.

**3.2.2. Research question 2**

*Will an overweight of buy-initiated (sell-initiated) transactions be followed by positive (negative) returns, and are there intraday differences?*

It may be more reasonable that traders make interference from the aggregated order flow rather than single traders, because of the difficulty associated with interpreting a single trade. Empirical studies finds that order flow explain a large part of return variation in indices (e.g. Blume, MacKinlay and Terker, 1989; Chordia and Subrahmanyam, 2004; Moberg, 2008; Dunne, Hau and Moore, 2010). Hence, we test if order flow is able to explain return in individual stocks.

The model we study has the aggregated sell-initiated ( $T_t^{\text{Sell}}$ ) and buy-initiated ( $T_t^{\text{Buy}}$ ) turnover in a 15 minute interval as the independent variable, i.e. normalized order flow ( $v_t^{\text{Normalized}}$ ). Normalized order flow is always between -1 and 1, where 1 (-1) mean that all trades in a period are buy-initiated (sell-initiated). We measure normalized order flows' effect the effect returns ( $r_t$ ) in the next 15 minute interval.

$$r_t = \log\left(\frac{q_{t+15 \text{ min}}}{q_t}\right) \quad v_t^{\text{Normalized}} = \frac{\sum_t^{t-15 \text{ min}} T_t^{\text{Buy}} - \sum_t^{t-15 \text{ min}} T_t^{\text{Sell}}}{\sum_t^{t-15 \text{ min}} T_t^{\text{Buy}} + \sum_t^{t-15 \text{ min}} T_t^{\text{Sell}}}$$

Given that individual trades contain information, we expect that the difference between buy-initiated transactions and sell-initiated transactions in a period to contain more information than individual trades. Furthermore, we expect traders with private information to split their orders to disguise their private information (Barclay and Warner, 1993). The order flow may capture this effect better than large individual transactions. When normalized order flow is greater than zero the value of buy-initiated transactions are greater than the value of sell-initiated transactions, i.e. a net buying pressure. Hence, we expect the impact on the future price ( $\beta_1$ ) to be positive.

Moberg (2008) measure how imbalance in the order flows explains OBX return. He find that foreign market return and local order flow jointly explains a large part of OBX return variation. His result suggests that return in the continuous trading session are affected by the return during opening auction. To capture this effect we include the overnight return ( $r_t^{\text{Overnight}}$ ) as a separate variable. We calculate overnight return as the difference between the last mid quote of the previous continuous trading session and first mid quote of the current continuous trading session. Overnight return is included as a slope dummy variable in the first interval, i.e. 09:05 to 09:20. We also include seven intercept dummies<sup>11</sup> to model intraday differences in return.

$$r_t^{\text{Lag}} = \log\left(\frac{q_t}{q_{t-15 \text{ min}}}\right)$$

---

<sup>11</sup> $D_1(10:05 - 11:05) = 1, D_2(11:05 - 12:05) = 1, \dots, D_8(16:05 - 17:05) = 1$  | other hours  $D_j = 0$

**Table 2 - Model and hypotheses for research question 2**

$$r_t = \alpha_0 + \beta_1 v_t^{\text{Normalized}} + \beta_2 r_t^{\text{Overnight}} + \beta_3 r_t^{\text{Lag}} + \sum_{j=1}^{j=7} \alpha_j D_j + \varepsilon_t$$

$$\mathcal{H}_0: \beta_1 \leq 0, \beta_2 = 0, \beta_3 = 0, \alpha_0 = 0, \alpha_1 = 0, \dots, \alpha_7 = 0$$

$$\mathcal{H}_1: \beta_1 > 0, \beta_2 \neq 0, \beta_3 \neq 0, \alpha_0 \neq 0, \alpha_1 \neq 0, \dots, \alpha_7 \neq 0$$

Alternative hypothesis – Positive (negative) normalized order flows are followed by positive (negative) returns the next 15 minutes. The overnight return and lagged return coefficients are different from zero.

## Chapter 4: Data

### 4.1. Data sample

The data material is extracted from the OBI<sup>12</sup> Continuous Data Feed (OCDF) and includes all trades and orders at Oslo Stock Exchange in the period the 1<sup>st</sup> of March 2007 to the 30<sup>th</sup> of March 2010<sup>13</sup>. We use the Perl programming language for parsing of the raw data. For each continuous trade session, we first track all changes of the best bid and ask quotes, and make a temporary series with the last mid quote for each second.

To facilitate the first step, the time for each observation is calculated in seconds from midnight, i.e. 09:05 and 17:20 (16:20) are 32700 and 62400 (58800) seconds after midnight respectively. For each 29700 (26100) seconds in the continuous trade session we save the best bid and ask quotes. In the same operation we also save each transaction with the best bid and ask quotes available at the same time, this means the last update of the best bid and ask quote before the transaction. Thereafter we extract the dependent and independent variables for each transaction in the temporary series. We use the open source program R to perform the econometric analysis (R Development Core Team, 2011). The R-packages used can be found in the references.

In the OCDF stocks are identified by ISIN numbers. Some stocks change ISIN number during our sample period. For this reason we select stocks that have the same name from March 2007 to March 2010. This excludes some stocks that may have gone bankrupt, merged with other companies or for some other reason are not listed under the same name. By selecting only

---

<sup>12</sup> Oslo Børs Information AS.

<sup>13</sup> The days; 04.07.2007, 15.03.2009, 08.05.2009, 04.06.2009, 16.06.2009, 21.06.2009 and 24.01.2010 was removed due to incomplete files. We have a total of 766 continuous trade sessions in our sample.

survivor stocks the average return may be higher than a sample including all stocks, this is called a survivorship bias. The price impacts we analyze are assumed to be both positive and negative. Thus, higher return in our sample cannot mean that we find price impacts because of survivorship bias, but the dynamics we find may not be representable for stocks in distress. Hindsight bias is possible in an analysis of price impacts, ergo when one includes information that is not available at the time of the trade. However, in our data processing we have only included explanatory variables that according to the OCDF occurred before the price impact we measure. Because this is a fully electronic market, it is reasonable to assume that the sequential data is correct.

We find 147 stocks that fulfill the name criteria. We extract all trades and changes of the order book for these stocks from the OCDF data. After excluding stocks which are traded on less than 70 % of the 766 days, we have 112 stocks available for analysis.

## **4.2. Data processing**

We analyze high frequency data and therefore is market microstructure noise an important aspect. Microstructure noise is a term that describes all price movements in the trading process, inter alia the bid-ask bounce, information of trades and discreteness (price ticks) (Aït-Sahalia and Yu, 2009). The ratio between the noise and the information added when reducing the intervals is called the noise to signal ratio, this ratio is lower for more liquid stocks (Aït-Sahalia and Yu, 2009). Much microstructure noise present in the series make it difficult to find any relationships about the price impact. Thus, it is important to control for and to reduce the noise in the dataset. We reduce this noise through both data handling and through our variables. Returns are calculated as the difference in the quote midpoint in order to remove the bid ask bounce. This reduces a large part of negative autocorrelation and hence noise in the return series.

**Table 3 - Descriptive statistics for the 112 stocks in the dataset**

<b>Description – daily figures</b>	<b>Ten most traded</b>		<b>Ten least traded</b>		<b>Sample</b>	
	<b>Average</b>	<b>Median</b>	<b>Average</b>	<b>Median</b>	<b>Average</b>	<b>Median</b>
<b>Trade value (in 1000 NOK)</b>	471 830	392 091	290	298	54 666	1 466
<b>Number of trades</b>	2 461	2 175	8	7	367	65
<b>Trade volume (in 1000 shares)</b>	192	180	35	43	149	58
<b>Trade size of total daily volume</b>	0.04 %	0.05 %	12.24 %	14.37 %	0.27 %	1.54 %

The 112 stocks in our sample have considerable different characteristics. Descriptive statistics are presented in Table 3. We see that the data has a positive skewness in terms of trade size and number of trades. For comparability of the results we prefer stocks with relatively homogenous characteristics. Furthermore, modeling stocks with large differences in frequency of trading require different time-series models. Ideally we would utilize as much of the dataset as possible, but since similar models and individuals make the estimation and interpretation more convenient we exclude less traded stocks. The OBX<sup>14</sup> index include the 25 most liquid stocks in the Oslo Stock Exchange Benchmark Index (OSEBX) and is a natural starting point for the selection. We include 19 stocks that both fulfill the name criteria and are included in the OBX index in the beginning of the sample period. Additionally, we include 11 additional stocks which are in the same turnover range. Hence, our analyses are limited to the 30 most traded stocks of the 112 stocks, measured by the total NOK turnover. Due to infrequent trading in certain periods for some stocks, we include the 20 most traded stocks in the panel data analysis.

We remove the opening and closing auction from the sample as Næs (2004) and Moberg (2008). In the opening and closing auctions all trades are executed at the same price, consequently there is no measurable price impact. Table 4 shows how many of the trades that are removed due to the opening and closing auctions.

**Table 4 - Decomposition of when trades occur**

<b>Trades</b>	<b>Opening auction</b>	<b>Closing auction</b>	<b>Continuous auction</b>
29959576	819305	721419	28418852
100 %	3 %	2 %	95 %

<sup>14</sup> Tickers for the 25 stocks in the OBX index the 1<sup>st</sup> of March 2007; ACY, AKER, AKVER, AWO, DNB NOR, DNO, FOE, FRO, MHG, NHY, NSG, OCR, ORK, PGS, PRS, SDRL, STL, STB, SUB, TAA, TAT, TEL, TGS, TOM and YAR.

The data is arranged as one long return series from the continuous trading period for each stock. E.g. the last trade before 16:20 (17:20 from the 1<sup>st</sup> September 2008) is followed by the first trade after 09:05 in the next trading session as in Moberg (2008). The variables for price impacts do not contain overnight returns. The transactions where the quote midpoint cannot be observed the next 5 (10) seconds (minutes) will be treated as missing observations and hence excluded. We also require the mid quotes in the dependent and independent variables to be within the continuous trade session. Thus, the number of observation included in the temporary and permanent impact model differs.

**Figure 2 - Intraday distribution of turnover in fifteen minute intervals**

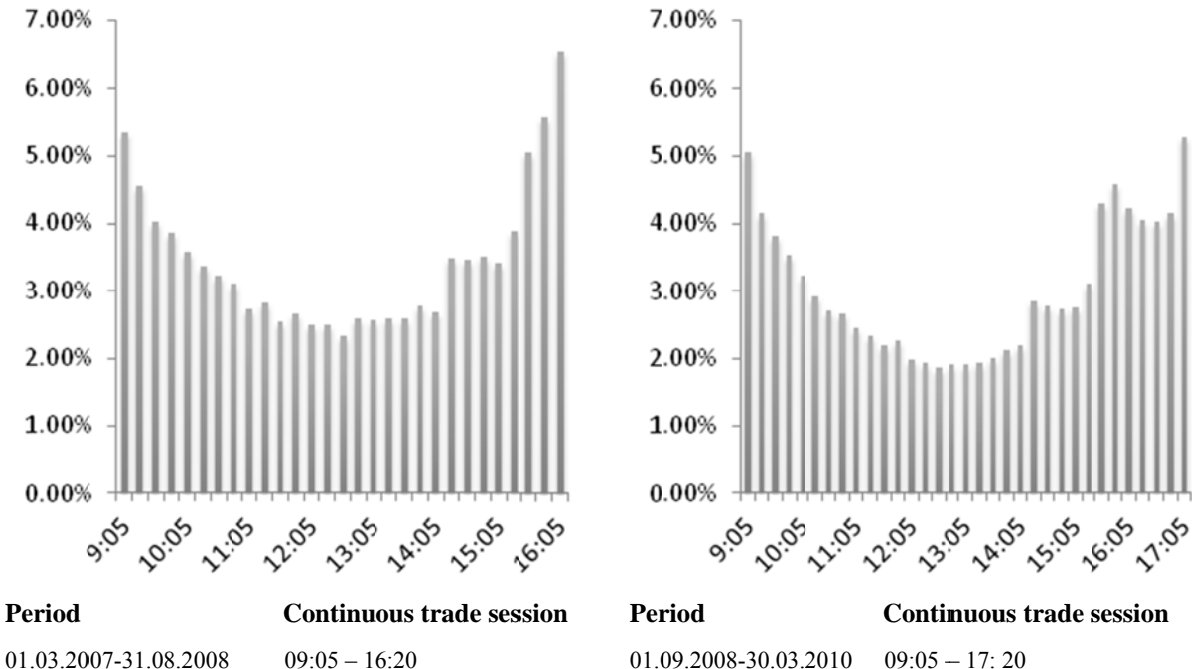


Figure 2 shows the turnover for each 15 minutes intraday calculated for the 30 most traded stocks, measured in percentage points of the total turnover. We see that it has the characteristic U-shape with higher turnover at the beginning and end of the day. Theory of market microstructure predicts that informed traders will trade more when there is higher trading activity (Kyle, 1985; Admati and Pfleiderer, 1988). With the observed volume pattern in our dataset we assume that informed trading is more frequent in the opening and closing hours. The 30 stocks total turnover for the first and the second period is NOK 2987 and NOK 1741 billion respectively.

In all the models we make distinction between buy-initiated and sell-initiated transactions and hence we must classify each transaction. Transactions at the current ask (bid) quote is buy-initiated (sell-initiated). The problem is related to classifying executed at a price between the bid and ask quotes.

**Table 5 - Transactions in our sample**

<b>Bid</b>	<b>Ask</b>	<b>≠ Bid, ≠Ask<sup>15</sup></b>
46 %	47 %	7 %

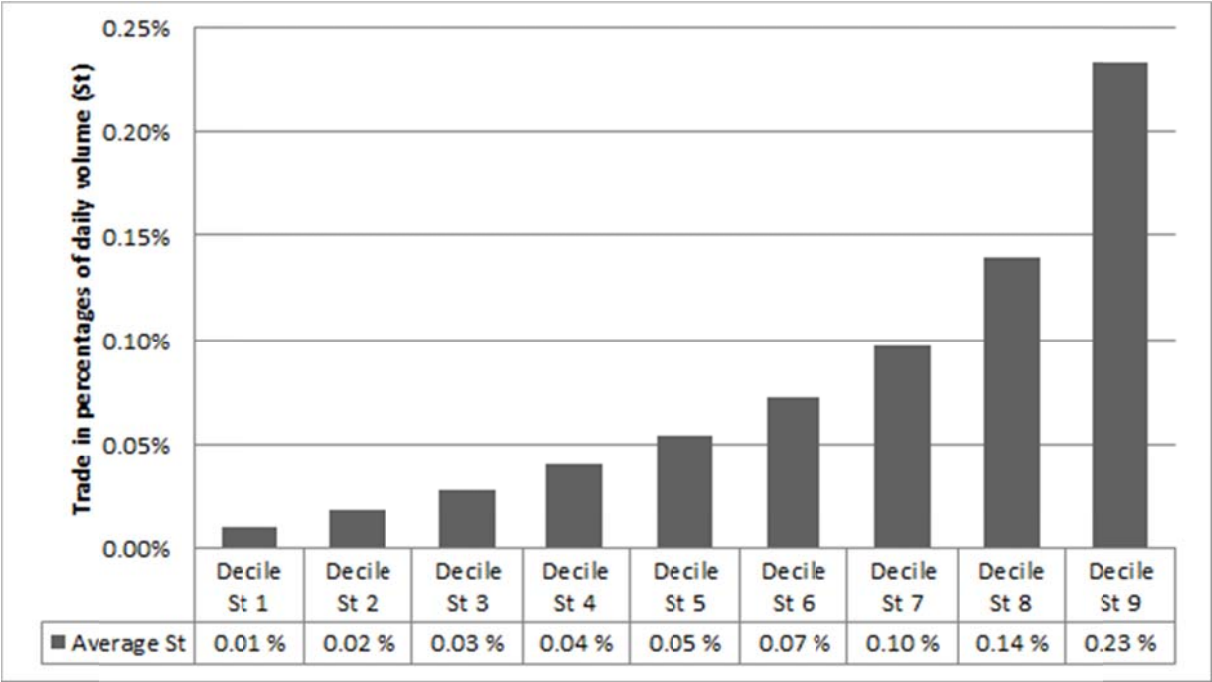
As Table 5 shows, most of the transactions in our sample are traded at the current ask or bid price. Some studies identify the sell-initiated or buy-initiated trades by simply comparing the price of the trade with the last trade, this is known as the “tick test” (Lee and Ready, 1991). This test is typically used by researchers who do not have quote data, only transaction data. The most robust method to determine if a transaction is buy-initiated or sell-initiated, is to track all orders and find which of the orders participating in a transaction that was submitted last. As an example, if the last submitted order is a sell order, the transaction is classified as sell-initiated. Trades between the bid and ask quotes are mainly caused by internal trades within member firms of OSE (Moberg, 2008). Hence, there are no corresponding orders in the order book. Odders-White (2000) studies various classification methods and find that the tick and midpoint method misclassifies 21.4 % and 9.1 % respectively. Because we cannot classify these transactions as buy-initiated or sell-initiated with certainty, we discard these observations.

---

<sup>15</sup> Some trading sessions have different closing hours, e.g. Christmas Eve, resulting in more trades in this classification.



**Figure 3 - The average deciles of the trade size in percentages of daily volume**



We measure the price impact from large trades, disturbances from small trades are considered to be microstructure noise. We therefore exclude small trades in the estimation of the models in RQ1. Removing small trades reduces noise in the data and allow for better analysis of the price impact from large trades. Almgren, et al. (2005) exclude small trades and define trades larger than 0.25% of daily traded volume as large, we use 0.2%. The measure is approximately the average of the 9<sup>th</sup> deciles for the stocks in our sample. The drawback with this measure is that we keep a higher percentage of the trades in the least liquid stocks, resulting in more microstructure noise.

**Table 6 - Sample after removing small trades and trades between bid and ask quotes**

Continuous auction	Undetermined	<0.02%	Sample size
28418852	1984951	25033668	1400233
100 %	7 %	88 %	5 %

Tick size is the minimum increment traders can offer over (under) the current bid (ask). Therefore, the smallest price change possible is one tick and hence the minimum bid ask spread. Discreteness is present in the data due to the tick sizes and affects the future outcomes of the price series. This result in a distribution concentrated around zero with a high kurtosis (Engle and Russel, 2010). Related studies (e.g. Næs, 2004; Moberg, 2008) exclude stocks

with a low nominal share price, because the tick size will cause the percentage price changes to be larger than for other shares. We do not exclude these stocks, but since we select the most traded stocks most of them are removed indirectly.

## **Chapter 5: Empirical analysis**

### **5.1. Research question 1**

#### **5.1.1. Time dimension**

The transactions are by nature irregularly spaced in time because the time between trades can occur at any fraction of a second. Thus, we should consider the time dimension in the models. Wall clock time and event time are two alternative methods for arranging of the data (Hasbrouck, 2007).

Wall clock time means that the researcher arranges the return series in intervals based on the time of registration. An example of the use of wall clock time is 15 and 30 minute intervals. In order to make these intervals, the researcher typically calculate the difference in mid quote from a point in time and the mid quote 15 or 30 minutes later. The observations in this interval are omitted, just as the intraday observations are in daily return series. One advantage of applying wall clock time in a study is that microstructure noise can be reduced through the use of longer intervals. Furthermore, all observations have the same interval length which means it is more convenient to estimate models and to compare results.

Event time, on the other hand, enables the use of all the observations regardless of when they occur. Including all the observations increases the presence of market microstructure noise in the sample compared to wall clock time. However, there is a trade-off between reduction of noise and actually measuring the price impact. E.g. in RQ1 we measure the price impact from large trades, predefined intervals may omit many large trades and hence event time is better to capture all of the large trades. We use event time for registering large trades, but to measure the price impacts ( $r_t^{\text{Temp}}$ ,  $r_t^{\text{Perm}}$ ) from the transaction we use wall clock time. Measuring the price impact a certain number of transactions forward in time decrease comparability, because of the differences in frequency of trading. The use of event time for the transactions is another argument for exclusion of the small trades, because this reduces the microstructure noise.

Since price impact coefficients are the difference in mid quote over an interval and the transactions occur in event time, some of the price impact observations are overlapping. E.g.

we measure the temporary impact as the mid quote return the next five seconds, this give a perfect overlap for trades the same second and an imperfect overlap for the trades the next four seconds. Overlapping observations result in persistent and non-converging autocorrelation in the dependent variable and is more present in the model of the permanent impact. If not controlled for the overlapping problem can give a persistent moving average specification in the models' residuals (Harri and Brorsen, 2009). Alternatively we could measure the temporary and permanent price impact over a longer horizon and reduce the microstructure noise. However, increased length of the price impact also increases the overlapping problem.

### 5.1.2. Estimated models

With shorter intervals in financial time-series there are often autocorrelation present, but due to the overlapping problem the autocorrelation is larger than for a normal return series. In order to model the autocorrelation in our data we apply the Box-Jenkins methodology. We explain the methodology and show some graphical output from our model fitting process in Appendix C.

We estimate the following models for each stock (i):

$$\begin{aligned}
 1) \quad \Delta r_{i,t}^{\text{Temp}} &= \beta_{i,1} \Delta S_{i,t} + \beta_{i,2} \Delta S_{i,t} D_{i,t}^{\text{Sell}} + \sum_{j=1}^8 \delta_{i,j} \Delta D_j S_{i,t} + \sum_{j=1}^8 \phi_{i,j} \Delta D_j S_{i,t} D_{i,t}^{\text{Sell}} + \Delta u_{i,t}^{\text{Temp}} \\
 2) \quad \Delta r_{i,t}^{\text{Perm}} &= \beta_{i,1} \Delta S_{i,t} + \beta_{i,2} \Delta S_{i,t} D_{i,t}^{\text{Sell}} + \beta_{i,3} \Delta v_{i,t}^{\text{Normalized}} + \beta_{i,4} \Delta r_{i,t}^{\text{Perm Lag}} + \sum_{j=1}^8 \delta_{i,j} \Delta D_j S_{i,t} + \\
 &\quad \sum_{j=1}^8 \phi_{i,j} \Delta D_j S_{i,t} D_{i,t}^{\text{Sell}} + \Delta u_{i,t}^{\text{Perm}}
 \end{aligned}$$

$$u_{i,t}^{\text{Temp}} = \varepsilon_{i,t} + \sum_{q=1}^3 b_{i,q} \varepsilon_{i,t-q}$$

$$u_{i,t}^{\text{Perm}} = \varepsilon_{i,t} + \sum_{q=1}^6 b_{i,q} \varepsilon_{i,t-q}$$

We have assumed that the stocks in the sample have similar characteristics after testing on a selected sample. Our conclusion is that MA (3) and MA (6) specifications, integrated at level one, for temporary and permanent impact models respectively. These specifications remove most of the autocorrelation present in the residuals. Although by studying the squared residuals from MA specifications we identify autocorrelation which indicate volatility clustering in the series. Alternatively we could model the conditional heteroscedasticity with a

generalized autoregressive conditional heteroscedasticity model (GARCH) (Bollerslev, 1986). However, we do not believe heteroscedasticity to be a severe problem in our data at large and uneven spaced observations make estimation of conditional volatility problematic.

### 5.1.3. Summary results

Coefficients with their significance level from model 1 and 2 can be found in Appendix A. We estimate one model for each of the stocks and this give a total of 480 estimated dummies. To test for intraday differences in price impact we use a likelihood ratio (LR) test<sup>16</sup>. The results from the LR test are shown in Appendix B (Wooldridge, 2008). We use maximum likelihood to estimate the models and hence we use a LR test instead of an F-test. In Table 7 we present a summary of the hypotheses. We use a standard significance level of 5 %.

**Table 7 - Result summary**

Model	Explanatory variables	Alternative hypotheses	Sum of $H_0$ rejections
1	$S_{i,t}$	$\beta_1 > 0$	27
1	$S_{i,t}, D_{i,t}^{Sell}$	$\beta_2 < 0$	28
1	$S_{i,t}, S_{i,t}D_{i,t}^{Sell}$	$\beta_1 + \beta_2 < 0$	29
1	$D_j S_{i,t}, D_j S_{i,t} D_{i,t}^{Sell}$	$\delta_1, \delta_2, \dots, \delta_8, \phi_1, \phi_2, \dots, \phi_8 \neq 0$	30
2	$\sqrt{S_{i,t}}$	$\beta_1 > 0$	25
2	$\sqrt{S_{i,t}} D_{i,t}^{Sell}$	$\beta_2 < 0$	30
2	$S_{i,t}, S_{i,t} D_{i,t}^{Sell}$	$\beta_1 + \beta_2 < 0$	21
2	$v_{i,t}^{Normalized}$	$\beta_3 > 0$	29
2	$r_{i,t}^{Perm Lag}$	$\beta_4 \neq 0$	30
2	$D_j S_{i,t}, D_j S_{i,t} D_{i,t}^{Sell}$	$\delta_1, \delta_2, \dots, \delta_8, \phi_1, \phi_2, \dots, \phi_8 \neq 0$	27

The results support the alternative hypotheses, for most of the stocks, that large trades are associated with a temporary impact and a permanent impact, i.e. large buy-initiated (sell-initiated) transactions are associated with a positive (negative) return the next 5 seconds and

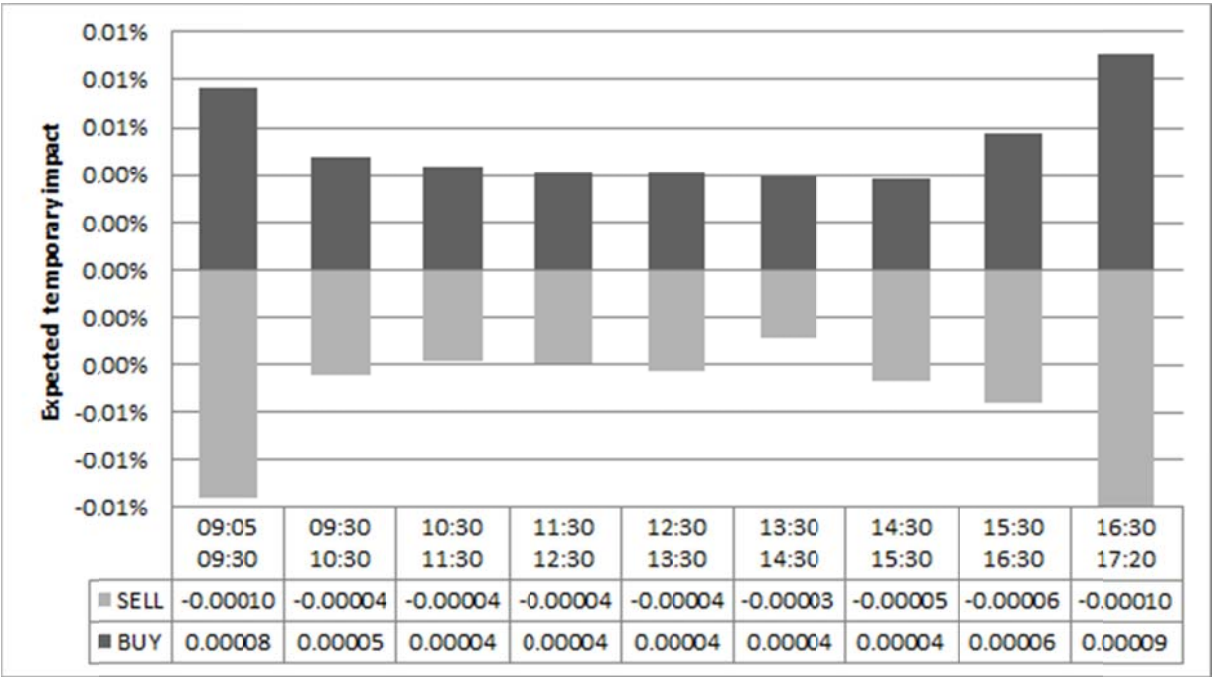
<sup>16</sup> Test statistics:  $LR = 2(L_{UR} - L_R)$  L is the log-likelihood for the unrestricted ( $L_{UR}$ ) and restricted ( $L_R$ ) model. Reject  $H_0$  if  $LR > \chi_{0.05,16}^2$

10 minutes. The results also supports that there are intraday differences for these price impacts.

**5.1.4. Temporary impact**

The estimated temporary impact coefficients indicate that a large buy order affects the price positively the next five seconds, with a coefficient ranging from -0.0085 to 0.0901. The negative coefficient may seem puzzling, but there is only one stock with a negative coefficient. This particular stock (Norwegian Property) has a very sharp decline (-86.6 %) in the share price during the sample period. A large sell-initiated transaction is associated with a negative return ( $\beta_{i,1} + \beta_{i,2}$ ), the sum of the coefficients range from -0.1217 to -0.0046 for the stocks in the sample.

**Figure 4 - Expected temporary impact**



Average of the estimated coefficients.  $S_{i,t} = 0.2 \%$

In Figure 4 we see that the average estimated temporary impact is larger in the first 25 minutes and the last 50 minutes of the continuous trade session. In a part of the sample the continuous auction close 16:20. The intraday differences cannot be a result of generally higher or lower return in these hours, because buy-initiated and sell-initiated price impacts

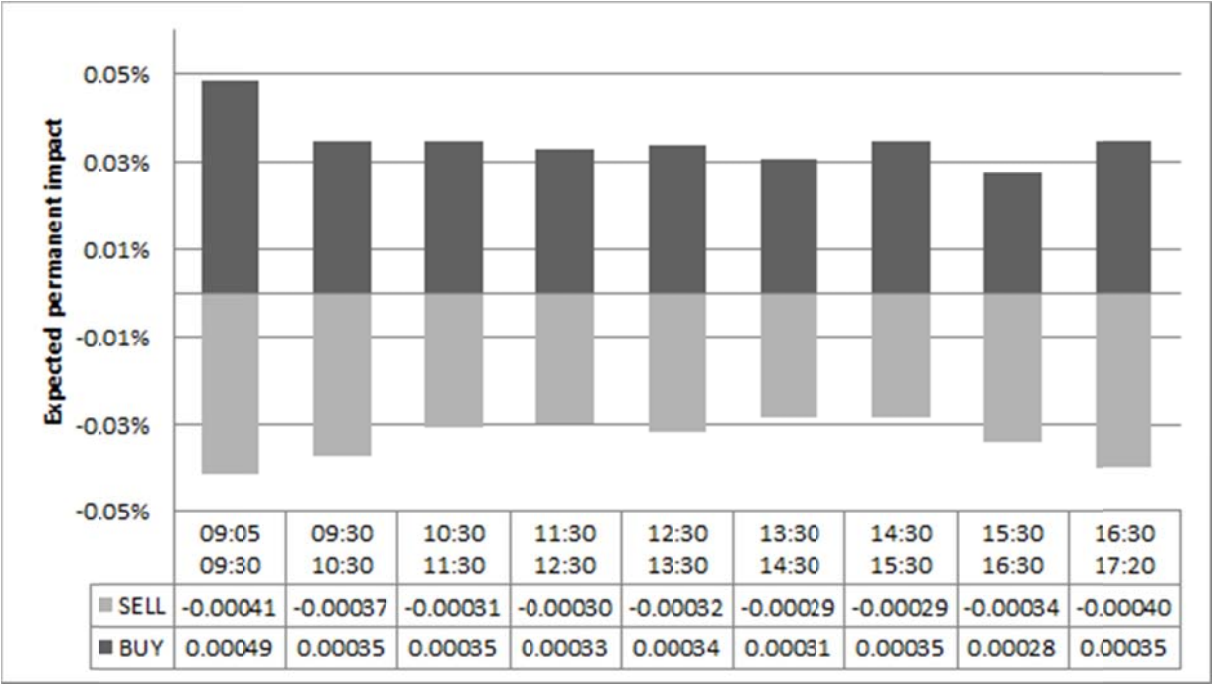
have opposite directions. However, turnover is generally higher in opening and closing hours which might suggest that these differences are related to the volume and volatility.

The results for the temporary impact are what we would expect from a liquidity perspective, since a large trade might temporarily consume a large part of the current available supply or demand. Furthermore, a trade can be a part of a larger order that “walk the book” and consume available quantities at multiple levels. This can mean that each part of the order has a partial impact on return.

**5.1.5. Permanent impact**

For the permanent impact the estimated buy coefficients ( $\beta_{i,1}$ ) ranged from 0.0022 to 0.0231 and the sell coefficients ( $\beta_{i,1} + \beta_{i,2}$ ) ranged from -0.0258 to 0.0053. In general the permanent impact has the same direction as the temporary impact. However, estimated price impacts for sell-initiated transactions are positive for 4 stocks. This clearly contradicts the alternative hypothesis that sell-initiated transactions are followed by a negative return the next 10 minutes.

**Figure 5 - Expected permanent impact**



Average of the estimated coefficients.  $S_{i,t} = 0.2 \%$

In Figure 5 we assume that the lagged return ( $r_{i,t}^{\text{PermLag}}$ ) and the normalized order flow ( $v_{i,t}^{\text{Normalized}}$ ) are zero. The estimated ten minute lagged return coefficient ranged from -0.5247 to -0.0296, i.e. positive return in a ten minute period is associated with negative return the next ten minutes. Figure 5 has an upward bias if large buy-initiated (sell-initiated) transactions are associated with positive (negative) return the previous ten minutes. This bias makes us reluctant to interpret the relatively larger permanent impacts in comparison with the smaller temporary impacts. Nevertheless, the interval from 09:05 to 09:30 differs from the rest of the day as buy-initiated and sell-initiated transactions have a larger estimated price impact than all other hours of the continuous trade session. However, the differences seem smaller than for the temporary impact.

The normalized order flow and the estimated coefficient range from 0.0001 to 0.0004. This mean that an overweight (underweight) in buy transactions is associated with a positive (negative) return the next 10 minutes. These results are significant for 29 of the 30 stocks, and indicate that not only the large trades impact prices.

The permanent impact is consistent with microstructure theories that suggest that there is an information component in large transactions (Easley and O'Hara, 1987). However, there is no way to separate the information components when measuring the impact, thus we do not claim that the estimated impact is solely information related.

#### **5.1.6. Economic significance**

An important question is how to interpret the economic significance of the coefficients. In our case this is dependent on the size of trades and the current shape of order book. To illustrate how the coefficient can be interpreted we provide an example. Recall from Chapter 1.2 how a trader at OSE can submit limit orders that “walk the book”. Assume that we wish to buy a large quantity (TQ) of stocks. We can either buy all the stocks we want with a market order immediately. Assume that we alternatively can buy equal quantities at the first level ( $P_n^{\text{ask}1}$ ) in fixed number (N) of 10 minutes intervals. This might not be optimal, e.g. Bertsimas and Lo (1998) propose a solution to an analogous optimization problem that give different quantities for each period, given a fixed price impact. Multiple orders may also be associated with an additional fixed cost to the broker ( $\alpha$ ) per transaction. Given that we are risk neutral the preferable alternative is that with the lowest expected average share price, i.e. buy now if the following condition is true:

$$\frac{\sum_{i=1}^I P_0^{\text{ask } i} \times Q_0^i}{TQ} < \frac{P_0^{\text{ask } 1} + \sum_{n=1}^N [E(P_n^{\text{ask } 1}) + \alpha]}{N}$$

The left side of the above expression is the average share price if we buy all stocks at the first level or we “walk the book” immediately, i.e. we pay  $P_0^{\text{ask } i}$  at each level (i) where we buy  $Q_i$  stocks. If we can buy all the stocks at the first level now the average share price is equal to the current best ask price. Hence, we should buy all the stocks now because our estimated price impacts are positive and submitting multiple orders are associated with a fixed cost. On the other side, if we have to submit an order that “walk the book”  $i$  levels and hence the average price is higher than the current best ask and the solution to the optimization problem is not straight forward. The current order book is known, but the expected ask quote is uncertain. Given that the spread is fairly constant, the estimate percentage change of the ask price can be approximated by the percentage change in the mid quote. Thus, we estimate the change in the ask quote with model 2 ( $r_t^{\text{Perm}} = \beta_1 \sqrt{S}_t + \beta_2 \sqrt{S}_t D_t^{\text{Sell}} + \beta_3 v_t^{\text{Normalized}} + \beta_4 r_t^{\text{Perm lag}} + \sum_{j=1}^8 \delta_j D_j \sqrt{S}_t + \varepsilon_t$ ). We ignore the time dummies and assume that the normalized order flow and lagged return are zero, hence we can estimate the future ask price conditional on our past trades:

$$E(P_1^{\text{ask } 1}) = P_0^{\text{ask } 1} e^{\beta_1 \sqrt{S}_0}$$

$$E(P_n^{\text{ask } 1}) = E(P_{n-1}^{\text{ask } 1}) e^{\beta_1 \sqrt{S}_{n-1}}$$

Where  $S_n$  is a transaction  $n$  10 minutes periods after the first trade ( $S_0$ ) as a fraction of the estimated total volume during this trade session. As we see from the decision problem above, economic significance depends on the quantity required and the shape of the current order book. Furthermore, future prices are conditional on our current and future trades and hence the price impact should be considered during execution of large orders. It is clear that if the estimated price impacts are correct, the optimization problem has a solution for all quantities and stocks. However, to determine whether this approach to execution gives better outcomes would require out-of-sample empirical analysis. This would be too extensive for this thesis and are left to future research. Nevertheless, lower values for the estimated price impacts ( $\beta_1 \sqrt{S}_{n-1}$ ) associated with our trades, favors splitting the order if there is insufficient volume available at level one.



## 5.2. Research question 2

Panel data has both a cross sectional and a time dimension. There are alternative methods available that can be used if there are unknown individual factors that affect the residuals. We use a pooled regression because there are no significant fixed effects. After we include the lagged return there is no autocorrelation present in the residuals. We use heteroscedasticity robust standard errors to account for differences in variance. Relevant tests that support these approaches are given in Appendix D.

Some of the stocks are not traded in each 15 minute interval of the continuous trading sessions. The reason for this is that some of the most traded stocks do not have a constant volume in the sample period. We observe that some stocks are traded infrequently in the beginning of the period and hence we get an unbalanced panel. Other reasons for missing observations might be trading halts imposed by the exchange because of suspicious trading activity or news announced by a company. These periods are treated as missing observations. However, we have arranged the data in such a way that the lagged return and order flow always corresponds to the return the next 15 minutes, i.e. the first return variable each day is the return from 0920 to 0935, with explanatory variables from 0905 to 0920, i.e. the last explanatory variable for each continuous trading session is order flow from 16:50 to 17:05.

**Table 8 - Results from the panel data regression model**

	<b>Unbalanced Panel: n=20, T=20298-22900 N=451704</b>			
	<b>Estimate</b>	<b>Std. Error*</b>	<b>Coeff.</b>	<b>P-value</b>
$\alpha_0$	-0.0001	0.0000	$\beta_0$	0.0000
$v_t^{\text{Normalized}}$	0.0002	0.0000	$\beta_1$	0.0000
$r_t^{\text{Overnight}}$	-0.0028	0.0018	$\beta_2$	0.1186
$r_t^{\text{Lag}}$	-0.0183	0.0104	$\beta_3$	0.0798
$D_1$ (10:05-11:05)	0.0001	0.0000	$\alpha_1$	0.0242
$D_2$ (11:05-12:05)	0.0001	0.0000	$\alpha_2$	0.0026
$D_3$ (12:05-13:05)	0.0001	0.0000	$\alpha_3$	0.1142
$D_4$ (13:05-14:05)	0.0001	0.0000	$\alpha_4$	0.0534
$D_5$ (14:05-15:05)	0.0001	0.0000	$\alpha_5$	0.0099
$D_6$ (15:05-16:05)	-0.0001	0.0001	$\alpha_6$	0.0780
$D_7$ (16:05-17:05)	0.0000	0.0001	$\alpha_7$	0.6301
<b>R-squared</b>	0.0007	* White heteroscedasticity robust std. errors		
<b>Adj. R-squared</b>	0.0007			
<b>P-value</b>	0.0000			

**Table 9 - Hypotheses for research question 2**

$\mathcal{H}_0$	$\mathcal{H}_1$	Result
$\beta_1 \leq 0$	$\beta_1 > 0$	Reject $\mathcal{H}_0$
$\beta_2 = 0$	$\beta_2 \neq 0$	Fail to reject $\mathcal{H}_0$
$\beta_3 = 0$	$\beta_3 \neq 0$	Fail to reject $\mathcal{H}_0$
$\alpha_0 = 0, \alpha_1 = 0, \dots, \alpha_7 = 0$	$\alpha_0 \neq 0, \alpha_1 \neq 0, \dots, \alpha_7 \neq 0$	Reject $\mathcal{H}_0$

In Table 9 we see that we fail to reject the hypothesis that overnight return and the lagged return coefficient are different from zero. Normalized order flow is significant and support that positive (negative) normalized order flows are associated with a positive (negative) return the next 15 minutes. We test the jointly significance of the time dummies with an F-test of a restricted model without the dummies over an unrestricted model. We reject the null hypothesis which states that there are no intraday differences in returns. The results indicate that returns are higher at the middle of the day. The dependent variable is 15 minutes after the explanatory variable which means that the estimated higher return is between 10:20 and 15:20. However, the coefficients are not all individually significant and we cannot claim that they are positive or negative.

The R-squared is very low for this regression compared to the regressions on the OBX index in Moberg (2008). This seems reasonable to some extent, since an index will have a lower standard deviation, due to correlations between securities. However, the low R-squared compared to studies at indices (e.g. Blume, et al., 1989; Chordia and Subrahmanyam, 2004; Dunne, et al., 2010), may suggest that modeling individual stock returns at such a short horizon is less sensible than modeling an index.

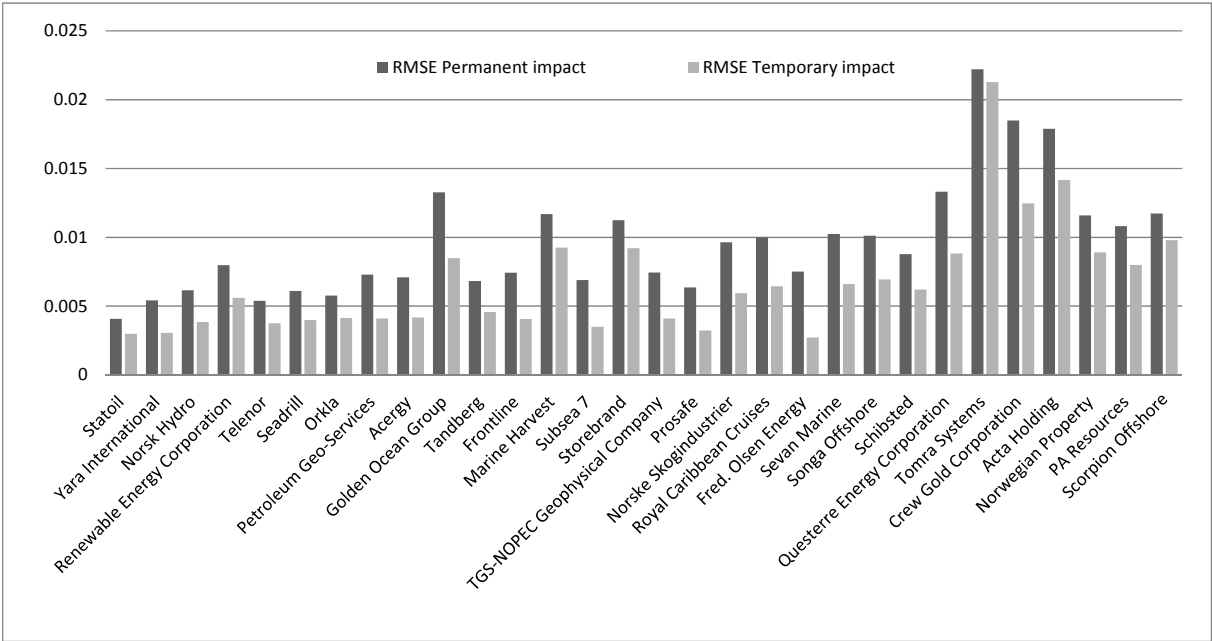
### 5.3. Robustness

We have only considered securities with a high turnover in the sample period, thus the results cannot be generalized for illiquid securities. It is reasonable to assume that volatility in our sample period differs from normal market conditions due to the financial crisis starting in 2008. We have not modeled volatility and hence this might influence our results. However, our sample is in microstructure terms relatively large, and the squared residuals seem acceptable.

Omitted variables are an important issue that can cause bias in a model. One might suggest that security prices are influenced by the general economic outlook, which might be measured by including an index in the model. However, calculating the index returns for each five seconds are considerably more complicated and hence not prioritized.

Figure 6 shows the root mean squared error (RMSE) for each of the 30 stocks in RQ1. The stocks are arranged descending after turnover. A large RMSE means that the in-sample predicted returns have large average deviations from the observed return. We see that the models for temporary and permanent impact have a better fit for the most trades stocks.

**Figure 6 - Root mean squared error**



**Chapter 6: Conclusions and further research**

In the empirical analysis we find support for a temporary and permanent price impact from trades larger than 0.2 % of daily traded volume. Theory suggests that impacts from large trades are related to private information (Easley and O'Hara, 1987). However, we cannot know if the impact is due to private information or not. We observe that large buy-initiated (sell-initiated) transactions are positively (negatively) correlated with returns the next 5 seconds and 10 minutes. This observation of price impact is consistent with several studies at other exchanges (e.g. Holthausen, et al., 1987; Hasbrouck and Schwartz, 1988; Almgren, et al., 2005).

The test for intraday differences support the alternative hypothesis that price impacts do vary. This finding suggest that optimization of execution costs, such as proposed by Bertsimas and Lo (1998), should take into account the differences in the intraday price impacts. Intuitively these results seem reasonable because the intraday volume has a characteristic u-shape.

We find that normalized order flows are positively correlated with returns. The interpretation is that a positive (negative) imbalance between the value of buy-initiated and sell-initiated transactions in a 15 minutes interval, are associated with positive (negative) returns the next 15 minutes. The R-squared is very low compared to similar studies performed with indices, and may suggest that its difficult to explain short term returns in individual stocks. This is what we would expect from a random walk perspective. On the other hand, all our models suggest that there are a negative correlation in returns, which indicate mean reversal.

If we had more time available we would consider conducting a panel data analysis of large transactions. E.g. we could take the largest daily trade in each stock and estimate the price impact. Then we might better control for any differences in the stocks' characteristics. In further empirical research it would be interesting to see if the price impacts are different in periods with more or less asymmetric information. An example can be before and after financial statements are issued from companies. More comprehensive theories of limit order markets are also an area of research with several possibilities. New theories could shed light over the market dynamics without the presence of market makers.

## References

- Admati, A. & Pfleiderer, P., 1988. A Theory of Intraday Patterns: Volume and Price Variability. *The Review of Financial Studies*, 1(1), pp. 3-40.
- Aït-Sahalia, Y. & Yu, J., 2009. High frequency market microstructure noise estimates and liquidity measures. *The Annals of Applied Statistics*, 3(1), pp. 422-457.
- Almgren, R. & Chriss, N., 2000. Optimal execution of portfolio transactions. *Journal of Risk*, 3(2), pp. 5-39.
- Almgren, R., Thum, C., Hauptmann, E. & Li, H., 2005. Direct Estimation of Equity Market Impact. *Risk*, July.
- Andersen, T. G. & Bollerslev, T., 1997. Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4(2-3), pp. 115-158.
- Bagehot, W., 1971. The only game in town. *Financial Analyst Journal*, March/April, Volume 27, pp. 12-14.
- Barclay, M. & Warner, J., 1993. Stealth trading and volatility. *Journal of Financial Economics*, Volume 34, pp. 281-305.
- Bertsimas, D. & Lo, A. W., 1998. Optimal control of execution costs. *Journal of Financial Markets*, Volume 1, pp. 1-50.
- Black, F., 1986. Noise. *The Journal of Finance*, July, 41(3), pp. 529-543.
- Blume, M. E., MacKinlay, A. C. & Terker, B., 1989. Order Imbalances and Stock Price Movements on October 19 and 20, 1987. *The Journal of Finance*, 44(4), pp. 827-848.
- Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3), pp. 307-327.
- Chakravarty, S., 2001. Stealth-trading: Which traders' trades move stock prices?. *Journal of Financial Economics*, 61(2), pp. 289-307.
- Chan, L. K. C. & Lakonishok, J., 1995. The Behavior of Stock Prices Around Institutional Trades. *The Journal of Finance*, 50(4), pp. 1147-1174.

- Chordia, T. & Subrahmanyam, A., 2004. Order imbalance and individual stock returns: Theory and evidence. *Journal of Financial Economics*, 72(3), pp. 482-518.
- Dunne, P., Hau, H. & Moore, M., 2010. International order flows: Explaining equity and exchange rate returns. *Journal of International Money and Finance*, March, 29(2), pp. 358-386.
- Easley, D. & O'Hara, M., 1987. Price, trade size and information in securities markets. *Journal of Financial Economics*, 19(1), pp. 69-90.
- Enders, W., 2010. *Applied Econometric Time Series*. 3rd ed. Hoboken NJ: John Wiley & Sons, Inc..
- Engle, R. F. & Russel, J. R., 2010. Analysis of High Frequency Financial Data. In: Y. Ait-Sahalia & L. P. Hansen, eds. *Handbook of financial econometrics: Tools and Techniques*. Amsterdam: Elsevier, pp. 383-426.
- Glosten, L. & Harris, L., 1988. Estimating the Components of the Bid-Ask Spread. *Journal of Financial Economics*, 21(1), pp. 123-142.
- Harri, A. & Brorsen, B. W., 2009. The Overlapping Data Problem. *Quantitative and Qualitative Analysis in Social Sciences*, 3(3), pp. 78-115.
- Hasbrouck, J., 1991. Measuring the Information Content of Stock Trades. *The Journal of Finance*, March, 46(1), pp. 179-207.
- Hasbrouck, J., 2007. *Empirical Market Microstructure*. New York: Oxford University Press, Inc..
- Hasbrouck, J. & Schwartz, R. A., 1988. Liquidity and execution cost in equity markets. *Journal of Portfolio Management*, 14(3), pp. 10-16.
- Holthausen, R. W., Leftwich, R. W. & Mayers, D., 1987. The effect of large block transactions on security prices. *Journal of Financial Economics*, 19(2), pp. 237-267.
- Holthausen, R. W., Leftwich, R. W. & Mayers, D., 1990. Large-Block transactions, the speed of response, and temporary and permanent stock-price effects. *Journal of Financial Economics*, 26(1), pp. 71-95.

- Kissell, R. & Malamut, R., 2006. Algorithmic Decision-Making Framework. *Journal of Trading*, 1(1), pp. 12-21.
- Koski, J. L. & Michaely, R., 2000. Prices, Liquidity and the Information Content of Trades. *The Review of Financial Studies*, 13(3), pp. 659-696.
- Kyle, A. S., 1985. Auctions and Insider Trading. *Econometrica*, 53(6), pp. 1315-1335.
- Lakonishok, J., Shleifer, A. & Vishny, R. W., 1992. The impact of institutional trading on stock prices. *Journal of Financial Economics*, 32(1), pp. 23-43.
- Lee, C. M. C. & Ready, M., 1991. Inferring Trade Direction from Intraday Data. *The Journal of Finance*, 46(2), pp. 733-746.
- Milgrom, P. & Stokey, N., 1982. Information, Trade and Common Knowledge. *Journal of Economic Theory*, 26(1), pp. 17-27.
- Moberg, J.-M., 2008. *Essays on empirical market microstructure*, Ph. D. thesis: Norwegian School of Economics.
- Næs, R., 2004. *Essays on the microstructure of stock markets: Empirical evidence from trading arrangements without dealer intermediation*, Ph. D. thesis: Norwegian School of Economics.
- Odders-White, E. R., 2000. On the occurrence and consequence of inaccurate trade classification. *Journal of Financial Markets*, 3(3), pp. 259-286.
- O'Hara, M., 1995. *Market Microstructure Theory*. Malden: Blackwell Publishing.
- Oslo Børs ASA, 2008. *Oslo Børs*. [Online]  
Available at: [http://www.oslobors.no/ob\\_eng/Oslo-Boers/About-us/Press-room/Press-releases/Oslo-Boers-extends-its-opening-hours-to-17.30](http://www.oslobors.no/ob_eng/Oslo-Boers/About-us/Press-room/Press-releases/Oslo-Boers-extends-its-opening-hours-to-17.30)  
[Accessed 14 October 2011].
- Oslo Børs Informasjon AS, 2006. *OCDF Technical Specification, Version 3.0.14*, Oslo: Oslo Børs.
- Perold, A. F., 1988. The implementation shortfall: Paper versus reality. *Journal of Portfolio Management*, Spring, 14(3), pp. 4-9.

R Development Core Team, 2011. *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing (<http://www.R-project.org/>).

Wooldridge, J. M., 2008. *Introductory Econometrics: A Modern Approach*. 4th ed. s.l.:South-Western.

## **References R-packages**

Croissant, Y. & Millo, G., 2008. Panel Data Econometrics in R: The plm Package. *Journal of Statistical Software*, 27(2). Available at: <http://www.jstatsoft.org/v27/i02/> [Accessed 17 December 2011].

Zeileis, A. & Hothorn, T., 2002. Diagnostic Checking in Regression Relationships. *R News*, 2(3), pp. 7-10. Available at: <http://CRAN.R-project.org/doc/Rnews/> [Accessed 17 December 2011].

Zeileis, A., 2004. Econometric Computing with HC and HAC Covariance Matrix Estimators. *Journal of Statistical Software*, 11(10), pp. 1-17. Available at: <http://www.jstatsoft.org/v11/i10/> [Accessed 17 December 2011].



## Appendix A: Estimated coefficients research question 1

Estimated coefficients from research question 1		Temporary			Permanent					
Alternative hypotheses		$\beta_1 > 0$	$\beta_2 < 0$	$\beta_1 + \beta_2 < 0$	$\beta_1 > 0$	$\beta_2 < 0$	$\beta_3 > 0$	$\beta_4 \neq 0$	$\beta_1 + \beta_2 < 0$	
Explanatory variables →	Obs.	$S_{i,t}$	$S_{i,t}D_{i,t}^{Sell}$	$S_{i,t}S_{i,t}D_{i,t}^{Sell}$	Obs.	$\sqrt{S_{i,t}}$	$\sqrt{S_{i,t}}D_{i,t}^{Sell}$	$v_{i,t}^{Normalized}$	$\Gamma_{i,t}^{Perm\ lag}$	$\sqrt{S_{i,t}}, \sqrt{S_{i,t}}D_{i,t}^{Sell}$
<b>Stocks:</b>										
Aceryg	36174	0.0641***	-0.0937***	-0.0296***	34083	0.0092***	-0.0109***	0.0004***	-0.2869***	-0,0018
Acta Holding	44470	0.067***	-0.0874***	-0.0203***	42265	0.0173***	-0.0315***	0.0001***	-0.4409***	-0.0142***
Crew Gold Corporation	57650	0.0476***	-0.1271***	-0.0795***	54928	0.0195***	-0.0453***	0.0001***	-0.4771***	-0.0258***
Fred. Olsen Energy	59614	0.0404***	-0.0533***	-0.0129***	56288	0.0087***	-0.0104***	0.0002***	-0.3827***	-0,0017
Frontline	49728	0.0277***	-0.0632***	-0.0355***	46956	0.0171***	-0.0199***	0.0002***	-0.0297***	-0.0027*
Golden Ocean Group	49109	0.0848***	-0.1372***	-0.0523***	47014	0.0198***	-0.0175***	0.0003***	-0.3961***	0,0022
Marine Harvest	52371	0.0901***	-0.2118***	-0.1217***	49677	0.0145***	-0.0222***	0.0003***	-0.3842***	-0.0077***
Norsk Hydro	20665	0.0305***	-0.0788***	-0.0483***	19635	0.0180***	-0.017***	0.0004***	-0.3184***	0,001
Norske Skogindustrier	55405	0.0894***	-0.1928***	-0.1034***	52706	0.0096***	-0.0257***	0.0001***	-0.3546***	-0.0162***
Norwegian Property	48005	-0.0085	-0.0133	-0.0219***	45197	0.0025	-0.0274***	0.0001*	-0.4359***	-0.0249***
Orkla	29138	0.0471***	-0.0796***	-0.0325***	27463	0.0037**	-0.0086***	0.0002***	-0.3393***	-0.0049***
PA Resources	46205	0.0793***	-0.1197***	-0.0405***	43848	0.0118***	-0.0293***	0.0001**	-0.3436***	-0.0175***
Petroleum Geo-Services	34347	0.0511***	-0.1292***	-0.0781***	32636	0.0231***	-0.0179***	0.0003***	-0.3099***	0,0053
Prosafte	50034	0.0653***	-0.1284***	-0.0631***	47150	0.0119***	-0.0234***	0.0002***	-0.3235***	-0.0115***
Questerre Energy Corporation	54830	0.031***	-0.0612***	-0.0303***	52407	0.0121***	-0.0271***	0.0002***	-0.3777***	-0.015***
Renewable Energy Corporation	21610	0.0616***	-0.1155***	-0.0539***	20521	0.0052*	-0.0078**	0.0003***	-0.3327***	-0,0026
Royal Caribbean Cruises	56410	0.0276***	-0.1464***	-0.1188***	53008	0.0076***	-0.0314***	0.0002***	-0.3622***	-0.0238***
Schibsted	61093	0.0152***	-0.0893***	-0.0741***	57584	0.0123***	-0.0273***	0.0002***	-0.375***	-0.0149***
Scorpion Offshore	26858	0.0203***	-0.045***	-0.0247***	25125	0.0086***	-0.0205***	0.0001**	-0.4896***	-0.0119***
Seadrill	33657	0.0173**	-0.0635***	-0.0463***	31774	0.0058***	-0.0146***	0.0003***	-0.3083***	-0.0087***
Sevan Marine	54522	0.0014	-0.006*	-0.0046**	51800	0.0088***	-0.0228***	0.0002***	-0.375***	-0.0139***
Songa Offshore	65316	0.0284***	-0.0442***	-0.0158***	61792	0.0167***	-0.0297***	0.0002***	-0.3967***	-0.013***
Statoil	16237	0.0235***	-0.045***	-0.0215***	15476	0.0083***	-0.0102***	0.0001**	-0.2844***	-0,0019
Storebrand	46409	0.0116	-0.0823***	-0.0708***	43499	0.0033	-0.0173***	0.0003***	-0.3935***	-0.0141***
Subsea 7	52805	0.0526***	-0.0907***	-0.038***	49721	0.0136***	-0.0201***	0.0001***	-0.3118***	-0.0065***
Tandberg	40485	0.0059**	-0.0242***	-0.0182***	38313	0.0056***	-0.0122***	0.0002***	-0.3145***	-0.0066***
Telenor	31071	0.0158**	-0.0736***	-0.0578***	29479	0.0022	-0.0112***	0.0001**	-0.342***	-0.009***
TGS-NOPEC Geophysical Company	49162	0.0054**	-0.0114***	-0.006*	46400	0.0142***	-0.0235***	0.0002***	-0.3244***	-0.0093***
Tomra Systems	56024	0.0336***	-0.1091***	-0.0755***	52613	0.0119***	-0.0105***	0.0003***	-0.5247***	0,0015
Yara International	23121	0.0242***	-0.0715***	-0.0473***	22161	0.0030	-0.0097***	0.0001***	-0.2624***	-0.0068***

P-value: \*\*\*  $\leq 0.01$ , \*\*  $\leq 0.05$ , \*  $\leq 0.1$

Temporary time-dummies coefficients for buyer initiated transactions.  $\sum_{j=1}^8 \delta_{ij} D_j S_{i,t}$

Stock	09:30 - 10:30	10:30 - 11:30	11:30 - 12:30	12:30 - 13:30	13:30 - 14:30	14:30 - 15:30	15:30 - 16:30	16:30 - 17:20
Aceryg	-0.056***	-0.0507***	-0.0449***	-0.0441***	-0.0405***	-0.046***	-0.0386***	-0.0123
Acta Holding	-0.0569***	-0.0371***	-0.064***	-0.0635***	-0.0532***	-0.064***	-0.0526***	-0.0255**
Crew Gold Corporation	-0.0138	-0.0218	-0.0225	-0.0085	0.0002	-0.004	0.0069	0.1175***
Fred. Olsen Energy	-0.0302***	-0.0353***	-0.0266***	-0.0338***	-0.036***	-0.0318***	-0.0277***	-0.0268***
Frontline	-0.0012	-0.0053	-0.0078	-0.0058	-0.0025	-0.0028	0.034***	0.0655***
Golden Ocean Group	-0.0256	-0.027	-0.0415**	-0.0375*	-0.0486***	-0.0513***	-0.028*	-0.0508***
Marine Harvest	-0.0286**	-0.0763***	-0.043***	-0.0662***	-0.0702***	-0.0521***	-0.0526***	-0.0123
Norsk Hydro	-0.0148	-0.0163	-0.0121	-0.02	-0.0264*	-0.0209	-0.0183	-0.0482**
Norske Skogindustrier	-0.0341***	-0.0357***	-0.0289**	-0.042***	-0.0544***	-0.0576***	-0.0476***	-0.0251*
Norwegian Property	0.0195**	0.0104	0.0118	0.0138	0.0183*	0.0248***	0.0181**	0.0184*
Orkla	-0.0234*	-0.0326**	-0.0415***	-0.036***	-0.0277**	-0.0447***	-0.0219*	-0.0198
PA Resources	-0.0453***	-0.0581***	-0.069***	-0.053***	-0.0443***	-0.0592***	-0.0459***	-0.0176**
Petroleum Geo-Services	-0.0197	0.0028	-0.0404***	-0.0082	-0.0288**	-0.0115	-0.009	0.0151
Prosafe	-0.0554***	-0.0549***	-0.0409***	-0.0469***	-0.0598***	-0.0585***	-0.0513***	-0.0376***
Questerre Energy Corp.	-0.0117*	-0.0234***	-0.008	-0.0195**	-0.022***	-0.0136*	0.0078	0.0685***
Renewable Energy Corp.	-0.0182	-0.0376*	-0.0569***	-0.0169	-0.0415*	-0.0316	-0.0006	0.0119
Royal Caribbean Cruises	0.0072	0.0104	0.0062	0.0019	0.0072	-0.0046	0.0336***	0.0934***
Schibsted	-0.0038	-0.0011	-0.0036	0.0005	0.021***	-0.0118**	0.0013	-0.0118**
Scorpion Offshore	-0.0159*	-0.0133	-0.0179**	-0.0151	-0.0038	-0.0061	-0.0141*	0.005
Seadrill	0.0147	0.004	-0.0038	-0.0048	-0.0055	0.0005	0.0141	0.0057
Sevan Marine	0.0036	0.0177***	0.0112*	0.0085*	0.0023	0.0068	0.0078**	0.0058
Songa Offshore	-0.0209***	-0.0217***	-0.0197**	-0.0235***	-0.0205***	-0.021***	-0.0097	0.0046
Statoil	0.0379***	0.0163	0.0283*	-0.0074	-0.0195*	0.0141	0.0017	0.0235
Storebrand	-0.0006	-0.004	0.0166	0.0003	0.0167	-0.0062	0.0038	0.0499***
Subsea 7	-0.0222**	-0.019*	-0.0119	-0.0054	-0.0321***	-0.0331***	-0.0319***	-0.0086
Tandberg	0.0025	-0.0002	-0.0011	0.0027	-0.0014	-0.0015	0.0069*	0.0007
Telenor	-0.0069	0.0259**	0.0057	-0.0063	0.0081	0.0161	0.0271***	-0.0077
TGS-NOPEC	0.0044	0.0022	0.0147**	-0.0001	0.0091*	0.0053	0.0086**	0.0165***
Tomra Systems	-0.0327**	-0.0142	-0.0305**	0.004	-0.0154	-0.0159	-0.02	-0.0191
Yara International	0.008	-0.0092	0.0068	-0.0038	0.0083	0.0127	0.0112	0.025

P-value: \*\*\*  $\leq 0.01$ , \*\*  $\leq 0.05$ , \*  $\leq 0.1$

Temporary time-dummy coefficients for seller initiated transactions.  $\sum_{j=1}^8 \phi_{ij} D_j S_{i,t} D_{i,t}^{Sell}$

Stock	09:30 - 10:30	10:30 - 11:30	11:30 - 12:30	12:30 - 13:30	13:30 - 14:30	14:30 - 15:30	15:30 - 16:30	16:30 - 17:20
Acergy	0.0439***	0.0699***	0.0485***	0.0459***	0.0453**	0.0238	0.0494***	-0.0009
Acta Holding	0.0729***	0.0537***	0.0802***	0.0743***	0.0575***	0.0805***	0.0665***	0.0115
Crew Gold Corporation	0.0722***	0.0661***	0.0818***	0.0662***	0.0613***	0.0531***	0.0382**	-0.0902***
Fred. Olsen Energy	0.0227***	0.0413***	0.0355***	0.0397***	0.0314***	0.035***	0.0251***	0.0231***
Frontline	0.0198	0.0227*	0.0183	0.0159	0.0122	0.0157	-0.0329***	-0.1068***
Golden Ocean Group	0.0386**	0.0366*	0.0793***	0.0509**	0.0639***	0.0535***	0.0262	0.0266
Marine Harvest	0.1193***	0.1548***	0.1158***	0.1434***	0.1745***	0.1284***	0.1216***	0.0634***
Norsk Hydro	0.051**	0.0442**	0.041**	0.0658***	0.07***	0.0584***	0.0529***	0
Norske Skogindustrier	0.1053***	0.0889***	0.1044***	0.1139***	0.1508***	0.1045***	0.1013***	0.0445***
Norwegian Property	-0.0122	-0.0006	0.0003	0.0013	-0.004	-0.0086	-0.005	-0.0455***
Orkla	0.0438***	0.0561***	0.0426***	0.0604***	0.063***	0.0446***	0.0439***	0.043***
PA Resources	0.0522***	0.091***	0.0962***	0.0336***	0.0752***	0.068***	0.0723***	0.0363***
Petroleum Geo-Services	0.0854***	0.0545***	0.091***	0.0464**	0.0791***	0.0784***	0.0417**	0.0297
Prosafe	0.1051***	0.1168***	0.0855***	0.0849***	0.1045***	0.1093***	0.1004**	0.0623***
Questerre Energy Corp.	0.0271***	0.0275***	0.0216**	0.0181	0.0357***	0.0266***	-0.0103	-0.1224***
Renewable Energy Corp.	0.0028	0.0599**	0.0861***	0.0551*	0.1326***	0.0535*	0.0051	-0.0169
Royal Caribbean Cruises	0.0708***	0.0715***	0.0851***	0.0716***	0.071***	0.0959***	0.0056	-0.1053***
Schibsted	0.0682***	0.067***	0.0645***	0.0583***	0.0433***	0.0698***	0.0629***	0.046***
Scorpion Offshore	0.0323***	0.0371***	0.0309**	0.0377***	0.0215*	0.0224*	0.0303**	0.0155
Seadrill	0.0299*	0.0171	0.0271*	0.0005	0.0408***	0.0244*	0.0047	-0.0128
Sevan Marine	-0.0064	-0.0217***	-0.0192***	-0.0073	0.0006	-0.0135**	-0.029***	-0.0582***
Songa Offshore	0.0288***	0.0299***	0.0284***	0.0322***	0.0314***	0.0209**	0.0133	-0.0537***
Statoil	-0.0411***	-0.027	-0.0324*	0.0054	-0.0138	-0.0233	-0.0207	-0.0413*
Storebrand	0.0287	0.0644***	0.0246	0.0596***	0.0369*	0.0553***	0.0491***	0.0151
Subsea 7	0.031***	0.0463***	0.0383***	0.0339***	0.0595***	0.0398***	0.042***	0.0085
Tandberg	0.0091	0.0109	0.0101	0.0095	0.0153**	0.0162***	-0.0048	0.0119*
Telenor	0.0258**	-0.0084	0.0202	0.0456***	0.0258*	-0.0016	-0.0106	0.0155
TGS-NOPEC	-0.0175**	-0.0228***	-0.0222***	-0.0043	-0.0193**	-0.0264***	-0.0237***	-0.0489***
Tomra Systems	0.079***	0.0856***	0.0881***	0.0627***	0.0847***	0.0846***	0.0807***	0.0439**
Yara International	0.0305**	0.0292*	0.0096	0.0142	0.0224	0.0141	-0.0086	-0.0419

P-value: \*\*\*  $\leq 0.01$ , \*\*  $\leq 0.05$ , \*  $\leq 0.1$

Permanent time-dummies coefficients for buyer initiated transactions.  $\sum_{j=1}^8 \delta_{i,j} D_j S_{i,t}$

Stock	09:30 - 10:30	10:30 - 11:30	11:30 - 12:30	12:30 - 13:30	13:30 - 14:30	14:30 - 15:30	15:30 - 16:30	16:30 - 17:20
Aceryg	-0.0028	-0.0006	-0.0004	-0.0050	-0.0009	0.0004	-0.0059**	-0.0014
Acta Holding	-0.008***	-0.0027	-0.0063*	-0.0089***	-0.0098***	-0.0071**	-0.0109***	-0.0073*
Crew Gold Corporation	-0.0031	-0.0037	-0.0092***	-0.0097***	-0.0071**	-0.0072**	-0.0016	0.0144***
Fred. Olsen Energy	-0.0063***	-0.0044*	-0.0023	-0.0030	-0.005**	-0.0049**	-0.0084***	0.0015
Frontline	-0.0076***	-0.0082***	-0.0059**	-0.0051*	-0.0084***	-0.0101***	-0.0088***	-0.0089***
Golden Ocean Group	-0.0104***	-0.0121***	-0.0163***	-0.0054	-0.0102**	-0.009**	-0.0106***	-0.028***
Marine Harvest	-0.0018	-0.0011	-0.0012	-0.0042	-0.0046	-0.005*	-0.0095***	-0.0153***
Norsk Hydro	-0.0096***	-0.0137***	-0.0118***	-0.0122***	-0.0118***	-0.0121***	-0.0149***	-0.0208***
Norske Skogindustrier	0.0030	0.0014	-0.0002	0.0027	-0.0032	-0.0027	-0.0067**	0.0003
Norwegian Property	0.0058**	0.0020	0.0027	0.0021	0.0034	0.0043	0.0044	0.006*
Orkla	-0.0013	-0.0005	0.0004	0.0010	0.0033	0.0003	-0.0012	-0.0033
PA Resources	-0.0008	-0.0046**	-0.004**	-0.0066***	-0.0019	0.0002	-0.0035*	0.0107***
Petroleum Geo-Services	-0.0124***	-0.0108***	-0.0134***	-0.0158***	-0.0175***	-0.0124***	-0.0138***	-0.0064
Prosafe	-0.0069***	-0.0059**	-0.0081***	-0.0064**	-0.0069**	-0.0044*	-0.0055**	-0.006**
Questerre Energy Corp.	0.0001	0.0017	0.0064**	0.0008	-0.0018	0.0030	0.0047*	0.0141***
Renewable Energy Corp.	0.0009	-0.0002	-0.0024	0.0008	-0.0034	0.0027	0.0001	0.0060
Royal Caribbean Cruises	-0.0008	0.0002	-0.0014	0.0004	0.0001	0.0039	0.0033	-0.0030
Schibsted	-0.0052**	-0.0023	-0.0053**	-0.0034	-0.0043*	-0.006***	-0.0066***	-0.0102***
Scorpion Offshore	-0.0024	0.0009	-0.0043	-0.0028	-0.0039	-0.0063**	-0.0047*	0.0076**
Seadrill	-0.0009	0.0021	0.0011	0.0001	0.0011	-0.0002	-0.0039*	-0.0111***
Sevan Marine	-0.0037	-0.0017	-0.0018	0.0017	-0.0009	0.0008	-0.0052**	0.0116***
Songa Offshore	-0.007***	-0.0073***	-0.0066***	-0.0083***	-0.0083***	-0.0083***	-0.0085***	-0.0016
Statoil	-0.0003	-0.0037	-0.0033	-0.0059**	-0.004*	-0.0022	-0.0054**	-0.0096***
Storebrand	-0.0001	-0.0033	-0.0008	0.0007	0.0010	-0.0006	0.0044	0.007**
Subsea 7	-0.0046**	-0.0048*	-0.0052**	-0.0059**	-0.0074***	-0.0056**	-0.0066***	-0.0044
Tandberg	-0.0006	-0.0012	-0.0012	0.0001	-0.0013	0.0009	0.0011	-0.0072***
Telenor	0.0053**	0.0069***	0.0032	0.0057**	0.0062**	0.0045**	0.0004	-0.0051*
TGS-NOPEC	-0.0041*	-0.0076***	-0.0044	-0.0079***	-0.0089***	-0.0072***	-0.0089***	-0.0026
Tomra Systems	-0.0112***	-0.0072***	-0.007***	-0.0039	-0.0065***	-0.0053**	-0.0098***	-0.0089***
Yara International	0.0047*	0.0001	0.0059**	0.0045	0.0044	0.0039	0.0023	-0.0099**

P-value : \*\*\*  $\leq 0.01$ , \*\*  $\leq 0.05$ , \*  $\leq 0.1$

Permanent time-dummy coefficients for seller initiated transactions.  $\sum_{j=1}^8 \Phi_{i,j} D_j S_{i,t} D_{i,t}^{Sell}$

Stock	09:30 - 10:30	10:30 - 11:30	11:30 - 12:30	12:30 - 13:30	13:30 - 14:30	14:30 - 15:30	15:30 - 16:30	16:30 - 17:20
Aceryg	-0.005	-0.0025	-0.0044	0.0016	-0.0051	-0.005	0.0003	-0.0083*
Acta Holding	0.0105***	0.0061*	0.0093***	0.0117***	0.0112***	0.0148***	0.0182***	0.013***
Crew Gold Corporation	0.014***	0.0163***	0.0203***	0.0219***	0.021***	0.0261***	0.0222***	0.0052
Fred. Olsen Energy	-0.0019	0.001	0.0013	0.002	0.0009	0.0012	-0.0024	-0.0019
Frontline	0.0077***	0.0055**	0.0024	0.0032	0.0043	0.0021	0.0015	0.0026
Golden Ocean Group	-0.0019	-0.0004	0.006	-0.004	-0.0008	0.0054	0.0006	-0.0072
Marine Harvest	0.0008	0.0026	0.0014	0.0051*	0.0061**	0.0066**	0.0063**	0.0028
Norsk Hydro	0.0035	0.0065	0.0064	0.0071*	0.0092**	0.0061	0.0045	-0.0012
Norske Skogindustrier	0.0028	0.0085***	0.0086***	0.0039	0.0091***	0.0079***	0.008***	0.0059*
Norwegian Property	0.0072**	0.0133***	0.0174***	0.0171***	0.015***	0.0139***	0.0144***	0.0116***
Orkla	0.0002	0.0008	-0.0028	-0.0009	-0.0006	0.0004	-0.0001	0.0025
PA Resources	0.006***	0.014***	0.014***	0.0104***	0.0121***	0.0092***	0.0139***	0.0029
Petroleum Geo-Services	0.0032	0.0036	0.0049	0.004	0.0021	0.0008	0.0066*	0.0045
Prosafé	0.0065**	0.0103***	0.0091***	0.0096***	0.01***	0.0107***	0.012***	0.0091***
Questaerre Energy Corp.	0.0025	0.0045*	0.0003	0.0044	0.0094***	0.0002	0.003	-0.0004
Renewable Energy Corp.	-0.0046	-0.0025	-0.0001	-0.0051	-0.0024	-0.0048	-0.0148***	-0.0024
Royal Caribbean Cruises	0.0142***	0.0144***	0.0178***	0.0115***	0.0148***	0.0096***	0.0106***	0.0108***
Schibsted	0.0099***	0.0094***	0.0137***	0.0144***	0.0156***	0.0147***	0.0146***	0.0196***
Scorpion Offshore	0.0054*	0.0072**	0.0109***	0.0139***	0.0134***	0.0094***	0.0129***	0.002
Seadrill	0.0041	0.0034	0.0037	0.0024	0.0031	0.0019	0.0068***	-0.0005
Sevan Marine	0.0099***	0.0068**	0.0095***	0.005	0.0105***	0.0078***	0.0102***	0.0029
Songa Offshore	0.0105***	0.0135***	0.0129***	0.0118***	0.0144***	0.0129***	0.0124***	0.0015
Statoil	0.002	0.001	-0.0002	0.0026	-0.0003	0.0025	0.002	-0.0019
Storebrand	0.0044	0.0067**	0.0051	0.0061*	0.0064*	0.0064**	0.0056*	0.0018
Subsea 7	0.0062**	0.0055**	0.0057**	0.0061**	0.0081***	0.0046*	0.0076***	0.0015
Tandberg	0.0018	0.0034	0.002	0.0021	0.0031	0.0015	0.0007	0.0067**
Telenor	-0.0007	-0.0027	0.0002	-0.0035	0	-0.001	0.0011	0.0088**
TGS-NOPEC	0.0048*	0.0089***	0.0071**	0.0108***	0.012***	0.01***	0.0112***	0.0084**
Tomra Systems	-0.0012	0.0013	0.0008	-0.0016	0.0014	0.0016	0.0018	-0.0039
Yara International	-0.005*	-0.0028	-0.0046	-0.0105***	-0.0005	-0.0019	-0.0035	0.0037

P-value: \*\*\*  $\leq 0.01$ , \*\*  $\leq 0.05$ , \*  $\leq 0.1$

## Appendix B: Likelihood ratio test

We test the unrestricted models (model 1 and 2) and the restricted models without time-dummies. The alternative hypothesis is that the unrestricted model has a significant better fit measured with log-likelihood. Test statistics for the Likelihood ratio test (LR):  $LR = 2(L_{ur} - L_r)$  where “L” is the log-likelihood for the unrestricted ( $L_{ur}$ ) and the restricted ( $L_r$ ) model. Reject  $H_0$  if  $LR > \chi_{0.05,16}^2$ . The P-value for the LR test is:  $\text{Prob}(LR \leq \chi_{0.05,16}^2)$ .

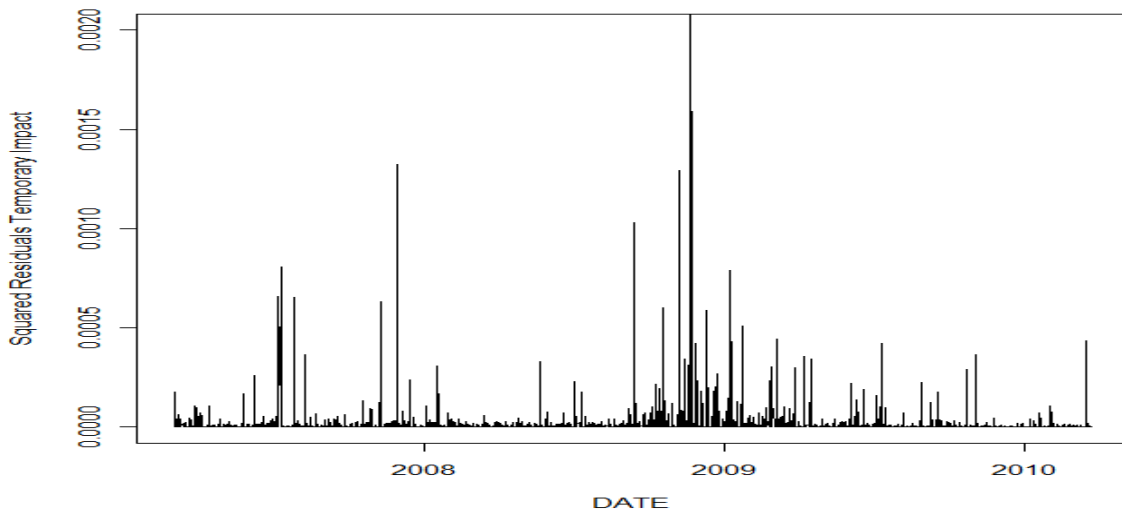
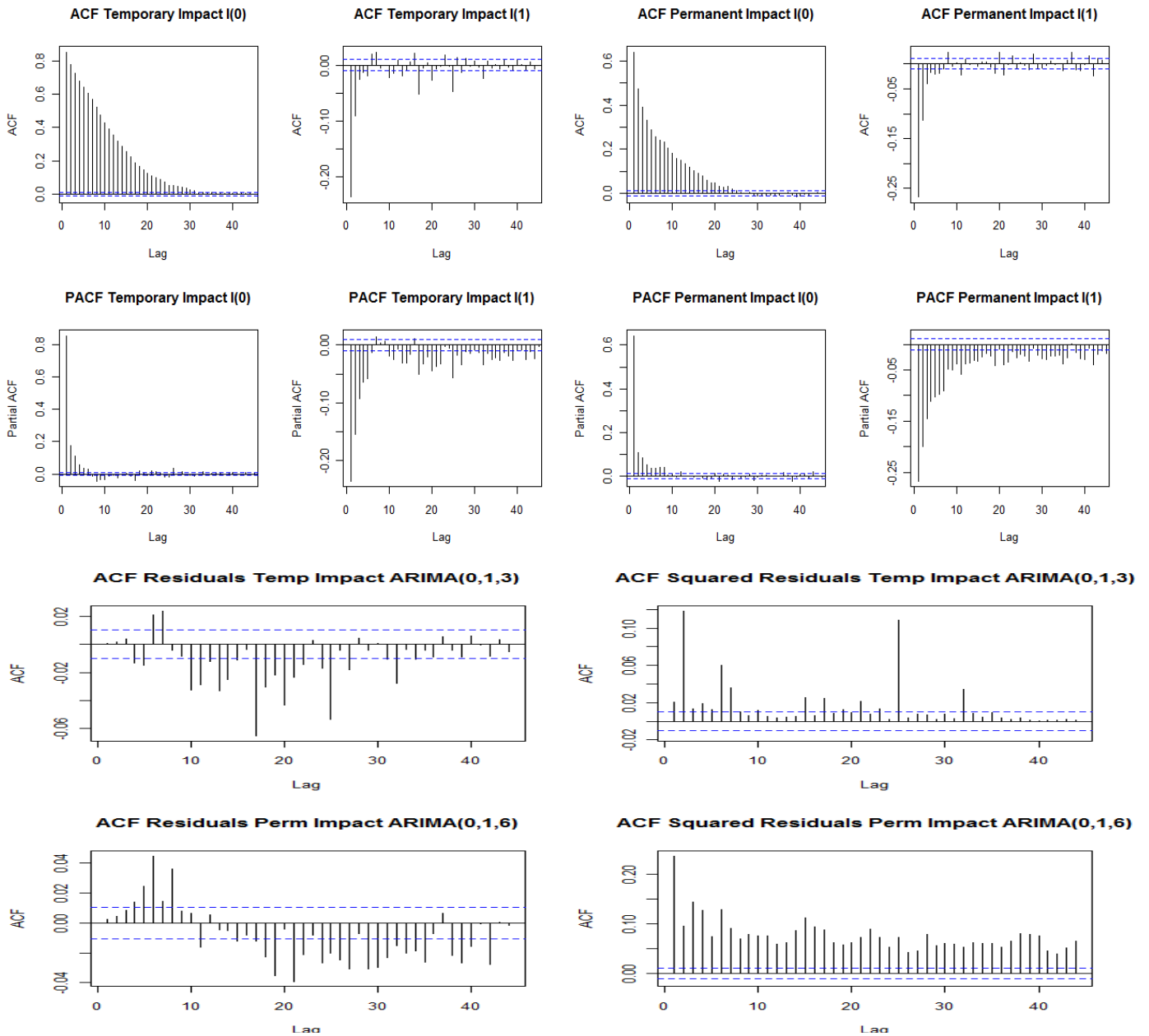
Stock	Temporary		Permanent	
	LR	P-Value	LR	P-Value
Acergy	69	0.0000	27	0.0395
Acta Holding	99	0.0000	48	0.0000
Crew Gold Corporation	184	0.0000	143	0.0000
Fred. Olsen Energy	101	0.0000	55	0.0000
Frontline	178	0.0000	39	0.0009
Golden Ocean Group	37	0.0019	73	0.0000
Marine Harvest	155	0.0000	50	0.0000
Norsk Hydro	41	0.0005	54	0.0000
Norske Skogindustrier	235	0.0000	45	0.0001
Norwegian Property	43	0.0003	84	0.0000
Orkla	74	0.0000	14	0.6228
PA Resources	334	0.0000	151	0.0000
Petroleum Geo-Services	99	0.0000	42	0.0004
Prosafe	131	0.0000	40	0.0007
Questerre Energy Corporation	185	0.0000	50	0.0000
Renewable Energy Corporation	110	0.0000	32	0.0097
Royal Caribbean Cruises	412	0.0000	79	0.0000
Schibsted	121	0.0000	87	0.0000
Scorpion Offshore	53	0.0000	95	0.0000
Seadrill	55	0.0000	37	0.0020
Sevan Marine	74	0.0000	53	0.0000
Songa Offshore	152	0.0000	79	0.0000
Statoil	60	0.0000	40	0.0007
Storebrand	66	0.0000	28	0.0356
Subsea 7	95	0.0000	20	0.2420
Tandberg	27	0.0376	23	0.1040
Telenor	87	0.0000	38	0.0016
TGS-NOPEC Geophysical Company	50	0.0000	37	0.0019
Tomra Systems	85	0.0000	650	0.0000
Yara International	34	0.0061	36	0.0032

$\mathcal{H}_0: \delta_1, \delta_2, \dots, \delta_8, \phi_1, \phi_2, \dots, \phi_8 = 0, \mathcal{H}_1: \delta_1, \delta_2, \dots, \delta_8, \phi_1, \phi_2, \dots, \phi_8 \neq 0$

## **Appendix C: Output from the model fitting (Acergy)**

The Box Jenkins framework aims to estimate parsimonious, stationary and invertible models with residuals that approximate a white noise process (Enders, 2010). It is proven mathematically that the autocorrelation function (ACF) for an autoregressive (AR) of order  $p$  process is exponentially decaying and the partial ACF dies after lag  $p$ . We also have that the ACF dies after lag  $q$  for a moving average (MA) process of order  $q$  and the partial ACF is exponentially decaying. The Box-Jenkins methodology relies on these facts in order to identify the best model. This methodology is often referred to as an art rather than a science, since there is a balance between fit and parsimoniousness and other ad hoc choices (Enders, 2010).

Some of the output (Acergy stock) used for the model estimation is shown on the next page. We can see that both of the temporary and permanent impact variables have high non-converging ACF. The non-convergent series problem is solved by integrating the series at level one. After differencing, the autocorrelation in the dependent variable has an exponential declining partial ACF and the ACF is shortly lived. Based on the Box-Jenkins methodology these characteristics are similar to a moving average specification (MA). ARMA specifications for our data give unstable solutions since the lagged dependent variable is often the same as the previous, resulting in a unit root. An MA specification is invertible and a better fit for our data.





## Appendix D: Panel data analysis

We include the most important tests from our panel data analysis in research question 2. We use a 5% significance level. We initially run an F-test to test whether our panel has significant fixed effects. The test statistic equals 1 ( $F_{q=19}^{\text{Critical}} = 1.58$ ), so we fail to reject the null hypothesis of no fixed effects. This means we can run a pooled OLS.

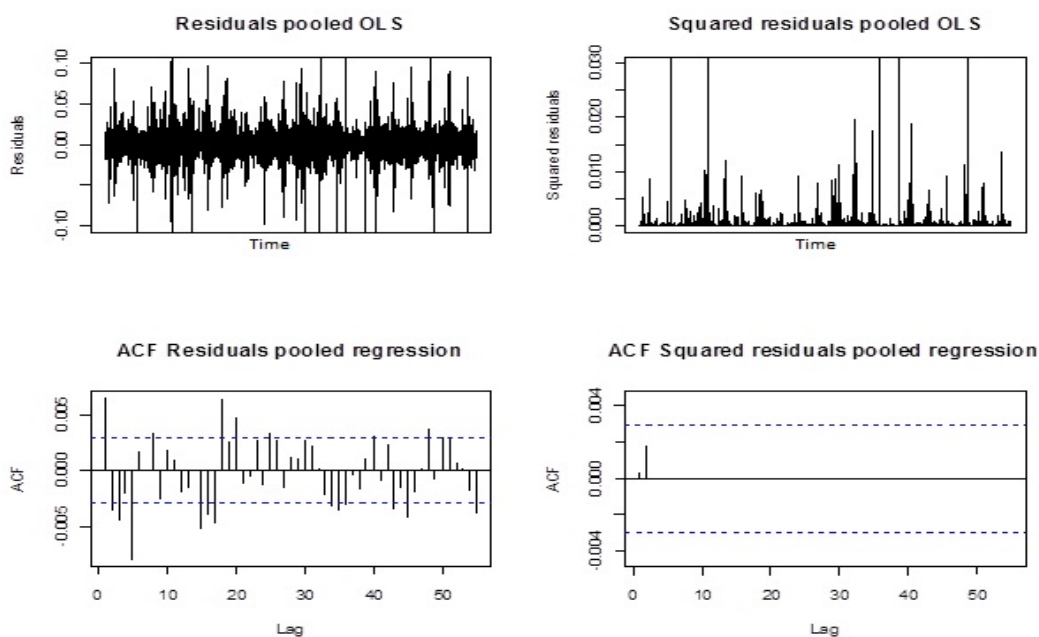
Initially we have two alternative measures for the 15 minutes order flow, normalized ( $v_t^{\text{Normalized}}$ ) and relative ( $v_t^{\text{Relative}}$ ):

$$v_t^{\text{Normalized}} = \frac{\sum_t^{t-15 \text{ min}} T_t^{\text{Buy}} - \sum_t^{t-15 \text{ min}} T_t^{\text{Sell}}}{\sum_t^{t-15 \text{ min}} T_t^{\text{Buy}} + \sum_t^{t-15 \text{ min}} T_t^{\text{Sell}}} \quad v_t^{\text{Relative}} = \log \left( \frac{\sum_t^{t-15 \text{ min}} T_t^{\text{Buy}}}{\sum_t^{t-15 \text{ min}} T_t^{\text{Sell}}} \right)$$

First we determine whether both measures for the order flow should be included. The model gave insignificant results for the relative order flow. Hence, we test whether relative order flow should be included in the model or not. We use an F-test with the restricted model where only normalized order flow is included and the unrestricted with both measures. The test statistics is 1.45 ( $F_{q=1}^{\text{Critical}} = 3.58$ ), which means that we cannot reject the null hypothesis stating that the unrestricted model has the same explanatory power as the unrestricted.

Then we test if model 3 (unrestricted) should include the seven intercept dummies. The test statistic equals 5.57 ( $F_{q=7}^{\text{Critical}} = 2.01$ ), this means that we can reject the null hypothesis.

To control for different variance across stocks and non-constant variance we use White heteroscedasticity (not autocorrelation) robust standard errors for the pooled regressions.



## **Appendix E: Data processing in Perl.**

We received 3 years of data from Oslo Stock Exchange. The data included all security transactions, orders and security changes, except data that are considered private, e.g. hidden orders. Because of the size and complexity of the dataset we use the Perl programming language<sup>17</sup> to prepare the data for the econometric analyzes. Form a starting point with little programming experience this took us a while to figure out. Therefore, to ease the effort for future scholars that use high-frequency data we have included some of the Perl scripts. Because the scope of this section is to guide the future researcher, the scripts are simplified to fit in an appendix. Be aware that if you study data from another period the field codes and data structure may be different from our sample, but the general approach may be applied. In this example we extract on file for a selection of stocks. Each file contains all stock transactions (price, quantity, time-stamp) and the current bid and ask quote.

We walk through the processing of the real time files from OBI OCDF. All the input and output files are semicolon separated. The main challenge is that field codes do not have a constant column for each trading session and hence we have to search each line for the relevant field codes. Perl is ideal for this task, because of its effectiveness in processing one line at a time and swiftly recognize combinations of expressions (pattern matching).

We found most effective to divide the extraction of data in a three step process:

- 1) Filter the data and keep security changes and security transactions**
- 2) Extract data for each security by ISIN number to separate folders**
- 3) Extract transactions with time and current bid and ask quotes for each stock**

---

<sup>17</sup>We use Strawberry Perl for windows in our data processing ( <http://strawberryperl.com/> )

## Example from part of the feed the 1<sup>st</sup> of March 2007:

Sc;401642;iNO0000000021;Mc1;La375.76735784;CqNOK;t103802;1ARt1172741882:497765;2ARt1172741882:520013  
 Sc;401643;iNO0007042299;Mc1;La461.87629415;CqNOK;t103802;1ARt1172741882:518829;2ARt1172741882:583480  
 Sc;401644;iNO0007035376;Mc1;La365.42004334;CqNOK;t103802;1ARt1172741882:537783;2ARt1172741882:585849  
 Ob;401645;OBId33817;iNO0010112675;Mc1;XIBt1172741883:148000000;1Vb5300;1Bo5;1OBCbGLI,DBL,LBI,SBN,NTF;t103802;1ARt1172741882:67294  
 6;2ARt1172741882:693363  
 o;401646;OBId33817;iNO0010112675;Mc1;XIBt1172741883:148000000;OBsB;MtD;OPId20070301103705004058;OId20070301103802600955;1ARt117274  
 1882:692125;2ARt1172741882:764134  
 Ob;401647;OBId33817;iNO0010112675;Mc1;XIBt1172741883:148000000;1Vb8300;1Bo6;1OBCbGLI,DBL,LBI,SBN,NTF;t103802;1ARt1172741882:69286  
 2;2ARt1172741882:766607  
 o;401648;OBId33817;iNO0010112675;Mc1;XIBt1172741883:148000000;OBsB;MtI;Op132.5;Oq3000;OcESO;OId20070301103802000279;  
 OCIR;1ARt1172741882:702135;2ARt1172741882:769510  
 t;401649;OBId33817;iNO0010112675;Mc1;XIBt1172741883:148000000;TId610;t103802;DTa20070301103802;DTd20070301103802;Tp132.5;Tq100;Tgl;Ttl  
 ;ULa;UTo;UVWp;TCbESO;TCsNON;1ARt1172741882:703064;2ARt1172741882:772175  
 Sc;401650;OBId33817;iNO0010112675;Mc1;CqNOK;La132.5;t103802;TUVo1034845;TUVa135431326.51999998;VWp130.98440359;1ARt1172741882:717  
 104;2ARt1172741882:777566  
 Sc;401651;OBId15290;iNO0003074809;Mc1;CqNOK;1a90.6;2a91.2;3a91.3;4a91.4;5a91.5;t103802;1ARt1172741882:766875;2ARt1172741882:800116

For explanation of all fields we refer to the technical documentation that is enclosed with the data (Oslo Børs Informasjon AS, 2006). For our purpose to type of lines are of relevance; security transactions [t] and security changes [Sc] (henceforth, the field codes we search for is embedded in []).

### General approach

For each step we have an input folder with data and an empty output folder prepared in advance. The reason for the folder structure is that we keep the filenames in the same date format (yyyymmdd.txt). We keep the date format to check dates at later stages, and when processing all files in a folder the data is processed in the right order. For illustration purposes we state the folders and content that are necessary before each step.

Operators used in pattern matching	Meaning of expression
/[]/	Contain []
[]\$	Ends with []
^[]	Starts with []
&&	And
	Or
	Either
\d	Number 0-9

## Step 1: Filter the data and keep security changes and security transactions

Each time [t] the best ask [1a] or bid [1b] price changes there will be a security change [Sc] line in the OCDF. For each stock transaction there is a security transaction [t] line where we can find the traded quantity [Tq], the traded price [Tp] and the date and time of the transaction [DTd]. All our Perl scripts start with the following three lines:

```
#!/usr/local/bin/perl
use warnings;
use strict;
```

### Required before running script 1:

```
d:/inputdata/*yymmdd.txt → real time files from OSE
d:/outputstep1/ → empty folder
```

### #Script 1 discards all lines except security changes and security transactions from input files

```
my $InputFolder, $file, $OutputFile, = ( "d:/inputdata", "", "d:/outputstep1" );
chdir "$InputFolder" or die;

while (<*.txt>) {
    $file = $_;
    print "Processing $file\n";
    $outputfile = substr $file, length($file) - 12, 12;
    open( INPUTFILE, "$file" ) or die "$!";
    open( OUTPUTFILE, ">", "$OutputFolder$OutputFile" ) or die "$!";
    while (<INPUTFILE>) {

        if ( /Sc;/ && /(1a|1b)\d/ && /t\d/ ) || ( /t;/ && /Tq\d/ && /Tp\d/ && /DTd\d/ )
        {
            print OUTPUTFILE "$_";
        }
        close INPUTFILE;
        close OUTPUTFILE;
    }
}
```

## Step 2: Extract data for each security by ISIN number to separate folders

For each trading day there are files with fixed data. From them we extract one file for each day placed them in a separate folder (with equivalent names as step 1), e.g. one day if there were only 3 stocks:

```
Kongsberg Automotive Holding;iNO0003033102
Ekornes;iNO0003035305
Kongsberg Gruppen;iNO0003043309
```

To extract stock names and ISIN numbers it is only necessary to modify the first script slightly for it to work with the fixed data files (change the folders and if statement). However, the fixed data files you can also import to e.g. Excel, hence we do not state the script here. After this we made a list over stock names that we extract data for (one name at each line).

**Required before running script 2:**

d:/stocknames.txt → list with stocks that we want to extract data from.  
d:/fixeddata/yyyymmdd.txt → one file for each trading session with stock name;ISIN for each line.  
d:/outputstep1/ yyyymmdd.txt → output from Step 1  
d:/securities/ → empty folder

**#Scripts 2 makes one folder for each stock name and creates one file for each trading session with all security changes and security transactions for this stock (d:/securities/stock name/yyyymmdd.txt).**

```
my ($isin,$file,$outputfolder,$outputfile,$stock) = ("", "", "d:/securities", "", "");
open (STOCKS,"d:/stocknames.txt") or die "$!";
my @stocks = <STOCKS>;
chomp (@stocks);
close STOCKS;
foreach $stock (@stocks) {
print "Processing $Stock\n";
mkdir("$outputfolder$stock\"); #make a folder for this stock
while (<d:/outputstep1/* .txt>) { #loop for each trading day, i.e. for each yyyymmdd.txt
$file = $_;
$outputfile = substr $file, length($file)-12,12; #remove folder path, get yyyymmdd.txt
#Retrive the ISIN number for the stock this day:
open (FIXEDINPUTFILE, "d:/fixedfiles\$outputfile") or die "$!"; #Open file with stockname and isin
while (<FIXEDINPUTFILE>) {
if (/^$stock/) {
($stock,$isin) = split(';',$_); #save the stock name and isin from the semi colon separated file
last; #exit loop, since we have found stock name and isin
}
}
close FIXEDINPUTFILE;

open (INPUTFILE, "$file") or die "$!";
open (OUTPUTFILE, ">", "$outputfolder$stock\$outputfile") or die "$!";

while (<INPUTFILE>){
if (/^$isin/) {
print OUTPUTFILE "$_ ";
}
}
close INPUTFILE;
close OUTPUTFILE;
}
}
```

### Step 3: Extract transactions with time and current bid and ask quotes for each stock

#### Required before running script 3:

d:/securities/ → output from step 2

d:/r-input/ -> empty folder

#### #Script 3 makes one file with all trade for each stock

```
my $folder = "d:/securities/";
my @Stocks = ();
chdir "$folder" or die;
while (<*>) { push( @Stocks, "$_" ); } #make list of stocks from names of folders
my $outputfolder = "d:/r-input/";
my ($TradedPrice, $TradedQuantity, $DateTime) = ( 0, 0, 0);
my ($ask,$bid, $Stock,$file,$field, $extention) = ("NA","NA","","","",".txt");
my @SLine = ();

foreach $Stock (@Stocks) {
print "Processing $Stock\n";
open( OUTPUT, ">", "$outputfolder$Stock$extention" ) or die "$!";
chdir "$folder$Stock" or die;

while (<*.txt>) {
$file = $_;
open( INPUTFILE, "$file" ) or die "$!";
#New trading session, reset bid and ask:
( $ask, $bid) = ("NA", "NA");

while (<INPUTFILE>) {
if ( /Sc/ && /(1a|1b)\d/ && /\d/) { #True if line is a security change
chomp; #Remove newline characters
@SLine = split( ',', $_ );

foreach $field (@SLine) {
if ( $field =~ /^1b\d/ && /\d$/ ) { #True if best bid has changed
$field =~ s/1b//; #Remove 1b from string
$bid = $field;
}
elsif ( $field =~ /^1a\d/ && /\d$/ ) { #True if best ask has changed
$field =~ s/1a//; #Remove 1a from string
$ask = $field;
}
}
}
else { #True if line is a security transaction
chomp;
@SLine = split( ',', $_ );
#Search for relevant fields and remove the field codes:
foreach $field (@SLine) {
if ( $field =~ /^DT\d\d/ && /\d$/ ) {
$field =~ s/DT\d\d//;
$DateTime = $field;
}
elsif ( $field =~ /^Tp\d/ && /\d$/ ) {
$field =~ s/Tp//;
$TradedPrice = $field;
}
}
elsif ( $field =~ /^(Tq)\d/ ) {
```

```

$field =~ s/Tq//;
$TradedQuantity = $field;
}
}
print OUTPUT "$DateTime;$TradedPrice;$TradedQuantity;$bid;$ask\n";
}
}
close INPUTFILE;
}
close OUTPUT;
}

```

## Excluding stocks outside the continuous trading session

One might want to exclude trades outside the continuous trading session in script 3. However, logical conditions with time are not straight forward. Additionally time is given on different format for security transactions and security changes, but they all have in common that they end with hour, minutes and seconds (hhmmss). Hence, we do this by calculating time in seconds after midnight from any string that ends with hhmmss. The time sub routine can be included in any script. The following is an example of how one can exclude all trades outside the continuous trading session in script 3:

Define the opening and closing of the continuous and trading session and a variable for time in the beginning of the script.

```

my $StartTradingSession = TimeToSecAfterMidNight("090500");
my $EndTradingSession = TimeToSecAfterMidNight("172000");
my $TradeTime = 0;

```

Add the following two line before the “print OUTPUT” statement and add a bracket after (}):

```

$TradeTimeSec = TimeToSecAfterMidNight ($DateTime);
if ($TradeTimeSec < $StartTradingSession || $TradeTimeSec > $EndTradingSession) {

```

Add the Sub routine at the end of the script:

```

sub TimeToSecAfterMidNight {
#Sub routine converts time (*hhmmss) and returns seconds from midnight
my $HHMMSS = shift;
my ( $hh, $mm, $ss ) = ( 0, 0, 0 );
$HHMMSS =~ s/(\D)+//g;
$HHMMSS = substr $HHMMSS, length($HHMMSS) - 6, 6;
$hh = substr $HHMMSS, 0, 2;
$mm = substr $HHMMSS, 2, 2;
$ss = substr $HHMMSS, 4, 2;
$hh =~ s/^0//;
$mm =~ s/^0//;
$ss =~ s/^0//;
$HHMMSS = $hh * 3600 + $mm * 60 + $ss;
return ($HHMMSS); }

```