UNIVERSITY OF VAASA

FACULTY OF BUSINESS STUDIES

DEPARTMENT OF ACCOUNTING AND FINANCE

Menno Gijzen

THE HIGH-VOLUME RETURN PREMIUM: EVIDENCE FROM THE FINNISH STOCK MARKET

Master's Thesis in

Accounting and Finance

VAASA 2017

TABLE OF CONTENTS page

LIST OF FIGURES **page**

LIST OF TABLES **page**

UNIVERSITY OF VAASA
Faculty of Business Studies
Author: Menno Gijzen
Topic of the Thesis: The high-volume return premium: evidence from the Finnish stock market
Supervisor: Vanja Piljak
Degree: Master of Science in Economics and Business Administration
Department: Department of Accounting and Finance
Major Subject: Finance
Line: Finance
Year of Entering the University: 2015
Year of Completing the Thesis: 2017 **Pages: 56**

ABSTRACT

This thesis tests whether the high-volume return premium is present in the Finnish stock market during a time horizon from April 1994 to May 2016. It provides a theoretical insight into the importance of trading volume, and how it may be used to predict future price movements of stocks. The thesis is thus testing whether the weak-form of the efficient market hypothesis holds in the Finnish stock market.

Stocks are classified as high (low) volume based on their previous 49 days trading volume, and placed in long (short) portfolios, which are then held for up to 100 days without rebalancing. Stocks are then classified as small, medium, and large sized firms based on end year market capitalization.

The results show weak evidence to support the existence of the high-volume return premium based on daily data. Weekly data does show significant returns for medium and large sized firms. This is not due to return autocorrelations or extreme values. However, it does turn out that the returns are purely driven by shifts in systematic risk.

KEYWORDS: High-volume return premium, Trading volume.

1. INTRODUCTION

The goal of an investor is to generate returns on their investments. Investors therefore search for the right stocks to place in their portfolios, and try to time the market to generate better results than the market does. The big players, financial institutions varying from hedge funds to banks, employ researchers and portfolio managers in order to beat the market, that is, to generate excess returns. However, per theory, an investor should not be able to consistently beat the market. One may obtain significant returns on their strategy for a short period, but in the long run it should not outperform the markets on a risk-adjusted basis. The only way for an investor to generate higher returns is by taking on more risk. This theory is fully described by the efficient market hypothesis (EMH), originally developed by Eugene F. Fama (1965).

The EMH dictates that stock behavior and characteristics should not have any predictive power over an according risk measure. In fact, the hypothesis states that future stock returns may not be forecasted with any tools since asset prices fully reflect all available information, including projections. This implies that it is impossible to beat the market by selecting specific stocks, on a risk-adjusted basis. Nevertheless, throughout history, investors have constructed various investment strategies which have proven to provide excess returns over a risk-adjusted basis, and thus have beaten the market. These strategies typically revolved around finding an anomaly or a systematic event in the market on which excess returns can be generated when selecting the right stocks. This thesis however, will examine an investment strategy that is based on a stock characteristic: trading volume. It should be clear here that in the academic literature various measures of trading volume have been defined depending on the relationships that are being studied. This thesis defines trading volume as share volume as this has been the main measure in previous studies, and different measures of trading volume have yielded no different results.

The trading volume of stocks have long been an indicator of stock movement and information flow for practitioners. It is basic economics that a shift in demand will subsequently cause for a shift in price. Therefore, it should be of no mystery that if there is a shift in the trading volume of a stock, that this goes hand in hand with a change in the price of the stock. In recent years, this concept has been investigated more and more in the academics, and it is now accepted that trading volume may serve as a proxy for various factors such as liquidity, trend, and information flow. These factors drive price changes of stocks, and may therefore hold predictive power on the future price movements. This insight into the meaning of trading volume has given path to trading strategies that attempt to beat the market by selecting stocks based on the behavior of the underlying trading volume.

This thesis will examine whether there exists a way to generate excess returns in the Finnish stock market by using abnormal daily trading volume as a tool to predict future returns, and thus be able to generate a portfolio which consistently beats the market. The gains that are made from abnormal trading volume stocks is called the high-volume return premium, which was originally documented by Gervais, Kaniel, and Mingelgrin (2001) on the United States (US) stock market. They found that when a stock experiences abnormal high (low) trading volume it is subsequently followed by large (small) stock returns. This effect has shown to generate returns more than the index and were economically significant with annual returns of up to 11.0 percent. The high-volume return premium has shown to typically lasts for at least 20 days and can last for up to 100 days, and is persistent through all stock sizes.

1.1 Purpose of this study

This thesis examines whether there exists a high-volume return premium in the Finnish stock market, which consequently tests whether the efficient market hypothesis holds. This is the null hypothesis of the thesis, and will be further discussed in Section 3.1. Even though the high-volume return premium has been studied in previous papers, which will be further discussed in Section 2.2, the Finnish stock market has been remained untouched for unclear reasons. It could be because the Finnish stock market is relatively small. In comparison to the New York Stock Exchange, about 2800 companies are listed with over 1.46 billion trading volume each day, the Finnish stock market only has about 140 stocks listed with on average 231 thousand trading volume each day. This thesis will thus not only provide insight into the information content of trading volume in the Finnish stock market, but it will also show whether the high-volume return premium can exist in a small market.

The high-volume return premium has shown to exist in most developed stock markets and even some emerging stock markets (Kaniel, Ozoguz, Starks 2012). This thesis will therefore not only examine whether the return premium exists in the Finnish stock market, but also see whether the Finnish stock market reacts similarly to the information content of volume as other developed stock markets do. Martikainen, Puttonen, Luoma, and Rothovius (1995) had found that the relationship between stock returns and trading volume in the Finnish stock market has been behaving accordingly to the relationship as found in the US stock markets, however this was not held for an earlier period during 1977-1982. It was explained by the structural changes in the Finnish stock market during that period, and thus it appeared that the Finnish stock markets became more developed and started to behave more like the US stock markets. Hence the Finnish stock market is defined as developed market and is expected to show similar behavior as other developed markets.

Figure 1 displays the main results of what is researched in this thesis. It shows how the average cumulative returns generated by three different types of stocks: high, low, and

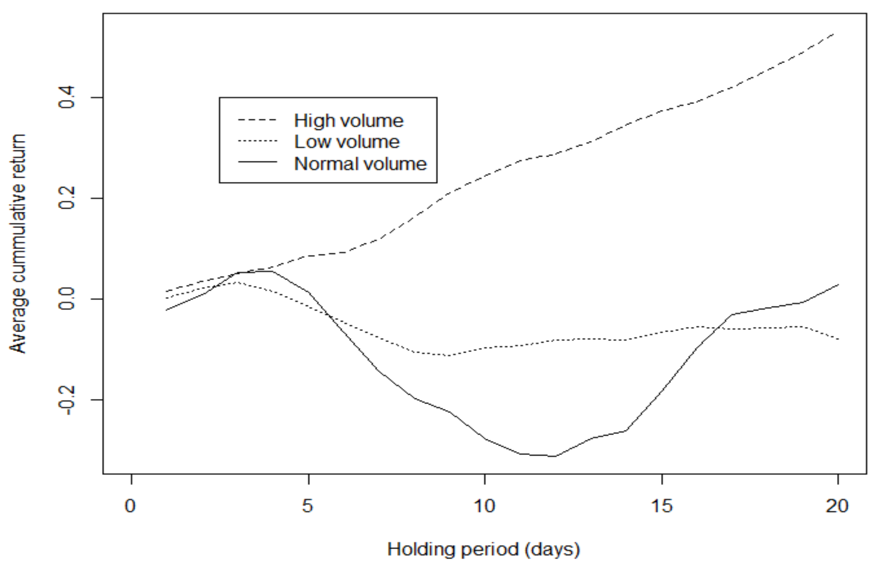


Figure 1: The average cumulative return of stocks which are separated on basis of their one day trading volume during a formation period from April 1994 to May 2016. Stocks are classified as high (low) volume when they are in the top (bottom) percentile of trading volume in respect to the previous 49 days. Stocks are then held for a certain holding period, given on the horizontal axis.

normal volume stocks on the Finnish stock market. It is clear to see that stocks that experienced abnormally high volume outperform both the abnormally low and normal volume stocks. Even though the actual returns for high volume stocks are relatively small, peaking out a little above 0.4 percent return after twenty days, the effect appears to be steadily growing over time. This thesis will further study this effect to see if it is of actual significance and if it comes due to systematic risk, or whether high (low) volume stocks do indeed contain information for future returns in the Finnish stock market.

1.2 Structure of the thesis

The structure of the thesis is as follows: Section 1 has introduced the topic and provided a description on what will be researched. Section 2 will go over the literature background, where a description will be given about the EMH. It will describe its implications and how it affects the view of investing in capital markets. It will also describe the visibility hypothesis, which is an important hypothesis for the high-volume premium. This will be followed by a literature overview of studies that have had a significant impact on the way trading volume is viewed as a key characteristic for stocks. It will focus on the information content of trading volume, and how this holds predictive power over the movement of stock prices. Section 3 describes the data and methodology. A similar approach as that of Gervais et al. (2001) will be used to detect the high-volume return premium. Additionally, certain screening procedures will be applied on the data and described here. This is to avoid any computational biases. Section 4 shows the empirical results and describes the implications. Section 5 has robustness checks and discusses alternative explanations. A test on the autocorrelations of returns and volume will be performed to check whether this drives the price movements seen in high and low volume stocks. Also, any outliers will be removed from the samples to see how strong the influence of extreme values are on the outcomes. This will be followed by testing whether returns are generated by shifts in systematic risk, or are indeed due to the high-volume premium. Section 6 concludes the results and describes what the implications are in regards to the empirical evidence and theory.

2. THEORETICAL FRAMEWORK

This chapter will go over previous studies on trading volume, and lays the foundation of some theories which are important to understand for this thesis. Section 2.1 goes over the theoretical background such as the efficient market hypothesis, and the visibility hypothesis. Also, a short introduction to the development of trading volume research in academics is given. Section 2.2 goes over important studies about trading volume and the relationship it holds with price movements.

2.1 Theoretical background

2.1.1 The efficient market hypothesis

> “A professor and a student are walking down a street and the student sees a €10 bill lying on the street. As the student is reaches down to pick up the money the professor says “Don’t bother; if that really was a €10 bill it would already be picked up.” – An old economist joke.

This joke is often given by finance professors during the first lecture, and gives a clever insight on the perception on financial markets. As financial markets are highly competitive any money-making opportunities should be rare to nonexistent. Hence, there are no such things as “free lunches”. This intuition is what is the basis for the EMH, and lays at the foundation of many financial theories and has probably been the most influential concept in finance.

The EMH states that the current price of a stock is the best price. An investor that investigates a company’s public information should not conclude to assign a higher or lower price to a stock, on average. This is because previous investors have already used the information to set the price to the right level (Fama, Fischer, Jensen, Roll 1969). Thus, there is no such thing as an overpriced or underpriced stock. Current market prices must therefore be trusted as these represent the true value of a given stock; they are priced fairly. This means that it should not be possible to consistently outperform the market through specific stock selection or timing the market. Higher returns than the market can only be obtained by taking on more risk (Shleifer 2000).

For the EMH to hold several market assumptions are placed. These assumptions are the so called "perfect market assumptions". These assumptions, noted below, describe a perfect market where prices fully reflect all available information.

i) Market participants have homogenous expectations.

ii) There are no transaction costs.

iii) All information is costless and received by all market participants.

The EMH knows three forms that describe the efficiency of the market through different levels. The following three paragraphs will go over each of these forms; weak, semi-strong, and strong.

The weak form states that using historical data and past prices do not have weight when forecasting future prices and price movements. It is impossible to beat the market in the long run by selecting stocks based on historical data and past prices, this type of analysis is also known as technical analysis. Further on, prices do not hold serial correlation, hence there are no patterns to the prices. So even though they are priced at a fair level, they are expected to change randomly and unpredictably. This brings forth the theory of random walk. The theory implies that stock price fluctuations are independent over time and may be described by a random process (Horne and Parker 1967). The random walk process formally states that stock returns are serially independent, and their probability distribution are constant over time. What this mean is that the best estimate for tomorrows stock price is todays actual price since the movement should be unpredictable and random. The first finding of this conclusion goes over a century back when Bachelier (1900) had concluded that commodity prices follow a random walk. During the 20th century many studies had similar findings, namely that the serial correlation between prices was close to, if not completely, zero.

Accordingly, a strategy revolving around historical prices can therefore not generate excess returns. There is a remarkable amount of empirical evidence supporting this form of efficiency. One may find studies that obtain a significant excess return over the market, however when taking transaction costs into account these returns quickly become obsolete. Later in this section a popular study is pointed out that shows a violation of this form, and has been accepted as an anomaly of the EMH.

The semi-strong form is concerned with whether the prices fully reflect all publicly available information (Fama 1970). Public information consists out of historical price data, financial statements, earning and dividend announcements, stock splits, plans for mergers or takeovers, new issues, competition analysis, and all sorts of macro-economic factors. It is not only limited to direct economic information, for example: the approval of a drug in the pharmaceutics to a scandal regarding a CEO are also regarded as public information that are incorporated in prices. The meaning behind this form of EMH is that an individual should not be able to make gains on stock selection with information that everyone else has. This immediately becomes rather difficult to accept. Think for example about an analyst that is required to select profitable stocks at an investment bank. Not only should this analyst be able to forecast present values of stocks based on financial statements, and relate this to the dynamics of global economic factors. The analyst should also be able to interpret news announcements, of perhaps even completely different fields, and analyze the quantitative impact of such event on the price of the stock. Even though it is difficult to conceive that this form of EMH can hold, there has been strong empirical evidence to show it does, which will be given later.

The strong form is concerned with whether all available information is fully reflected in prices. That is, no individual, regardless of a monopolistic access to information, generates higher returns than others (Fama 1970). What is typically meant with monopolistic access to information is perhaps better understood as inside information. This implies, for example, that managers of a firm are unable to make gains on buying or selling their own company's stock with the detailed information they have in regards to profitable or loss making projects. This form was already proven to not hold by Niederhoffer and Osborne (1966), who showed that monopolistic access to information on unfiled limit orders can generate monopolistic profits.

The three forms of EMH are interdependent from the strong form towards the weak form. Figure 2 below shows the relationship dependence. It shows that the strong form cannot hold without the semi-strong and weak form holding. However, the semi-strong and strong form do not need to hold for the weak form to hold. What can be seen from the requirements from the weak form is that it is in direct conflict with the technical analysis conducted by many practitioners. Selecting specific stocks based on analyzing data should not result in a superior investing strategy that beats the simple buy and hold strategy. One could argue that if the weak form truly holds there is no reason for investment banks and asset management firms to hire portfolio managers, as the buy hold strategy will generate equal returns relatively to the risk undertaken.

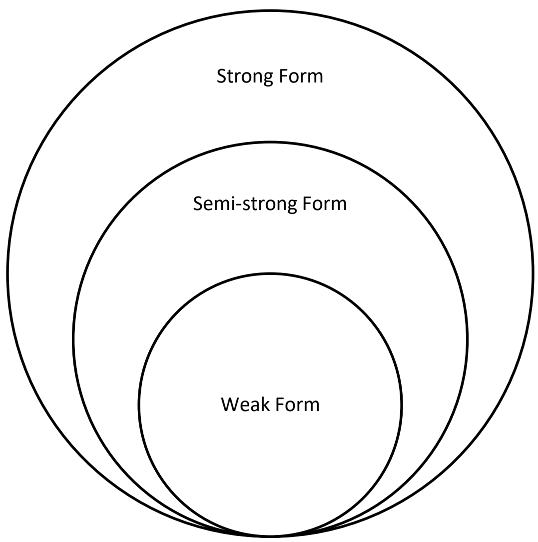


Figure 2: The dependence structure of the Efficient Market Hypothesis.

There is an abundance of literature in the field of finance that challenges the weak form. Various trading strategies have been developed which have shown that excess returns can be generated by selecting certain stocks based on events occurring, stock characteristics, or anomalies. Arguably, the most famous strategy that showed the weak form to not hold is the momentum strategy. Jegadeesh and Titman (1993) had shown that zero-cost portfolios based on buying top decile winners (highest returns) and selling lowest decile losers (lowest returns) generates returns in excess to that of the market. The portfolios that look at 6 month historical prices as reference period and hold the stocks for 6 months had shown to generate annual returns of 12.01 percent. Due to the persistent results this strategy had widely become accepted as the anomaly that shows evidence against the weak form of EMH.

As was described before, the semi-strong form is quite a difficult form to accept. It requires analysts to comprehend the relevance of a vast range of information and change the price of a stock price accordingly. Jensen (1969) had investigated the profitability of mutual funds. He found that on a risk-adjusted basis, the average return was close to zero percent annually. When considering the cost of investment managers, the performance fell to a significant negative 0.9 percent annually. This shows that top investment managers, whom are considered to be able to beat the market based on interpretation of events, are unable to generate profits and even lead to losses when taking their costs into account.

There has been strong evidence that violates the semi-strong form, however. One of the most well-known studies is the one of Bernard and Thomas (1990). They had uncovered the “post-earnings announcement drift”, which means that stock prices, after a company has announced positive (negative) results, tend to drift upwards (downwards) for about a year long. This shows that analysts are not capable of fully reflecting the value of information into stock prices. Still this anomaly has not been explained.

In recent decades, the EMH has received more and more critique due to the narrow assumptions that are required. For example, it assumes that markets are rational and create unbiased forecasts about the future. But when talking about human nature, we never are truly rational and carry many different types of biases. This has been the main intuition for what is known as behavioral finance, which is the paradigm of studying financial markets without the strict assumptions (Ritter 2003). As giving a full description of behavioral finance goes beyond the scope of this thesis, it is advised to read Ritter (2003) for an introduction to the topic. One of the reasons why behavioral finance has received so much attention in academics is due to strength in explaining mispricing in the stock market, and how some investors achieve either significant gains or losses. For example, it has been found that investors have certain cognitive biases such as the heuristic bias as documented by Benartzi and Thaler (2001), who had found that investors simplify investing decisions. Barber and Odean (2001) had found that people tend to be overconfident, especially male investors. Behavioral finance, however, is typically an ex post study of events in which it basically becomes a situation of "what cognitive bias or psychological trait fits this event?". It is and has been a challenge for behavioral finance to predict what bias will dominate in the nearby future. The basic implication that can be derived from behavioral finance studies is that stock prices are not always priced "fairly", and that investor sentiment, for example, can drive prices away from their true values (Baker and Wurgler 2007).

So, given this very small introduction of behavioral finance, it should be clear that it would not be a strange phenomenon if the weak-form of EMH does not hold, which is exactly what this thesis will examine.

2.1.2 The visibility hypothesis

As will be seen in Section 2.2, the visibility hypothesis may be a solid argumentation against the EMH, and in particular be of significance in describing the effects that abnormal trading volume may have over future price movements.

Miller (1977) argues that risk and uncertainty cause for a difference amongst opinions, and that when a market has restrictions on short selling, the demand for a stock will come from a small portion of optimistic investors. Since difference amongst opinions will increase when risk increases, it is expected that risky stocks exhibit lower expected returns. As an increase in well-informed investors should diminish the chance of undervalued stocks, there may be a push of overvalued stocks due to the small portion of optimistic investors who are not well-informed. This imbalance goes against the EMH. This is further recognized by Mayshar (1983), who states that holders of a stock are typically be more optimistic about future developments.

Arbel and Strebel (1982) argued that an increase of investors in the market should increase the value of a stock due to a reduction in uncertainty and risk. They showed that there exists a strong negative relationship between excess returns and the number of institutional holders of the stock. Merton (1987) went on to develop a model of incomplete information to show that stocks with low market capitalization, small investor base, and less known stocks tend to have relative large expected returns. He argues that if such stocks can be identified, an accurate and low-cost analysis can be performed on its value, then an institutional investor can obtain an excess return by developing a strategy that invests in such stocks. It goes without saying, that such actions would drive the price upwards and with a certain increase of investors it would reach an equilibrium value. Investors whom had originally held this stock have therefore seen strong returns on their stock due to it becoming on the radar of institutional investors.

This hypothesis is the backbone of the high-volume return premium. If stocks experience an increase in visibility it is associated with an increase in daily trading volume. An abnormal change in daily trading volume may therefore be considered as a proxy for a shift in visibility, which subsequently leads to an increase in returns. This is exactly what has been found in the NYSE by Gervais et al. (2001), and confirmed by further studies. By studying the returns generated by high (low) volume stocks this thesis will examine if this is indeed the case for the Finnish stock market. It will test whether the returns, as shown in Figure 1, are driven by abnormal daily trading volume, or whether other effects can explain the price movement behavior.

2.1.3 Trading volume theory

Trading volume can also be seen as the number of exchanges that occur in the market. This happens when investors assign different values on the given stock. For example, company XYZ has a stock price of 10 euro, and investor A thinks that the correct price is 11 euro. Investor A will then purchase the stock of investor B who holds the stock, but thinks that the real price should be 9 euro. Investor A will buy the stock for anywhere between 10 to 11 euro, depending on the spread. An exchange has occurred, and trading volume has increased for that day. Depending on several liquidity factors, such an exchange would drive the price of the stock upwards, as there is a demand for higher priced stocks. This sounds like a very simple concept, and is quite a basic economical principle. However, when one spreads their scope and looks at the amount of products stocks being traded on exchanges this quickly becomes a very complex process. What most common financial models do to avoid this problem is just simply removing the possibility of such event. This brings forth the assumption of homogeneity.

When investors are assumed to be homogenic, it is assumed that they have an equilibrium understanding and interpretation of information, and react in a similar fashion. One will see that many economic models have a key assumption that investors are homogenic. However, this does not make sense since there is a lot of trade going on in the financial markets. Not only does trade happen between investors when new information is released, it should also happen when investors revise their liquidity and speculative desires as is argued by Karpoff (1986).

This idea was not fully acknowledged in the past. Verrecchia (1981) had pointed out that there existed no theoretical framework to create a connection between trading volume and the reaction of investors to public information is at most an ambiguous relationship. So he created a model to show that when investors interpret information in a homogeneous fashion there could still be trading volume. For example, consider that the market is in an equilibrium state where all investors have made any adjustments to their portfolios and where supply meets demand. There is no desire to trade in this situation. Now, when new information arrives into the market, all investors should change their expectations in a similar way. This drives the prices accordingly, until it has met the consensus expectation. When prices are driven to the right amount, one could say that prices fully reflect the information. However, in practice, prices fail to reflect this. There are many different investors, each with a different risk apatite, and a price simply reflects the geometric mean of each investor’s expectation that is weighted by their risk apatite (Pratt 1964). So even when there is no new information, investors may still cause for trade as they value a different price in accordance to their portfolio formation.

Academic literature has come a long way since then. A substantial amount of research has been performed on trading volume, and the consensus has become that trading volume is an important aspect in finance. Not only does it provide useful information for practitioners, but it is also a fundamental aspect for academic research. For example, Karpoff (1987) pointed out that, when performing an event study, trading volume is used to infer whether the event had an informational content and whether investors used the information equally. In such event study, price changes would measure whether information has been correctly understood by the market, where trading volume measures investors agree or disagree with one and another. A large change in trading volume would mean that there is disagreement amongst investors. Additionally, when the behavior of changes in prices and trading volume are jointly determined, integrating such joint measure may empower statistical tests of event studies as was shown by Richardson, Sefcik, and Thompson (1986).

Another reason why a deep understanding of the relationship between stock prices and trading volume is a crucial component for event studies comes from the mixture of distribution hypothesis (MDH). When one would sample rates of return over time, the distribution would have strong kurtosis over the normal distribution (Karpoff 1987). Two hypothesis have been capable to explain this. The first one, the Stable Paretian hypothesis, basically states that rates of return fall in a different time of distribution with an infinite variance. The other hypothesis, the MDH, tries to explain the kurtosis by arguing that rates of return data are sampled from a mixture of distributions with different conditional variances (Karpoff 1987).

The proposition of the MDH is that changes in daily prices and trading volume come from the same information flow. When good (bad) news arrives to the market, it will increase (decrease) the price of the stock. Both good and bad news events are associated with abnormal daily trading volume. Thus, return volatility and trading volume must be positively correlated. So, if prices are driven by the random walk process, also known as a stochastic process, and the variance, which is the volatility, is a non-static parameter, then trading volume can be used as a proxy for information flow. This leads to the conclusion that a deep understanding of the stock price and trading volume relationship is key in event studies to measure the changes in volatility from non-event to event window.

Bamber, Barron, and Stevens (2011) had brought the studies on the behavior of trading volume around financial news events under an umbrella to direct future studies into the right way. It had shown that studies focused on price behavior had evolved much stronger than those on trading volume. They argue that the homogeneity assumption has been key to this lack of development since disagreements amongst investors is a key driver for trading volume. Nevertheless, empirical evidence in recent decades has proven that investors have heterogenous expectations. One key empirical study by Gillette, Stevens, Watts and Williams (1999), had shown in an experimental market that investors develop various forecasts based on similar available information, even with cash incentives. They had also found that even when investors are in an equilibrium of expectations, trades still occur. This confirms the original theories mentioned earlier, that trading volume occurs due to dispersion of beliefs and that investors have incentive to trade purely based on their own risk and speculation desires.

This section has provided some of the key developments in the theory of trading volume. There is a vast amount of more theory on trading volume, focusing on various measures of trading volume and on different economical fields. However, as this would go way beyond the scope of this thesis it is recommended to refer to Bamber et al. (2011), and Karpoff (1986) for an overview. The next section will go deeper into the relationship between trading volume and stock price movements, and how the information content of trading volume can be a predictor for price behavior.

2.2 Literature review

The relationship between stock market prices and trading volume has long been studied. A well-known paper from Ying (1966) had shown that an increase (decrease) in volume on the New York Stock Exchange (NYSE) is associated with an increase (decrease) in the price of the S&P 500 Index. Additional findings were that a volume increase (decrease) of five trading days has a tendency to be followed by an increase (decrease) in the price for the following four days. This paper, however, received criticism as the data was not comparable since Ying had used two different indices for the measurements.

The volume to index relationship has also been tested earlier by Granger and Morgenstern (1963). They had shown that in the short-run stock prices do confirm with the random walk hypothesis, however in the long-run other components are at play that are not fully described by the hypothesis. Further on, they found that there was very little impact of trading volume and stock price movements by analyzing the Securities Exchange Commission price index and NYSE trading volume. The consensus on the results from these papers was that there did not exist a useful theoretical framework for the analysis of speculative markets, and that one had to be developed before the importance of volume could be measured.

Epps (1975) constructed such a theoretical framework by grouping transactions into "bulls" and "bears" to predict the behavior of bond price changes and volume. The model showed that stock prices can predict changes in bond price behavior and trading volume. The resulting theory implied that there exists a relationship between volume and price changes on individual transactions. This verified the old Wall Street saying that bull markets tend to have high volume while bear markets observe low volume. It was further reinforced by models developed by Copeland (1976), who studied the demand curve shifts of investors as new information is exposed to them. By studying the increasing function of trading volume in relation to the number of traders Copeland had shown that the subsequent information model predicted a positive correlation between absolute price changes and trading volume, and a positive skewness in the distribution of trading volume. Tauchen and Pitts (1983) further reinforced the theory by developing a model that describes general trends in price variability and volume. Their findings show that the variance of daily price changes and the mean daily trading volume rely on information flow, interpretation of information, and the number of market participants. These models have in common that they show a positive relationship between stock price movements and trading volume.

The theory described above mainly discusses the relationship between trading volume and price movements. This makes perfect sense given that trading occurs when there is a misalignment in the valuation of the asset by buyers and sellers. However, most financial models have the assumption that investors have homogeneous expectations and live in a changing environment. This is difficult to reason with, because, if in fact it would be the case that investors have homogeneous expectations, then there should be little trading activity.

The study on the relationship between trading volume and information by Karpoff (1986), as described in the previous section, assumes that investors have heterogeneous expectations and that trading activity may occur without new information due to individual demands. The findings confirm with the results of Epps (1975) that trading volume increases with the number of shares, decreases with the bid-ask spread, and that trading volume does indeed contain information which may be used to study the disagreement of valuation of stocks by investors.

Smirlock and Starks (1985) had empirically tested whether there is indeed an asymmetrical relationship between stock price movements and trading volume as described in the model of Epps (1975). Using data from all transactions occurring on the NYSE for a 49-day period they had shown that there is strong support on the claim that bull markets carry more volume than bear markets on days of new information arrival. There was, however, no support for trading days when no new information flowed into the market. Overall, the empirical results support the original hypothesis of Epps. The model of Epps was further empirically tested with cross-sectional data and transactional data by Harris (1986 & 1987). Both studies comply with the theoretical model and show that trading volume is a good estimate for the evolution rate of information, and that positive relations exist among measures of stock return and trading volume.

The empirical relationship between trading volume and stock prices are further analyzed by Harris and Raviv (1993). They had built a model to measure the dynamics of trading in a speculative market based on differences of opinions. Their motivation for doing so is that trading occurs when cumulative information switches from positive to negative, or the other way around. Given the assumption that investors and analysts utilize the same information, this implies that there is a difference of opinion on the available information that generates trades. The main results show a positive correlation between absolute price changes and trading volume. Additionally, overestimates (underestimates) of the value changes from new information induces negative (positive) changes in the serial correlation of prices. Together with the study on dispersion of beliefs by Shalen (1993), who had found that dispersion of beliefs explains for the stock price behavior and trading volume, it is concluded that large trading volumes are an indicator for return autocorrelations and absolute price changes.

The relationship of trading volume, volatility and stock returns was further studied on the US and Chinese stock markets by Lee and Rui (2000). The motivation behind this study was to find relationships between trading volume, volatility and stock returns not only in the markets themselves, but also whether one market's trading volume has predictive power over another market's stock returns. The results indicated that there exists very little predictive power in the US stock market on the Chinese stock market. Nevertheless, within the Shenzen B stock market, trading volume does hold predictive power over stock returns. Lee and Rui (2002) investigated the cross-country relationship of trading volume and stock returns in the three largest stock markets: NYSE, London Stock Exchange, and the Japanese NIKKEI 225. According their findings, trading volume does not have predictive power over all the three stock markets. There does exist, however, a clear relationship between trading volume and volatility in all three markets. Additionally, the US market's volume was shown to have strong predictive power on both the London Stock Exchange and the Japanese NIKKEI 225. This implies that information flows from the US markets to other markets and that trading volume related cross-country relationships exist.

The fact that trading volume carries unique valuable information has been well documented by Blume, Lawrence, Easley and O'hara (1994). They studied the strength of technical analysis, mainly whether stock price and trading volume data can be used to derive information that predicts future returns. Under the assumption that markets are efficient, as in accordance with the EMH, the process of price adjustments of stocks may still have a small delay and their movement may therefore be forecasted slightly. By studying the statistical properties of volume, they showed that trading volume provides an informational role that is unlike the informational role stock prices provide. Trading volume had shown to provide information on the quality of information signals. As the distribution of trading volume differs from stock price distributions a trader may use trading volume to distinct information that is not captured by stock prices. Traders who do so have a higher performance than traders who do not.

Campbell, Grossman, and Wang (1993) investigated the effects of large trading volume on return serial correlations. They had documented that daily serial correlations of returns are lower on high-volume days than on low-volume days, persistent through multiple stock indexes and individual stock returns. This paper has not investigated long-term effects; however, results had shown that on the short-term there exists a negative relationship between trading volume and return serial correlations. Using Lehmann's (1990) contrarian trading strategy, Conrad, Hameed, and Niden (1994) investigated the relationship between volume and return patterns as proposed by Blume et al. (1994), and Campbell et al. (1993). Their findings show a strong evidence for a relation between trading volume and serial autocorrelations in weekly returns, as high-transaction securities show price reversal while low-transactions securities show positive return serial correlations.

A study by Lo and Wang (2006) examined the effects of trading volume in an Intertemporal Capital Asset Pricing Model (ICAPM) framework. They argue that any asset pricing model that applies economic factors to their model should also use trading volume due to the structural link between stock prices and trading volume, as described in the above paragraph with the study of Campell et al. (1993). Many studies on asset pricing have been focusing on stock returns and prices, while omitting the trading volume variable and its information content have been ignored. By using an ICAPM framework, Lo and Wang constructed a hedging portfolio based on weekly trading volume for stocks listed on the NYSE and AMEX. The hedging portfolio had shown to outperform other hedging portfolios based on different predictors. The hedging portfolio held predictive power on future market returns. Its explanatory power on cross-sectional variations was of par to that of market betas, the SMB and HML factor (Fama and French, 1992).

Lu and Lee (2016) examined the effect of a volume based trading rule during downtrends at toughs and uptrends at peaks. Their results show that their trading rule consistently achieved higher average returns, implying that abnormal trading volume provides information. Cooper (1999) researched the overreaction hypothesis on stocks of large firms by forming portfolios that only invest in stocks which have experienced movements in lagged returns and a growth in volume. The results indicated that high growth in volume stocks show weaker price reversals, and even positive serial correlations, the reverse was true for low growth in volume stocks.

Furthermore, momentum studies have also seen a relationship between volume and stock returns. Lee, Charles and Swaminathan (2000) had found that volume predicts the intensity and longevity in their momentum strategies. In a study on the impact of noise on price returns Hoitash and Krishnan (2008) had constructed a specific measure based on autocorrelation in daily volume and found that the measure has positive and significant impact on stock returns in momentum strategies.

Miller (1977) argued that high volume does not predict positive price movements, and that the occurrence of high volume should not induce anyone to act. However, as volume brings attention more investors are exposed to the stock and may buy it if they assign different values based on their analysis. This brings forth the visibility hypothesis, further described in Section 2.3, which states that when a stock suddenly gains attention by a volume shock more analysts will research the stock and adjust the price per their findings. This hypothesis is strongly reinforced by the results of by Gervais et al. (2001) who examined the high-volume return premium in the NYSE, which is the main motivator for this thesis. They argued that when a stock experiences a shock in its trading activity, its visibility is affected. It will show up on more radars and more analysts will be drawn to study any relevant information to stock. The findings may therefore shift the demand of this stock, and subsequently alters the price. Effects such as return autocorrelations, liquidity, and firm announcements did not appear to have influence on this effect. Additionally, systematic risk could not explain for the results obtained by investing in abnormal volume stocks. The findings brought forth the notion of high-volume return premium.

The high-volume return premium received attention and was further investigated on a global level by Kaniel, Ozoguz, and Starks (2007). They had investigated 41 countries for the existence of high-volume return premium. Their results showed that the return premium exists in most of the developed stock markets and several emerging markets. Their conclusions were, too, that the return premium is not caused by changes in systematic risk or liquidity. It is also persistent during a wide range of economic conditions. Further conclusions were that this strategy would be profitable for a retail investor due to the implicit transaction costs. Zhou (2008) observed that the return premium exists in the Chinese stock market, and that it is favored for small-size stocks. Furthermore, he constructed volume momentum portfolios by sorting on high trading volume and positive (negative) returns as winners (losers), but finds that this strategy generates negative returns. Tang, Zou, and Li (2013) showed that it also exists in the Australian stock market, however it is persistent in large firms, which contradicts the findings of Gervais et al. (2001), and Zhou (2008), who find it to be strong in small to medium size firms and weak in large firms.

3. METHODOLOGY AND DATA

In this chapter the methodology and data used for analysis is described. Section 3.1 will state the null hypothesis that will be tested. Section 3.2 describes how the trading intervals are constructed, and Section 3.3 describes the data that goes into the trading intervals. As there may be some computational biases from using the data straightforwardly, a screening will take place. This is described in Section 3.4. The portfolio formation is described in Section 3.5.

3.1 Hypothesis

The objective is to see whether trading volume provides any forecasting power in predicting returns on the Finnish stock market. High (low) volume is expected to provide high (low) returns. If buying high (low) volume stocks provide high (low) returns it shows that, within the Finnish stock market, a high-volume return premium is present. This implies that the weak form of the EMH does not hold as historical data is used to forecast the movement of future stock prices, which leads to the null hypothesis:

H0: Trading volume does not contain predictive power on the movement of future returns in the Finnish stock market.

H1: Trading volume does contain predictive power on the movement of future returns in the Finnish stock market.

3.2 Trading interval

This strategy will take on long (short) positions on high (low) volume stocks. To do so, a trading interval is created. The trading interval is divided up into three parts; a reference period (49 days), a formation period, and a test period (50 days). On the formation period the trading volume of a stock is measured if it is in the top (bottom) decile of trading volume of the reference period and formation period combined. The stock will then be placed in a long (short) portfolio, without rebalancing, for a period of 1, 10, 20, 50, to 100 trading days. There will be two portfolio formation procedures: zero-cost portfolios and reference return portfolios. The formation procedure of these portfolios is described in Section 3.5. Figure 3. displays how the trading interval is constructed.

During the trading interval, a stock will be classified to be of high (low) volume by measuring whether the trading volume of that day was in the top (bottom) decile of the 49 days before the trading interval, which is the reference period. The stock will then be placed in a portfolio on the formation period, which is held for a period of 1, 10, 20, 50, or 100 trading days, without rebalancing, which is called the test period. The returns will then be sorted on firm sizes, determined by end year market capitalization.

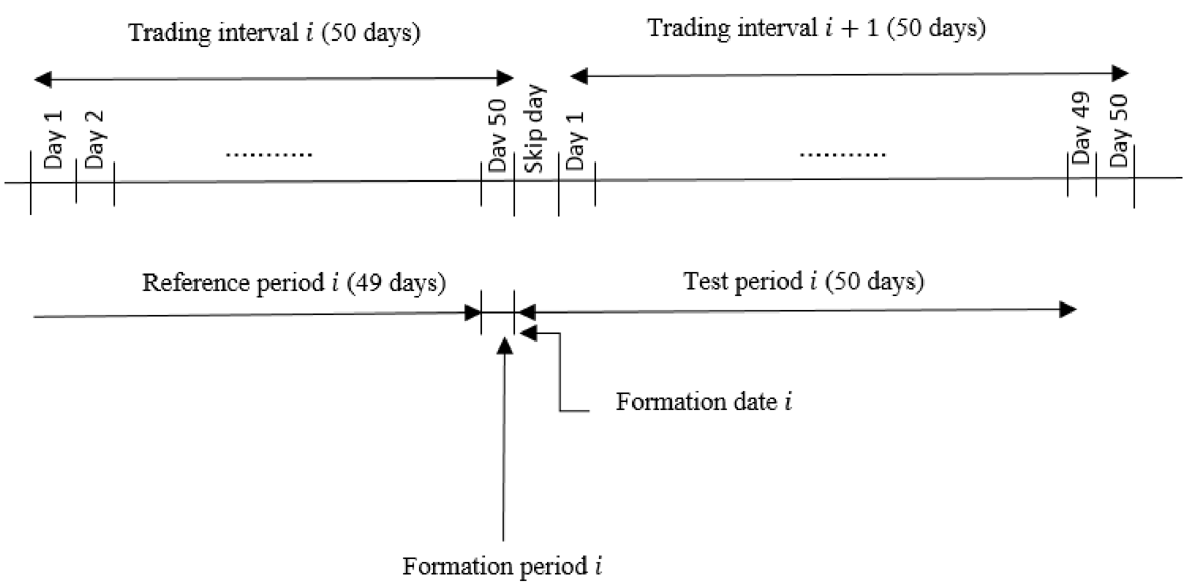


Figure 3: Trading interval formation. The trading interval consists out of 50 days, of which the first 49 days are used as a reference period for measuring trading volume levels, and on the final day a stock is compared to the reference period to place it in to the portfolio. Portfolios are then held for the test periods' length to measure the returns. Trading intervals do not overlap and skip 1 day after formation.

Evidence by Gibbons and Hess (1981), Dubois and Louvet (1996) show that the distribution of stock returns varies according to what day of the week it is. The main motivation has been that Monday's return is computed over three days, as on Saturday and Sunday the exchanges are closed. Zhang, Lai, and Lin (2017) further show that the distribution of returns varies not only on Mondays, but other days may also display different distributions. They had shown that in the 25 countries of their study all countries show different return distributions on a specific day of the week. Therefore, to avoid the day of the week bias, the trading interval will skip one day after each formation period to ensure that the formation period does not fall on the same day as pervious'.

3.3 Data

The data is obtained via DATASTREAM and consists out of all (de)listed stocks on the Helsinki OMX from April 1994 to June 2016, which accounts for a total of 228 stocks, and total of 5706 daily observations. This provides for a total of 111 nonintersecting trading intervals of 50 days. Weekly data will also be studied, which is described later. During every trading interval, every single stock that is listed on the Finnish stock market, the Nasdaq Helsinki, will be considered. Summary statistics for these stocks are given in Table 1. Stocks are categorized into three types: small, medium, and large firms. This is done based on the end year market capitalization, which is calculated by multiplying the end year outstanding shares by the share price. Stocks that fall in decile nine and ten are categorized as large firms, whereas firms from the sixth to the eight deciles are medium firms, and firms that fall below decile six are small firms.

Table 1 shows that the average stock price for small firms is not only higher than medium sized firms, but almost equal to that of large priced firms. This is due to a few outliers on the stock exchange. The initial listing of a few stocks had been overvalued, which lead to their respective stock price dropping from over 100 euro down to 10 euro in a matter of a few weeks. When removing these outliers, the average stock price for small firms lies around 5 euro. The data shows that large firms are traded more frequently in the beginning of the sample period, as can be seen from the number of large firms in the first trading interval.

There are 35 large firms, versus only 6 small and 5 medium sized firms. During the progressing years trading intervals start to see a more evenly distributed amount of firm sizes. With a final distribution of 44 small, 54 medium, and 42 large sized firms.

Table 1: Descriptive statistics of the daily data of the Finnish stock market. The data contains a total of 111 nonoverlapping trading intervals, all 50 days long. Stocks are classified into three types of stocks based on the market capitalization of end year, in which the trading interval was constructed. Large firms are those in decile ten and nine, medium firms are those in decile eight, seven, and six, all bellow that are allocated to the small firm group.

	Small firms	**Medium firms**	**Large firms**
Overall sample of 111 intervals			
Average stock price	€ 12.53	€ 7.92	€ 12.64
Median stock price	€ 4.93	€ 4.46	€ 8.92
Average daily volume	7619	7559	114806
Median daily volume	430	840	17350
First interval 25-04-1994			
Number of stocks	6	5	35
Average stock price	€ 6.28	€ 3.87	€ 9.75
Median stock price	€ 6.73	€ 1.68	€ 5.66
Average daily volume	1980	150	38218
Median daily volume	1980	150	1355
Last interval 14-05-2016			
Number of stocks	44	54	42
Average stock price	€ 11.74	€ 8.53	€ 19.49
Median stock price	€ 5.18	€ 5.94	€ 17.29
Average daily volume	6637	6134	60578
Median daily volume	760	995	18210

From Table 2 it can be seen what the characteristics are for the trading intervals. On average, there are more medium sized firms than larger and smaller firms in the trading intervals, respectively. It occurs at least once for every size category that there are no firms of that size in a trading interval. Small firms see a maximum of 5 low and 10 high volume stocks for a certain interval. Whereas medium firms see a maximum amount of 35 stocks for low volume, and 16 stocks for high volume. Large firms have a maximum of 6 stocks for both low and high volume.

There exists a positive correlation between low and high volume stocks for small firms, while it is negative for medium and large firms. As this correlation is relatively small, in comparison to Gervais et al. (2001) and Tang et al. (2013), no additional measures are applied. Tkac (1999) argues that to avoid computation biases generated by these correlations that a different measure of volume ought to be applied. Namely the fraction of market volume of which a given stock accounts for. However, when this was applied by Gervais et al. (2001) it yielded no significant changes, even on correlations three times as high.

Table 2: Summary statistics of trading intervals. Stocks are classified into three types of stocks based on the market capitalization of end year, in which the trading interval was constructed. Large firms are those in decile ten and nine, medium firms are those in decile eight, seven, and six, all bellow that are allocated to the small firm group.

	Small firms		**Medium firms**		**Large firms**	
	Low	High	Low	High	Low	High
Average	1.34	1.91	5.79	3.89	2.75	3.05
Median	1	1	4	3	3	3
Standard deviation	1.24	2.09	5.77	3.12	1.34	1.54
Minimum	0	0	0	0	0	0
Maximum	5	10	35	16	6	6
Correlation high-low	0.29		-0.23		-0.19	

Weekly data follows the same procedures as daily data. Daily trading volume data is aggregated to weekly points, consistently on Wednesdays. The trading intervals are therefore constructed by using 10 weeks, which equals 50 trading days, for the reference period. The holding period will exist out of 1, 10, 20, 50, and 100 trading days. There is a total of 101 trading intervals using weekly data. The stocks are again sorted into small, medium and large firms. On the interval day stocks that are in the top (bottom) ten percent of the previous 9-week reference period will be placed in the high (low) volume classification. The stocks are again classified into small, medium, and large sized firms based on firms' year end market capitalization. The returns are then computed in accordance with the respective holding period.

3.4 Screening of data

It has been well documented in the past that low stock prices can cause for biases when computing returns. Blume and Stambaugh (1983), Conrad and Kaul (1993), have shown that computing returns may be biased by the bid-ask spread bounce and price discreteness that low-priced stocks can experience. The bid-ask spread bounce is a condition of volatility, where a stock price bounces, in high frequency, back and forth within the range of the bid-ask spread. Low priced stocks exhibit higher exposure to this event due to the relative impact of bid-ask spread on the price of low priced stocks. Out of the 23.7 percent total computational bias, low priced stocks below five dollars contributed to 18.7 percent of the bias.

Further on, Bhardwaj and Brooks (1992) had found that the January effect is primarily a low stock price effect. Therefore, to avoid any return computational bias, stocks that experience a drop below a five-euro threshold during the reference period will be exempted from the formation period. This procedure does eliminate a significant amount of stocks from the sample. On the last trading interval, there are a total of 140 stocks, but 17 are eliminated as they fall below the five-euro threshold. This is indeed a significant reduction in the total number of stocks used in the sample, however when not excluding these the results reached fifty times as high, clearly showing that small stocks bias the overall sample.

Since stocks that are newly listed on the stock exchange may experience unusual price and volume behavior, any stocks that are younger than 252 trading days will be dropped from the trading interval. Similarly, any stocks that experience any delisting within the preceding 252 trading days when the trading interval is constructed are also exempted as they may bias the results.

3.5 Portfolio formation

Two different type of portfolios will be used to investigate whether trading volume carries predictive power on future returns. These portfolios will be created by taking on a long (short) position in high (low) volume stocks on the formation date, and will be held during the extend of the holding period (1, 10, 20, 50, 100 days) without rebalancing, which is in line the approach of Gervai et al. (2001).

The first portfolio type is the zero-cost portfolio. This portfolio will take on a total position of one euro long in all high-volume stocks, and one euro short in all low volume stocks. This effectively cancels out the long position cost by the short position gain, hence zero cost. The stocks in each portfolio carry equal weight. The returns generated during the test period are denoted as R_i^H for high volume and R_i^L for low volume, where i stands for trading interval {1, 2, 3, ..., I }. The net returns are denoted as below.

$$(1) \quad NR_i = R_i^H + R_i^L$$

This represents the returns generated by the zero-cost portfolio. These net returns should actually be seen as profits rather than returns, as the initial investment equals zero, thus any percentage return becomes infinite by default. Nevertheless, for simplicity of reading, it will be referred to as net return. The null hypothesis will be tested by verifying whether the average net return is significantly positive:

$$(2) \quad \overline{NR}_i = \frac{1}{I} \sum_{i=1}^{I} NR_i$$

The calculation $\frac{1}{I}$ is the equal weight that is assigned to each portfolio, where I represents the total amount of trading intervals.

The second portfolio is the reference return portfolio. Which has size-adjustments in line with the numbers of extremely traded stocks. The weight of this portfolio is adjusted by the number of stocks that are classified as high or low volume stocks for the given interval. This means that an interval experiencing a large quantity of different stocks that have high volume will be weighted more heavily. The cost of going long in high volume stock is still offset by investing an equal amount, one euro, into low volume stock. Therefore, the cost of setting up the portfolio remains zero. Like the zero-cost portfolio, this portfolio is not rebalanced during the test period. The weight calculation for the reference return portfolio is as follows:

(3) $$W^H = \frac{1}{I} * [1 + (M_i^H - \overline{M}^H)/M^H$$

(4) $$W^L = \frac{1}{I} * [1 + (M_i^L - \overline{M}^L)/M^L$$

Where W^H and W^L represent the size-adjusted weight for high and low volume stocks respectively. The calculation between the brackets is the coefficient for adjustment, where M_i^H and M_i^L are the number of high and low volume stocks in interval i respectively. The notation of $\overline{M}^H$ and $\overline{M}^L$ represents the average number of high and low volume stocks for each interval, M^H and M^L represent the total number of high and low volume stocks, respectively. When the size-adjusted weight is obtained, the returns for the high and low volume stocks can be obtained by the following equations:

(5) $$\overline{R}^H = \frac{\sum_{i=1}^{I} \sum_{j=1}^{M_i^H} R_{i,j}^H}{\sum_{i=1}^{I} M_i^H}$$

(6) $$\overline{R}^{L} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{M_i^L} R_{i,j}^{L}}{\sum_{i=1}^{I} M_i^L}$$

And for the net return:

(7) $$\overline{NR} = \frac{\sum_{i=1}^{I}(\sum_{j=1}^{M_i^H} R_{i,j}^{H} + \sum_{j=1}^{M_i^L} R_{i,j}^{L})}{\sum_{i=1}^{I}(M_i^H + M_i^L)}$$

Where the null hypothesis will be tested whether the net return of the reference return period is significantly greater than zero.

4. RESULTS

The main results for the daily data are provided in Table 3, and for the weekly data it can be found in Table 4. They both show the average cumulative returns for the zero-cost and reference return portfolio for the given holding period, categorized as small, medium, and large firms. Between the brackets, for the net return rows, the t statistics are presented. As the long and short positions are not of interest to study, these t-statistics are not presented.

As can be seen from the daily data in Table 3. small sized firms have a return around zero for high and low volume stocks. This holds for all holding periods in the zero-cost portfolio, and for the reference return portfolio. The net return of the small sized firms is therefore, too, floating around zero. Additionally, none of the holding periods in both type of portfolios show any significance. It can therefore be concluded that there exists no high-volume return premium in the Finnish stock market for small sized firms.

Medium sized firms display a small positive return, on average, in high volume stocks during the first day. This turns into negative returns onwards, and is more severe than small sized firms during the first 20 days. Low volume stocks tend to be generating slightly negative returns throughout all holding days, in an increasing fashion. Thus, the net returns for medium sized firms vary from positive to negative across the holding days. But, again, these returns are not significant, and therefore there exists no high-volume return premium for medium sized firms.

Large sized firms with high volume show an increasing positive return until a 100-day holding period, where it has decreased slightly from 0.14 percent on 50 days to 0.13 percent on 100 days in the zero-cost portfolio. The reference return portfolio follows a similar pattern, where returns drop from 0.21 percent on 50 days to 0.12 percent on 100 days. The net returns from large sized firms are positive for all zero-cost portfolios, and negative in the first 20 days for the reference return portfolio. Only one portfolio shows a positive significant return of 0.16 percent, which is the 50-day holding period large sized firms' portfolio. This implies that there may exist a high-volume return premium for large sized firms on the 50-day horizon.

Table 3: Average returns for the zero-cost and reference return portfolio for daily data. Stocks are classified into three types of stocks based on the market capitalization of end year, in which the trading interval was constructed. Large firms are those in decile ten and nine, medium firms are those in decile eight, seven, and six, all bellow that are allocated to the small firm group. Stocks are held for 5 test periods: 1, 10, 20, 50, and 100 holding days, without rebalancing. The symbols, *, **, and *** denote a statistical significance at the 10%, 5%, and 1%, respectively.

	Zero-cost					Reference return				
Holding period (days)	1	10	20	50	100	1	10	20	50	100
Small firms										
High volume ($\overline{R}^H$)	-0.001	0.006	-0.005	-0.074	-0.135	0.004	0.004	-0.026	-0.179	-0.335
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Low volume ($\overline{R}^L$)	0.000	0.008	-0.034	-0.032	-0.101	-0.002	0.004	-0.067	-0.095	-0.314
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Net return ($\overline{NR}$)	-0.001	0.014	0.029	-0.042	-0.034	0.006	-0.001	0.041	-0.084	-0.021
	(-0.008)	(0.953)	(0.990)	(-0.848)	(-0.451)	(0.325)	(-0.016)	(0.670)	(-0.609)	(-0.088)
Medium firms										
High volume ($\overline{R}^H$)	0.009	-0.016	-0.023	-0.010	-0.010	0.016	-0.025	-0.037	-0.009	-0.102
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Low volume ($\overline{R}^L$)	-0.001	-0.004	-0.017	-0.031	-0.048	0.026	0.061	0.030	-0.027	-0.053
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Net return ($\overline{NR}$)	0.010	-0.012	-0.006	0.021	0.038	-0.009	-0.086	-0.067	0.018	-0.050
	(1.008)	(-0.513)	(-0.182)	(0.377)	(0.469)	(-0.365)	(-1.247)	(-0.610)	(0.092)	(-0.287)
Large firms										
High volume ($\overline{R}^H$)	0.008	0.041	0.070	0.143	0.133	0.027	0.091	0.131	0.216	0.122
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Low volume ($\overline{R}^L$)	0.004	0.026	0.026	0.020	-0.036	0.092	0.228	0.266	0.084	-0.095
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Net return ($\overline{NR}$)	0.003	0.015	0.044	0.163	0.169	-0.065	-0.137	-0.135	0.132	0.217
	(0.182)	(0.330)	(0.737)	(2.196)**	(1.494)	(-0.926)	(-0.533)	(-0.416)	(0.421)	(0.486)

Table 4: Average returns for the zero-cost and reference return portfolio for weekly data. Stocks are classified into three types of stocks based on the market capitalization of end year, in which the trading interval was constructed. Large firms are those in decile ten and nine, medium firms are those in decile eight, seven, and six, all bellow that are allocated to the small firm group. Stocks are held for 5 test periods: 1, 10, 20, 50, and 100 holding days, without rebalancing. The symbols, *, **, and *** denote a statistical significance at the 10%, 5%, and 1%, respectively.

	Zero-cost					Reference return				
Holding period (days)	1	10	20	50	100	1	10	20	50	100
Small firms										
High volume ($\overline{R}^H$)	-0.005	0.004	0.016	0.005	-0.052	-0.024	0.005	0.026	-0.031	-0.129
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Low volume ($\overline{R}^L$)	0.001	-0.007	-0.020	-0.060	-0.083	0.008	-0.020	-0.048	-0.161	-0.230
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Net return ($\overline{NR}$)	-0.006	0.010	0.036	0.065	0.031	-0.032	0.025	0.073	0.131	0.101
	(-0.913)	(0.790)	(1.940)*	(2.023)**	(0.652)	(-1.223)	(0.739)	(1.723)*	(1.603)	(0.797)
Medium firms										
High volume ($\overline{R}^H$)	0.016	0.024	0.035	0.023	0.027	0.011	0.041	0.045	0.010	-0.003
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Low volume ($\overline{R}^L$)	-0.009	-0.011	-0.020	-0.059	-0.121	-0.011	-0.011	-0.040	-0.134	-0.255
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Net return ($\overline{NR}$)	0.025	-0.035	0.055	0.081	0.148	0.022	0.053	0.085	0.144	0.253
	(1.984)**	(1.647)*	(1.940)	(1.364)	(1.768)*	(0.951)	(1.451)	(1.590)	(1.246)	(1.562)
Large firms										
High volume ($\overline{R}^H$)	0.035	0.097	0.132	0.136	0.099	-0.089	0.201	0.264	0.269	0.192
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Low volume ($\overline{R}^L$)	-0.001	0.014	-0.003	-0.102	-0.048	0.022	0.048	-0.034	-0.316	-0.101
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Net return ($\overline{NR}$)	-0.035	0.083	0.136	0.238	0.147	-0.111	0.153	0.299	0.585	0.294
	(-1.646)*	(1.656)*	(2.006)**	(1.827)*	(0.905)	(-1.743)*	(0.906)	(1.341)	(1.338)	(0.587)

The results from applying weekly data to the procedure are shown in Table 4. High volume stocks from small sized firms display a return close to zero during the first 50 holding days. At a 100-day holding period they achieve a negative return of -0.05 percent. Low volume stocks, besides on the initial day, display negative and decreasing returns all the way up to a 100-day holding period. For a 20-day and 50-day holding period, the zero-cost portfolio generates significant positive net returns of 0.04 and 0.07 percent. The reference return portfolio generates significant net returns of 0.07 percent at the 20-day holding period. This implies that for small sized firms there exists a high-volume return premium for a 20-day holding period based on weekly trading volume.

Medium sized firms have positive returns across the high-volume stocks during all holding periods. Similarly, low volume stocks have negative returns across all holding periods. The first day for the zero-cost portfolio a significant positive net return of 0.03 percent is obtained. This turns to a significant negative net return of -0.04 percent. For a holding period of 20 and 50 days there appears no significant net return. At a 100-day holding period, however, the net return becomes positive and significant at 0.15 percent. It appears that a high-volume return premium is present in medium sized firms for a 1, 10, and 100 day holding period.

High volume stocks of large sized firms show an increasing positive return up to 0.14 percent at a 50-day holding period, this slightly decreases to 0.1 percent at a 100-day holding period. Low volume stocks display a barely negative return on the initial day, which is followed by a positive return of 0.01 percent at the 10-day holding period. The consecutive holding periods display negative returns, as expected. Large sized firms display the strongest results with significant net returns across a 1, 10, 20, and 50-day holding period. The first day shows a negative return of 0.04 percent, however this is followed by positive returns of 0.08, 0.14, and 0.24 percent for the 10, 20, and 50-day holding period, respectively.

The results show that there is weak evidence for a high-volume return premium in daily data. Small sized and medium sized firms appear to follow a random pattern of returns around zero. Large sized firms do seem to follow an increasing trend across the holding days, with a 50-day holding period showing significantly positive returns. These findings are inconsistent with those of Gervais et al. (2001), Zhou (2009), Kaniel et al. (2012), and Tang et al. (2013). However, as suggested by Kaniel et al. (2012), who had found some inconsistencies in the presence of high-volume return premium in markets where there are fewer listed entities. Countries with similar demographics, such as those in the Nordics, showed no significant net returns on their 20 day holding period portfolios. Nevertheless, weekly data on trading volume shows much stronger support for the high-volume return premium. This is in line with the previous mentioned papers, where they find improvements of the returns on weekly data. Although the daily sample has shown weak support for the high-volume return premium, the weekly data does support this, thus the null hypothesis is not yet accepted or rejected. Further analysis is required to test whether the returns are truly driven by high volume, or if other effects are into play. This is exactly what the next chapter will cover. First an analysis is performed on the short horizon effect of returns, namely the autocorrelations in returns and volume. This is followed by examining the effect extreme values have on the returns, and finally a test on whether shifts in systematic risk drive the returns.

5. OTHER POTENTIAL EXPLANATIONS

In this chapter, other potential explanations for the results are tested. Section 5.1 will test whether the returns from the trading intervals are driven by autocorrelations in returns. Section 5.2 will remove extreme values from the trading intervals to see how this impacts the net returns, and Section 5.3 tests the returns against systematic risk.

5.1 Autocorrelations in returns and volume

To examine whether short term autocorrelations in returns and trading volume influence the short-term horizon the following analysis is performed. From the original trading intervals constructed from the daily and weekly sample, any extreme returns are omitted, which gives room for “normal” returns. By comparing the returns on the formation date to the returns generated in the reference period, any return which has an extreme value in comparison to the reference period, is omitted. Hence the remaining returns that are in the trading interval are called normal returns. By doing so, the effect of extreme returns on the returns generated during the trading intervals are removed.

There are two types of normal returns: the middle 40 percent, and the middle 20 percent. For the middle 40 percent returns, any return which is higher (lower) than the top (bottom) 30 percent of that stocks’ return in the reference period will be omitted. For the middle 20 percent, the cutoff is at the top (bottom) 40 percent of that stock’s return in the reference period. This procedure is applied to all the previous constructed trading intervals for both the daily and weekly sample. Table 5. displays the net returns of these trimmed portfolios. These results will be compared to the original results from Table 3 and 4.

Unsurprisingly, the data shows that all the net returns generated for the daily sample are insignificant, just as in those of the original sample. The original sample did show a significant net return for large firms at a 50-day holding period, however. But since this analysis is interested on the short-term effects, it does not include any portfolio exceeding

Table 5: Net returns of normal stocks for daily and weekly sample. Same procedure for trading intervals as described in Tables 3 and 4 is applied. Stocks are then further screened based: stocks experiencing extremely high (low) returns in each interval are omitted. This is done on basis of top (bottom) 30 percent, and top (bottom) 40 percent of the test period. The symbols, *, **, and *** denote a statistical significance at the 10%, 5%, and 1%, respectively.

	Zero			**Reference**		
Holding period (days)	**1**	**10**	**20**	**1**	**10**	**20**
			Daily sample			
Small firms						
Middle 40%	-0.011	0.012	-0.056	0.000	0.000	0.000
	(-0.384)	(0.427)	(-0.766)	(-0.242)	(0.796)	(0.327)
Middle 20%	-0.011	0.014	-0.060	0.000	0.000	0.000
	(-0.342)	(0.461)	(-0.781)	(-0.233)	(0.802)	(0.324)
Medium firms						
Middle 40%	0.009	-0.022	-0.028	0.000	0.000	0.000
	(0.678)	(-0.639)	(-0.739)	(0.313)	(-0.543)	(-0.853)
Middle 20%	0.009	-0.022	-0.027	0.000	0.000	0.000
	(0.642)	(-0.624)	(-0.732)	(0.277)	(-0.528)	(-0.846)
Large firms						
Middle 40%	-0.003	0.018	0.018	0.000	0.000	0.000
	(-0.267)	(0.503)	(0.342)	(-0.323)	(0.496)	(0.354)
Middle 20%	-0.007	0.003	-0.004	0.000	0.000	0.000
	(-0.624)	(0.076)	(-0.070)	(-0.657)	(0.084)	(-0.050)
			Weekly sample			
Small firms						
Middle 40%	-0.011	0.009	0.010	0.000	0.000	0.000
	(-0.850)	(0.211)	(0.199)	(-1.233)	(-0.099)	(-0.279)
Middle 20%	-0.001	0.015	0.013	0.000	0.000	0.000
	(-0.116)	(0.317)	(0.251)	(-0.610)	(-0.086)	(-0.271)
Medium firms						
Middle 40%	-0.015	0.012	-0.001	0.000	0.000	0.000
	(-0.825)	(0.423)	(-0.022)	(-0.546)	(1.115)	(-0.285)
Middle 20%	-0.016	0.021	0.005	0.000	0.000	0.000
	(-0.859)	(0.742)	(0.177)	(-0.584)	(1.759)	(0.017)
Large firms						
Middle 40%	0.020	0.014	0.085	0.000	0.000	0.001
	(2.273)**	(0.417)	(2.033)**	(2.281)**	(0.387)	(2.071)**
Middle 20%	0.014	0.017	0.072	0.000	0.000	0.001
	(1.689)*	(0.566)	(1.853)*	(1.707)*	(0.525)	(1.880)*

the 20-day holding period. For the weekly sample, however, there are more interesting results. On the initial day, normal returns in both the middle 40 percent, and middle 20 percent for large sized firms, generate a significantly positive net return of 0.02 percent and

0.01 percent, respectively. Also, the 20-day holding period shows significant returns with 0.09 percent in the middle 40 percent, and 0.07 percent in the middle 20 percent. These returns are lower than the originally documented 0.14 percent, so it shows that some of the extreme returns were indeed generated by return autocorrelations. Nevertheless, it does show that the net returns from the originally constructed trading intervals for the weekly sample are not completely driven by return autocorrelations.

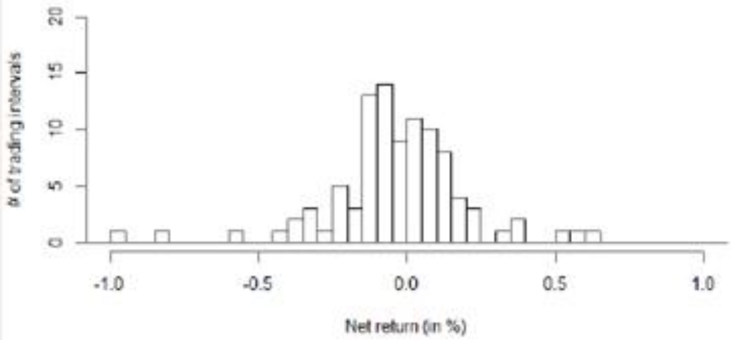
The net returns have been economically insignificant from the originally constructed trading intervals, and the net returns as shown in Table 5 show an even further reduction of these net returns. It is therefore clear that no capital gains can be made on selecting stocks with abnormal volume in the Finnish stock market. However, as large sized firms in the weekly sample have significant returns, a further analysis is performed to determine if these are indeed driven by abnormal trading volume, or if this is an effect caused by outliers.

5.1 Firm announcements and outliers

To examine the effect that extreme values have on the zero-cost strategy, a closer look is taken to the distribution of returns on the sample. Figure 4. displays the net returns obtained from the zero-cost portfolios from a 20-day holding period. The returns are categorized into the three firm sizes, of which the distributions are given in the histogram on the left, with summary statistics of the distribution on the right. It shows that the small and medium sized firm returns have an uneven distribution, as can be seen by the negative skewness of -2.8 and -2.1 respectively. Large size firm returns are relatively evenly distributed with an obsolete skewness value.

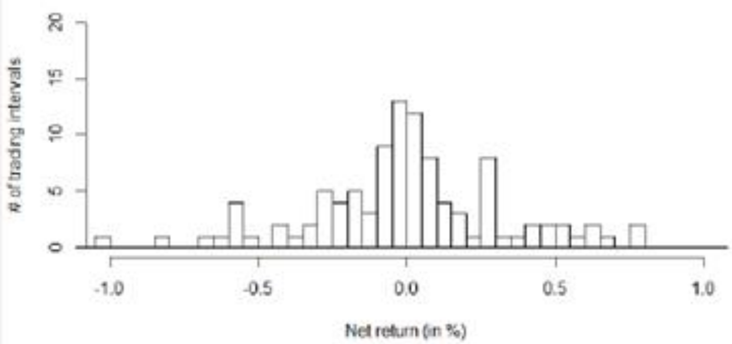
The trimmed mean represents the mean with top (bottom) 2.5 percent of the returns removed from the sample. It is of main interest to see how the mean value changes when removing these outliers. The mean changes from -0.04 to -0.02 for small sized firms, which is half as low. This shows that the average negative net returns generated by small sized firms are driven quite strongly by outliers. Medium sized firms see a change from -0.04 to -0.01, which is an even bigger change. Therefore, it can be concluded that the returns of medium sized firms are largely driven by outliers, too. Large sized firms, however, change from 0.1 to 0.6.

Small firms



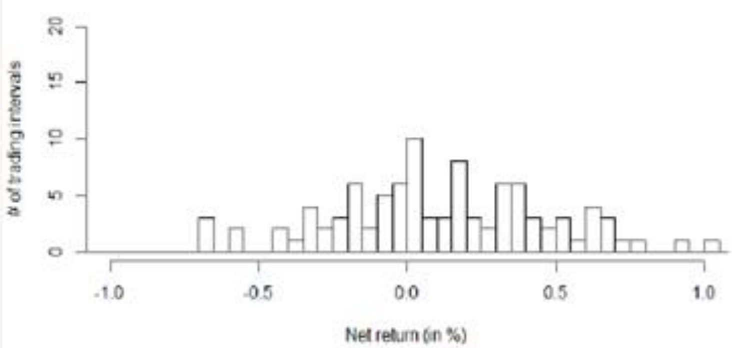
Minimum	-2.5090
Maximum	1.2438
Mean	-0.0445
Lower quantile	-0.1244
Median	-0.0223
Upper quantile	0.0862
Standard deviation	0.3425
Skewness	-2.8452
Trimmed mean	-0.0223

Medium firms



Minimum	-3.2605
Maximum	1.6123
Mean	-0.0417
Lower quantile	-0.1975
Median	-0.0100
Upper quantile	0.1707
Standard deviation	0.5384
Skewness	-2.0841
Trimmed mean	-0.0100

Large firms



Minimum	-2.8139
Maximum	3.1757
Mean	0.0971
Lower quantile	-0.1759
Median	0.0641
Upper quantile	0.3804
Standard deviation	0.6978
Skewness	0.0178
Trimmed mean	0.0641

Figure 4: Return distribution for small, medium, and large sized firms Stocks are classified into three types of stocks based on the market capitalization of end year, in which the trading interval was constructed. Large firms are those in decile ten and nine, medium firms are those in decile eight, seven, and six, all bellow that are allocated to the small firm group.

This implies that even though there is some effect of outliers in the returns of large sized firms, it is not fully explained by the effect of outliers. The next section will perform a final analysis on whether the effect of shifts in systematic risk drive the returns, or whether it is due to the existence of a high-volume return premium.

5.2 Systematic risk

As is known in finance, an increase in systematic risk subsequently drives an increase in returns due to the risk return trade-off. Therefore, to test whether the results are not driven by systematic risk a test is applied by using a linear regression of the joint market model. Originally developed by Zellner (1962), the seemingly unrelated regression equation (SURE) model, consists of multiple regression equations having its own depended variable. Each equation is a linear regression on its own and can therefore be estimated separately. What this model, as shown below, allows to measure are the effects of systematic risk on the returns generated by the zero-cost portfolio.

(8) $$R_i^H = \alpha^H + \beta^H R_i^{market} + \varepsilon_i^H$$

(9) $$R_i^L = \alpha^L + \beta^L R_i^{market} + \varepsilon_i^L$$

Here R_i^H and R_i^L represent the returns generated by the high and low trading volume stocks during interval i, respectively. The market returns are generated by creating a portfolio on the same intervals as the stock portfolios, and held for an equal amount of days. This tracks how much return the market has made during the time of the portfolios, and is noted as R_i^{market}. The coefficients β^H and β^L are the regression slopes of high and low trading volume for each size group, which represents the systematic risk of each stock. If the results are indeed generated by the systematic risk, it is expected that the coefficient for high volume stocks are higher than low volume stocks. Also, since the weekly sample has suggested the presence of a high-volume premium, this sample is included in the analysis.

The results are displayed in Table 6. For the daily sample, high volume small sized firms show to have a positive coefficient of 0.39, 0.58 and 0.61 for the 1, 10, and 20 day holding period respectively. These are higher than the respective low volume stocks. This would imply that higher returns are indeed driven by higher systematic risk. However, the actual difference between high and low volume stocks is insignificant across all holding day periods. The coefficients for high volume stocks of medium sized firms are higher than low volume stocks on the first day and during the 20-day holding period. For the 10-day holding

Table 6: Systematic risk test results for daily and weekly sample. Following the seemingly unrelated regressions equations (SURE) method the following regressions are tested:

$$R_i^H = \alpha^H + \beta^H R_i^{market} + \varepsilon_i^H$$

$$\mathrm{R_i^L = \alpha^L + \beta^L R_i^{market} + \varepsilon_i^L}$$

Where $R_i^H(R_i^L)$ denotes the returns for high (low) trading volume stocks in interval i {1,2,...I}. $\beta^H(\beta^L)$ is the regression slope for high (low) trading volume stocks, which represents the systematic risk effect. P-values are in brackets, and symbols, *, **, and *** denote a statistical significance at the 10%, 5%, and 1%, respectively.

	Daily			Weekly		
Holding period (days)	1	10	20	1	10	20
Small firms						
β^H	0.394	0.584	0.609	0.088	0.011	0.247
β^L	0.084	0.324	0.287	-0.014	-0.045	0.064
β^H - β^L	0.310	0.261	0.322	0.102	0.056	0.183
P-value	(0.164)	(0.195)	(0.223)	(0.174)	(0.392)	(0.227)
Medium firms						
β^H	1.078	0.985	2.325	0.033	0.321	1.208
β^L	0.816	1.959	1.925	-0.088	-0.028	0.030
β^H - β^L	0.261	-0.974	0.400	0.121	0.350	1.178
P-value	(0.305)	(0.978)	(0.218)	(0.273)	(0.131)	(0.002)***
Large firms						
β^H	5.241	3.217	4.293	0.712	2.736	5.717
β^L	1.422	3.000	2.538	0.233	0.449	1.299
β^H - β^L	3.819	0.217	1.756	0.479	2.287	4.417
P-value	(0.000)***	(0.387)	(0.008)***	(0.068)*	(0.000)***	(0.000)***

period it shows a 0.99 coefficient for high volume, against a 1.96 coefficient for low volume. Nevertheless, the respective p-value for this holding period suggests that this is strongly insignificant. Large firms do exhibit significant differences. However, as this is connected to a positive difference between high and low volume coefficients, it implies that any returns generated by high (low) volume stocks of large firms are driven by systematic risk.

From the weekly sample a similar conclusion is drawn. Across all holding periods and firm sizes, coefficients of high volume stocks are higher than those of low volume stocks. The difference is strongly significant for medium firms at a 20-day holding period with a p-value of 0.002. Large size firms show the strongest difference in coefficients, and in significance power. The conclusion can be drawn that returns generated by high (low) volume stocks are due to shifts in systematic risk.

6. **Conclusion**

In this thesis, the relationship between abnormal trading volume and future movements of prices in the Finnish stock market has been studied. Using a daily, and weekly, sample data of all (de-)listed stocks from the Nasdaq Helsinki over a time from April 1994 to May 2016 trading intervals are constructed. Trading intervals are constructed by placing stocks that have an unusually high (low) trading volume in respect to the previous 49 days in a long (short) portfolio. These portfolios are then held for a holding period of 1, 10, 20, 50, to 100 days without rebalancing. Trading intervals are non-overlapping, and stocks are categorized to small, medium, and large sized firms based on end-year market capitalization. The results show that, based on daily data, there exists significantly positive net returns for large sized firms at a holding period of 50 days. All other holding periods and firm sizes are not significant. Weekly data shows significant net returns for large sized firms up to 50 days, but not anymore at 100 days. Medium sized firms show a significant net return up to 10 days, and small sized firms show significant returns on day 20 and 50.

Returns are tested against return autocorrelations by creating two portfolios that have removed the top (bottom) 30 percent of the returns, and top (bottom) 40 percent of the returns of the originally constructed trading intervals. The returns from the weekly sample remain significant for the medium and large sized firms, concluding that the results are not driven by return autocorrelations. When removing outliers from the trading intervals the mean of small and medium sized firms changes drastically, concluding that these are largely driven by a few extreme values. Large sized firms, however, do not exhibit a strong change in the mean when excluding extreme values, concluding that the returns are not driven by extreme values. Nevertheless, when testing the results against systematic risk, that is market risk, it shows that the returns are very strongly driven by shifts in systematic risk. This concludes that there exists no high-volume return premium in the Finnish stock market, and so the null hypothesis that trading volume does not contain predictive power on the movement of future returns may not be rejected. This implies that is it not possible to predict price movements of stocks based on abnormal trading volume levels, and thus one cannot generate a trading strategy to obtain excess returns based on trading volume. Even so, abnormal levels of trading volume may still reflect heterogeneous understanding of information by investors, the returns associated with abnormal volume are just not significant to make gains on. Further on, the visibility hypothesis may persist in the Finnish stock market, but not at significant levels. The results do imply, however, that the efficient market hypothesis' weak-form holds.

Given the findings, it should be mentioned that when removing stocks below a 5-euro threshold, quite a large portion of the stocks were left out. Many of the stocks in the Finnish market are quite low priced, and due to the low amount of stocks the input towards trading intervals may have been limited. However, including such stocks would create a strong bias in the results due to the bid-ask bounce. Overall, the results are in line with previous high-volume return premium studies on Nordic countries, where weak to no evidence was found. What this may suggest is that relatively small sized stock exchanges have too few listings to show evidence of the high-volume return premium. It would be interesting for further research to study how the information content of trading volume differs across different demographics and stock exchange sizes. Also, different measures of volume could be tested on small stock exchanges, as perhaps different measures yield significant results.

APPENDIX

In this Appendix, a numerical example is given on how a trading interval works in the formation of the zero-cost portfolio and reference return portfolio. Suppose there are two trading intervals (trading interval 1, and 2), and stock A through H, as are given in Table 7. The calculations for high- and low-volume stocks, average trading interval returns, and net returns are provided in Table 8 for the zero-cost portfolio and Table 9 for the reference return portfolio. For the reference return portfolio, it is assumed that the average return for all stocks during trading interval 1 is 1.00 percent, and during trading interval 2 is 1.05 percent.

Table 7: Stock returns for two trading intervals.

Trading interval 1		**Trading interval 2**	
Stock	**Test period return**	**Stock**	**Test period return**
High-volume		**High-volume**	
A	1%	E	2%
B	1.5%	**Low-volume**	
Low-volume		F	0.5%
C	0.5%	G	0.25%
D	0.75%	H	0.75%

The average return of the zero-cost portfolio is the average net returns of both trading intervals, where stocks carry an average weight. This results in a net return of 0.63 percent for trading interval 1, and 1.50 percent for trading interval 2, giving a final net return of 1.06 percent for the zero-cost portfolio. The reference return portfolio is simply the average return of all 8 stocks, resulting in an average return of 0.48 percent.

Table 8: Calculations of the returns for the zero-cost portfolio.

Trading interval	Position	Returns from test period
High-volume		
1	0.5 euro long A and B	1.25%
2	1 euro long E	2.00%
	Average return high-volume	1.63%
Low-volume		
1	0.5 euro short C and D	-0.63%
2	1/3 euro short F, G, and H	-0.50%
	Average return low-volume	-0.57%
Net returns		
1	1.25% + (-0.63%)	0.63%
2	2.00% + (-0.50%)	1.50%
	Average net return (high-low)	1.06%

Table 9: Calculations of the returns for the reference return portfolio.

Trading interval	Position	Returns from test period
High-volume		
1	1 euro long A & 1 short ref-1	0.00%
2	1 euro long B & short ref-1	0.50%
3	1 euro long E & short ref-2	0.95%
	Average return high-volume	0.48%
Low-volume		
1	1 euro short C & long ref-1	0.50%
2	1 euro short D & long ref-1	0.25%
3	1 euro short F & long ref-2	0.55%
4	1 euro short G & long ref-2	0.80%
5	1 euro short H & long ref-2	0.30%
	Average return low-volume	0.48%

REFERENCES

Andersen, T. G. (1996). Return volatility and trading volume: An information flow interpretation of stochastic volatility. *The Journal of Finance, 51*(1), 169-204.

Arbel, A., & Strebel, P. (1982). The neglected and small firm effects. *Financial Review*, *17*(4), 201-218.

Bachelier, L. (1900). *Théorie de la spéculation*. Gauthier-Villars.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *The Journal of Economic Perspectives*, *21*(2), 129-151.

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The quarterly journal of economics*, *116*(1), 261-292.

Benartzi, S., & Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American economic review*, 79-98.

Bernard, V. L., & Thomas, J. K. (1990). Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics*, *13*(4), 305-340.

Bhardwaj, R. K., & Brooks, L. D. (1992). The January anomaly: Effects of low share price, transaction costs, and bid-ask bias. *The Journal of Finance*, *47*(2), 553-575.

Blume, M. E., & Stambaugh, R. F. (1983). Biases in computed returns: An application to the size effect. *Journal of Financial Economics*, *12*(3), 387-404.

Blume, L., Easley, D., & O'hara, M. (1994). Market statistics and technical analysis: The role of volume. *The Journal of Finance*, *49*(1), 153-181.

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

Conrad, J.S., & Kaul, G. (1993). Long-Term Market Overreaction or Biases in Computed Returns?. *The Journal of Finance*, *48*(1), 39-63.

Conrad, J. S., Hameed, A., & Niden, C. (1994). Volume and autocovariances in short-horizon individual security returns. *The Journal of Finance*, *49*(4), 1305-1329.

Cooper, M. (1999). Filter rules based on price and volume in individual security overreaction. *Review of Financial Studies*, *12*(4), 901-935.

Copeland, T. E. (1976). A model of asset trading under the assumption of sequential information arrival. *The Journal of Finance*, *31*(4), 1149-1168.

Dubois, M., & Louvet, P. (1996). The day-of-the-week effect: The international evidence. *Journal of Banking & Finance*, *20*(9), 1463-1484.

Epps, T. W. (1975). Security price changes and transaction volumes: Theory and evidence. *The American Economic Review*, 586-597.

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business*, *38*(1), 34-105.

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

Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The adjustment of stock prices to new information. *International economic review*, *10*(1), 1-21.

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

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

Gibbons, M. R., & Hess, P. (1981). Day of the week effects and asset returns. *Journal of business*, 579-596.

Gillette, A. B., Stevens, D. E., Watts, S. G., & Williams, A. W. (1999). Price and volume reactions to public information releases: An experimental approach incorporating traders' subjective beliefs. *Contemporary Accounting Research*, *16*(3), 437-479.

Granger, C. W., & Morgenstern, O. (1963). Spectral analysis of New York stock market prices. *Kyklos*, *16*(1), 1-27.

Harris, L. (1986). Cross-security tests of the mixture of distributions hypothesis. *Journal of financial and Quantitative Analysis*, *21*(01), 39-46.

Harris, L. (1987). Transaction data tests of the mixture of distributions hypothesis. *Journal of Financial and Quantitative Analysis*, *22*(02), 127-141.

Harris, M., & Raviv, A. (1993). Differences of opinion make a horse race. *Review of Financial studies*, *6*(3), 473-506.

Hoitash, R., & Krishnan, M. M. (2008). Herding, momentum and investor over-reaction. *Review of Quantitative Finance and Accounting*, *30*(1), 25-47.

Karpoff, J. M. (1986). A theory of trading volume. *The Journal of Finance*, *41*(5), 1069-1087.

Karpoff, J. M. (1987). The relation between price changes and trading volume: A survey. *Journal of Financial and quantitative Analysis*, *22*(01), 109-126.

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.

Lee, C. F., & Rui, O. M. (2000). Does trading volume contain information to predict stock returns? Evidence from China's stock markets. *Review of Quantitative Finance and Accounting*, *14*(4), 341-360.

Lee, B. S., & Rui, O. M. (2002). The dynamic relationship between stock returns and trading volume: Domestic and cross-country evidence. *Journal of Banking & Finance*, *26*(1), 51-78.

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

Lehmann, B. N. (1990). Fads, martingales, and market efficiency. *The Quarterly Journal of Economics*, *105*(1), 1-28.

Jensen, M. C. (1969). Risk, the pricing of capital assets, and the evaluation of investment portfolios. *The Journal of Business*, *42*(2), 167-247.

Lo, A. W., & Wang, J. (2006). Trading volume: Implications of an intertemporal capital asset pricing model. *The Journal of Finance*, *61*(6), 2805-2840.

Lu, T. H., & Lee, J. D. (2016). Is Abnormally Large Volume a Clue?. *International Journal of Economics and Finance*, *8*(9), 226.

Martikainen, T., Puttonen, V., Luoma, M., Rothovius T. (1994). The Linear and Non-Linear Relationships between Stock Returns and Trading Volume in the Finnish Stock Market, Applied Financial Economics, 4, 159–169.

Mayshar, J. (1983). On divergence of opinion and imperfections in capital markets. *The American Economic Review*, *73*(1), 114-128.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The journal of finance*, *42*(3), 483-510.

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. *The Journal of finance*, *32*(4), 1151-1168.

Niederhoffer, V., & Osborne, M. F. (1966). Market making and reversal on the stock exchange. *Journal of the American Statistical Association*, *61*(316), 897-916.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica: Journal of the Econometric Society*, 122-136.

Richardson, G., Sefcik, S. E., & Thompson, R. (1986). A test of dividend irrelevance using volume reactions to a change in dividend policy. *Journal of Financial Economics*, *17*(2), 313-333.

Ritter, J. R. (2003). Behavioral finance. *Pacific-Basin finance journal*, *11*(4), 429-437.

Shalen, C. T. (1993). Volume, volatility, and the dispersion of beliefs. *Review of Financial Studies*, *6*(2), 405-434.

Shleifer, A. (2000). *Inefficient Markets: An introduction to behavioural finance*. OUP Oxford.

Smirlock, M., & Starks, L. (1985). A further examination of stock price changes and transaction volume. *Journal of Financial Research*, *8*(3), 217-226.

Tang, T., Zou, L., & Li, J. (2013). The high-volume return premium: Evidence from the australian equity market. *Journal of Accounting and Finance, 13*(5), 74.

Tauchen, G. E., & Pitts, M. (1983). The price variability-volume relationship on speculative markets. *Econometrica: Journal of the Econometric Society*, 485-505.

Tkac, P. A. (1999). A trading volume benchmark: Theory and evidence. *Journal of Financial and Quantitative Analysis*, *34*(01), 89-114.

Van Horne, J. C., & Parker, G. G. (1967). The random-walk theory: an empirical test. *Financial Analysts Journal*, 87-92.

Verrecchia, R. E. (1981). On the relationship between volume reaction and consensus of investors: Implications for interpreting tests of information content. *Journal of Accounting Research*, 271-283.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American statistical Association*, *57*(298), 348-368.

Ying, C. C. (1966). Stock market prices and volumes of sales. *Econometrica: Journal of the Econometric Society*, 676-685.

Zhang, J., Lai, Y., & Lin, J. (2017). The day-of-the-Week effects of stock markets in different countries. *Finance Research Letters*, *20*, 47-62.

Zhou, Z. G. (2010). The high-volume return premium: evidence from the Chinese stock market. *Review of Quantitative Finance and Accounting*, *35*(3), 295-313.