ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
MSc Economics & Business
Master Specialisation Financial Economics

The price effects of concentrated trade volumes on the NYSE

Author:	D.T. Eggermont
Student number:	357654
Thesis supervisor:	Dr. J.J.G. Lemmen
Finish date:	09-2016

NON-PLAGIARISM STATEMENT
By submitting this thesis the author declares to have written this thesis completely by himself/herself, and not to have used sources or resources other than the ones mentioned. All sources used, quotes and citations that were literally taken from publications, or that were in close accordance with the meaning of those publications, are indicated as such.

ABSTRACT

In this thesis, it is researched what effects concentration in trading volume have on (future) returns on the NYSE. It is expected that a concentrated buying (selling) volume would cause positive (negative) returns at the moment of the trade and a short interval thereafter, but negative (positive) returns at a longer interval due to the recovery of the price to its fundamental value. For the daily tests this pattern has not been observed. For the intraday tests the recovery was observed, but it happens within one trade after the block trade, making it impossible for traders to profitably trade on information about trade concentration.

Keywords:
Block trade, Volume, NYSE, Stock market, Concentration

JEL Classification:
G12, G14, G15

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1 Introduction

When assets are traded on some form of exchange, many factors influence the prices at which these trades are performed. Some examples of these factors are market transparency, transaction costs, market liquidity, market intermediaries and trading volume. This research focusses on the latter and to be more specifically on trade concentration and its effects on (future) return. The difference between trade concentration and trading volume is that with trade concentration the trading volume relative to the amount of trades is measured and thus the absolute height of trading volume does not matter.

The relationship between trade concentration and (future) return is researched in two ways. Firstly, trade concentration is measured on a daily basis. This is done by dividing the total dollar volume traded on a day by the number of transactions during that day. A distinction is made between buy and sell initiated trades. After that, the relationship can be measured between the trade concentration on a day and the return over that day or later days. Secondly, intraday trade concentration is studied. This is done in a similar way as in existing literature, by detecting block trades and measuring the impact they have on prices at the moment the block is traded and certain intervals thereafter. A block is defined as a transaction in which more than 10,000 shares are traded or the dollar volume is larger than 200,000.- USD.

This testing of both daily and intraday data for the effects of trade concentration, has not been used before. Prior research focused only on trade concentration on an intraday level. These studies generally found that price effects caused by trade concentration exist, but that they are arbitraged away before they can be traded on. This makes it interesting to see if it is possible to detect price movements caused by trade concentration on a daily level and see if they provide possibilities to trade profitably on them. Additionally, it is also interesting to test the sample for intraday price effects caused by trade concentration as has been done in previous research, to see if the same results are found for this sample.

The first hypothesis that is tested is that a concentrated buying (selling) volume (block trade) indicates a price increase (decrease) in the very near future due to misinterpretation of other traders. The reasoning behind this is that a large buy (sell) initiated trade, causes the price to go up (down) due to market liquidity effects. This price increase is observed by irrational traders. They respond to this price change by following the momentum. The second hypothesis is that a concentrated buying (selling) volume (block trade) indicates a price decrease (increase) in the near future due to overreaction. This hypothesis follows from the fact that if the price pattern described in the first hypothesis is observed, rational traders will realise that the price is above the fundamental value. They will respond to this by trading in the opposite direction which turns the price pattern around. Irrational investors see this new price direction and again follow the momentum.

The following steps will be undertaken to discuss and test these two hypotheses. In the next chapter an overview of the existing literature on both trading volume and trade concentration will be given. Next, the data and methodology used for the empirical tests of the two hypotheses will be discussed. This section is divided in the data and methodology for the daily trade concentration tests and for the intraday trade concentration tests. In chapter 4, the results from these tests are presented and discussed, which will again be divided between daily and intraday trade concentration. Finally, a conclusion and further research recommendations will be given in chapter 5.

CHAPTER 2: Literature Review

In this chapter an overview of the existing literature about trading volume will be discussed and its main findings will be presented. In the first section, there is no focus on concentration in trading volume yet. Instead it will present an overview and discussion about the most important papers that focus on trading volume effects in general. In the second section, literature that focusses on concentration in trading volume is presented.

2.1 The power of trading volume

A lot of research has tried to examine the effects of trading volume on stock prices. Various results have been found. This means that it might be possible to devise a trading strategy to take advantage of this but that it is certainly worthwhile to be aware of the effects trading volume has in the market. These effects found in prior trading volume studies can broadly be divided in three main categories: market liquidity, momentum and information.

The relationship between trading volume and market liquidity of a stock seems logical. Market liquidity is defined as the possibility to trade an asset in high quantity, at low cost and with little impact on the price of the asset in the market (Brown, Crocker, & Foerster, 2009). An active and transparent market is required to achieve this. Trading volume appears to be a good indicator of this and therefore might serve as a proxy for the market liquidity of an asset. This proxy might be useful when trading assets since there is the possibility that there are additional costs when trading in an illiquid market. This could be a risk investors should consider. That liquidity matters to investors was first documented by Amihud and Mendelson (1986). They studied what effect differences in bid-ask spreads have on asset prices. The hypothesis they use is that market-observed expected return is an increasing and concave function of the bid-ask spread. In others words, if the bid-ask spread of an asset goes up, they expect the expected return of that asset also to rise. They say that trading volume has a negative relation with the bid-ask spread and that trading volume is a proxy for market liquidity. However the bid-ask spread is not only dependant on market liquidity. Broker fees for example also play a role. Amihud and Mendelson find that assets traded on the NYSE between 1961 and 1980 with a higher bid-ask spread, earn higher risk-adjusted returns. The reasoning behind this positive relationship between the bid-ask spread (and thus liquidity) and expected return is, according to Amihud and Mendelson, that investors require a higher return to compensate for the extra costs they face when buying or selling an illiquid asset.

Chordia, Subrahmanyam and Anshuman (2001) use this as a starting point. They try to find out how the market deals with (il)liquidity, but use trading volume as a direct proxy for market liquidity by using dollar trading volume and share turnover (number of shares traded divided by number of shares outstanding). Additionally they also look at the relationship between fluctuations in trading volume and expected returns. The reasoning behind this is that if the expected return of an asset is dependent on the height of its trading volume, highly volatile trading volume could be a risk when holding that asset. They find a significant negative relationship between trading activity and asset return. Thus, the theory that investors require a compensation for the risk of holding an illiquid stock has been confirmed again. Furthermore, they find a surprising negative relationship between volatility in trading activity and return. This means that this is not perceived as a risk by the market. They cannot give a clear explanation for this based on their research. However, they expect that it is caused by investors perceiving high volatility in trading volume as a sign of extra market liquidity since it possibly means new investors can easily enter and exit the market.

Datar, Naik and Radcliffe (1998) performed an alternative test of the one Amihud and Mendelson did. Instead of using the bid-ask spread as a proxy for market liquidity, they used like Chordia et al. (2001) asset turnover as a market liquidity indicator. The hypothesis they use is that stock returns have a negative relationship with market liquidity as the Amihud and Mendelson research found. After controlling for the Fama and French factors they find strong evidence that their hypothesis should not be rejected and this negative relationship indeed exists in their data.

The level of trading volume for a stock is not only related to market liquidity. It can also be used to detect momentum in stock prices. This relationship was first discovered by DeLong, Shleifer, Summers and Waldmann (1990). They study whether there exists momentum in the financial markets that causes asset prices to significantly diverge from their fundamental value. In order for this to be true, noise traders must have significant price impact and this effect should not be neutralised by rational arbitrageurs. They found that there is indeed a momentum effect because arbitrageurs are often not willing to correct prices to their fundamental value. This is mainly caused by various forms of risk that the arbitrageur needs to bear in order to do this. They also found that due to the noise traders entering the market for a stock and driving its price up, the momentum effect is linked to trading volume.

Later this was proven by Lee and Swaminathan (2000) who focussed on this relationship. The starting point for their research was a paper of Jegadeesh and Titman (1993) in which they found that stocks that performed well in the recent past continue to do so for the next 12 months, after which a reversal takes place. Lee and Swaminathan tried to link this finding to trading volume and see if this follows the same movement. In other words, they try to see whether glamour stocks (stocks that currently have high returns) have high trading volumes and value stocks (stocks that currently have low returns) have low trading volumes and whether trading volumes also go down when glamour stocks seize to be glamour stocks and become value stocks. They find that this is indeed true and that past trading volume can predict the magnitude and persistence of price momentum.

This means however that the negative relationship between trading volume and expected return found in the liquidity studies is now a positive relationship in the short run (1 to 12 months). Brown, Crocker and Foerster (2009) have tried to explain this difference. They focus on portfolios of larger stocks which are usually more liquid and compare the results to portfolios consisting of smaller, more illiquid stocks. For these portfolios they establish a relationship between trading volume and expected return and try to determine whether this was caused by (il)liquidity, momentum or information (discussed later in this chapter) effects. They find a significant difference between the two types of portfolios. For the portfolios consisting of small stocks they find the relationship between expected return and trading volume to be negative which is in line with the liquidity hypothesis. For the large and liquid portfolios however, they find a positive relationship on the short run (1 to 12 months). This confirms the momentum and information hypothesis.

More recently a third important relationship has been found. Namely, that between the level of trading volume in a stock and the way a stock responds to information. This has been discovered by Chordia and Swaminathan (2000). The hypothesis they use is that stocks with low trading volumes are slower to respond to a new piece of market wide information than stocks with larger trading volumes. Furthermore, they try to find out if returns on stocks with low trading volume follow the returns of stocks with high trading volume with some delay. They find that both hypotheses cannot be rejected and that it appears that trading volume is related to the speed of adjustment to news. This result should obviously be of great importance to investors. They could use it to accurately predict the return patterns of stocks with low trading volume. However, this would also mean that the effect would be arbitraged away quickly but this is not the case. Chordia and Swaminathan looked into this as well. They found that due to the added costs of trading the low volume stocks, strategies based on this relationship become quickly unprofitable. However, the results are still interesting. It indicates the importance of trading volume and shows that it can be used to predict the speed of adjustment to new information.

Gervais, Kaniel and Mingelgrin (2001) research what the impact of short term changes in trading volume on subsequent returns is. They test the hypothesis that a short term increase (decrease) in trading volume is being followed by a positive (negative) return. The reasoning behind this is that this increase (decrease) makes the stock more (less) visible to investors and thus a higher (lower) return is required. They find that this is indeed the case. Stocks that had an unusual increase (decrease) in trading volume over the last day or week earn a high (low) return over the following month. This effect appears to be the same as the momentum effect. However, in this case this holds for winners and for losers. In other words, both losers and winners earn a positive return after an increase in trading volume. The effect is even larger for losers.

This overview demonstrates that there is definitely valuable information that can be deducted from the height of trading volume. Trading volume is related to market liquidity, momentum and information. All of this is in violation with the weak form of the efficient market hypothesis (Fama, 1970). The research discussed thus far has all focussed on standard trading volume. Research on concentration in trading volume will be discussed in the next section.

2.2 Concentration in trading volume

The previously discussed research focussed on trading volume in general. This research focusses specifically on the information that can be gained from looking at concentrations in trading volume. The goal is to find out whether one investor buying (or selling) a large portion of the total stocks of a company, or many investors buying (or selling) a lot of small portions, has a different impact on returns.

Prior research that is closest to this, is research on the effects of block trades. A block trade is usually defined as a single trade in which at least 10,000 shares are traded or that has a value greater than 200,000.- USD (Phadnis, 2015). On most exchanges block trades have to be registered. A block trade could be considered as a trade that heightens the concentration in trading volume. The first research that looked into the relationship between block trades and the return path after such a trade, has been performed by Grier and Albin (1973). They investigate price movements that occur at the moment a block is traded and in the period shortly before and after a block is traded. The hypothesis of their paper is that the price behaviour around the trade of a block is significantly different from the price behaviour associated with normal trades. To study this hypothesis they start by looking at the ratio of reversals to continuations. This is a measure to detect if the market is random. A reversal is defined as a trade that has a different sign (negative/positive) than the sign of the trade prior to it and a continuation is the opposite. In a random market a ratio of one would be expected. Grier and Albin find for the sample of block trades they use, a ratio of reversals to continuation of 6.0, which is significantly different from that of a random market. However, this has no economic significance yet, since this assumes it would be possible to trade directly after the moment the block trade takes place. In reality, this is impossible. Since 15 minutes is generally accepted as a realistic period for the completion of a trade, they recalculate the reversal to continuation ratio using this period. This time they find a ratio of 5.5 which is not significantly different from the first ratio but is significantly different from the ratio of a random market (1.0). The next step of Grier and Albin is to see if they can come up with a trading rule based on this finding, that can yield excessive profits after the subtraction of transaction costs. The idea behind this trading rule is that the price movement prior to the trade of a block is reversed in the period after the block trade. Their rule therefore states that a stock should be bought right after the sale of a block for those stocks that had the largest price decline prior to the block trade. This strategy proved to yield significantly higher profits than a benchmark portfolio. A possible explanation for this could be that a block sale provides the stock with extra publicity which drives the price up in the short run. It could also be caused by extra information specialists have. They are usually involved in a block trade and often are better informed than the average market participant. It could therefore be that the market reacts in a positive way to a specialist buying (part of) the block. Grier and Albin also look at differences in volatility between block trades and a combination of ordinary trades that when combined have the same value. They find that a block trade causes the price to be significantly less volatile than a combination of ordinary trades. This indicates that the market reacts differently to the trading volume that is traded in one block than to the same trading volume in multiple smaller transactions. This goes against what the efficient market hypothesis predicts. This finding that the effect a block transaction has on volatility is smaller than that of multiple smaller transactions that trade the same volume, might be explained by the study of Avramoc, Chordia and Goyal (2006). Part of their study focuses on the distinction between the effects trades performed by informed and uninformed traders have on volatility. Informed traders are defined by them as traders that recently gained additional insights which are not priced by the market yet and uninformed traders, are traders that trade for liquidity purposes. This means they perform sell transactions to gain liquidity and buy transactions to employ excess liquidity. They find that trades performed by informed traders have less impact on future volatility (and returns) than those performed by uninformed traders. The reasoning behind this is that a trade performed by an informed trader moves prices closer to their fundamental value and therefore no reverse price movement is expected afterwards. A trade performed by an uninformed trader is likely to move prices away from their fundamental value and is expected to be followed by a reversed price movement afterwards. When this finding is combined with the theory of Easly and O'Hara (1987) that block trades are usually trades based on information, it can explain the finding of Grier and Albin that block trades have relatively less impact on volatility than regular trades. In that case a block trade would drive prices closer to their fundamental value and is thus less likely to be followed by a subsequent price reversal.

Kraus and Stoll (1972) focus on the price effects of block trades and wonder what causes this. They try to examine in what way and why block trading is causing markets to be inefficient. In order to do this, a block trade must cause a price change that cannot be ascribed to new information coming onto the market. This is because an efficient market is defined as a market where many small buyers and sellers with equal access to information cause prices of securities to only change to new information and thus prices always reflect the fundamental value of the asset. It is possible however that a market is not efficient due to too few participants for example, or investors with different investor preferences. In that case, it is possible that trading on that market causes (temporary) price changes. Kraus and Stoll find that a block trade on average has this effect on prices. If a block is sold this happens usually at a price lower than the current market price. This causes the market price of the asset to go down. After the block trade is executed the market price partially recovers during the rest of the day but it does not fully reach the old equilibrium price it had prior to the block trade. The reversed happens when a block is bought but in a less severe way. In that case, there is also no significant price reversal. This could indicate that for blocks that are bought, this price increase is based on new information because these trades are usually performed by institutional investors that have the ability and resources to do additional research and thus trade on extra information only they have. This theory is the focus of the research of Easly and O'Hara (1987). Their hypothesis is that a trader with extra information about a security price, usually wants to exploit this as much as possible. Therefore, he prefers to make a large trade so he profits from his information before the market catches up on this. This correlation between information and trade size is also known by market makers. They can profit from this by quoting a higher price for larger trades because they know the trader is willing to pay more than the average investor, due to the extra information only he has. This provides evidence for the theory that the effect trade size has on security prices, is caused by information asymmetry. However, Kraus and Stoll also found evidence that for blocks that are sold the price effect can probably be ascribed to market inefficiency. This price effect is composed of two parts. The first is liquidity. This is caused by a shortage of investors that are willing to buy (part of) the block at the current price. This causes the price to go down until the entire block has been sold. The second is trading costs because these are typically higher in the case of a block trade. If they are incorporated in a bid-ask spread then this spread rises, causing prices to lower. Kraus and Stoll conclude by saying that their findings indicate two practical implications. Firstly, it shows that block trades indeed have an influence on prices, although usually this is just a temporary effect. Secondly it indicates that (at the time of their research at least) the way a block is sold might not be the most efficient way since the seller of a block pays higher trading costs than usual.

According to Attig, Fong, Gadou and Lang (2006), blocks usually also trade against higher bid-ask spreads. They argue that the height of the bid-ask spread is positively related to the concentration of ownership of a company. Thus, if there are a few parties that own a large amount of the total number of outstanding stocks, the bid-ask spread would be higher. This is caused by information asymmetry between the holders of the large blocks and the holders of a small part of the total outstanding stock quantity. Since it is more likely that for stocks in which there are relatively many investors that hold a big block a block transaction is made, it is more likely that a block trades at a higher bid-ask spread.

High trading costs might also be a reason for the continuation of the violation of the efficient market hypothesis. This is one of the topics discussed in the paper of Dann, Mayers and Raab (1977). They reassessed the trading rule devised by Grier and Albin which has been discussed earlier. They disagree on the methodology used by them and think that this has had a big effect on the results they found. The main criticism Dann et al. have on Grier and Albin their research is that they have not incorporated trading costs in the right way. Grier and Albin used proportional estimates for their transaction costs. Dann et al., retest the trading rule but instead they use actual transaction costs. They find that this causes the profits of the trading rule to lower. They continue to outperform the benchmark portfolio, which is a capitalisation weighted portfolio with all the S&P 500 stocks in it. But this time this outperformance is not high enough to be economically significant for non-members of the NYSE. This is evidence that is consistent with the efficient market hypothesis in its weak form. Dann et al. do note however that the results they found only hold for investors that are not a member of the NYSE. Members of the NYSE often have substantially lower trading costs which could make returns earned by following the trading rule on block trades significantly higher than those earned on the benchmark portfolio. This would implicate that the efficient market hypothesis in the strong form should be rejected. They cannot test this however, because they do not have data on the height of the trading costs for NYSE members. An additional reason why Dann et al. find the trading rule to not outperform the benchmark portfolio is that in their research they find that prices adjust back to the equilibrium price level within 15 minutes after a block trade. Since this is often considered as the response time an investor needs to trade on a block, it makes it very hard to trade on information about blocks according to them.

This shows that the time for prices to readjust after a block trade, plays an important role in the ability to earn a profit by trading on information about block trades. Holthausen, Leftwich and Mayers (1990) performed additional research on this. They investigated how long it takes for prices to move back to their equilibrium level after a block. Additionally, they tried to distinguish if price effects are temporary or permanent. A temporary price effect is the adjustment of the price right after the block transaction to the new equilibrium price. A permanent price effect is the price change from the equilibrium price before the block trade to the new equilibrium price after the block trade. They find that prices after a block trade recover quickly to an equilibrium price and thus the temporary price effect only lasts a short time. For seller-initiated block trades this recovery happens for the biggest part in one trade and is fully completed within three trades after the block is sold. This is in line with most previous research. There is however a substantial permanent price effect after a block is sold. This is something that previous research has not covered extensively yet. For buyer-initiated block trades the recovery to the equilibrium price and thus the temporary price effect is even shorter. In that case it takes only one trade for the price to get to its new equilibrium level. However, there is again a permanent price effect. For both seller- and buyer-initiated block transactions this permanent price effect increases when the trading volume of the block is bigger.

This shows that in order to successfully trade on block trade information it should be done as quickly as possible after the execution of the block. Since block trades have to be registered on most exchanges, information about an executed block trade is almost immediately available to all investors. The speed at which this information is communicated could therefore potentially have a large impact on the speed at which prices recover after a block trade. Gemmil (1996) focussed his research on this. He uses the fact that the London Stock Exchange has had three different regimes for communicating executed block trades over three different periods. From 1987 to 1988 block trades were published immediately. From 1991 to 1992 they were communicated after 90 minutes and from 1989 to 1990 they were communicated after 24 hours. If it is necessary that a block trade is communicated for prices to recover, there should be a significant difference in adjustment speed between the three regimes. This would give insiders that work on the block trade an advantage since they own private information that is valuable since they can predict which way the market is going to move when the information becomes publicly available. If on the other hand there is no significant difference and prices start to recover quickly after the block trade, this would be in line with the efficient market hypothesis in its strong form and information would be incorporated in the prices before becoming public. Gemmil finds that there is not a significant difference in the speed at which prices adjust to a new level after a block trade between the three regimes. Under all three of the regimes prices adjust back to their equilibrium level in a rapid way. This is a surprising result because this means that insiders on the trade have no informational advantage because information leaks out at such a high speed that they would not have enough time to trade on this information. Thus, the efficient market hypothesis in its strong form should, on basis of this study, not be rejected for the London Stock Exchange. The focus of this study on the London Stock Exchange yields some additional interesting insights since it is the first that studies a different market than the US stock markets. They find that prices adjust in a similar way and speed as they do on US stock markets, namely within three to five trades after the block trade. However, they do also find that prices start to anticipate an impending block trade two to three trades before it is executed. This implies that information about a block trade leaks before it is executed. They also find contrary to studies on the US stock exchange, only a temporary price effect and no (significant) permanent price effect.

A similar research to the study of Gemmil (1996) is that of De Jong and Van Beusichem (2006). They also studied a different stock market than the US stock markets, namely the Dutch stock market. On the Dutch stock market, block trades also have to be reported. However they are again not published immediately. This enables them to look at price effects on the trading date and on the publication date and compare these effects. They find that the price effects are the largest on the trading date. On the publication date there are only small and insignificant price effects. This means that information about the block trade leaks out again and prices go back to their equilibrium level before information about the block trade has been officially published.

The studies discussed above, all focus on trading on an exchange, also called the downstairs market. The study of Keim and Madhavan (1996) focuses on the upstairs market. In an upstairs market, trades are performed with the use of an intermediary or broker that matches two or more parties before executing the trade. Data is less easily available for the upstairs market than for the downstairs market. Therefore, it is not possible to estimate the amount of blocks traded on the upstairs market. However, it is likely that a large part of blocks traded, are traded on the upstairs market due to their characteristics. They are usually traded by institutional investors which are more likely to use the upstairs market than a non-institutional investor. Furthermore, due to their size it is often not possible to trade a block as a whole on the downstairs market. In that case the upstairs market could be used to have the intermediary find (multiple) counterparties that are willing to trade the block. Keim and Madhavan find that because of this, information about block trades leaks out on the upstairs market prior to the trade. They find that prices anticipate the trade of a block and that a large part of the price effects found in studies that focus on the downstairs market, are apparent on the upstairs market but are observed prior to the block transaction. They reason that this is caused by the search of the intermediaries for counterparties willing to trade the block. They find this effect for up to 4 weeks prior to the block transaction.

This literature overview shows that a block trade, which can act as a proxy for trade volume concentration, carries interesting information and that prices do not just rise (fall) for a short period of time when a block is bought (sold) due to liquidity effects after which they go back to their original equilibrium level. It is not entirely clear whether it is possible to profitably trade on block trade information. This is highly dependent on the level of transaction costs a trader incurs and on the exact speed at which prices recover after the block trade and the time it takes the trader to execute a transaction. All research finds that prices recover fairly rapidly. It takes at most a day for them to recover and usually happens within five trades. There is a possibility that there are permanent price effects on which could be traded. Later research finds them to be more significant than prior research thought. However they have not been found for the London Stock Exchange.

Chapter 3: Data and Methodology

In the first section of this chapter an overview of the source and composition of the used data in this thesis will be presented. After that, the methodology used for this research will be explained. The methods used to measure trade volume concentration and the used statistical tests will be discussed.

3.1 Data

To be able to measure trading volume concentration and fluctuations in it, intraday data is required. To be more specific, data on trade level is needed. This limits the possibilities to US exchanges. For this research the choice has been made to use NYSE stocks. The NYSE records and publishes data on multiple variables for every trade that is performed on their exchange. To access this data, the Trade and Quote (TAQ) database is being used.[1] Before data can be gathered, a sample of stocks needs to be selected that will be used to test the hypotheses used in this thesis. To do this, a list of all stocks traded on the NYSE in 2013 is downloaded from Bloomberg. From this list 100 stocks are randomly selected. This list is then used to gather data from the TAQ database. The downloaded data has the following form. For each trade it lists the Ticker symbol, the time and date at which the trade took place, the price at which the trade was performed and the volume traded.

After collecting the intraday data some alterations are made. Firstly, the data is sorted on ticker symbol. This makes it possible to look at intraday effects per stock. Next, dollar volume is calculated for every trade. After that, block trades are identified. The criteria used in previous literature and the NYSE itself, are being used. This means a trade is defined as a block trade if more than 10,000 shares are traded or the total value of the trade is more than 200,000.- USD (Phadnis, 2015). This gives a total of 179,087 blocks for the selected sample in 2013 and an average of 1,809 blocks per stock. Lastly, for every block trade it is determined whether it was a buy or a sell initiated trade, since they are expected to have opposite effects on return. According to previous literature the best way to do this is by comparing the price at which the block is traded to the price of the previous trade. If the block trades at a higher price (uptick), it can be assumed that this is a buy initiated trade (the increase in demand raises the price). If the block trades at a lower price (downtick), it is likely a sell initiated trade (the increase in supply lowers the price). The last possibility is that the block price is

[1] TAQ presents their data in two ways. The first is a live feed of trading activity on the NYSE and NASDAQ, which is called a ticker tape. With this, trading activity can be monitored. The second form is a database that contains data on all trades performed on the NYSE and NASDAQ over the period 1993-2016. The latter is used for this research.

equal to the price of the previous trade. In that case it is impossible to determine whether the trade is buy or sell initiated and will not be taken into account for the remainder of the research. The uptick, downtick method has proven to provide the best approximation to determine whether a trade is buy or sell initiated (Lee & Ready, 1991).

These alterations make it possible to compute daily data out of the intraday data. To do this, the intraday data needs to be summarized in a few daily variables. This is done for every ticker on every trading day. This yields for every variable a table with on the x-axis the individual tickers and on the y-axis the individual trading days (252 days). The first variable that is computed daily, is the dollar volume. This is the sum of the total dollar volume traded for every ticker on each day. Next, the number of trades, number of blocks, number of buy intitiated trades and number of sell initiated trades are calculated. This is done by simply counting them for every ticker on each day. Finally, closing prices need to be gathered for all tickers on each day. For this CRSP[2] is being used.

[2] CRSP stands for Center for Research in Security Prices. It is a provider of historical stock market data for the NYSE, AMEX and NASDAQ over the period 1926-2016.

3.2 Methodology

After gathering the data, it needs to be structured to be able to test the hypotheses. The testing can be divided into two parts. The first part is testing on daily data and the second part is testing on intraday data. The data has to be prepared differently for both categories. Below an overview will be given of the steps taken to prepare the data and the methods used for testing, divided into the two categories.

3.2.1 Daily trade concentration

The goal of doing tests on daily trade concentration, is to evaluate whether days that have a high average trade concentration, have different subsequent returns than days that have a low average concentration in trading volume.

3.2.1.1 Data preparation

To do this, the intraday data needs to be structured further. First, for every day and for every ticker, the total trading volume and total number of trades per day needs to be calculated. There is not yet literature that looks at trade concentration on a daily frequency. Therefore, a formula needs to be formulated that can be used to calculate a variable that measures daily trade concentration. This is done with the following formula:

$$Trade\ concentration_{i,t} = \frac{\sum_{n=1}^{N}(Dollar\ volume_{i,t,n} * Sign_{i,t,n})}{number\ of\ trades_{i,t}} \tag{1}$$

Where *i* indicates the ticker, and *t* the day. *Sign* gets a value of 1 if the trade was buy initiated and a value of -1 if the trade was sell initiated. Because the effects of trade concentration on return are expected to be opposed to each other, this is necessary to measure this. Thus, a day that has a positive and high trade concentration is expected to have a high return over that day but a day that has a negative and high trade concentration is expected to have a low or even negative return over that day.

The next step is to calculate daily returns for every ticker by using the gathered closing prices. To calculate the returns the following formula is used:

$$Return_{i,t} = \frac{(price_{i,t0} - price_{i,t-1})}{price_{i,t-1}} \tag{2}$$

Weighting

The previous steps produce for every individual ticker daily data on the return, trade concentration and trading volume. This enables the possibility of testing for every ticker individually whether trade concentration plays a significant role on return, but that is not the goal. The goal is to create a portfolio of the stocks in the sample and to test whether the trade concentration of the portfolio has a significant role on return. To do this, the individual data has to be aggregated. Therefore, a weighted average portfolio is constructed. This is done by weighting the return, trade concentration and trading volume by the market capitalization of every individual stock[3]. The following formulas are used to do this:

$$Return_{P,t} = \frac{\sum(return_{i,t} * market\ cap_{i,t})}{\sum market\ cap_{i,t}} \tag{3}$$

$$Trade\ concentration_{P,t} = \frac{\sum(trade\ concentration_{i,t} * market\ cap_{i,t})}{\sum market\ cap_{i,t}} \tag{4}$$

$$Trading\ volume_{P,t} = \frac{\sum(trading\ volume * marke\ \ cap_{i,t})}{\sum market\ cap_{i,t}} \tag{5}$$

In these formulas, the return, trade concentration and trading volume for every stock *i* in the sample at time *t* are weighted against the market capitalization. This produces a single average portfolio return, trade concentration, and trading volume for every trading day in the sample (252). The fact that it has to be assumed that market capitalization stays constant throughout the month, creates noise in the data because this is obviously not the case in reality. However, it should not influence the results dramatically since it has the same effect on the dependant variable (return) as on the independent variables (trade concentration and trading volume).

3.2.1.2 Regression testing

To test the effects of average trade concentration on the return over that day, regression tests are used. Firstly, the effects of average trade concentration on the return over that day are measured. The following formula is used to do this:

[3] CRSP is being used to collect monthly market capitalization data for every stock. It is being assumed that the market capitalization stays constant throughout the month. For example, all the return data of stock *i* over January is weighted against the market capitalization of stock *i* on the first of January.

$$Return_{P,t} = \alpha + \beta_1 * trade\ concentration_{P,t} + \beta_2 * retu\quad _{p,t-1} + \varepsilon_t \qquad (6)$$

In this formula, trade concentration on day *t* is regressed on the return over day *t*. The coefficient β_1 measures the relationship between the average trade concentration of the value weighted portfolio on day *t* and the return of the value weighted portfolio over day *t*. Because it might be possible that the level of trade concentration is influenced by the return and not the other way around, the return over day *t-1* is also added as an independent variable to this regression and all other regression formulas (except formula 9). By doing so, it becomes possible to determine whether the other independent variable(s) cause the level of the dependant variable and not the other way around. Thus, a significant positive value for β_1 means that a high average trade concentration causes a high return over that day.

After measuring the impact of trade concentration on return, the relationship between trading volume on return is measured. For that the following formula is used:

$$Return_{P,t} = \alpha + \beta_1 * trading\ volume_{P,t} + \beta_2 retu\quad _{p,t-1} + \varepsilon_t \qquad (7)$$

This regression is used to determine what the effects of trading volume are on return. This should be taken into account because it is possible that part of the relationship found in formula 6 is actually caused by differences in trading volume between the days. If the trade concentration on a certain day is very high, it means that there have been transactions during that day where high volumes were traded. These transactions could cause a high trading volume on that day. If the coefficients of trade concentration in formula 6 and of trading volume in formula 7 have the same sign, it is likely that they measure partly the same effect. To test whether there are indeed interaction effects between trade concentration and trading volume, the following regression is used:

$$Return_{P,t} = \alpha + \beta_1 * trade\ concentration_{P,t} + \beta_2 * trading\ volume_{P,t} + \beta_3 return_{p,t-1} + \varepsilon_t \qquad (8)$$

In this formula, trade concentration on day *t* is regressed on the return over day *t*. Total trading volume on day *t* is added as control variable. This measures what part of the return can be explained by the height of trading volume. This prevents that the height of trading volume has an impact on the estimated coefficient of trade concentration and thus makes sure that this coefficient only measures the effects of trade concentration on return.

Timing

In formulas 6, 7 and 8, trade concentration and/or trading volume on day *t* are regressed on the return over the same day *t*. It would also be interesting to see whether the effects of a high trade concentration persist also on later days, or possibly have a delayed effect. It has to be taken into account that days with a high trade concentration might be followed by more days with a high trade concentration. It therefore should first be tested whether this is the case. This can be established using the following formula:

$$Trade\ concentration_{P,t} = \alpha + \beta_1 * trade\ concentration_{P,t-1} + \varepsilon_t \tag{9}$$

A significant positive value for β_1 indicates that days with a high trade concentration are likely to be followed by more days with a high trade concentration and that days with a low trade concentration are followed by more days with a low trade concentration. This means that it is not random whether a day has high or low trade concentration.

Even if days with a high trade concentration are distributed randomly across the days, it could still be possible that a day with a high trade concentration is followed by another day with a high trade concentration. Therefore, formulas 6, 7 and 8 are adapted in the following way to measure the effects of trade concentration on the return over later days:

$$Return_{P,t+x} = \alpha + \beta_1 * trade\ concentration_{P,t} + \beta_2 * trade\ concentration_{P,t+x} + \beta_3 + return_{p,t-1} + \varepsilon_t \tag{10}$$

$$Return_{P,t+x} = \alpha + \beta_1 * trading\ volume_{P,t} + \beta_2 * trading\ volume_{P,t+x} + \beta_3 * return_{p,t-1} + \varepsilon \tag{11}$$

$$Return_{P,t+x} = \alpha + \beta_1 * trade\ concentration_{P,t} + \beta_2 * trade\ concentration_{P,t+x} + \beta_3 * trading\ volume_{P,t} + \beta_4 * trading\ volume_{P,t+x} + \beta_5 * return_{p,t-1} + \varepsilon_t \tag{12}$$

With *x* being the number of days between the day of the return and the day of the trade concentration. The trade concentration and/or trading volume on day *t+x* has to be added to the regression, because this captures the effects a high trade concentration/trading volume has on the return on day *t+x*. By doing so, the coefficient β_1 measures solely the effects of the trade concentration of day *t* on the return over day *t+x*. If the coefficient is for example positive and significant, it indicates that the effect of a high trade concentration on day *t* still persists *x* days after day *t*.

Lastly, a regression is performed that tests whether the results that are found for trade concentration and dollar volume, are robust. To do this, an interaction term is added. This term measures how dollar volume and trade concentration are related to each other and whether the effect one of them has on return, is dependent on the height of the other. This is done with the following formula:

$$Return_{P,t+x} = \alpha + \beta_1 * trade\ concentration_{P,t} + \beta_2 * trade\ concentration_{P,t+x} + \beta_3 * trading\ volume_{P,t} + \beta_4 * trading\ volume_{P,t+x} + \beta_5 * trade\ concentra\quad_{p,t} * trading\ volume_{P,t} + \beta_6 * return_{p,t-1} + \varepsilon_t \qquad (13)$$

Regression validity

For the regressions stated above to produce valid results, they must adhere to four assumptions (Brooks, 2008):

1. The average value of the errors is zero: $E(\varepsilon_t) = 0$
2. The errors are homoscedastic: $var(\varepsilon_t) = \sigma^2 < \infty$
3. There is no autocorrelation in the errors: $cov(\varepsilon_t, \varepsilon_{t-1}) = 0$
4. The errors are normally distributed: $\varepsilon_t \sim N(0, \sigma^2)$

The tests that are used to test these four assumptions will now be discussed.

Assumption 1: $E(\varepsilon_t) = 0$

If this assumption is violated, it indicates that there is a set of coefficients that fits the data better than the coefficients estimated in the regression. This is only possible if there is no constant term in the regression. Since all the regression formulas stated above contain a constant term, this will not be an issue. Therefore, this assumption does not need to be tested.

Assumption 2: $var(\varepsilon_t) = \sigma^2 < \infty$

If this assumption is violated, the errors are heteroscedastic instead of homoscedastic. This means that the variance of the errors is not constant over time. To test whether this is the case, White's test is used. The null hypothesis of this test is that the errors are homoscedastic. Thus, if the test produces significant results and the null hypothesis gets rejected, it indicates that the errors are heteroscedastic.

Assumption 3: $cov(\varepsilon_t, \varepsilon_{t-1}) = 0$

This assumption states that the errors over time are uncorrelated with each other. If they are correlated, it is said that there is autocorrelation in the errors. The first step that is undertaken to detect autocorrelation, is to plot ε_t against ε_{t-1} and see whether there is an observable pattern. If this is the case the errors are probably auto correlated. After that the Breusch-Godfrey test is used to formally test for autocorrelation. The null hypothesis states that the errors at time *t* and *t-1* are independent. A violation of the null hypothesis indicates that they are correlated over time and that there is autocorrelation in the errors.

Assumption 4: $\varepsilon_t \sim N(0, \sigma^2)$

To test whether the errors are normally distributed, the Jarque-Bera test is used. The null hypothesis of this test is that the error terms are normally distributed. Thus, if the null hypothesis needs to be rejected, assumption 4 is violated and the test results might not be reliable.

3.2.2 Intraday trade concentration

Besides testing on daily data, tests are also performed on intraday data. The goal of this, is to establish whether a large trade has a different influence on the price pattern over a specified time interval during that day, than a small trade has.

3.2.2.1 Data preparation

For this, the same intraday data as for the daily tests is used. Again, the data needs to be structured further. It is only interesting to compare differences for buy initiated block trades versus regular buy initiated trades and sell initiated block trades versus regular sell initiated trades, because the results are probably opposite. If this distinction between buy initiated and sell initiated is not made, it is unlikely that any result is found. Therefore, the first step is to delete all trades for which it cannot be determined whether it was a buy initiated or sell initiated trade (see section 3.1). This left 39,439,512 trades, of which 86,045 are block trades and 39,353,467 are regular trades. This method is in line with that used in previous literature on block trades. An example of this is the research performed by Kraus and Stoll (1972).

The next step is to collect for every trade the price *x* minutes after the trade was executed. The intervals chosen for *x* are: +1 trade, +1 minute, +5 minutes, +10 minutes and +15 minutes. Since it is generally accepted that 15 minutes is sufficient time to trade on an observed market event, a significant result for the 15 minute interval indicates that arbitrage fails and there is a case of market inefficiency (Grier & Albin, 1973). If it is not possible to collect a price at *t+x* because there were no trades executed at that time, the price of the trade closest to *t+x* is used. If it is not possible to collect a price at *t+x* because at *t+x* the market was closed, the closing price of that day is used.

The last step is to calculate the return and the log return over every time interval for every trade. This is done with the following formulas:

$$Return_{t,x} = \frac{P_{t+X}}{P_t} - 1 \qquad (14)$$

$$Ln\ Return_{t,x} = \ln\left(\frac{P_{t+x}}{P_t}\right) \qquad (15)$$

3.2.2.2 Testing

T-test

The testing for intraday results will be done by performing unpaired t-tests. The goal is to test whether the mean return of buy initiated (sell initiated) block trades over the interval *x* is significantly higher than the mean return of buy initiated (sell initiated) regular trades. To do this, first the mean returns need to be calculated for block trades and regular trades over all time intervals for both buy initiated and sell initiated trades. The following formula is used for this:

$$\overline{Return}_{b,x,s} = \frac{\sum return_{b,x,s}}{n_{b,x,s}} \tag{16}$$

Where, $\overline{Return}$ is the mean return, *b* indicates whether it concerns block trades or regular trades (block trade: *b*=1, regular trade: *b*=-1), *x* indicates the time interval over which the returns are measured and *s* indicates whether it concerns buy initiated or sell initiated trades (buy initiated: *s* = 1, sell initiated: *s* = -1). Now the following formulas can be used:

$$t_{x,s} = \frac{\overline{Return}_{b=1,x,s=1} - \overline{Return}_{b=-1,x,s=1}}{S_{b=1,b=-1}} \tag{17}$$

$$t_{x,s} = \frac{\overline{Return}_{b=1,x,s=-1} - \overline{Return}_{b=-1,x,s=-1}}{S_{b=1,b-1}} \tag{18}$$

Where:

$$S_{b=1,b=-1} = \sqrt{\frac{s_{b=1}^2}{n_{b=1}} + \frac{s_{b=-1}^2}{n_{b=-1}}} \tag{19}$$

Data requirements

To perform an unpaired t-test, the data needs to satisfy a few assumptions (Field, 2009). The first is that the dependant variable needs to be (roughly) normally distributed. This will be tested by using the Shapiro-Wilk test of normality:

$$w = \frac{(\sum_{i=1}^{n} a_i \, Return_{(i)})^2}{\sum_{i=1}^{n}(Return_i - \overline{Return})^2} \tag{20}$$

The next assumption is that the dependant variable needs to be measured on a continuous scale. Since the dependant variable is return, the data satisfies this assumption. The third assumption is that the independent variable consists of two categorical independent groups. This assumption is also satisfied since the independent variable can only be a block or a regular trade, which are two categorical groups. The fourth assumption is that the observations should be independent of each other. This means that there should be no relationship or overlap between the two categories of the independent variable. The data also satisfies this assumption. The last assumption is that the variances in return of the two categories should be roughly equal. This will be tested with Levene's test for homogeneity of variances.

3.2.3 Hypotheses

The tests described above, are used to test the following hypotheses:

Hypothesis 1: A concentrated buying (selling) volume (block trade) indicates a price increase (decrease) in the very near future due to misinterpretation of other traders.

Hypothesis 2: A concentrated buying (selling) volume (block trade) indicates a price decrease (increase) in the near future due to overreaction.

In the next chapter, the results of these tests will be presented and the hypotheses will be discussed. For all these tests a significance level of 5% will be used, unless stated otherwise.

Chapter 4: Results

In this chapter, the results of the tests described in the last chapter are presented. The results are used to determine whether the hypotheses need to be rejected or not. The results to the tests are presented in the same order as in the previous chapter. Thus, first the results to the daily tests are discussed and after that the results to the intraday tests are presented. For all tests a 5% significance level is used.

4.1 Daily trade concentration

The tests used for the daily trade concentration are all regression tests. Below, the results of each regression are discussed. For all regressions, except the regressions of formula 6 and 10, robust (White) standard errors were used because the errors were either heteroscedastic or not normally distributed.

4.1.1 Data overview

First, an overview of the descriptive statistics of the data used for the daily tests is provided.

Table 1: Descriptive statistics of the data used for the daily tests.

	N	*Range*	*Minimum*	*Maximum*	*Mean*	*Deviation*	*Skewness*	*Kurtosis*
Return	251	5.72%	-3.32%	2.40%	0.08%	0.008	-0.56	1.25
Trading volume (USD)	251	410,956,003	62,383,704	473,339,706	191,647,525	53,860,530	1.64	5.18
Trade concentration	251	4,531	-3,081	1,450	34	423	-1.42	12.40

A few things stand out when looking at this table. The first is that both the average daily return in the sample and the average trade concentration (corrected for buy and sell initiated trades) are positive. This is in line with the hypothesis that a high positive (negative) trade concentration causes high positive (negative) returns over that day. This obviously does not prove a significant relationship yet. For this, the test results discussed below are needed. When looking at the skewness, it can be seen that dollar volume and trade concentration have fairly high values for skewness. The dollar volume and trade concentration distributions also have high kurtosis. This means that there are relatively many observations with a value close to the mean.

4.1.2 Daily results

To test whether the trade concentration on day *t* has an effect on the return over day *t*, formula 6 is used. The results of this regression are found in the following table.

Table 2: Regression results of formula 6. Trade concentration on day t and return over day t-1 are regressed on return over day t. Additionally, p-values for the regression validity tests are supplied. These are the White test (H0: errors are homoscedastic), Breusch-Pagan test (H0: no autocorrelation in the errors), Jarque-Bera (H0: errors are normally distributed).

$$Return_{P,t} = \alpha + \beta_1 * trade\ concentration_{P,t} + \beta_2 * retu \quad _{p,t-1} + \varepsilon_t$$

			$Return_{p,t}$		
	Coefficient (%)	*Std. Error*	*t-statistic*	*Probability*	*N*
Trade concentration$_{p,t}$	0.0004	0.000001	3.07	0.002	251
Return$_{p,t-1}$	-11.384	0.062669	-1.82	0.071	251
α	0.1017	0.000497	2.05	0.042	251
R-squared	*0.0491*				
White	0.701				
Breusch-Pagan	0.978				
Jarque-Bera	0.224				

From the table it can be seen that the coefficient that measures the trade concentration is significant at the 1% level because the probability value is below 0.01. This means that the null hypothesis that trade concentration has no significant effect on return, can be rejected. The probability values of the validity tests are all above 0.05. This means that the null hypothesis should not be rejected and the errors are homoscedastic, not autocorrelated and normally distributed. This means the regression meets all four assumptions described in section 3.2.1 and no further corrections are required. From the coefficient it can be seen that for every increase of 1 in average trade concentration on a certain day, the return over that day also goes up by 0.00039%. Remember that when calculating the trade concentration, a correction has been made to differentiate between buy and sell initiated trade concentration. Thus, for days that have a negative trade concentration, the height of the trade concentration has a negative effect on the return over that day. For example, from the descriptive statistics table it can be seen that the day with the highest positive trade concentration, has a concentration of 1,450. On this day it is expected that the return over that day is 1,450 times 0.00039% is 0.57% higher. On the day with the most negative trade concentration, the return is expected to be 4,531 times 0.00039% is 1.77% lower. This finding is in line with the hypothesis that a concentrated trade volume causes prices to increase in the short run in case of buy initiated trades and to decrease in case of sell initiated trades. From this, it can be concluded that a higher concentration in trade volume on a day influences prices during that day, which is in line with the findings of Kraus and Stoll (1972). They say that this finding could be caused by information asymmetry because large trades, that cause the trade concentration to rise, are usually performed by institutional investors who potentially have more information than the average market participant. If this is the case, the finding that trade concentration on a certain day influences the return over that day, would be in line with market efficiency. It could also be the case that the price effects are caused by market inefficiency. This effect can be divided in two parts. The first is market illiquidity. In that case illiquidity in the market might cause prices to change when large trades are performed. The second part is trading costs. These are usually higher for large transactions. This could have an influence on the price. These last two effects would indicate market inefficiency. Therefore, they are expected to be priced away in the future by arbitrageurs. This is not the case with the information asymmetry hypothesis. Then, there is no inefficiency and the price after a large trade should be the correct price.

4.1.3 Trading volume

To see whether trading volume might be (part of) the cause behind the relationship found in table 2, trading volume on day *t* and the return over day *t-1* are regressed on the return over day *t*. The results of this can be found in table 3.

Table 3: Regression results of formula 7. Trading volume on day t and the return over day t-1 are regressed on return over day t with robust standard errors. Additionally, p-values for the regression validity tests are supplied. These are the White test (H0: errors are homoscedastic), Breusch-Pagan test (H0: no autocorrelation in the errors), Jarque-Bera (H0: errors are normally distributed).

$$Return_{P,t} = \alpha + \beta_1 * trading\ volume_{P,t} + \beta_2 retu\quad {}_{p,t-1} + \varepsilon_t$$

			$Return_{p,t}$		
	Coefficient (%)	*Robust Std. Error*	*t-statistic*	*Probability*	*N*
Trading $volume_{p,t}$	-3.1E-10	1.2E-11	-0.26	0.794	251
$Return_{p,t-1}$	-11.167	0.063	-1.78	0.077	251
α	0.1732	0.002	0.81	0.422	251
R-squared	*0.0127*				
White	0.000				
Breusch-Pagan	0.793				
Jarque-Bera	0.184				

From the above table it can be seen that trading volume is not significant at the 5% significance level (probability value is higher than 0.05). This means that it cannot be concluded that a high trading volume on a certain day causes the return over that day to be higher or lower. This finding is not in line with previous studies discussed in chapter 1. In those studies a relationship between trade volume and return was found. The absence of this effect in the data indicates that the effects found in table 2 are probably caused by concentration in trading volumes and are not (partly) caused by the height of trading volume. This will be more formally tested in formula 8.

However, it still could be the case that trading volume explains part of the effect trade concentration had in table 2. In that case, it could be possible that the coefficient for trade concentration is lower or that it is less significant. To test this, both trade concentration and trading volume are regressed simultaneously on return. The results of this are presented in the table below.

Table 4: Regression results of formula 8. Trade concentration on day t, trading volume on day t and return over day t-1 are regressed on return over day t with robust standard errors. Additionally, p-values for the regression validity tests are supplied. These are the White test (H0: errors are homoscedastic), Breusch-Pagan test (H0: no autocorrelation in the errors), Jarque-Bera (H0: errors are normally distributed).

$$Return_{P,t} = \alpha + \beta_1 * trade\ concentration_{P,t} + \beta_2 * trading\ volume_{P,t} + \beta_3 return_{p,t-1} + \varepsilon_t$$

			$\text{Return}_{p,t}$		
	Coefficient (%)	*Robust Std. Error*	*t-statistic*	*Probability*	*N*
Trade	0.0004	1.4E-06	2.51	0.013	251
Trading $\text{volume}_{p,t}$	4.4E-11	1.2E-11	0.04	0.970	251
$\text{Return}_{p,t-1}$	-11.375	0.0616	-1.85	0.066	251
α	0.0932	0.0021	0.44	0.663	251
R-squared	*0.0491*				
White	0.000				
Breusch-Pagan	0.978				
Jarque-Bera	0.230				

It can be seen that trade concentration becomes slightly less significant since the probability is lower than in table 3. However, the effects are very limited and it is still significant on the 1% level because the probability is still below 0.01. From table 3 and table 4 it can thus be concluded that trading volume is not explaining the results found in table 2.

4.1.4 Timing

To test what the effects of a high trade concentration are on returns over later days, it first needs to be established whether days with high trade concentration are followed by more days with high trade concentration. For this, formula 9 is used. The results can be seen in the table below.

Table 5: Regression results of formula 9. Trade concentration on day t-1 is regressed on trade concentration on day t with robust standard errors. Additionally, p-values for the regression validity tests are supplied. These are the White test (H0: errors are homoscedastic), Breusch-Pagan test (H0: no autocorrelation in the errors), Jarque-Bera (H0: errors are normally distributed).

$$Trade\ concentration_{P,t} = \alpha + \beta_1 * trade\ concentration_{P,t-1} + \varepsilon_t$$

	Trade concentration$_{p,t}$				
	Coefficient	*Robust Std. Error*	*t-statistic*	*Probability*	*N*
Trade concentration$_{p,t-1}$	-0.015	0.06	-0.28	0.784	251
α	33.92	26.76	1.27	0.206	251
R-squared	*0.0002*				
White	0.936				
Breusch-Pagan	0.960				
Jarque-Bera	0.000				

From this, it can be seen that there is no significant relationship between the trade concentration on day *t* and the trade concentration on day *t-1* because the probability for *Trade concentration$_{p,t-1}$* is higher than 0.05. This means that days with high trade concentration are distributed randomly across the days. Thus, if there is a relationship found between the return over day *t* and the trade concentration on day *t-1*, this indicates that this is caused by the high trade concentration on day *t-1* and not by the fact that days with high trade concentration are followed by more days with high trade concentration.

This relationship between the return over day *t* and the trade concentration on day *t-1* is tested in formula 10. The results are in the table below.

Table 6: Regression results of formula 10. Trade concentration on day t-1, trade concentration on day t and return over day t-1 are regressed on return over day t. Additionally, p-values for the regression validity tests are supplied. These are the White test (H0: errors are homoscedastic), Breusch-Pagan test (H0: no autocorrelation in the errors), Jarque-Bera (H0: errors are normally distributed).

$$Return_{P,t+x} = \alpha + \beta_1 * trade\ concentration_{P,t} + \beta_2 * trade\ concentration_{P,t+x} + \beta_3 + return_{p,t-1} + \varepsilon_t$$

			Return$_{p,t}$		
	Coefficient (%)	*Std. Error*	*t-statistic*	*Probability*	*N*
Trade concentration$_{p,t}$	3.6E-04	1.2E-06	0.07	0.942	251
Trade concentration$_{p,t-1}$	8.6E-06	1.2E-06	3.06	0.002	251
Return$_{p,t-1}$	-11.473	0.0640	-1.79	0.074	251
α	0.1015	5.0E-04	2.03	0.043	251
R-squared	*0.0491*				
White	0.919				
Breusch-Pagan	0.965				
Jarque-Bera	0.222				

The probability value for *Trade concentration*$_{p,t-1}$ is below 0.05. This means that there is no significant relationship between the return over day *t* and the trade concentration on day *t-1.* Since the trade concentration on day *t* was also added to the regression, this variable only measures the effect trade concentration on day *t-1* has on the return one day later. In all three regressions, only trade concentration on day *t* is significant (probability value below 0.05). This indicates that the effect a high trade concentration has on return, is only apparent on the day itself and that it has no effect on the return over later days. This finding indicates that the effects of trade concentration on prices are permanent. If they were temporary, a reversal would have been expected after the price change. This is not in line with previous studies discussed in chapter 1. They all find that prices (partly) revert back to their equilibrium price. The absence of this reversed price movement in this study indicates that prices go to a new equilibrium price and that an increase in trade concentration is caused by (institutional) investors that have extra information. It has to be noted however, that for this study the effects of trade concentration have only been taken into account for day *t* and *t+1 day*. It could be possible that a reversal takes place, but that this happens on later days.

To test whether trading volume has an effect on *t+1*, formula 11 is used. The result are presented in the table below.

Table 7: Regression results of formula 11. Trading volume on day t-1 and trading volume on day t are regressed on return over day t with robust standard errors. Additionally, p-values for the regression validity tests are supplied. These are the White test (H0: errors are homoscedastic), Breusch-Pagan test (H0: no autocorrelation in the errors), Jarque-Bera (H0: errors are normally distributed).

$$Return_{P,t+x} = \alpha + \beta_1 * trading\ volume_{P,t} + \beta_2 * trading\ volume_{P,t+x} + \beta_3 * return_{p,t-1} + \varepsilon$$

			$Return_{p,t-1}$		
	Coefficient (%)	*Robust Std. Error*	*t-statistic*	*Probability*	*N*
Trading $volume_{p,t}$	-7.4E-10	1.3E-11	-0.55	0.581	251
Trading $volume_{p,t-1}$	9.2E-10	1.1E-11	0.86	0.391	251
$Return_{p,t-1}$	-11.16	0.0622	-1.79	0.074	251
α	0.079	0.0023	0.34	0.733	251
R-squared	*0.016*				
White	0.000				
Breusch-Pagan	0.837				
Jarque-Bera	0.172				

Both trading volume on day *t* and trading volume on day *t-1* have no significant effect on the return over day *t* because their probability values are below 0.05. This is in line with what would be expected from the results of the regression of formula 7.

In the regression of formula 12 all variables are taken into account in order to correct for all possible effects which were tested for individually in the previous regressions. The results of this test are shown below.

Table 8: Regression results of formula 12. Trade concentration on day t-1, trade concentration on day t, trading volume on day t-1, trading volume on day t and return over day t-1 are regressed on return over day t with robust standard errors. Additionally, p-values for the regression validity tests are supplied. These are the White test (H0: errors are homoscedastic), Breusch-Pagan test (H0: no autocorrelation in the errors), Jarque-Bera (H0: errors are normally distributed).

$$Return_{P,t+x} = \alpha + \beta_1 * trade\ concentration_{P,t} + \beta_2 * trade\ concentration_{P,t+x} + \beta_3 * trading\ volume_{P,t} + \beta_4 * trading\ volume_{P,t+x} + \beta_5 * return_{p,t-1}$$

			$\text{Return}_{p,t-1}$		
	Coefficient (%)	*Robust Std. Error*	*t-statistic*	*Probability*	*N*
Trade	3.6E-04	1.5E-06	2.46	0.015	251
Trade	2.2E-05	1.1E-06	0.20	0.842	251
Trading $\text{volume}_{p,t}$	-4.1E-10	1.4E-11	-0.30	0.767	251
Trading $\text{volume}_{p,t-1}$	9.8E-10	1.1E-11	0.88	0.381	251
$\text{Return}_{p,t-1}$	-11.60	0.0624	-1.86	0.064	251
α	-0.009	0.0023	-0.04	0.970	251
R-squared	*0.053*				
White	0.001				
Breusch-Pagan	0.965				
Jarque-Bera	0.220				

From the table it can be seen that the results almost do not change when all variables are combined and all possible effects are taken into account. The trade concentration variable is still highly significant (probability value below 0.01). This means that trade concentration on day *t* still can explain part of the return on day *t*. Trade concentration on day *t-1* has a probability value higher than 0.05 and is again not significant. Thus, a high trade concentration today has no significant effect on the return over tomorrow. Trading volume on day *t* and day *t-1* are both not significant (probability value above 0.05). Thus, the height of trading volume has no significant influence on the return over day *t*. This means that trading volume only plays a role if it is viewed in relation to the number of trades on that day (trade concentration).

However, in the results above the interaction term has not yet been taken into account. When this is taken into account, the results change to those in the table below.

*Table 9: Regression results of formula 13. Trade concentration on day t-1, trade concentration on day t, trading volume on day t-1, trading volume on day t, the cross term trade concentration on day t*trading volume on day t and return over day t-1 are regressed on return over day t with robust standard errors. Additionally, p-values for the regression validity tests are supplied. These are the White test (H0: errors are homoscedastic), Breusch-Pagan test (H0: no autocorrelation in the errors), Jarque-Bera (H0: errors are normally distributed).*

$$Return_{P,t+x} = \alpha + \beta_1 * trade\ concentration_{P,t} + \beta_2 * trade\ concentration_{P,t+x} + \beta_3 * trading\ volume_{P,t} + \beta_4 * trading\ volume_{P,t+x} + \beta_5 * trade\ concentrati\ _{p,t} * trading\ volume_{P,t} + \beta_6 * return_{p,t-1} + \varepsilon_t$$

			$Return_{p,t-1}$		
	Coefficient (%)	*Robust Std. Error*	*t-statistic*	*Probability*	*N*
Trade concentration$_{p,t}$	4.9E-04	3.6E-06	1.36	0.177	251
Trade concentration$_{p,t-1}$	2.1E-05	1.1E-06	0.19	0.847	251
Trading volume$_{p,t}$	-4.9E-10	1.4E-11	-0.36	0.718	251
Trading volume$_{p,t-1}$	1.0E-09	1.1E-11	0.94	0.349	251
Trade concentration$_{p,t}$*Trading volume$_{p,t}$	-5.1E-13	1.4E-14	-0.35	0.724	251
Return$_{p,t-1}$	-11.859	0.0628	-1.89	0.060	251
α	-0.0063	0.0023	-0.03	0.978	251
R-squared	*0.053*				
White	0.001				
Breusch-Pagan	0.960				
Jarque-Bera	0.204				

It can be seen that adding the interaction term, makes trade concentration on day *t* less significant. It no longer is significant at the 5% level (probability value above 0.05). This mean that the earlier results found for trade concentration on day *t*, are not robust and are dependent on the height of trading volume on day *t.*

From these insights it can be concluded that when the interaction term is ignored, hypothesis 1 does not get rejected. The test results provide evidence that a concentrated buying (selling) volume indicates a price increase (decrease) in the very near future due to misinterpretation of other traders. However, this result no longer holds when the interaction term is taken into account. Then, both trade concentration and dollar volume are insignificant. In that case hypothesis 1 needs to get rejected. In both cases there is evidence that hypothesis 2 needs to get rejected. A concentrated buying (selling) volume does not indicate a price decrease (increase) in the near future due to overreaction. Thus, it can be concluded that with the interaction term taken into account, daily trade concentration has no significant impact on returns and cannot be used to trade on and earn significant abnormal returns.

4.2 *Intraday trade concentration*

The results of the daily tests show that trade concentration has an effect on the return over the same day but this effect is not apparent on later days and insignificant if an interaction term between trade concentration and trading volume is taken into account. Therefore, the daily results do not provide results on which can be traded. That makes it very interesting to look at the intraday effects. Because it might be possible that a sudden increase in trade concentration during a day still has an effect on the returns later that day. The results of the tests used to test this will be discussed below.

4.2.1 Data overview

First, an overview of the data used for the buy and sell initiated intraday tests is presented.

Table 10: Descriptive statistics of the data used for the intraday tests for buy initiated trades.

Buy initiated	N	Range	Minimum	Maximum	Mean%	Deviation
Return	19,709,612	9011	1.51E-07	9011	1.2826%	9.8085
Return t+1trade	19,709,612	85.68	-1.0000	84.6818	-0.1327%	0.0193
Return t+1minute	19,709,612	9.015	-0.8904	8.1241	0.0021%	0.0041
Return t+5minutes	19,709,612	9.018	-0.8905	8.1276	0.0091%	0.0072
Return t+10minutes	19,709,612	9.015	-0.8905	8.1247	0.0136%	0.0049
Return t+15minutes	19,709,612	9.014	-0.8904	8.1234	0.0188%	0.0059

Table 11: Descriptive statistics of the data used for the intraday tests for buy initiated trades.

Sell initiated	N	Range	Minimum	Maximum	Mean%	Deviation
Return	19,729,462	0.9999	-0.9999	0.0	-0.043%	0.0016
Return t+1trade	19,729,462	9010.00	-1.0000	9009.0	1.028%	0.0131
Return t+1minute	19,729,462	8771.74	-0.6312	8771.1	-1.504%	11.1919
Return t+5minutes	19,729,462	8816.09	-0.6234	8815.5	-1.380%	11.2485
Return t+10minutes	19,729,462	8816.09	-0.6234	8815.5	-1.368%	11.1248
Return t+15minutes	19,729,462	8816.09	-0.6234	8815.5	-1.266%	11.1238

What stands out from these overviews are the big ranges for the returns over all time intervals. This is caused by errors in the TAQ data. There are multiple price observations in the data, that are absolutely not realistic. This has an effect on the ranges, means and standard deviations for all return variables. Due to this, it is highly unlikely that the data for these variables is normally distributed. This has also been formally tested by using the Shapiro-Wilk test for normality. The results of this are presented in the table below.

Table 12: Results of the Shapiro-Wilk test on normality. H0: The data is normally distributed.

	Z	Prob
Return	33.07	0.000
Return t+1trade	33.06	0.000
Return t+1minute	32.75	0.000
Return t+5minutes	32.73	0.000
Return t+10minutes	31.51	0.000
Return t+15minutes	31.66	0.000

From the table it can be seen that the null hypothesis that the variables are normally distributed is rejected and that the variables are indeed highly non-normal due to the errors in the data. Because of this, it is no longer possible to use a t-test since this test requires the data to be normally distributed. This leaves two options. The first is to determine which observations in the data are errors and remove them from the data. The problem with this solution is that it is not easy to determine which observations are actually errors and which are valid observations of extreme events. Manually removing observations might even make the data less reliable. Therefore, the second option is preferable. This option is to accept that there are some errors in the data but use a Wilcoxon rank sum test. This test does not require the data to be normally distributed and thus the errors pose no problem. Furthermore, it can be assumed that the errors in the data are randomly distributed across the observations. This means that the data for both block trades as for non-block trades, are equally affected by the measurement errors. Therefore, the measurement errors should have no significant impact on the test results.

4.2.2 Intraday results

As described in chapter 3, the intraday test is repeated for 5 intervals. The first interval is *t+1 trade*, in which the return of a trade at time *t* is compared to the return of the next trade. The other four intervals are all fixed time intervals. Thus, the return at time *t* is compared to the return at time *t+x*, where *x* can be 1, 5, 10 or 15 minutes. The results of the tests are discussed below.

4.2.2.1 Buy initiated trades

The results of the Wilcoxon sum rank test for all 5 intervals after buy initiated trades, are presented in the table below.

Table 13: Results of the Wilcoxon sum rank test for buy initiated trades. H0: There is no significant difference between the return after a buy initiated block trade or a buy initiated non-block trade.

Return block - Return non-block	Z	Prob	Prob: Return Block>Return non-block	Mean%
Return t+1trade	60.26	0.00	0.420	-0.1327%
Return t+1minute	23.38	0.00	0.467	0.0021%
Return t+5minutes	18.46	0.00	0.474	0.0091%
Return t+10minutes	17.31	0.00	0.476	0.0136%
Return t+15minutes	17.40	0.00	0.476	0.0188%

The P value for all intervals is below 0.05 which means that the null hypothesis that the return of blocks is equal to the return of non-blocks over the specified intervals, is rejected. This means that the return of buy initiated block trades is significantly different than that of buy initiated non-block trades over all intervals. However, it would be expected that the return over the intervals is more positive for block trades than that for non-block trades due to the fact that the trades are buy initiated, but this does not seem to be the case. The probability that the return after a buy initiated block trade is higher over the intervals than the return after a buy initiated regular trade, is below 0.5 for all intervals. This means that on average the return after a buy initiated block trade is lower than the return after a buy initiated regular trade for all intervals. Therefore, hypothesis one which says that a concentrated buying volume indicates a price increase in the very near future due to misinterpretation of other traders, needs to be rejected for buy initiated trades. Even for the *t+1 trade* interval, which is the shortest interval possible, there is no evidence that the return over this interval is higher after block-trades than the return over this interval after non-block trades. Hypothesis two, which says that a concentrated buying volume indicates a price decrease in the near future due to overreaction, also needs to be rejected for buy initiated trades for all intervals except the *t+1* trade interval. Even though the return over all intervals is significantly lower after block trades than after non-block trades, it can be seen from the table that the mean return is positive for all intervals except *t+1* trade. This indicates that there is indeed an opposite price movement after a block trade, but that this happens immediately after the block trade. As said earlier, the estimated reaction time needed to be able to trade on block information, is 15 minutes. Therefore, these results do not provide possibilities to trade on information about buy initiated block trades. This finding is in line with the research of Dann, Mayers and Raab (1977) who also find that prices recover after a block trade within 15 minutes. The study of Holthausen, Leftwich and Mayers (1990) provides even more similar results. They also find that for buy initiated trades, the recovery takes place within one trade.

4.2.2.2 *Sell initiated trades*

The results of the Wilcoxon sum rank test for all intervals after sell initiated trades, are presented in the table below.

Table 14: Results of the Wilcoxon sum rank test for sell initiated trades. H0: There is no significant difference between the return after a sell initiated block trade or a sell initiated non-block trade.

Return block - Return non-block	Z	Prob	Prob: Return Block>Return non-block	Mean%
Return t+1trade	-67.64	0.00	0.589	1.028%
Return t+1minute	-9.64	0.00	0.513	-1.504%
Return t+5minutes	-8.54	0.00	0.512	-1.380%
Return t+10minutes	-8.29	0.00	0.512	-1.368%
Return t+15minutes	-7.23	0.00	0.510	-1.266%

Again, the P value is below 0.05 for all 5 intervals. This means that the return over all intervals is significantly different after sell initiated block trades than after sell initiated non-block trades. The probability that the return over the interval is higher after a sell initiated block trade than after a sell initiated regular trade, is greater than 0.5 for all intervals. This means that the return is significantly higher after block trades than after regular trades. For sell initiated trades, this is not what is to be expected according to hypothesis one. This states that a concentrated selling volume indicates a price decrease in the near future, due to misinterpretation of other trades. From the means it can be seen that the opposite happens. Over the shortest interval, a positive return is observed higher than after a sell initiated non-block trade. Over the longer intervals a price decrease is observed. This means also hypothesis 2 needs to be rejected. Because this hypothesis states that a concentrated selling volume indicates a price increase in the near future due to overreaction. Thus, there is an observed price path that is opposed to what was expected. Additionally, for the intervals longer than *t+1* trade, the effects are less severe after sell initiated block trades than after regular sell initiated trades. These results are in line with the results found for buy initiated block trades and the results found in earlier studies.

To conclude, there is a price pattern associated with block trades. The price goes up (down) right at the moment the buy (sell) initiated block trade is performed as would be expected. However, according to hypothesis 1 it would be expected that this price movement continues for some time after the block trade. This does not seem to be the case and therefore hypothesis 1 needs to be rejected. According to hypothesis 2 there is a reversed effect some more time after the block trade, in which prices revert back to their equilibrium price. This effect can be observed, but it happens almost instantaneously after the block trade. Therefore, based on these findings, it is not possible for traders to react on time on information about block trades and profitably trade on this information. The only exception would be if information about block trades leaks prior to the execution of the trade to the market. In the study of Gemmil (1996), the study of De Jong and Van Beusichem (2006) and the study of Keim and Madhavan (1996), it is found that it is possible that this information leaks to the market and traders can anticipate a block trade. This strategy could yield positive returns.

Chapter 5: Conclusion

Trade concentration is an important factor when studying asset prices on an exchange. It is observed in this and previous studies that the size of a trade has an influence on the price at which the trade is performed. The two main explanations for this are either based on liquidity effects or on information effects and the price effect can probably be explained by a combination of both. The liquidity effect is caused by a change in the supply and demand of a certain asset on the exchange. In the case of buy initiated block transactions, the demand increases suddenly but the supply remains the same. This causes the price for the trade to be higher. For sell initiated block trades, the effect is opposite. The reasoning behind the information effect is that large trades are usually performed by institutional investors that have gained additional information which the rest of the market does not have. By trading a large quantity, they can profit as much as possible from this information. The market knows this and adapts its prices to this information.

In this thesis the price pattern caused by trade concentration is observed. For the daily trade concentration, it was initially found that for days which had a high positive (negative) average trade concentration, the return over that day was higher (lower). However, when a cross term was introduced to the regression, this result no longer was significant. In both situations, the results for the effects of trade concentration on the return over a later day, were not significant. Therefore it has to be concluded that hypothesis one which says that a concentrated buying (selling) volume indicates a price increase (decrease) in the very near future due to misinterpretation of other traders, needs to be rejected for the daily tests. After introducing the cross term, this pattern was no longer significant. The second hypothesis which says that a concentrated buying (selling) volume indicates a price decrease (increase) in the near future due to overreaction, also needs to get rejected for the daily tests. This implies that it is not possible for an investor to profitably trade on information about daily trade concentration because the effects are traded away too quickly. What also stands out when looking at the daily results, is that dollar volume is not significant. This is not in line with most of the literature discussed in section 2.1. It is not entirely clear what could explain this. It could be caused by a difference in the tested sample or research setup.

For the intraday data, there is an observable price pattern. As expected, the price goes up (down) at the moment a buy (sell) initiated block trade is performed. After that however, the price appears to readjust almost instantaneously. This means that there appears to be no momentum effect that causes prices to rise (fall) further after the block trade. Therefore, it has to be concluded that there is no evidence supporting hypothesis one that says that a concentrated buying (selling) volume (block trade) indicates a price increase (decrease) in the very near future due to misinterpretation of other traders. There is however an observable pattern that follows hypothesis two that says that a concentrated buying (selling) volume indicates a price decrease (increase) in the near future due to overreaction. From the results it can be seen that the price effect of the block trade is later reversed. However, it has to be noted that this happens within one trade after the block is traded. This interval is too short for traders to trade profitably on information about block trades. Furthermore, due to the rejection of hypothesis one, the effect of a block trade on the price is limited and thus the correction in the period afterwards is also limited. This is in line with the findings in earlier literature from for example Holthausen, Leftwich and Mayers (1990).

From these findings it can be concluded that, as expected, trade concentration has a role on asset prices. This is likely to be caused by market liquidity effects or information effects. However, the transparency of the NYSE exchange and the instant publication of block trades, cause information about this to be spread across the market. Because of this, prices readjust almost immediately, making it impossible to profitably trade on trade concentration effects. From this, it can be concluded that market transparency about trade concentration causes the market to be more efficient and prices to remain closer to their fundamental value.

A few remarks have to be made to this research which might have had impact on the results. The first remark is that the focus of this paper has been on trades performed on one exchange. Therefore, one has to be cautious when applying these results to different markets. As documented by Keim and Madhavan (1996), a big part of blocks are not traded via a regular exchange like the NYSE, but are traded on a different market form on which an intermediary matches buyers and sellers before executing the trade. This is called the upstairs market. It is possible that the results would have been significantly different when this research methodology was applied to that market form. The second remark is that it is common for block trades to be divided in multiple smaller trades to improve the ease at which they can be traded on an exchange. This cannot be observed from the used data. Therefore, it is possible that block trades have been treated as multiple regular trades. It is possible that this has an influence on the results. Another possibility that has not been taken into account is that information about an upcoming block trade might leak to the market prior to the actual trade. This effect has been documented for the upstairs market by Keim and Madhavan (1996). In this research however, only intervals after the block trade are taken into account. It could be that there is a significant price effect prior to the trade, causing the effect at the moment of the actual trade and the intervals thereafter, to be less high since the information is already priced into the market. Another remark is that when calculating the returns, dividends were not taken into account. Since a stock price usually lowers by the dividend value after the dividend is payed, this could have an influence on the results. The last remark is that from the used data it was unobservable whether a trade was buy or sell initiated. Therefore, the tick method was used to determine this. Although this method is used in most prior research and is generally valued as a good alternative, it is probably not perfect and therefore causes some degree of noise to the data. This might have a (small) impact on the results. For further research it is recommended to assess these remarks and test whether the results still hold when one or more of these remarks are taken into account. Another interesting topic for future research would be to test whether it is possible to detect situations in which relatively many block trades are executed. It could be possible for example that there is a certain calendar effect that causes the likelihood for a trade to be a block trade to rise. If such a relationship is found this could be used to profitably trade on, since this would remove the reaction time as the block is already anticipated.

References

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics, 17*(2), 223-249.

Attig, N., Fong, W.-M., Gadhoum, Y., & Lang, L. (2006). Effects of large shareholding on information asymmetry and stock liquidity. *Journal of Banking and Finance, 30*, 2875-2892.

Avramov, D., Chordia, T., & Goyal, A. (2006). The impact of trades on daily volatility. *The Review of Financial Studies, 19*(4), 1241-1277.

Brooks, C. (2008). Classical linear regression model assumptions and diagnostic tests. In C. Brooks, *Introductory Econometrics for Finance* (pp. 129-179). Cambridge: Cambridge University Press.

Brown, J., Crocker, D., & Foerster, S. (2009). Trading volume and stock investments. *Financial Analysts Journal, 65*(2), 67-84.

Chordia, T., & Swaminathan, B. (2000). Trading volue and cross-autocorrelations in stock returns. *The Journal of Finance, 55*(2), 913-935.

Chordia, T., Subrahmanyam, A., & Anshuman, R. V. (2001). Trading activity and expected stock returns. *Journal of Financial Economics, 59*(3), 3-32.

Dann, L., Mayers, D., & Raab, R. (1977). Trading rules, large blocks and the speed of adjustment. *Journal of Financial Economics, 4*(3), 3-22.

Datar, V., Naik, N., & Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets, 1*(2), 203-219.

De Jong, A., & Van Beusichem, H. (2006). *The price impact of block trades in the Netherlands: liquidity or information? (Working paper).* Retrieved from www.researchgate.net: https://www.researchgate.net/publication/266574432_The_price_impact_of_block_trades_in_the_Netherlands_liquidity_or_information

De Long, B., Shleifer, A., Summers, L., & Waldmann, R. (1990). Noise trader risk in financial markets. *The Journal of Political Economy, 98*(4), 703-738.

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

Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383-417.

Field, A. (2009). Assumptions of the t-test. In A. Field, *Discovering Statistics Using SPSS* (p. 326). London: SAGE publications.

Gemmil, G. (1996). Transparency and liquidity: A study of block trades on the London Stock Exchange under different publication rules. *The Journal of Finance, 51*(5), 1765-1790.

Gervais, S., Kaniel, R., & Mingelgrin, D. (2001). The high-volume return premium. *The Journal of Finance, 56*(3), 877-919.

Grier, P., & Albin, P. (1973). Nonrandom price changes in association with trading in large block. *The Journal of Business, 46*(3), 425-433.

Holthausen, R., Leftwich, R., & Mayers, D. (1990). Large-block transactions, the speed of response, and temporary and permanent stock-price effects. *Journal of Financial Economics, 26*(1), 71-95.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance, 48*(1), 65-91.

Keim, D., & Madhavan, A. (1996). The Upstairs Market for Large Block Transactions: Analysis and Measurement of Price Effects. *The Review of Financial Studies, 9*(1), 1-36.

Kraus, A., & Stoll, H. (1972). Price impacts of block trading on the New York stock excange. *The Journal of Finance, 27*(3), 569-588.

Lee, C., & Ready, M. (1991). Inferring trade direction from intraday data. *The Journal of Finance, 46*(2), 733-746.

Lee, C., & Swaminathan, B. (2000). Price momentum and trading volume. *The Journal of Finance, 55*(5), 2017-2069.

Phadnis, K. (2015, 5 20). *Bloomberg Tradebook.* Retrieved from www.bloombergtradebook.com: http://www.bloombergtradebook.com/blog/block-trading-in-todays-electronic-markets/