TURUN KAUPPAKORKEAKOULU
Turku School of Economics

ABSTRACT

- [X] Master´s thesis
- [ ] Licentiate´s thesis
- [ ] Doctor´s thesis

Subject	Accounting and Finance	Date	17.12.2010
Author(s)	Ilja Ripatti	Student number	
		Number of pages	114pp.+ appendices
Title	Investor sentiment and stock market behavior: Empirical tests on the Finnish stock market		
Supervisor(s)	Professor Mika Vaihekoski D.Sc (Econ.) Antti Fredriksson		

Abstract

The neoclassical finance theory suggests that investors act rationally and even if some investors do not act rationally, their actions tend to offset each other. Sentiment is not considered to affect stock prices. Behavioral finance has tried to broaden the traditional view by incorporating the effects of human actions and psychology. In this view, people are not entirely rational and are regarded to be influenced by the actions of other market participants and by the prevailing investor sentiment. This thesis studies the effects of investor sentiment in the Finnish stock market in the new millennium. Factors that influence the behavior of investors and cause the formation of investor sentiment are presented, as well as the possible consequences of sentiment extremities. The empirical part of the thesis investigates whether sentiment is related to returns of various portfolios and whether sentiment can be used to forecast returns.

The data in this thesis consists of sentiment related and return related data spanning from 2000 to 2008. The investor sentiment index is composed by Principal Components Analysis, which enables us to reduce the common variation of multiple variables into a single measure. An additional index that accounts for macroeconomic factors is constructed by regressing the sentiment proxies by macroeconomic variables and performing a Principal Components Analysis for the residuals. The relationship between investor sentiment and portfolio returns is examined by conducting regression analyses. The results indicate that investor sentiment has an influence on the returns of stocks and especially among smaller stocks and less volatile stocks. Sentiment has a larger influence during extreme levels of sentiment and particularly during extreme pessimism.

The effects of investor sentiment were at least partially as proposed by the behavioral finance theory and in line with previous literature, although the effects on returns were not strong. Hence, it is evident that investor sentiment can hardly be used to form a systematically profitable trading strategy. Nevertheless, investor sentiment is related to the returns of portfolios in the Finnish stock market.

Key words	Investor sentiment, market sentiment, investor behavior, behavioral finance
Further information	

TURUN KAUPPAKORKEAKOULU
Turku School of Economics

TIIVISTELMÄ

X	Pro gradu -tutkielma
	Lisensiaatintutkielma
	Väitöskirja

Oppiaine	Laskentatoimi ja rahoitus	Päivämäärä	17.12.2010
Tekijä(t)	Ilja Ripatti	Matrikkelinumero	
		Sivumäärä	114 s. + liitteet
Otsikko	Investor sentiment and stock market behavior: Empirical tests on the Finnish stock market		
Ohjaaja(t)	Professori Mika Vaihekoski KTT Antti Fredriksson		

Tiivistelmä

Perinteinen rahoitusteoria olettaa, että sijoittajat käyttäytyvät keskimäärin rationaalisesti. Vaikka kaikki markkinaosapuolet eivät toimisikaan rationaalisesti, markkinavoimat palauttavat osakkeiden hinnat oikeille, fundamentaalisille tasoille. Kuitenkin useat pörssikuplat ja kriisit herättävät kysymyksen markkinoiden rationaalisesta luonteesta. Käyttäytymistieteellinen rahoitus on pyrkinyt laajentamaan perinteisen rahoitusteorian viitekehystä ottamalla huomioon ihmisten käyttäytymisen vaikutukset osakemarkkinoilla. Ihmisten oletetaan olevan alttiita markkinatunnelman vaikutuksille ja käyttäytyvän laumamaisesti epävarmuuden vallitessa. Tutkielmassa etsitään vastausta siihen, että vaikuttaako markkinatunnelma osakkeiden tuottoihin lyhyellä aikavälillä Helsingin pörssissä ja esiintyykö markkinoilla markkinatunnelmaan liittyvää osakkeiden yli- tai aliarvostusta. Tutkielmassa esitellään markkinatunnelmaan ja sijoittajien käyttäytymiseen vaikuttavat tekijät ja pohditaan äärimmäisen markkinatunnelman vaikutuksia.

Aineisto tutkielmassa koostuu vuosien 2000–2008 pörssituotoista ja markkinatunnelmamuuttujista samalta ajanjaksolta. Markkinatunnelmaindeksi muodostetaan pääkomponenttianalyysin avulla, jolloin useiden markkinatunnelmamuuttujien yhteisvaihtelu voidaan tiivistää yhdeksi muuttujaksi. Epärationaalista markkinatunnelmaa kuvaava indeksi saadaan, kun alkuperäisestä indeksistä poistetaan makrotaloudellisten tekijöiden vaikutus regressioanalyysin keinoin ja regressiojäännöksistä luodaan uusi indeksi. Markkinatunnelman ja portfolioiden tuottojen suhdetta tutkitaan regressioanalyysin avulla. Tulokset osoittavat, että markkinatunnelmalla on vaikutusta Suomen osakemarkkinoilla. Vaikutus on suurempi, kun markkinatunnelma on voimakkaampi ja erityisesti sen ollessa pessimistinen.

Markkinatunnelman vaikutukset osakkeiden tuottoihin olivat osittain käyttäytymistieteellisen rahoitusteorian mukaisia, mutta vaikutukset tuottoihin eivät olleet suuria. Täten markkinatunnelmaa ei pysty käyttämään systemaattisena kaupankäyntistrategiana, mutta tutkimustulosten perusteella markkinatunnelmaa voidaan pitää osakkeiden hintoihin vaikuttavana tekijänä.

Asiasanat	markkinatunnelma, sentimentti, sijoittajien käyttäytyminen, käyttäytymistieteellinen rahoitus, noise trader, arbitraasin rajoitteet
Muita tietoja	

Turun yliopisto
University of Turku

INVESTOR SENTIMENT AND STOCK MARKET BEHAVIOR

Empirical tests on the Finnish stock market

Master's thesis in accounting and finance

Author
Ilja Ripatti 50268

Supervisors
Professor Mika Vaihekoski
D.Sc (Econ.) Antti Fredriksson

17.12.2010
Turku

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 Background

The traditional finance theory assumes that investors make rational decisions based on careful probabilistic estimates and by weighing different options rationally. Even if not every market participant acts rationally, their actions tend to offset each other. The stock prices reflect the expectations of future cash flows and if the stock price deviates from the fundamental value, the rational arbitrageurs will take advantage of the situation and return the stock price to its balanced, fundamental price. One cannot earn abnormal profits without taking greater risk. Sentiment does not play any role in the traditional framework (Schmeling 2009, 395).

Numerous historical market bubbles and crashes plant a seed of doubt to the rationality of the stock market. The stock markets are comprised of people and people tend to get overly optimistic or pessimistic occasionally. The US stock market rose eightfold between 1982 and 2000 and fell by half between 2000 and 2008. These sorts of movements might be justifiable on individual stock level but hardly on an aggregate market level even if all the necessary variables in economic conditions are taken into account (Akerlof & Shiller 2009, 127).

Behavioral finance has tried to broaden the view of the traditional finance theories by taking human actions and behavioral tendencies into consideration. The behavioral finance researchers have augmented the traditional finance models with two assumptions: the first assumption is that investors are subject to sentiment. The second assumption is that betting against investors that are prone to sentiment can be costly and risky (Baker & Wurgler 2007, 129). Investor sentiment can be defined as the mood and expectations of investors or how bullish (bearish) they are. Investor sentiment can also be seen as a belief of future cash-flows and risks that cannot be justified by rational expectations and known facts (Schmeling 2007, 128; Baker & Wurgler 2007, 129).

Small investors are considered to be most affected by sentiment and are more likely to make their decisions on more emotional basis than larger or institutional investors. Investors that make their investment decisions based on irrelevant information and emotion are considered as “noise traders” (Black 1986;Lee, Shleifer & Thaler 1991).
If the effects of investor sentiment can be quantified, an increasing amount of knowledge about the influence of market behavior can be obtained. Whether investor sentiment is a factor that is needed to be taken into consideration in asset pricing and corporate finance is yet to be resolved.

Investor sentiment and its effects in the stock market have been of increasing interest of researchers in the new millennium. Studies such as Brown and Cliff (2004) and Sheung, Lu, Tung and Wei (2010) investigated the causal relationship between stock returns and sentiment. Brown and Cliff (2004) documented that returns cause sentiment, but Sheung et al. (2010) found out that causality would run from sentiment to market behavior even when taking different market scenarios into account.

Investor sentiment and equity returns on an event perspective have been studied by Wong and Lievano (2009), Drakos (2009) and Nikkinen and Vähämaa (2009) who studied the effects of the terrorist attacks on investor sentiment and stock returns. The discoveries were quite similar. The terrorist attacks had a significant impact on especially small investor sentiment and found empirical support for the sentiment effect. Kaplanski and Levy (2010) studied the effects of aviation disasters on equity returns and found out that aviation disasters are followed by negative rates of return. Edmans, Garcia and Norli (2006) investigated the stock market reaction to sudden changes in investor mood caused by sports results. They documented a strong negative stock market reaction after soccer match losses especially in the world cup elimination games and amidst small stocks.

The effects of sentiment on stock return predictability have been studied by e.g Neal and Wheatley (1998), Brown and Cliff (2005), Lemmon and Portniaguina (2006), Baker and Wurgler (2006), Schmeling (2009), Canbas (2009) and Finter, Niessen and Ruenzi (2010). These studies have used different types of proxies for sentiment such as consumer confidence, closed end fund discounts, turnover etc. All of these studies, excluding the studies of Schmeling, Canbas and Finter et al. reported significant effects of investor sentiment on stock returns. In line with the assumptions of behavioral finance and noise trader theory, the effects are most pronounced among stocks that are mainly traded by small investors. However, Schmeling and Canbas found out that sentiment is not associated with stock returns in every country. Schmeling compared the effects of sentiment on 18 different nations and discovered that the effects of sentiment vary in different nations. Sentiment measured by consumer confidence had only an impact on value stocks in Finland. Lemmon and Portniaguina (2006) also found out that sentiment had an impact on value stocks but not on growth stocks. Canbas did not find a relationship between stock returns and sentiment in Turkey. Finter et al. did not find a strong relationship between sentiment and returns when several control variables were included.

1.2 Objective of the thesis

This thesis studies the effects of investor sentiment on the Finnish stock market. The objective of this thesis is to provide a thorough research on the matter by forming a sentiment index and by assessing how sentiment affects the cross-section of stock returns. It is studied whether sentiment has a larger impact on stocks whose valuations are highly subjective and if any patterns of mispricing correction can be observed. Hence, the relationship between sentiment and the returns of subsequent months in relation to sentiment is scrutinized.

The research is conducted by forming a sentiment index in a similar manner to Baker and Wurgler (2006). The index is composed of different variables that should capture the essence of sentiment. These variables are closed-end fund discount, turnover, number of IPOs, average first day return of IPOs, equity share in new issues and the dividend premium. All of these variables may not be applicable in this thesis due to data availability and suitability. However, a wide range of different sentiment measures have been introduced in prior literature that may be used (Brown & Cliff 2005; Baker & Wurgler 2007). Each of these variables is likely to include a sentiment component as well as non-sentiment related components. Principal Components Analysis is used to isolate the common sentiment component. Then another index is composed that accounts for business cycle variation. This is done by regressing each of the proxies on a set of macroeconomic variables. In the study of Baker and Wurgler (2006) the index was formed on an annual basis. In this thesis the index is formed on a monthly basis. This is done in order to assess whether sentiment can be used as a basis of a short term trading strategy. In addition, data availability sets its own restrictions.

Whether the sentiment indices are related to the returns of a variety of portfolios comprised of stocks with different characteristics for size and investment style is evaluated. In addition portfolios are created on the basis of historical volatility. The effects of sentiment on returns are analyzed by performing regression analyses on the whole time period as well as during periods of extremely high and low sentiment. The main interest is directed towards detecting whether sentiment can be used to forecast returns and if sentiment related return patterns can be observed. In practice, sentiment is matched by the subsequent monthly and quarterly returns.

1.3 Focus of the thesis

This thesis will focus on the Finnish stock market and tries to determine the effects of investor sentiment on the returns of stocks and assess whether investor sentiment related mispricing correction patterns can be observed. Hence, it is identified how the returns of shares that are reckoned to be more subject to investor sentiment act in relation to sentiment compared to shares that are presumed to be more immune to investor sentiment.

Finland is chosen because the effects of sentiment have not been widely investigated in the Finnish stock market, the market is smaller and more illiquid than the American stock market and due to personal interests of the author who is of Finnish origin. According to previous research results and research on behavioral finance issues, stocks more subject to sentiment are companies with a more speculative nature or companies that are widely held by retail investors. The returns of various stocks are investigated on different time periods and sentiment levels by focusing mainly on the short term relation between sentiment and returns.

1.4 Structure of the thesis

The second chapter of this thesis will provide a theoretical framework. The chapter starts with the view of traditional finance theory regarding investor sentiment. The efficient market hypothesis, its relation to sentiment and the market efficiency in Finland are discussed. This is followed by a behavioral point of view of finance. The focus is directed towards two critical components that are considered to be essential in order for the sentiment to have an impact on the stock market: noise trader hypothesis and limits of arbitrage.

After that a variety of psychological and sociological factors that can cause irrational sentiment or at least influence its formation are introduced. By sentiment it is meant the extreme levels of sentiment or noise trader sentiment which is mainly overoptimism or pessimism about the future which cannot be linked to fundamentals. The last part of chapter two uncovers the possible consequences of sentiment and especially the effects of sentiment extremities.

Chapter three presents the data and the research method that is used in this thesis and provides insight why and how the methods are used in order to obtain the results of chapter four. The data availability and suitability is introduced and assessed first. A wide number of sentiment proxies are presented and evaluated in order to form the best available sentiment index. Afterwards, the method behind forming the sentiment indices is presented step by step. The last part of the chapter focuses on the regression analysis method used in this thesis.

Chapter four begins by presenting descriptive statistics of the return portfolios and performing preliminary tests. The main regression results are revealed and analyzed. It is also assessed whether sentiment has a bigger impact on returns during extreme sentiment levels. The last part of chapter four is related to conducting robustness tests. It is evaluated whether the results remain after introducing additional control variables.

Chapter five summarizes the thesis and the main findings and presents research topics for further research.

2 THEORETICAL FRAMEWORK

2.1 Neoclassical approach

2.1.1 Market efficiency and investor sentiment

Traditional finance theory assumes that agents or market participants are rational and that the law of one price holds. This means that stocks with similar future prospects, levels of risk and payoff have identical prices. Market efficiency may exist if agents in general are rational and the law of one price holds. Short term deviations of the fundamental price are balanced by arbitrageurs (Glaser, Nöth & Weber 2003).

Market efficiency can be divided into external and internal efficiency. External efficiency means that information is rapidly and broadly usable by all market participants, which facilitates a swift and correct reaction of stock prices to new relevant information (Nikkinen, Rothovius & Sahlström 2005, 80-82).

Internal efficiency means that the market is operationally efficient. Hence, there needs to be an adequate amount of competition between brokerage firms which keeps the transaction costs low and the processing of transactions quick. When people talk about market efficiency, external efficiency is usually meant (Nikkinen et al. 80-82). The main result of market efficiency is that speculative asset prices such as stocks always incorporate the best available knowledge about fundamentals and the prices only change due to sensible new information (Shiller 2003, 83). The stock prices follow a random walk with a drift and no predictions about future price movements can be made based on current prices or the historical development of the prices (Wärneryd 2001, 17-18). Tomorrow's price change will reflect tomorrow's news and will be independent of the price changes today. As news is unpredictable the price changes must be unpredictable and random as well (Malkiel 2003, 59).

Fama (1970) divided the market efficiency to three forms depending on what kind of information is compounded on the stock prices. The three forms are weak, semi-strong and strong form of efficiency. Weak form of efficiency means that all past knowledge is already in the stock price. Semi-strong means that all publicly available information is in the stock price and in the strong form efficiency even all private information has been taken into account.

In an efficient market no arbitrage opportunities exist. It is not possible to get above-average returns without taking greater risk. As all the stocks are correctly priced a blindfolded chimpanzee throwing darts at the stock pages could earn similar or better profits than the experts. A good indicator of market efficiency is that professional investors and wealth managers have had a hard time beating a passive index fund (Malkiel 2005, 1-2).

It has been later noted that not all investors might act rationally. However, this irrationality among investors will not affect the stock prices as their actions offset each other (Nesvetailova 2007, 31). Even if the irrational investors acted in a concert and their purchases and sells were correlated the market will be balanced by smart money investors that act as rational arbitrageurs. The movements of smart money would balance the market and the actions of those who are swept up in a frenzy that would otherwise lead to mispricing. This means that each stock would be at their rational price or immediately returns to its fundamental price after a short deviation (Hasset 2002). There can be various ways to calculate the fundamental or intrinsic value of a stock. Most commonly used measures by investors are the discounted cash flow or discounted dividend methods. The price of the stock should be the discounted value of future cash flows of the company. The expectations between investors about a company's future prospects, macroeconomic factors or discount rate may vary quite a lot but the aggregate market is expected to form the correct price level (Wärneryd 2001, 14-18, 48). If all the stocks are correctly priced, each investment is as good as any other investment. The choice between stocks is ultimately made on the amount a person is willing to take risk, and on the other hand this risk should be compensated by heftier profits. A person is unable to obtain abnormal risk adjusted profits from the market. As the price level of the stocks are at their correct level, companies and their managers are able to use market data in order to estimate their cost of capital and compare different investment options (Mishkin 1998).

If irrational or less informed investors make constantly poor decisions and the arbitrageurs take advantage of the situation, the irrational investors will lose money and their wealth to the extent that they can no longer participate in the stock market. It has a tone of Darwinism where only the fittest and most informed investors survive (Shleifer 2000; Hasset 2002).

Traditional finance theory does not recognize the role of investor sentiment in asset valuation. Rational investors will drive the prices to equilibrium where the prices reflect the discounted cash flows of the stocks and the expected returns depend on the amount of systematic risk. This means that by the commonly used Capital Asset Pricing Model a speculative stock, which is subject to more systematic risk, has a bigger expected return than a bond-like stock that is not prone to a high degree of systematic risk. Investor sentiment will not affect the prices as the trades of irrational traders offset each other or rational arbitrageurs will take advantage of the situation and the price returns to its correct level (Baker & Wurgler 2006). Investor sentiment should not be able to be used to predict future returns for certain types of stocks nor should you be able to use it in order to create a trading strategy that exploits any return patterns related to investor sentiment. There should not be any exploitable patterns (Barberis, Shleifer & Vishny 1998).

2.1.2 The efficiency of the Finnish stock market

The Helsinki Stock Exchange was founded in 1912. Even though there had been some sort of stock exchange activity before but it lacked proper regulations and an organization (Nyberg & Vaihekoski 2010). The financial system was tightly regulated until 1980s and the stock market was quite underdeveloped compared to other western nations. The process of financial market deregulation took nearly a decade and proceeded from the liberalization of the money market to the restrictions on all capital movements being abolished in Finland. The restrictions of foreign ownership were removed in 1993. The number and value of financial instruments and the amount of investors increased exponentially from the 1970s to the late 1990s. The Finnish legislation was amended to comply with the European Union directives and the Finnish stock market was harmonized with other Nordic exchanges to create OMX (Vaihekoski 1997; Pörssisäätiö 2007).

Korhonen (1977) was one of the first to study the Finnish stock market efficiency. He found out that the Finnish stock market was quite efficient at least in the weak form even though the market was very small and relatively illiquid. Prices did fluctuate in a random walk way and no leads or lags were found that could be exploitable. However, the sample size of his study was small as the data consisted of only 18 companies. Virtanen and Yli-Olli (1987) made a divergent finding in their study. They discovered many anomalies and deviations from the market efficiency and propose that the prices do not act in a totally random walk way.

Berglund, Wahlroos and Örnmark (1983) investigated the weak form efficiency in the Finnish stock market and found out that there is serial correlation in the stock returns. Although the serial correlation indicated that excess returns could be obtained by careful buy and sell orders, a profitable speculative strategy would be hard to create because of costs related to information gathering and trade execution.

Even later research papers have noticed that significant autocorrelation can be found from both stock and option returns meaning that today’s returns can be forecasted by yesterday’s returns (Kallunki, Martikainen, Martikainen & Yli-Olli 1997). In the paper of Knif and Löflund (1997) first order serial correlation among Finnish stocks was found. They proposed that the friction in the trading process and limited trading is an important source of serial correlation. Another reason for serial correlation can also be the predictability of a time-varying risk premium utilizing a set of past information. Knif and Löflund pointed out that serial correlation can arise due to changing information and serial correlation of the stock returns may be a reflection of the serial correlation in the economic fundamentals, thus not violating the market efficiency.

Kallunki et al. (1997) presented that there have been some anomalies found in the Finnish stock market. The returns of stocks have been higher in the turn of the month than during the rest of the month and this effect has been the strongest for the most traded stocks. However, Nikkinen, Takko, Sahlström and Äijö (2009) did provide an explanation for this anomaly. They found out that the US macroeconomic news is the cause of the turn of the month returns. After controlling for the US macroeconomic news data, the anomaly seizes to exist.

The other found anomalies are related to the size of the company and the P/E-ratios of the companies. Small companies have tended to vastly outperform the largest companies even if systematic risk has been accounted for. In addition, portfolios comprised of companies with low P/E-ratios have realized heftier returns than other test portfolios. This anomaly may be at least partially explained by the risk estimation problems when using market based data in Finland. (Kallunki et al. 1997).

Schadéwitz and Blevins (1998) investigated the speed of adjustment of share prices to the announcement of unexpected interim earnings and tried to determine whether the Finnish stock market could be regarded as semi-strong efficient. The main finding of the study was that the adjustment of the share prices to new information was delayed which suggests inefficiencies in the semi-strong form.

The information content of earnings announcements and cash dividends from year 1983 to 1994 was studied by Heikkilä (1997). The information content of earnings announcements appeared to be significantly lower than that of dividend announcements. This could be interpreted in a way that dividend announcements convey more information than earnings announcements and earnings announcements do not provide much new information as the information has been already discounted on the stock price from other information sources. Kallunki et al. (1997) pointed out that the former Finnish accounting rules have provided the firms with exceptionally good opportunities to smooth income intentionally. Thus earnings information might not have been as informative as the accounting numbers would have to be adjusted to describe the economic reality.

Ekholm (2002) studied the stock market reactions to new information and discovered that the stock market reaction to new interim information is lagged. In his paper he tried to specify whether this was due to market inefficiency or an expected return specification problem. The finding was that some of the investigated financial statement signals seemed to proxy for risk while other financial statement signals appeared to contain information that was priced with a delay.

Some research has been conducted on the behavior of different types of investors around earnings announcements. Booth, Kallunki and Martikainen (1999) found out that small and large investors do behave differently and that small investors increase their sell orders after negative earnings surprises. According to Ekholm (2002) large investors are seen as antagonists to the majority of investors regarding the reaction to new earnings information. Majority of investors are biased towards increasing their holdings after new financial statement information. Positive earnings surprises increase the probability that institutional investors increase their holdings as opposed to the reaction of small investors and foreign investors who liquidate at least a part of their winnings after positive earnings announcements.

As a conclusion of this chapter regarding Finnish stock market efficiency it can be noted that although some anomalies and deviations from the market efficiency have been found it is hard to point out whether it is a deviation of market efficiency itself or just the unsuitability of the expected return model such as the CAPM. In addition, many of the found anomalies are quite hard to exploit and do not offer a consistent chance of abnormal profits. The Finnish stock market seems to be efficient at least in the weak form. The availability and timeliness of information has exponentially increased in the new millennium and the stock market can be presumed to act more efficiently. As this thesis is about testing how the investor sentiment affects the stock market it is also a test of market efficiency or whether the asset pricing models should include an investor sentiment related component in them.

2.2 Behavioral finance

2.2.1 Behavioral finance and sentiment

The main theme in many behavioral finance models and theories is that investors are not fully rational and that they make biased decisions based on irrelevant information. Many mental and psychological factors affect the decision making process of a person.

Behavioral finance provides a broader social science perspective by incorporating sociology and psychology in finance (Nofsinger 2008, 4; Shiller 2003, 83). The irrational decisions among investors can be correlated and persistent even for long periods of time. The market price can deviate from the fundamental value as there are limits to arbitrage which deters informed investors from eliminating the mispricing. Betting against irrational traders or “noise traders” as they are often called can be risky as one cannot predict how long the mispricing is going to last (Schmeling 2009; Neal & Wheatley 1998).

Shleifer (2000) has stated that especially noise traders are prone to investor sentiment or the aggregate mood of the market and that investor sentiment affects noise traders the most. However, institutional investors, that are usually considered as rational or smart money traders, have been found to herd in and out of stocks and use momentum-related investment strategies. This means that even institutional investors might reinforce the market movements when investor sentiment is overly optimistic or pessimistic and try to profit from the market over or underreactions. Institutional investors become more pessimistic when individual investors become more optimistic as they recognize that the prices have been driven above their fundamental value and might be corrected soon. As investors become even more overly optimistic, the institutional investors become optimistic as well as they notice that noise traders might drive the prices even higher (Schmeling 2007).

Baker and Wurgler (2007) propose that investor sentiment can be defined as the propensity to speculate by the marginal investor. Sentiment drives the relative demand for speculative investments, thus causing cross-sectional effects despite arbitrage forces. The characteristic that makes some stocks more speculative than others is due to the difficulty and subjectivity of determining their true values. A company that has immense growth potential, intangible assets or is heavily leveraged is harder to valuate than a company that has a steady cash flow and tangible assets. Also young and small companies are regarded to be more influenced by sentiment. Investor sentiment can also be viewed as simply optimism or pessimism about stocks in general. As arbitrage is seen to be particularly risky and costly on companies that are considered speculative and harder to value, the aforementioned optimism or pessimism will have a more pronounced impact on these stocks. Stocks that are speculative and harder to arbitrage and value have higher relative valuations when the sentiment is high. Thus, there should be a contrarian relation between sentiment and future returns of speculative stocks.

The demand for speculative or safer stocks can vary over time. Individuals demand for stocks that have the bundle of salient characteristics compatible with their sentiment. Sentiment fluctuations can affect the demand for speculative or stable and profitable stocks when for example there are flights to quality (Baker & Wurgler 2006; Baker & Wurgler 2007).

The building up of sentiment can be related to price feedback models. This means that when speculative assets' prices go up they create success stories for some investors, attract public and media attention, induce word-of-mouth enthusiasm and heighten the expectations for further price increases. The prices are high due to expectations of even higher prices which can produce a bubble or in a contrarian situation a bust. Past price changes fuel the expectations of people although it might not be rational in the efficient market sense. Psychologists Andreassen and Krauss tested the perceptions of people when they were showed real historical stock prices in a sequence and asked to trade in a simulated market. People tended to chase trends and behave as if they extrapolated past price changes (Shiller 2003, 91-94). This can be interpreted that the overall trend and atmosphere of the market does influence the expectations of future returns.

One of the most notable and first papers to address the noise traders and their effect on the stock market was done by De Long, Shleifer, Summers and Waldman (1990). They presented that noise traders can have an impact on security prices and that they create an additional noise trader risk in the market which should be taken into consideration in asset valuations. The unpredictability of noise traders' sentiments creates additional risk in the market. The risk that noise traders' beliefs will not revert to their mean for a long time and might become even more extreme. De Long et al. also argue that the idea of noise trader risk can explain a number of financial market anomalies.

Baker and Wurgler (2006) found out that investor sentiment did have a larger impact on speculative stocks and stocks that were more prone to noise trader risk. The effects were especially significant when a portfolio was formed on the basis of volatility. Highly volatile stocks exhibit characteristics of difficulties of valuation and limits of arbitrage. In addition, extreme growth and value companies seemed to earn high returns following periods of low sentiment and earn low returns following periods of high sentiment. Lemmon and Portniaguina (2006) find similar patterns for value stocks but not for growth stocks. However, they do report that small stocks are overvalued in relation to large stocks during periods of high sentiment. Finter, Niessen and Ruenzi (2010) suggest that sentiment prone stocks provide higher returns than sentiment insensitive stocks in the short run from one to three months and afterwards the effect reverses and sentiment insensitive stocks outperform sentiment prone stocks.

Next we will take a deeper look into noise trading and the limits of arbitrage which enable the possibility that investor sentiment might have some effects on security prices.

2.2.2 *Noise trading*

The participants of the stock market can be divided into rational arbitrageurs or smart money investors, noise traders and passive investors. Noise traders are regarded as a lower income group compared to especially smart money investors. Smart money investors are often seen as either institutional investors or wealthy private investors that have a greater incentive and knowledge to gather meaningful information (Kelly 1997).

Noise traders are investors that do not behave as a rational investor would in an efficient market. They do not buy and hold a market portfolio or diversify adequately but rather hold a single stock or a small number of stocks. Many of their investment decisions are based on a gut feeling or different pseudo-signals. These pseudo-signals can be either investment advice from brokers or even friends, buy and sell signals created by means of technical analysis or almost any other information that is not relevant and is already compounded in the stock price (Shleifer & Summers 1990). Previous ascension of stock prices lures uninformed investors to participate in the stock market. These noise traders focus on historical price patterns and expect trends to continue. It is common for them to buy stocks when the prices are rising and sell when prices are declining (Tokic 2007; Wärneryd 2001, 52). Noise traders are often interested in the stocks of small companies that are cheaply priced as it is easier to visualize a bigger profit from a stock that is less costly. They do not often grasp the extent of risk of the stocks or the portfolio they own and are too optimistic about the performance of their stocks (Shleifer 2000).

According to Bazerman (2006, 108-109) people had forecasted a 50 % better profit for their portfolio on a 6 month period of time than what was the actual profit during years 1985-1994. And when the investors thought that they had earned a profit similar to the market portfolio they had actually earned 8 % below it. Nofsinger (2008, 89) points out that optimism justifies the former investment decisions of investors and that overoptimistic investors conduct less critical analysis of the stocks they own and disregard negative information about their stocks.

Many trading strategies based on pseudo-signals, noise, financial models and price patterns tend to be correlated due to the fact that different judgment biases afflicting investors in processing information tend to be the same. Subjects in psychological exams have the tendency to make the same mistakes. These judgment biases can lead to aggregate demand shifts and thus influence the price of a stock (Shleifer & Summers 1990).

Although it might seem that uninformed noise traders would detract market efficiency, they might actually be essential for the functionality of the stock market. According to Black (1986, 531-532) "noise makes the financial markets possible, but also makes them imperfect". He has argued that although noise may make the market seem imperfect at times, it is crucial for the financial market and to its liquidity. If there are no noise traders then there will be very little trading on the market. As differences in beliefs must be created by some differences in information, a trader with a special piece of information acknowledges that other traders might have their own special pieces of information and will not rush out to trade. However, noise traders trade on noise as if it was valuable information and they are willing to trade even though they might be better off not trading. As the amount of noise increases it becomes more profitable for the rational traders to trade on information and try to gather new meaningful information and conduct research.

De Long et. al (1990) suggest that noise trading creates an additional risk element and that it should be priced in a security and that securities that are subject to greater noise trader risk tend to exhibit greater volatility and mean reversion. Noise traders can also garner larger returns than rational investors by bearing more of the risk that they themselves create. Noise trader risk can make assets less appealing to risk-averse rational arbitrageurs and thus the assets might be traded at a discount. If noise traders in general overestimate returns or underestimate risk, they invest more in the risky asset and can earn higher average returns. On the other hand noise traders can create price pressure when sentiments drive noise traders to buy a security, which in turn drives the price of the security up and lowers the future return. Investor sentiment can affect both noise traders and rational arbitrageurs in a way that drives the asset prices too high or too low in the short term which is followed by mean reversion in the longer term.

Smart money traders might also be affected by the noise and they might not always even try to make balancing efforts. Institutional investors and fund managers might be tempted to make noisy investment decision in order to attract new customers. The people that invest in mutual funds usually expect the fund manager to have superior knowledge about the stock market or certain stocks and that new information about the prospects of stocks triggers the trading in the fund. Trading can be seen as a signal of the fund manager's private information and expertise. The person who invests in the fund cannot recognize which trades are based on information, thus the fund manager may have an incentive to trade without new meaningful information (Trueman 1988).

Many investment banks have corporate finance services, such as services for IPOs and mergers, trading services and research services. Trading services try to maximize the amount of transactions and this can be done by giving more positive suggestions to buy shares as many private investors are unfamiliar with short sales or dislike to sell stocks short. The bank might also make biased analysis of a company that is its corporate finance customer as the bank tries to please the company and ensure the continuity of their business relationship (Michaely &Womack 1999).

The bonus schemes of analysts might be tied to the profit the bank makes which might increase the biased behavior of them. It is more profitable to give positive buy suggestions because the stock prices rise and more companies are in need of different corporate finance services (Marttila 2001, 146-148). Issuing sell suggestions is more risky to an analyst since they are rarer and more often deviate from the consensus opinion. Wrong suggestions that deviate from the consensus might cost the analyst his/her job (Womack 1996, 165).

Imam, Barker and Clubb (2008) investigated the valuation methods that analysts use in Great Britain and found out that analysts use valuation models that their customers expect them to use. The formation of target prices was quite subjective and usually the target prices were set according to what the market was ready to hear. The aggregate sentiment of the market did have an influence in their price targets.

Womack (1996) discovered that the buy and sell suggestions of the analysts did have an impact on share prices and the impact was more pronounced on small firms. This finding is aligned with the previous noise trader characteristics, that noise traders follow investment advice and that they are more interested to invest in small companies. In a way investment banks can have an influence in noise trader sentiment and presumably an impact in the stock market.

2.2.3 Limits of arbitrage

The reason why irrational noise trader sentiment might have an impact on security prices is due to limits of arbitrage. Arbitrageurs are likely to be risk averse and have short investment horizons. As a consequence, their willingness to take positions against noise traders is limited. One of the greatest risks for a rational arbitrageur is that the noise traders' beliefs will not revert to the mean for a long time and might become even more extreme. If the arbitrageur has to liquidate his position before the price recovers, he makes a loss (De Long et al. 1990).

Rational arbitrageurs are usually considered as some form of institutional investors such as fund managers and hedge funds. The time periods of these investors are quite limited because their clients might be in need of money or just want to liquidate their own share of the fund. When the arbitrageur manages other people's money and these people are unaware what the fund manager is doing and all they observe is whether the manager is making them a profit or a loss at the specific time, they might liquidate their share before the arbitrage position creates a profit. If a client wants to liquidate his share of the fund, the fund manager must execute the transaction. (Shleifer & Vishny 1997).

As the initial investors or clients of the fund might act in an unpredictable way and because the fund manager's job might be at risk if the fund performs poorly, the fund managers are not willing to take big opposite positions as there is rarely a perfect arbitrage without any risk (Gemmill & Thomas 2002). Arbitrageurs must borrow cash or securities to implement their trades and therefore have to pay their lenders per period fees. These fees cumulate over the period that the position remains open and can add up to large amounts for long term arbitrage. In addition, the performance of money managers is usually evaluated once every few months. These factors also make it reasonable to assume that arbitrageurs have relatively short horizons (Shleifer & Summers 1990). In the stock market there is no riskless hedge for the arbitrageurs. The arbitrageur cannot just buy certain stocks and sell the absolute substitute portfolio. When the stock is underpriced or overpriced the arbitrageur can only buy or sell the stock and hope that the price corrects to its fundamental level. When a stock is overpriced compared to its competitors the arbitrageur might short sell the overvalued firm and buy the undervalued one. However, if some unexpected positive news arrives about the overvalued company the stock price might rise even higher or if some negative news arrives about the undervalued company the stock price might plummet even more (Shleifer & Summers 1990; Malkiel 2007, 241).

There are many costs to short selling apart from the obvious financial risks of almost infinite loss opportunities that short sales entail. These costs can be bureaucratic, psychological and social. Some countries' governments might have restricted the possibility of short sales or even forbidden it. Even in countries where short sales are allowed, the institutions that are supporting them may not work well. One of these reasons is caused by political pressure towards short sellers. In many societies short sellers are blamed for different types of incidents and financial catastrophes. There might be a widespread antipathy towards them which causes social pressure (Shiller 2005, 182-183). The psychological costs can be inferred by the wide risks of short selling. The almost infinite loss potential and the possibility that the real owner of the stocks wants to sell the shares and that you might be forced to return the shares incur emotional pressure to the arbitrageur. There has been ample evidence that people are far more upset by losses than they are pleased by equivalent gains and that people try to avoid losses by taking even greater risks. This pain of regret ought to cause short sellers to want to avoid covering their shorts in an unfavorable situation. People want to avoid psychologically difficult decisions and might not be that keen to take large short positions. In addition, in some rare cases short sales might not be possible as there are no available shares to lend (Shiller 2003, 97-102).

The most recognizable rational arbitrageurs are usually hedge funds but they might not always act in a way that restores the security prices to their justified or fundamental level. Brunnermeier and Nagel (2004) found out that the hedge funds did not act as a corrective force during the Internet boom in the beginning of the new millennium. They actually rode the internet bubble and the average hedge fund was heavily tilted toward internet stocks although they were ridiculously overpriced. Brunnermeier and Nagel conclude that their findings were consistent with the view that the investor sentiment driving the stock market bubble was predictable and the hedge funds seemed to exploit it. Buying a 30 dollar stock with a "correct" or fundamental value of 15 dollars is a good investment if some other fool is willing to pay 60 dollars for it in the future (Malkiel 2007, 241).

There have been numerous occasions of prolonged arbitrage possibilities in several stock markets. De Jong, Rosenthal and Dijk (2008) studied the limits of arbitrage in real-world equity markets by comparing the theoretical values of dual listed companies and looking at possible mispricing. A Dual listed company structure is an agreement of two companies to operate as a unified enterprise while retaining their separate legal identity and their existing stock exchange listing. The most well-known examples are companies such as Shell/Royal Dutch and Unilever NV/PLC. As the companies are perfect substitutes to each other the prices should be quoted to be at the same price level. De Jong et al. found out that the average mispricing of all 13 known dual listed companies ranged from 2.5 to almost 12 percent, while maximum deviations were from 15 to nearly 50 percent. They also found out that the standard deviations among dual listed companies are very high and that the relative return of a twin is strongly affected by the fluctuations of the domestic market indices.

Another popular example of limits of arbitrage and the pricing irrationalities is the case of Palm and 3Com. 3Com was a profitable information technology company that sold different network services and systems. One of their products was a hand held computer, Palm pilot and it was produced by their subsidiary, Palm. 3com decided to spin off Palm into its own company and distribute 95 percent of the shares to the owners of 3com and issue the rest on an initial public offering. For every 3Com stock owned the owners would get 1.5 new Palm stocks in addition to their 3com stocks. On March 2, 2000, 3Com sold the 5 percent of Palm in the IPO and the rest was to be distributed to the owners of 3Com later in the year. By the end of the IPO day, the newly issued shares of Palm were quoted at 95.06 $. As the owners of 3Com would receive 1.5 shares of Palm, the stock price of 3Com should be at least 1.5 x 95.06$ + the value of 3Com which should equal to a minimum of 142.59 $ in a case where the operations of 3Com have zero value. However, 3Com stock was valued at 81.81 $ by the end of that day meaning that the operations of 3Com were calculated to have a negative value of 54.54 $ per stock in respect to Palm. According to the previous income statement 3Com made a profit of 750 Million. The mispricing between these stocks continued for months. It is likely that optimistic and irrational investors drove the prices through the roof and the relative illiquidity of the shares had an arbitrage limiting impact (Nofsinger 2008, 90).

Speculative stocks are both difficult to value and difficult to arbitrage because of the subjective attributions of the stocks as there is greater uncertainty of their future. Uncertainty means that the effect of many psychological factors such as overconfidence, representativeness and conservatism is more pronounced. The next chapter is dedicated to various psychological factors that can have an impact in the formation of investor sentiment and have an impact in the behavior and decision making of investors.

2.3 Psychological factors influencing the formation of irrational sentiment

2.3.1 Biases

Psychological and cognitive biases have an impact on a person’s decision making process and can lead to suboptimal decisions. In uncertain circumstances, the assessments of people tend to differ from rationality in systematic ways. For changes in investor sentiment to have a significant impact on returns, the buy and sell decisions of individuals must be correlated. The trading decisions of individuals have been recognized to be highly correlated and one possible explanation for this is the same shared psychological biases among people (Barber, Odean & Zhu 2009).

The most common biases are related to the tendency to trust one’s abilities, beliefs and traits over extensively and the tendency of overoptimistic evaluations of matters concerning the future. When people trust their own personal judgment and skills too much they are often referred to as being overconfident (Malkiel 2005, 224). Good examples of overconfidence are that over 80 percent of people regard themselves as better than average drivers or when people were asked to estimate a 99 percent confidence interval of stock prices there was a an actual surprise rate of 20 percent, meaning that people chose too narrow confidence intervals (Kahneman & Riepe 1998).

Psychologists have determined that overconfidence can cause people to overestimate their knowledge, underestimate risks and exaggerate their ability to control different situations. People seem to be more overconfident when they feel like they have control of the outcome (Nofsinger 2008, 10). This is related to "magical thinking" which means that people have occasional feelings that certain actions will make them lucky or make a better result even if they know logically that certain actions cannot have an effect on their fortunes. It has been shown that people will place larger bets on a coin that has not yet been flipped than on a coin that is already in the air. In addition people tend to think that a lottery ticket that they have selected is better than a random lottery ticket and demand more money in order to sell their ticket to someone else compared to a ticket they did not choose. In the stock market this can be related to the "hot hand syndrome" which means that previous successful investments are considered as signs of great personal skills, knowledge and intuition about the market and this will fuel the willingness to make more trades (Shiller 2005, 152-153). An overconfident investor is a person who overestimates the precision and quality of his private information. When an investor receives public information that confirms the previous trade, his confidence rises. Nevertheless, when the public information is negative in relation to the trade, his confidence falls only moderately if at all as the failure is seen as bad luck or bypassing external noise. This is known as biased self-attribution. Such continuing overreaction can cause momentum in prices and leads investors to buy more of the asset when the traders receive good signals and sell more of the asset with bad signals. And when the sentiment rises, noise traders hold more of the risky asset than rational investors (Daniel, Hirshleifer & Subrahmanyam 1998; Wang 2001).

Success leads to overconfidence and overconfidence leads to excessive trading and risk taking. The feedback of previous successful investment decisions increases the investor sentiment and fuels the overconfidence even more. In a way overconfidence can help the propagation of a speculative bubble when many investors suffer from overconfidence at the same time. During bull markets individual investors will attribute too much of their success to their own abilities making them overconfident. Thus, overconfidence will be more pronounced in bull markets. Overconfidence has been seen as a fundamental factor promoting the high volume of trade observed in speculative markets (Nofsinger 2008, 11-18; Shiller 2005, 152-154).

Even though overconfidence is usually seen as a trait of an unsophisticated investor, overconfidence has also been observed among professionals such as investment bankers and managers. Overconfidence was found to be strongest for tasks of moderate to extreme difficulty and seemed to increase with the personal importance of the task. People were also more confident of their predictions in areas where they have selfdeclared expertise. Overconfident investors trade too much, do not make careful probabilistic estimates and overreact to their private signals and underreact to public signals (Dittrich, Güth & Maciejovsky 2005).

Allen and Evans (2005) have suggested that overconfident traders might even dominate rational investors at certain circumstances and have an impact on trading volume, market depth and distribution of wealth. Wang (2001) made a finding that moderate overconfidence or bullish sentiment can survive in the long run and affect the stock prices.

Men have been recognized to be more overconfident than women in many aspects and the behavior in the stock market is no exception. Overconfidence leads to excessive trading that is not awarded by heftier profits. In the study of Barber and Odean (2001) women turned their portfolios over approximately 53 percent annually. In comparison the turnover of the average male portfolio was 77 percent. These trades were not justifiable by more information as the portfolios would have earned a better return without the trades. Men seemed to lower their returns more through excessive trading. The difference between men and women was most pronounced between single men and women.

Overconfidence is seen to originate from a person's ability to process information, biased evaluation of evidence and reactions to imperfect signals about the reliability of information. In other words, people have cognitive capacity limits and tend to overweight the reliability of highly unreliable information (Allen & Evans 2005). Nofsinger (2008) proposed that overconfidence stems from the illusion of knowledge which is the tendency for people to believe that the accuracy of their forecasts increases with more information even though they actually lack the ability to interpret the information correctly.

The overconfidence is believed to have an impact on the overvaluation of growth stocks and speculative stocks as overconfident investors might be extremely certain of the accuracy of their estimates of a company's growth prospects which tend to be overoptimistic. Hindsight bias promotes overconfidence and leads to the illusion that past events were foreseeable (Malkiel 2007, 227).

Hindsight bias means that a person perceives his or her own performance better than it actually was. It is the inability to correctly remember one's prior expectations after observing new information. Hindsight biased investors will form inaccurate beliefs about asset returns, volatility and hinders the investor's learning (Biais & Weber 2008). Shiller (2005, 154) carried out a survey after the stock market crash of October 1987 and asked investors whether they thought on October 19th, 1987 that they had a pretty good idea when a rebound was to occur. Of both individual and institutional investors nearly 50 percent said they knew what the market would do that day. This question was followed by a question that tried to chart out the reason why these people thought that the rebound was due to occur that day. References were made to intuition, gut feeling, common sense and historical evidence or market psychology. Mentions of concrete facts were rare even among institutional investors.

Overconfident investors tend to gamble more and a form of conscious risk and thrill orientation is known as sensation seeking in the psychological literature. Sensation seekers search for novel, intense and varied experiences. In investing, new trades bring excitement as trading is entertaining per se. These trades should be focused on stocks that can be considered more risky or speculative in nature as bond-like stocks are not as interesting. Men are considered to be more sensation seeking than women and men are more intrigued by action-oriented forms of gambling (Grinblatt & Keloharju 2009).

These sensation seeking traders are a textbook example of irrational noise traders that trade because they like to trade or find it exciting. When the aggregate mood or investor sentiment of the market is positive, more noise traders that have a sensation seeking attribute are withdrawn to the market. As certain stocks can be considered more exciting than others, these stocks are more likely to gain the attention of sensation seekers which can lead to an even strengthened sentiment and affect the stock prices.

2.3.2 Heuristics

Psychological research has shown that the human brain uses shortcuts in order to reduce the complexity of analyzing information. Using these shortcuts allows the brain to organize and rapidly process large quantities of information. However, these shortcuts also make it hard for investors to analyze new information in a proper way and can lead to inaccurate conclusions (Nofsinger 2008, 63). Individuals develop rules of thumb which are also known as heuristics to reduce the information-processing demands of decision making under uncertainty. Heuristics produce good decisions a significant proportion of the time but occasionally lead people to make systematically biased mistakes (Bazerman 2006, 13).

The first and most recognized psychologists to address the issue of decision making under uncertainty by the use of heuristics were Kahneman and Tversky (1974). In their study they conducted several tests on decision making and contradicted the prevailing perception of probabilistic or Bayesian decision making where people make rational decisions based on careful probabilistic evaluations. Kahneman and Tversky introduced three heuristics: *representativeness, availability and anchoring*.

The essence of the *representativeness heuristic* is that things that share similar qualities are considered quite alike. Representativeness is judgment based on stereotypes rather than on probabilities (Nofsinger 2008, 63). People understand the relevance of base-rate information, but tend to disregard such data when descriptive data are also at hand. Sample size is rarely a part of our intuition and it is often ignored when making heuristics based decisions (Bazerman 2006, 22-23).

Consider this type of example of representativeness: Tom is a shy person who enjoys quiet nights at home reading his favorite books. He likes to keep his personal belongings in order and albeit being helpful he has little interest in meeting new people. When people are advised to choose Tom's occupation from a list that might contain work positions such as a farmer, an airline pilot, a librarian and a firefighter people tend to choose the occupation that fits Tom's description even though farmers accounted for 75 % of the population (Kahneman & Tversky 1974).

According to representativeness heuristic people expect small samples and short time series of data to be representative of the whole population or distribution and make decisions based on very little information. Representativeness heuristic leads investors to overweight the importance of past returns when determining future returns (Barber, Odean, Zhu 2009). Investors tend to predict outcomes that seem most representative of the evidence. This can lead to extreme predictions (Marsden, Veeravaragan & Ye 2008). Barber et al. (2009) investigated the effects of psychological biases among investors and found out that the buying decisions of individual investors are highly correlated with past returns and in line with the representativeness heuristic.

Representativeness can cause people to have misconceptions of chance. Patterns that seem more representative of the probability distribution are regarded as more likely than more peculiar patterns. For example a sequence of coin flips of H-T-H-T-H-T is considered more likely than a sequence of H-H-H-H-H-H even though the probability of these is exactly the same. If you were asked whether a heads or a tails was more probable in the latter sequence many people would answer that tails is more probable. This is known as gambler's fallacy (Tversky & Kahneman 1974).

Representativeness errors are not scarce in financial markets. Individual investors often confuse a good company with a good investment. Investors often overestimate the company's ability to sustain high levels of growth and think that the most recent past is representative of what will happen in the future. On the aggregate market level bull markets and bear markets are considered to continue for extensive periods. Investors might be interested in spotting trends or chasing mutual funds that have performed well in the past and are regarded as "hot" (Malkiel 2007, 227-231; Nofsinger 2008, 64-66).

When people use the *Availability heuristic* they assess the probability of an event or the frequency of a class by the ease which different instances or scenarios can be brought to mind. For example, one may evaluate the risk of heart attack among middle aged men by recalling such occurrences among acquaintances (Tversky & Kahneman 1974). Events that are easier to recall are experienced as more important and more probable. When an instance is more emotionally arousing and easier to imagine in great detail, it is considered more likely (Hirshleifer 2001). The availability heuristic can stem from a person's own experience, memory or imagination. Imagination or explanation processes can lead to heightened availability of an event because a person who has imagined an event has constructed a mental image of that event. As this image has already been formed and is readily available in memory, any subsequent consideration will lead to the same image reconstructed more easily. In addition an initial construction of an event may create a cognitive set that impairs the ability to picture the event in alternative ways. When an event is easier to imagine it is considered to be more frequent (Wärneryd 2001, 127-129; Sherman, Cialdini, Schwartzman & Reynolds 1985).

Increased reporting and news coverage makes events easier to bring in mind and elevates the considered likelihood of a similar event happening. Attention grabbing news is likely to affect people's perceptions of the probability of certain instances occurring (Harvey 2007, 9-11). In the stock market the use of availability heuristic can lead to an individual investor bias to select well-known companies and brands or companies that get a lot of positive news coverage. This can lead to the overvaluation of glamour stocks when the market is bullish and when new unsophisticated investors pour to the market constantly. According to Barber and Odean (2008), individual investors have limited attention and cognitive capabilities and limit their search of stocks to those that have caught their attention. Investors' attention can be attracted by news, high trading volume and extreme past returns. Attention based purchases could temporarily inflate the price of the stock and yet withdraw more unsophisticated investors to the market thus increasing noise trader sentiment due to the positive feedback of past returns.

Anchoring means that people often develop estimates by starting from an initial anchor, based on whatever information is provided, and adjusting from the anchor to produce a final answer. Adjustments away from the anchors are rarely sufficient and the existence of an anchor leads people to access information that is consistent with that specific anchor and disregard information that is inconsistent with the anchor. People often place too much emphasis on first impressions and fail to adjust their opinions appropriately at a later date. When the anchor is externally set, the anchor leads to a biased search for information compatible with the anchor. In contrast, when someone develops his own anchor, he will use that anchor as a starting point and insufficiently adjust away from it (Bazerman 2006, 29-31).

In a famous test conducted by Tversky and Kahneman (1974) people were asked to estimate the percentage of African countries in the United Nations. This was done after the test subjects had to spin a wheel of fortune which had values ranging from 0 to 100 percent. Then they were asked whether the actual amount of African nations was higher or lower than that number and requested to make a specific estimate of the actual amount of African member nations in the UN. Even though the participants knew that the anchor was random and unrelated to the judgment task, the anchor did have an astonishing effect on their judgment.

Extrapolating from past information about a variable is probably the most common type of forecasting. Results from many studies have shown that people tend to use the last data point in a time series as a mental anchor and then adjust away from the anchor to take into account the major feature of the series. Typically the last point on a series lies on a trend line. People anchor on this last point and then make an adjustment for the trend. Anchoring leads to under-adjustment and investors are slow to modify their own perceptions of the market or a single stock. New relevant information might not be fully incorporated in the price or the price formation process might be lagged. The initial forecast serves as a mental anchor and people regard their abilities, opinions and judgments better than those of most people. As a conclusion, they are reluctant to fully revise their former estimates (Harvey 2007, 18-20).

Although the previous characteristics of investors that are influenced by anchoring are in line with our previous descriptions of noise traders, also investment professionals are shown to be subject to anchoring. This has been shown to be most prevalent among analysts. In the studies of both American and Australian analysts' behavior, it has been shown that analysts are subject to anchoring and their prior forecasts are powerful anchors and they are slow to make relevant adjustments in the face of new information (Marsden, Veeraravaghan & Ye 2008).

Heuristics can have an impact on investor sentiment as they guide the behavior of many unsophisticated noise traders. Search for distinctive patterns, overweighing of personal and often irrelevant information and oversimplified assumptions of the future are all traits of a noise trader and are in line with the use of heuristics in decision making. Noise traders both create and are most prone to the noise trader sentiment.

2.3.3 Emotions

Modern theories in cognitive psychology and neuroscience indicate that there are two ways how people comprehend risk. The analytic system of the brain uses algorithms, logic, probability calculations and risk assessment. The experiental system is intuitive, automated and not easily accessible to conscious awareness. It has facilitated the survival of human beings under the long period of evolution. It relies on images and associations linked by experience to emotion and affect. People do not only base their judgments on an activity, technology or a company solely on what they think about it, but also on how they feel about it. If their feelings about an activity are favorable they are leaned toward judging the risks as lower and the benefits as higher than in opposite situations (Slovic, Finucane, Peters & MacGregor 2004).

Although emotional responses are typically considered as being irrational, recent research suggests that rational decision making and emotion are complementary. Emotions and intuition can help a person create a fast and adequate decision when there is an abundance of information related to the task and the decision process would otherwise be paramount (Ackert, Church & Deaves 2003).

The use of feelings increases when the activity or judgmental task induces hedonistic pleasure such as imagining a future level of wealth and how that wealth can be applied. People who have a higher experience regarding a task are less likely to rely on their feelings. The impact of feelings in decision making is more pronounced under conditions of low processing capacity (Schwarz 2010).

Two recent areas of stock market research have addressed the influence of feelings and emotions in investor decision making. The first area known as mood misattribution investigates the impact of environmental factors, such as the weather, body's biorhythms and social factors, on stock prices. Mood fluctuations due to external conditions and shocks are argued to have an impact on equity purchasing decisions as people in a good mood are presumed to make more optimistic evaluations and vice versa. In addition, a person who is in a good mood is more prone to use simplistic heuristics. The second research area looks at the impact of image on investor decision making, the argument being that the image of a stock induces emotions that drive the investment behavior. For example, a company that has a positive publicity image and is touted in the media is regarded as a good company that is a good investment choice (Lucey & Dowling 2005).

Hirshleifer and Shumway (2003) investigated the relationship of sunshine and weather conditions and stock market returns in 26 cities around the world during a time period of 1982-1997 and found out that sunshine is highly significantly correlated with daily stock returns. An annualized nominal market return between perfectly sunny and perfectly cloudy days in the New York stock exchange was 24.8 percent and 8.7 percent per year respectively. Creating a profitable trading strategy seems to be difficult due to active trading and transaction costs. However, an investor with very low transaction costs could modestly improve his portfolio's Sharpe-ratio by trading on the weather.

Kamstra, Kramer and Levi (2003) studied the effect of seasonal affective disorder in several stock markets around the globe. Seasonal affective disorder is related to the amount of daylight and the response of a human being to the varying amounts of daylight during different seasons of the year. It is a condition that affects people during the seasons of relatively fewer hours of daylight. Kamstra et al. suggest that seasonal affective disorder does exist in stock markets around the world. The effect is more pronounced in northern countries where the amount of daylight can be quite dismal during winter periods.

Even sports results have been proved to influence the aggregate investor mood or investor sentiment. There has been a significant effect on investor sentiment the day after a loss in the soccer world cup elimination games. This was especially pronounced in soccer intensive countries. Wins do not have a significant impact on the market; only losses in the critical matches do (Edmans, Garcia & Norli 2006).

People in good mood react to new information and news events more favorably and are more subject to act on pseudo-information and irrelevant news (Hirshleifer & Shumway 2003). Investors often gamble both on an event outcome (e.g earnings announcement) and on the anticipated price appreciation of a stock. Anticipation of a reward generates a positive affect state. Positive affect motivates both added risk taking and added purchasing behaviors. Affective assessments of risk and rewards are influenced by the vividness of the imagined consequences, personal exposure with outcomes and a past history of conditioning. The vividly imagined possibility of wealth and investment success will lead to a strong drive to invest. Understanding the effects of affect can give insight in the origins of security price trends and reversals and why affect is a vital ingredient in the formation of investor sentiment (Paterson, 2002).

Lucey and Dowling (2005) state that there is a wide array of evidence that emotions and feelings do significantly influence decision making in matters involving uncertainty and risk. Extremes in emotions such as fear and greed lead to increasingly bounded rationality as emotions cloud the decision maker's judgments. Mehra and Sah (2002) argue that feelings of investors can have an impact on security prices if investor's risk aversion and judgment of an appropriate discount factor fluctuates over time in accordance with the fluctuations of mood, these fluctuations in mood are widely experienced and investors do not acknowledge that their decisions are influenced by their fluctuations in mood. This relies on the noise trader and limits of arbitrage assumptions.

In conclusion it can be stated that most people are affected by their emotions and feelings and this factors might have an influence in decision making. Inexperienced investors are more prone to be influenced by their feelings. Emotions and feelings are an essential ingredient in the formation of noise trader sentiment and by gaining knowledge of the impact of them in investor decision making we are getting a better understanding of the market mechanisms under extreme conditions.

2.4 Sociological factors influencing the formation of irrational sentiment

2.4.1 Herding

Herd behavior is often said to occur when many people take the same action by mimicking each other (Sornette 2003, 94). Herding is especially common in markets with less publicly available information where the value of the private information is larger. Individuals with access to information that is less accurate tend to follow the lead of those who are expected to have more detailed information (Zhou & Lai 2009).

Psychological research has shown that differences of opinion create anxiety and individuals try to reach a consensus of opinions. But the search of consensus may lead one to reconsider the validity of one's own opinions and adopt the consensus opinion. By a slight exaggeration, a herding investor acts like a prehistoric man who grouped together in order to feel more secure as he had inadequate knowledge of the environment. Herding is often related to an inefficient market situation which is characterized by sentiment extremes and things like speculative bubbles (Caparrelli, D'Arcangelis & Cassuto 2004).

The suppression of private information as herding gathers pace may lead to a situation where the market price no longer reflects all relevant fundamental information which moves the market towards inefficiency in an informational cascade (Hwang & Salmon 2004). An informational cascade occurs when individuals choose to ignore or downplay their own private information and imitate the actions of individuals that acted before them. This is likely to happen when the existing aggregate information becomes so overwhelming that an individuals' single piece of information is not enough to reverse the decision of the crowd. If this scenario holds for a single individual it is likely to hold for a majority of people acting after the person (Sornette 2003, 95).

In an uncertain situation, herding is more likely to be increasingly apparent and can become dominant in the market creating market bubbles and crashes. Herding is closely related to investor sentiment as market extremities require a large mass of people to act in a concert in the same direction (Hwang & Salmon 2004).

Investment decisions are rarely made in a vacuum and all the traders in the world are organized into a network of family, friends, colleagues and other contacts that are sources of opinion and influence each other. Other sources of influence involve all sorts of media (Sornette 2003, 99). According to an American study over half of the people get interested in a specific stock because some other person has mentioned that specific stock (Nofsinger 2008, 74). If the majority of people think that it’s the right time to buy or sell stocks it is very likely to have an impact on a person that is unsure of what to do. Especially in uncertain situations, the impact of social interaction has the largest effect on decision making (Wärneryd 2001, 205). The ability to communicate in the internet via e-mails, chat rooms, message boards and internet messengers has allowed us to extend our social network and reach a great number of people rapidly and with low costs. In finance, these social linkages play an important role in facilitating the exchange of information and investment ideas and spread sentiment which ultimately reflects in the share prices (Tay 2009, 207).

Herding is not likely to just affect non-informed individual investors. Mutual fund managers and other institutional investment managers have been observed to herd in and out of stocks. Reasons for this might be that the traders use similar models and trading strategies and fear of losing one’s job. If the perception of the trader was against the perception of the majority of traders and the trader realized significant losses, he might be in danger of losing his vocation. In addition, if the incentives of the manager are not based on relative terms but in absolute terms, there is no inducement to deviate from the actions of other managers (Grinblatt, Titman & Wermers 1995; Scharfstein & Stein 1990).

Capparelli et al. (2004) found that herding is most prevalent in extreme market conditions in the Italian stock market and fades away when the market is in “normal” conditions. Zhou and Lai (2009) make similar findings for the Hong Kong stock exchange. They also find that herding is more apparent when the sentiment is poor which indicates correlated sell situations labeled by fear. Small stocks and stocks with high P/E-ratios experience more herding than other stocks. Thus companies that are more prone to noise trader risk and have a more speculative nature (small companies with higher growth prospects) intrigue the herds more. Hwang and Salmon (2004) found out that herding does occur both in the American and South Korean stock markets and was especially prevalent among small stocks and stocks with high market-to-book ratios. Macroeconomic variables could not explain herding in the stock markets and herding was not limited just to market extremes but instead it was found to occur in more normal market conditions as well. It can be concluded that herding and sentiment are closely intertwined and herding has been observed to have a greater impact on small and high growth stocks which are expected to be more prone to noise trader risk and investor sentiment.

2.4.2 Media

A large part of our social environment is the media with a wide array of media shows and venues competing for our attention. Business and investment writers keep us interested by telling a good and captivating story. Most of the time, the media directs our attention towards storytelling instead of formal investment analysis (Nofsinger 2008, 78). Unfortunately, fear catches our attention better than any other emotion and by capturing more viewers or readers the company that provides the news is likely to get more revenue from advertisements (Read 2009, 216-217).

The news media are naturally attracted to financial markets because the markets provide constant news in the form of daily price changes. The general public considers the stock market as the big casino with people making or breaking fortunes. Many media accounts tell stories about the satisfaction felt by those who invested in the stock market in the past years and how the stock market is your chance to get truly rich. It is not surprising that financial news and sports news account for roughly half of editorial content of the newspapers. The content of the papers are directed towards the interests of the public in order to stay competitive. The level of public interest towards stock markets does vary over time and goes through fads from time to time (Shiller 2005, 61, 66, 86). The interest of the general public towards stock markets increased exponentially in the late 1990's and in the beginning of the 2000's. Tales of people who got extremely rich by buying stocks gained a lot of space in magazines and newspapers. In the beginning of the new millennium a lot of new economic and stock market related newspapers and TV shows were founded. As the stock prices rose more people got interested in the stock market day by day and investing started to be a sensible topic of conversation. The stock markets could be monitored in real time from news channels where investment advisors and analysts shared their views on the market. The media bulged with appraising comments of the IT-stocks and a common investor could not have remained unaffected by it (Marttila 2001, 214-217).

Significant market events generally occur when there is similar thinking among large groups of people, and the news media are essential vehicles for the spread of ideas. News events usually have the effect of setting a sequence of public attentions. Media attention through a cascade of attentions can lead investors to act on news that would normally be considered as irrelevant or noise. The media actively shape public attention and categories of thought and create an environment where speculative market events are played out. The news media can be fundamental propagators of speculative price movements through their efforts to make news interesting to their audience. They sometimes strive to enhance such interest by attaching news stories to price movements that have already been observed, thereby enhancing the salience of these movements and focusing increased attention on them (Shiller 2005, 85, 91-93, 105).

Large amounts of optimistic or pessimistic news headlines and stories can reflect sentiment extremities and provide a real time picture of popular sentiment and serve as anecdotal support for other measures of stock market psychology even though media headlines and cover stories should not be used as the basis of investment decisions alone (Davis 2003, 42-43). Tetlock (2007) investigated whether news media content could predict movements in broad indicators of stock market activity and the results were consistent with theories of noise and liquidity traders. He found out that high levels of media pessimism predict downward pressure on market prices and that high media optimism or pessimism leads to increased market trading volume and low market returns lead to media pessimism. The media content was detected to be related to investor sentiment and especially small stocks were prone to the effects of media pessimism which is in line with the noise trader theories that stocks of small companies are disproportionately owned by uninformed individual investors.

2.5 Possible consequences of irrational sentiment

2.5.1 Bubbles and market crashes

Episodes of speculative frenzy followed by a market crash seem the norm in our history. The advent of a new technology usually acts as a trigger to such episodes. A common element is the belief that the world has entered a new bright dawn of history that transforms both society and the productive possibilities open to mankind (Forbes 2009, 89, 101). Although each bubble has its own features and evident reasons that lead to the bubble, there are some common elements among them. It is critical to remember that the stock markets consist of people executing trades and people are inclined to short sightedness and making decision on an emotional basis. When the stock prices simultaneously sky rocket, a collective euphoria grasps the investors. However, when their greed turns to fear, the bubble can burst quite rapidly. It is essential to note that stock markets reflect the collective mood of the investors and investors are prone to excessive optimism and pessimism from time to time (Ferguson 2009, 108).

When people are gripped by fear, their investment horizons tend to get shorter and in extreme market situations the effect of investor sentiment has a larger impact on investment decisions. Overreacting is a part of human nature when life changing events may occur and your personal wealth and wellbeing is at stake. In a bull market people tend to overreact to good news and give little weight to bad news as during a bear market all the focus is directed towards bad news. Greed and fear have always moved the markets (Marttila 2001, 10-13, 25).

The behavioral finance literature has tried to provide an answer to why people are so easily gripped by fear and greed and do not act rationally. The main foundations can be found from psychology and sociology. People tend to be guided by behavioral biases, are overconfident and follow trends, do positive feedback trading and herd with the crowd (Bitmead, Durand & Ng 2004). In addition to mass psychology, many bubbles have been created in conditions of easy and inexpensive credit. There might have even been a financial innovation that contributed to the formation of the bubble by increasing the volume of trade as in the crash of 1929. Investors were allowed to buy stocks by just 10 percent of the stock price as collateral when the security certificate was stored at the bank (Ferguson 2009, 110; Forbes 2009, 106).

A crash occurs because the market has entered an unstable phase and any small disturbance or process may have triggered the instability. A crash may be caused by local self-reinforcing imitation among traders. A financial collapse has never happened when things look bad. On the contrary, macroeconomic flows seem strong before crashes. This explains why crashes catch many people by surprise (Sornette 2003, 3, 14, 23). Market bubbles and crashes are not a new phenomenon although they have been more frequent in the last hundred years. Tulip mania in the Netherlands in the 17th century was one of the first reported bubbles. A new kind of tulip sort was imported from Turkey that was more vibrant in color. The demand for these tulips increased exponentially and the scarcity of tulips and their beautiful colors made them a must for all the upper classes of society. Forthcoming tulip bulbs were sold in advance as like options. An option entitled its holder to buy a predetermined amount of tulip bulbs at a set price. As the prices of the tulips seemed to increase all the time, the tulip bulb options were traded furiously. Many people that participated in this mania mortgaged their houses and businesses to trade tulips because the trade of tulips was regarded to be a "sure thing". The tulip mania lasted from year 1634 until 1637 and during the last year the prices of tulips got twentyfold. The ingredients of a bubble were present in the Netherlands: an increasing currency, a new economy and an increasingly prosperous country had created the optimistic atmosphere in which bubbles are grown (Marttila 2001, 54-55; Sornette 2003, 7-9).

The railroads created a bubble in the United States and in Britain in the 19th century, and a large stock crash was experienced in 1929. The crash of 1929 was created by the irrational behavior of the investors, by the extensive speculative demand and investor optimism. Techniques of mass production, mass advertisement and the expansion of investment banking had created fertile conditions for a bubble. Added by easy and low credit and financial innovations the stock market was bound to overheat (Marttila 2001, 60-61; Shiller 2005, 112-113). The 50's and 60's experienced two booms. The first directed towards new electronics companies and the latter towards large and profitable stocks known as the nifty fifty (Baker & Wurgler 2006). Another stock market crash was experienced in 1987. The famous date when the Dow Jones stock market plummeted by 23 percent is known as the black Monday. The stock market has fallen by over 10 percent in one day four times in the history. If the stock market followed a normal distribution stock market falls by over 10 percent in a year would be extremely rare and falls of over 20 percent a year should not have happened (Ferguson 2009, 145).

One of the reasons for the black Monday has been acclaimed to be computer trading. This means that computers automatically execute large stock trades when certain market trends prevailed. In particular sell orders after losses. The illiquidity of the market increased the price change (Sornette 2003, 5-6). Nowadays 60 percent of the trades executed at the New York Stock Exchange are computer automated trades where as in Finland around 30 percent of the trades executed are computer automated (Meklari 2010, 28). These computer automated trades can strengthen the price fluctuations especially during uncertainty.

The IT-bubble in the late 1990's and in the beginning of the 2000's filled all the characteristics of a classic bubble. A new technological innovation that seemed to launch a new era of productivity increased the optimism of investors and consumers all around the world. The effects of the bubble were particularly strong in the small but technologically intensive country of Finland. Individual investors got interested in the new technology stocks as the stocks experienced fast growth in stock value. People queued to buy these stocks. By the previous success of Nokia many American investment funds grew an interest in Finnish stocks and the prices of stocks seemed to continuously increase (Marttila 2001, 77-79).

Media interest towards investing increased during the IT-bubble and many highly respected investment analysts suggested people to buy stocks despite they might have seemed to be expensive as the former valuation methods could not be applicable anymore. Many investment shows and related magazines were founded and investing was a generally acceptable topic of conversation. Companies that changed their names to dot-com experienced stock price increases and companies that made almost any kind of disclosures about their future goals made a price increase in the stock. Yahoo had a Price-earnings ratio of over 600 and people fought over the stock. Rational traders tried to exploit the mania rather than making stabilizing trades. The extreme sentiment of individual investors and herd behavior created the IT-bubble (Bitmead, Durand & Ng 2004; Saario 145, 294-342).

The financial crisis of 2008 was due to careless lending of money. Money was lent to people who could not afford it. The crisis was developed by overly aggressive mortgage lenders, compliant appraisers, overly optimistic borrowers and obscured incentives. The loans were packaged into securities that seemed to be well diversified when in fact the loans were distributed to people from poor areas and the chance of default was a lot higher and the loan defaults were more correlated than expected. People thought that the prices of houses keep going up and it does not matter if the debtor is unable to pay the debt as the houses act as collateral. Banks sold the collateralized debt obligations to other banks and to private investors and as banks were uncertain of the amount of other banks had of these instruments, interbank trust crumbled leading to the financial crisis (Shiller 2009, 3-11).

2.5.2 Sentiment and corporate decision making

Nofsinger (2005) proposes that the general level of optimism or pessimism in society affects the decisions of financial decision makers and can lead to market wide phenomenon. High general sentiment or social mood will lead to more optimistic investors and corporate managers and influences their perspectives of risk of decision outcomes and their estimates of probabilities of success. Investor sentiment is seen as the subset of the larger social mood. Over-optimism can cause corporate overinvestment and increase the amount of investment projects taken and acquisitions made. Gombola and Marciukaityte (2007) suggest that managerial overoptimism leads to undertaking negative net present value projects. Managers might also be inclined to use more debt financing as they are confident of the profitability of the future projects. Some managers may however exhibit a different type of behavior. They may take advantage of the general optimism of the market and issue equity when the market is high.

Baker and Wurgler (2002) studied equity market timing in their article. Equity market timing refers to the practice of issuing shares at high prices and repurchasing them at low prices. According to their findings, firms tend to issue equity when the market value is high and when investors are too enthusiastic by the earnings prospects. In addition, several managers have admitted market timing behavior in anonymous surveys.

Lamont and Stein (2006) made similar findings in their article. They concluded that both equity issues and merger activity is related to market values of companies and thus market timing is of importance. According to Shleifer and Vishny (2003), corporate policies such as equity issues, investments, share repurchases and level of dividends are driven by market mispricing. Companies are seen to adjust their policies according to investor sentiment and the relative valuation of the stock market compared to the fundamental value. Baker and Wurgler (2004) suggest that corporations cater for the investor demand for dividends and pay heftier dividends when they are at a premium and smaller dividends when they are not appreciated to the same extent. When investor sentiment is low, investors appreciate safer stocks that pay a reasonable dividend and managers are regarded to take this into account.

When the sentiment is high more companies make an initial public offering. Part of this can be explained by the fact that usually macroeconomic conditions are favorable at the same time. Companies can also get a higher price for their shares and be confident that there is sufficient demand for the stocks. As sentiment grows IPO offer sizes grow and lower quality companies are taken public. In addition, companies become more likely to raise funds for non-investment purposes (Ljunqvist, Nanda & Singh 2003).

Companies might even adjust their voluntary corporate disclosure in relation with the prevailing investor sentiment. During periods of high sentiment managers reduce the frequency of long horizon earnings forecasts to the market while during low sentiment they are prone to increase these forecasts. The reason why managers report more during low sentiment may be due to their attempt to increase firm value as the firm is seen to be undervalued and in order to correct the analysts' pessimism (Bergman & Roychoywdury 2008). In some cases companies have exploited investor overoptimism by increasing the amount of voluntary disclosures about the future. Several high tech companies exploited the market situation during the IT-bubble in Finland and painted rosy pictures about the future without any concrete actions. The CEO of Balansor has later said they noticed that every time they issued a voluntary disclosure about the future, the stock price would rise by about 15 percent (Marttila 2001, 97-98).

2.6 Hypothesis formulation

In order to crystallize the main research questions of this thesis, the following hypotheses are formulated based on the introduced behavioral finance theory:

H1: Investor sentiment has an impact on the aggregate market returns in the Finnish stock market

H2: Investor sentiment affects the returns of speculative and non-speculative stocks in different ways

H3: Sentiment can be used to forecast returns on sentiment prone stocks. Sentiment prone stocks fare worse than bond like stocks when sentiment is high and fare better when sentiment is low.

It is tested whether sentiment is related to stock returns as predicted by the behavioral finance theory. The relationship between stocks and sentiment for the same time period is analyzed first. It is expected that sentiment is at least related to the current returns of stocks. The impact is regarded to be more powerful for the small stocks and stocks held by predominantly individual traders although even institutional traders have been reported to herd and act in an irrational manner from time to time.

The investors' demand for speculative and non-speculative stocks is regarded to fluctuate with sentiment. Thus, investor sentiment is seen to affect different types of stocks in dissimilar ways. At times investors are keen to invest in speculative stocks and at times they may prefer safer stocks.

We expect that investors' tend to overreact and the effects of overoptimism and overpessimism drive sentiment prone stocks to be overvalued when sentiment is particularly high and undervalued when sentiment is low. Investor sentiment is regarded to be related to these patterns of over and undervaluation. Hence, sentiment related mispricing correction should be observed when sentiment is matched by the subsequent returns of sentiment prone stocks. In addition, it is assessed whether investor sentiment could be used as a profitable trading strategy by exploiting sentiment related patterns of mispricing.

3 DATA AND RESEARCH METHODS

3.1 Data and sample selection

3.1.1 General

The purpose of this study is to find out whether investor sentiment has an influence on the returns of stocks and if the effect is larger on stocks that are considered to be more sentiment prone. The largest emphasis is placed on forecasting returns with sentiment and on spotting sentiment related return patterns. The behavioral finance theory proposes that sentiment leads to mispricing that is corrected in the subsequent periods. Hence, sentiment should be able to be used in order to forecast returns if sentiment leads to systematical mispricing of stocks.

This thesis revolves around two types of data: sentiment related data and returns data. Various sentiment proxies are considered to be used in order to create an index that captures the levels of investor sentiment in Finland. However, not every proxy that has been used in previous literature works in a similar fashion in every time period and every market. A sentiment proxy that seemed to grasp the essence of investor sentiment in the 1970's in the American stock market might not necessary do the same thing when applied to the 21st century and the dismal Finnish stock market.

The sentiment index is constructed in a similar manner as in Baker and Wurgler (2006) who used Principal Components Analysis in order to create a sentiment index from six sentiment proxies that were NYSE turnover, closed end fund discount in percents, number of IPOs and their first day returns, equity share in new issues and dividend premium which is the log ratio of value weighted average market-to-book ratios of dividend payers and non-payers. These measures were regressed on different macroeconomic factors such as growth in industrial production, growth in durable goods, nondurable goods and services consumption and growth in employment in order to form an index that reflects sentiment that is not warranted by fundamentals.

The sample period of this thesis is from January 2000 until December 2008. This time period is chosen as there has been both market peaks and market crashes in this quite short time horizon that have been regarded to be somewhat influenced by market psychology and investor sentiment. Another reason for selecting this time period is due to data availability and an interest to discover whether investor sentiment has an impact in the 21st century. The data of many sentiment proxies is quite fresh and do not go far back in the history. The sentiment index was contemplated to be formed on a quarterly basis but as the proxies that were only available quarterly did not seem to correlate with the other variables sufficiently or act in an expected manner, the time interval could be reduced to monthly sentiment data. This is more appealing as when the size of the sample increases, the reliability of our tests increases as well. In both Principal Components Analysis and in regression analysis the sample size should be sufficiently large (Metsämuuronen 2008, 29, 88).

The investor sentiment proxy data and returns data was collected from various sources such as Datastream, Bank of Finland and Eurex exchange. The data was first processed in Microsoft Excel for both sentiment related and return related data. After the data had been processed and modified in Excel, the data was preliminarily analyzed in SPSS. However, SPSS is not able to perform certain procedures especially related to correcting autocorrelation and heteroscedasticity in regression analysis. Eviews was used for these purposes and all of the regressions were run on Eviews.

3.1.2 Return data

The monthly levels of investor sentiment index or monthly changes in the index are matched by returns of the subsequent time period of a specific portfolio. The time periods used in this thesis regarding the returns are monthly and quarterly returns following the sentiment of the previous month. In other words, the sentiment of January is used to explain the return of February when the forecast horizon is one month and when the forecast horizon is three months, the sentiment of January is used to explain the return of February to April. Baker and Wurgler (2006) used yearly sentiment levels in order to explain both monthly and annual returns of portfolios. Schmeling (2009) used monthly levels of sentiment in order to explain the subsequent monthly, semi-annual and annual returns. However, it is possible that a sentiment changes index is a more suitable measure as it is likely to be less autocorrelated. In such a case, the return of e.g March would be matched by the sentiment change in February.

The return data of all indices and portfolios spans out from February 2000 to January 2009 or March 2009 depending on the return forecast horizon. The returns are total returns that account for splits and changes in share capital, dividends and other value relevant transactions. The returns are calculated as logarithmic differences because they are more normally distributed than percentage returns and are more symmetrical. The attributes of logarithmic returns are usually considered more suitable for financial research. The equation for the logarithmic return is simply (Vaihekoski 2004, 193-194):

$$r_{it} = \ln \frac{P_{it} + D_{it}}{P_{it-1}} \quad (1)$$

where r_t is the logarithmic return, *ln* is the natural logarithm, P_t and P_{t-1} are returns at time period *t* after the dividend and D_t is the dividend at time *t* for stock *i*. All return data is gathered from Datastream.

The sample of Finnish companies consists of 154 companies of which 27 companies have been delisted, gone bankrupt or been merged and have made an exit of the Finnish stock exchange during the sample period. If a company is delisted, it is removed from the sample and if a new company is listed, it is included in the sample. This is done to mitigate survivorship bias.

3.1.3 Defining the test portfolios

The return portfolios can be classified as benchmark index returns and speculative stock portfolios based on volatility. The benchmark indices are gathered from Datastream and the indices in this thesis include MSCI Finland Value stocks, MSCI Finland Growth stocks, MSCI Finland Small value, MSCI Finland Large value, MSCI Finland Small growth, MSCI Finland Large growth and SHB Nordix Finland Small cap. The speculative stock portfolio returns are calculated from all the Finnish companies that are quoted on the Finnish stock exchange. The portfolios are sorted by historical volatility and the portfolios are equal-weighted. Portfolios that are harder to value and arbitrage are considered to be more speculative. More speculative stocks are considered to more prone to investor sentiment and volatile stocks are regarded to be both hard to value and arbitrage. Additional measures of speculativeness can be for example amount of intangible assets, indebtness, the company's time of operation, growth prospects etc. (Baker & Wurgler 2007).

The MSCI indices are total return indices that reinvest cash dividends in indices the day the security is quoted ex-dividend. The indices are calculated and distributed daily and are value-weighted (MSCI Barra 2010). The MSCI Finland Value and Growth portfolios are generated in a way that the portfolio covers 99 percent of the free float adjusted market capitalization. Thus, the companies are sorted in descending order by market capitalization in the desired investment style. Those companies that account for 99 percent of the total free float market capitalization are included in the portfolio. For the MSCI Large Value and Large Growth portfolios the free float adjusted market capitalization coverage needs to be 70 percent of total market capitalization, and those companies that account for 70 percent of total cumulative market capitalization are included in the portfolio. The MSCI Small value and growth portfolios are created in a way that the portfolios consist of companies that are between 85 and 99 percent in the free float adjusted market capitalization coverage. The minimum size of the company that is included in the portfolio is the size of the company that reaches the cumulative market capitalization of 99 percent of total market capitalization (MSCI Barra 2008).

The value and growth investment style characteristics are measured by a handful of ratios. When determining the value characteristics of a company, these three ratios are calculated: book value to price ratio, 12 month forward earnings to price ratio and dividend yield. These measures are standardized and winsorized and given equal weight to determine the value Z-score. The growth investment style characteristics are defined by these ratios: long term forward earnings per share growth rate, short term forward earnings growth rate, current internal growth rate, long term historical EPS growth trend and long term historical sales per share growth trend. These ratios are also standardized, winsorized and given equal weight to form a growth Z-score. If a company has a positive value in value Z-score and zero or negative value in growth Z-score, it is labeled as a value company and vice versa. If a company has both characteristics, the weights of each characteristic are defined, and those weights are used when calculating how much a certain company contributes to either a value or growth index. This weight is referred to as an inclusion factor (MSCI Barra 2007).

The speculative stock portfolios are formed on the basis of previous historical volatility of the companies listed in the Finnish stock market. Baker and Wurgler (2007) propose that high volatility is characteristic for highly speculative stocks and low volatility is predominant in stocks that have bond-like features. Highly volatile stocks are also considered to be harder to value and to arbitrage. As in the aforementioned article of Baker and Wurgler, a number of portfolios are formed on the basis of the previous 12 month annualized historical volatility. The annualized volatilities are calculated by multiplying the daily volatilities by the square root of 251 which is the average number of trading days in Finland (Vaihekoski 2004, 277).

The portfolios are sorted by the volatilities of the stocks and the portfolios are equally weighted. Equally weighted portfolios are used instead of value-weighted portfolios because empirical results and the theory we are interested in testing predict that large firms are less affected by investor sentiment and hence value weighting tends to obscure the relevant patterns (Baker & Wurgler 2006, 1646). Portfolios are formed on the basis of previous volatility and the portfolios are sorted in to high, medium and low volatility portfolios where the high portfolio consists of the first 20 percent of companies that have the highest volatility. The mid portfolio consists of companies that are in the range of 40-60 percent sorted by volatility or in the mid quintile. The low volatility portfolio consists of companies that are in the last quintile and account for companies that are in the 20 percent smallest volatility range. In addition, two portfolios that consist of 10 percent of the companies that have the highest and smallest volatility are formed in order to determine whether the effect of sentiment is even larger in the utmost volatile and least volatile companies. Companies that are delisted are removed from the portfolios and companies that have been listed are included in the portfolios. Thus, the group of companies is constantly up to date and accounts for survivorship bias.

Each portfolio is resorted monthly and the returns are total returns that account for dividends, splits and other financing activities. As proposed by the theory and by the finding of Baker & Wurgler (2006, 2007), it is expected that especially volatile stocks should be more prone to the effects of sentiment as more volatile stocks are harder to arbitrage, harder to value and can be regarded as more speculative. Thus, it is expected that noise traders and speculative traders are more interested in these stocks and that the market forces are unable to correct the mispricing, which leads to misvaluations that are later mean reverted. So it is expected that the returns of the highly volatile stocks should exhibit a negative relationship with sentiment as with the less volatile stocks sentiment should have a smaller impact.

3.1.4 Sentiment proxies

A wide array of different investor sentiment measures can be found from existing scientific literature. Still, there are no definitive or uncontroversial sentiment proxies that could capture the essence of sentiment in every market and time period (Baker & Wurgler 2006). That is why multiple sentiment proxies are considered in this thesis.

Baker & Wurgler (2006) used NYSE turnover, closed end fund discount in percents, number of IPOs and their first day returns, equity share in new issues and dividend premium. All of the same proxies cannot be applied in this thesis due to data availability. In the Finnish stock market IPOs are quite rare in number and there is not enough deviation in the yearly or monthly amount of them that could be used as a sentiment proxy. In the previous decade there had been just a handful of initial public offerings. Due to lack of volume the IPO first day returns are also ruled out. In addition to these sentiment measures, a range of measures such as consumer confidence, mutual fund flows, retail investor trades, insider trading, option implied volatility and put/call ratio are regarded to contain a sentiment related component. (Baker &Wurgler 2007, 135-137; Brown &Cliff 2004, 11)

Brown and Cliff (2004) divide investor sentiment measures to direct and indirect measures. Direct measures of investor sentiment are mainly based on surveys and questionnaires. These include consumer confidence surveys, industrial confidence surveys and other similar polls that try to determine how a specific group sees the future. The indirect measures are further grouped into four subgroups: market performance proxies, trading activity proxies, derivatives based proxies and other miscellaneous sentiment proxies. Market performance proxies relate to recent market performance, such as advance/decline-ratio. Trading activity proxies relate to trading activity among specific groups, such as the trades of retail investors. Examples of derivatives based proxies are put/call-ratio and option implied volatilities series. The miscellaneous proxy group consists of sentiment measures that cannot be solely categorized into the previous groups. An example of a proxy that is included in this group is closed-end fund discount. The ideal situation would be to create the sentiment index from a variety of different sorts of sentiment proxies.

Nine different sentiment proxies that can be divided into different groups are considered in this thesis. These proxies are closed-end fund discount, consumer confidence, industrial confidence, OMXH turnover, dividend premium, equity share, put/call-ratio, VDAX option implied volatility and advance/decline-ratio. Five of them seem to function in an expected manner. This means that the sentiment proxy itself seems to act in an unbiased manner and does not exhibit any unexplainable patterns, has a high correlation and in the expected direction with the other sentiment proxies and acts in a way that has been described in the previous literature.

The investor sentiment proxies' data were collected from various sources. Turnover, equity share and retail investor data was collected from the website of Bank of Finland. Option related proxies such as the put-call ratio and implied option volatility data was collected from Eurex exchange and survey related sentiment proxies were gathered from DG ECFIN which stands for Directorate General for Economic and Financial Affairs. Dividend premium data was garnered from Datastream and the closed end fund discount data was collected from the websites of Neomarkka and Norvestia and the missing data was obtained directly from the companies. All of the proxies' data is on monthly basis. Next every sentiment proxy is explained individually.

Closed-end fund discount: A closed-end fund is a mutual fund that holds other publicly traded securities. A closed-end fund issues a fixed number of shares that are traded on the stock market. In order to liquidate ones owning in the fund, the shares must be sold to other investors rather than redeem the shares from the fund itself. The closed-end fund discount is the difference between the net asset value of a fund's security holdings and the fund's market price. The closed-end funds are usually traded at a discount of 10-20 percent compared to their net asset value. A number of reasons are suggested as reasons of the discount. Agency costs, net asset illiquidity, tax reasons and the impact of noise traders are considered to be reasons that cause the discount. Holding the closed-end fund discount is riskier than holding the similar portfolio directly and the fund needs to be sold at a discount to attract investors. However, the discount widens and shrinks in relation how optimistic or pessimistic investors are about the future (Lee, Shleifer & Thaler 1991).

When investors are irrationally optimistic, closed-end fund discounts are smaller and when investors are pessimistic, the closed-end fund discount gets larger (Lin, Rahman & Yung 2009). In other words, there is supposed to be a negative relation between investor sentiment and closed-end fund discount. Neal and Wheatley (1998) found out that closed-end fund discount had a relation with the returns of small firms but did not find a relation between the discount and the returns of large companies which is in line with the noise trader theory. Baker and Wurgler (2006) used closed-end fund discount as one of their proxies and calculated the discount as the value-weighted average discount of available closed-end funds. As formula it can be expressed as following:

$$CEFD_t = \sum_{j=1}^{n} w_{jt} \times \left(\frac{NAV_{jt} - P_{jt}}{NAV_{jt}} \right), \tag{2}$$

where w is the weight of the closed-end fund of company j as a proportion of the total market capitalization of all closed-end funds at time t, NAV is the net asset value of the closed-end fund j at time t and P is the stock price of the closed-end fund.

Neomarkka and Norvestia are used as closed-end funds in this thesis as they are investment companies that do not have any other type of business activity and are listed on the Finnish stock exchange. The closed end fund discount is formed as the value-weighted sum of the closed-end fund discounts of Norvestia and Neomarkka. NAV is the fair value of the holdings of the investment company. The closed-end fund data regarding net asset values was obtained from the websites of Norvestia and Neomarkka and the price related data was gathered from Datastream. The proxy is formed on a monthly basis.

Consumer confidence and industrial confidence: Baker and Wurgler (2007) propose that different investor surveys can be used as sentiment proxies by straightly asking how optimistic or pessimistic investors seem to be about future. Lemmon and Portniaguina (2006) used consumer confidence as a sentiment indicator in their article and found a relation between stock returns and consumer confidence especially in the last 25 years. They suggested that in addition to consumer confidence being a measure of investor optimism, it responds to the tone and volume of economic news rather than content. This fits to the theoretical framework of the thesis on investor sentiment formation affected by noise traders. Also Schmeling (2009) used consumer confidence as an investor sentiment measure and found out that it had a relation with future returns in some countries. Industrial confidence measure is considered as well because it is very similar to the consumer confidence measure apart from that it is surveyed from industrial companies. Industrial confidence survey is a direct measure of sentiment of companies where as equity share and IPO volume can be regarded as indirect measures of company activities that are seen to be affected by the general level of investor sentiment.

The Finnish industrial confidence survey and consumer confidence surveys are conducted every month during the second or third weeks of the month. The total sample of companies surveyed is 850 and the total number of consumers surveyed is 2200 monthly. Both surveys are largely qualitative and the main questions in the industrial confidence survey relate to assessments of trends in production and stock levels and to expectations about future production, selling prices and employment. The consumer confidence survey tries to gain information about the spending and saving intentions of consumers and to assess the perceptions of consumers of factors affecting these decisions. The questions are organized around four topics: the financial situation of the household, the general economic situation, savings and intentions regarding major purchases (DG ECFIN 2007). The data is garnered from the website of DG ECFIN.

OMXH turnover: Baker and Stein (2004) claim that market liquidity measured as turnover can indicate investor optimism or pessimism and turnover can be regarded as a sentiment indicator. An unusually liquid market is a market which is dominated by irrational investors that can affect the prices of shares and causing the future returns to be lower. Their model is based on two assumptions that have already been introduced to the reader: the limits of arbitrage and investor behavior that is based on investor overconfidence. The overconfident investors can cause sentiment shocks by participating in the market. When their sentiment is overly positive the market is overvalued relative to fundamentals and when their sentiment is negative they do not participate in the market to the same extent. Thus, measures of liquidity provide an indicator of the relative presence or absence of these investors and of the level of prices compared to fundamentals.

Baker and Wurgler (2006) used NYSE turnover per share as a sentiment proxy in their article. Turnover had a positive trend in their time series and they detrended the turnover time series by a 5 year moving average. There appears to be a short term trend in our time series that spans from January 2000 to December 2008. However, using a 5 year moving average does not seem to work well with our series as the development in the Finnish stock exchange has been quite rapid and the increase in public interest towards securities markets has been tremendous in the late 1990's and in the new millennium. Turnover is defined as the ratio of current month's value of (euro) turnover per listed share series divided by the 24 month rolling average of turnover per share. By this way, we can account for the trend in the turnover and compare whether the turnover is higher or lower than in the recent past. As a formula it can be expressed as following:

$$TURN_t = \frac{Tps_t}{24m.avgTPS_t} \quad , \tag{3}$$

where Tps is turnover per share, which is the total turnover divided by the number of listed shares. 24m.avgTPS is the previous 24 month average turnover per share. As each of the monthly observations is detrended by the previous 24 month rolling average, the monthly turnover of January 2000 is divided by the average monthly turnover of 1998-1999 and so on. Thus, the first turnover observation in the denominator is from January 1998.

Dividend premium: Baker and Wurgler (2004) suggest that dividends are highly relevant to the stock price but in different directions at different times. Managers of companies cater for investors' demands for dividends and when higher dividends are appreciated among investors, the managers are more inclined to adjust their dividend policies towards those demands. Dividend premium variables may also be driven by sentiment and when the dividend premium is high, investors are seeking for companies that are regarded safer and pay larger dividends and when the premium is low investors prefer companies with large growth opportunities and potential for large capital appreciation. They also found a significant correlation between dividend premium and the closed end fund discount.

Later on Baker and Wurgler (2006) used dividend premium as a sentiment proxy. The dividend premium was calculated as the log difference of the value weighted average market-to-book ratios of dividend payers and non-payers which can be expressed as the following formula:

$$DIVPREM_t = \ln\left(\frac{\sum_{j=1}^{n} w_{jtp} \times mtb_{jtp}}{\sum_{j=1}^{n} w_{jtg} \times mtb_{jtg}}\right) \quad , \tag{4}$$

where DIVPREM is the dividend premium at time *t*, *ln* is the natural logarithm, *w* denotes the weight of the company as a portion of the total market capitalization of all similar companies at time *t*, *mtb* is market to book ratio, *p* is a dividend paying company and *g* is a non-dividend paying company.

Equity share: Equity share proxy relates to the market timing hypothesis of capital structure. Managers are perceived to act opportunistically and issue equity instead of debt when the market value is high relative to book value and past market values. They tend to repurchase equity when the market price is low and prefer debt as a choice of capital raised (Baker & Wurgler 2002). As a sentiment proxy, the total amount of equity issues are divided by the total amount of both debt and equity issues. When sentiment is high, the ratio should have high values and when the sentiment is low, the ratio should have lower values (Baker & Wurgler 2006). The Finnish stock market is quite small and equity issues are not very frequent. Due to the small size of equity and debt issues in Finland and their infrequency it is quite likely that Equity share proxy is not the best possible sentiment measure.

Put/Call-ratio: The put/call-trading volume ratio is a measure of sentiment that is derived from options trading. It equals the trading volume of put options divided by the trading volume of call options. When market participants are bearish they buy put options either to speculate or to hedge their position and when the sentiment is bullish, the volume of calls traded should clearly outweigh the number of puts traded (Wang, Keswani & Taylor 2006, 112). Although a put/call-ratio of 1 would seem to be a neutral reading, it has been empirically observed that more call options are traded on average. An average put/call-ratio lies usually between 0.8-0.9 (Bandopadhyaya 2006).

The put/call-ratio in this thesis is constructed from all of the Finnish equity options traded at the Eurex exchange. The number of Finnish companies that contribute to the index is about ten during our time period. Nokia has a dominating impact in the ratio. The next biggest companies in relation to trading volume of calls and puts are Stora Enso, UPM Kymmene, Neste Oil and Sampo.

VDAX option implied volatility: Baker and Wurgler (2007) suggest that option implied volatility could be used as a sentiment proxy. Option implied volatility is extracted from the current option prices to indicate what the option market expects the future volatility to be. The market volatility index measures the implied volatility of options in a specific market; VIX measures the S&P 500 stock index and VDAX the German DAX index. These market volatility indices have a tendency to surge during sharp sell offs as investors hedge their portfolios by buying index puts. Volatility indices are regarded as investor fear gauges (Simon & Wiggins 2001, 451).

VDAX measure was decided to be used instead of VIX as the DAX index had a higher correlation with the OMXH CAP index than the SP index. DAX had a correlation of 0.766 with the OMXH CAP as SP had a correlation of 0.747. The correlations were both significant at the 0.01 level. OMXH CAP is used rather than OMXH because our theory posits that sentiment has a larger effect on smaller companies and CAP limits the effect of Nokia.

The VDAX volatility index is an indication of the expected volatility of the DAX index for the next 30 days. It measures the implied option volatility using at the money options (Deutsche Börse 2005). Although the VDAX index is not a direct measure of volatility in Finland is does have a quite high correlation with the Finnish equity indices. It is possible that it at least partially reflects the same shifts in sentiment as many other proxies.

Advance/Decline-ratio: According to Brown & Cliff (2004), the Advance/Decline-ratio is a commonly used technical indicator used by many technical analysts in order to gain a picture of the prevailing sentiment. The indicator itself is quite simple; it is just the number of advancing stocks divided by declining stocks. When the advance decline ratio is high, the market is considered to be bullish and vice versa. In addition to these sentiment proxies that have been introduced, we also considered proxies relating to retail investor trades but suitable data set its own limitations and no such data could be used in this thesis.

3.2 Research methods

3.2.1 Constructing the investor sentiment index

In order to create the best possible sentiment index of our available sentiment proxies it must be first assessed which proxies seem to reflect investor sentiment at least partially. After that, the proxies that exhibit a sentiment component can be combined to create the sentiment index. If a sentiment proxy acts as expected and it has a significant correlation with other sentiment proxies, it is included in the sentiment index.

As there are no uncontroversial measures and as each measure is likely to contain idiosyncratic, non-sentiment related components, Principal Components Analysis is used in order to isolate the common component of these measures. A preliminary evaluation of whether a measure is suitable in the sentiment index is crucial, although Principal Component Analysis also indicates which components are scrap and should not be included in the index (Baker & Wurgler 2006, 1655-1657). Before conducting a Principal Components Analysis we need to define what variables should be included. Thus, the first task is the calculation of a correlation matrix which helps to determine which variables have common variation (Hair, Anderson & Tatham 1987). Table 1 provides descriptive statistics and Table 2 correlations of each sentiment measure and Figure 1 plots each measure.

Table 1 Descriptive statistics of the sentiment proxies

This table presents the descriptive statistics of the contemplated sentiment proxies. Turn denotes Turnover, CEFD closed end fund discount, VDAX the option implied volatility of the VDAX index, Cons. and Ind. denote the consumer and industrial confidence indices, DIVP is the dividend premium, PC-R is the put-call-ratio, ADV is the Advance/decline-ratio and EQ is the equity share of new equity and debt issues. AC denotes autocorrelation with lags of 1, 2 and 3 months. Values exceeding 0.2 signal significant autocorrelation.

	Turn	CEFD	VDAX	Cons.	Ind.	DIVP	PC-R	ADV	EQ
Mean	1.133	0.286	23.800	13.979	3.730	-0.125	0.774	1.621	0.539
Std.Dev	0.372	0.101	10.303	4.737	12.092	1.230	0.218	0.376	0.376
Max	2.550	0.506	62.050	22.500	23.800	1.237	1.815	8.000	1.000
Min	0.343	0.072	11.670	-6.400	-35.500	-2.927	0.302	0.031	0.006
Kurtosis	0.617	-0.746	1.398	0.901	0.050	-0.176	4.739	3.095	-1.630
Skewness	0.599	0.250	1.354	-0.588	-0.519	-1.124	1.270	1.535	0.014
AC (1)	0.627	0.911	0.863	0.792	0.860	0.944	0.283	0.217	0.045
AC (2)	0.467	0.855	0.666	0.610	0.734	0.887	0.296	0.265	0.008
AC (3)	0.390	0.823	0.518	0.468	0.667	0.837	0.174	0.289	0.005

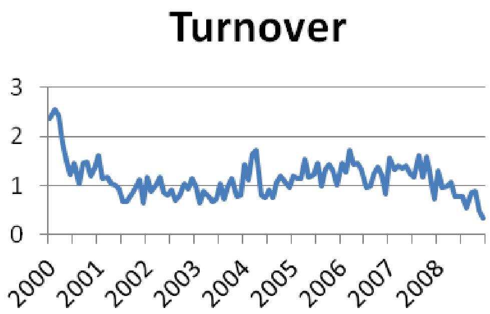
It can be observed from Table 1 that each sentiment proxy has experienced both quite high and low values during the time series. This indicates that there have been sentiment peaks and lows that are reflected in the sentiment proxies. Nearly all of the proxies are quite strongly autocorrelated. This is not surprising as sentiment is likely to be persistent. Sentiment is related to macroeconomic conditions and previous stock prices and the economy does experience cyclical patterns. The closer the autocorrelation coefficient is to one, the stronger the autocorrelation. Values of over 0.2 indicate significant first-order autocorrelation for a sample size of 100 (Eviews 2007). Some of the variables seem to be somewhat skewed or exhibit kurtosis. The normality of variables and their residuals is of greater concern when conducting regression analysis.

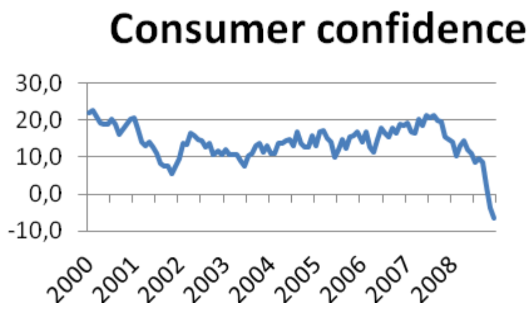
Sentiment is expected to be especially high in the beginning of the new millennium during the Internet bubble. Sentiment is expected to hit low numbers during years 2001-2004 and to gradually incline towards a new peak in 2007 before the financial crisis and total panic in the stock market in the end of 2008. By plotting each sentiment proxy, and eyeball test can be carried out to assess which measures fluctuate as expected.

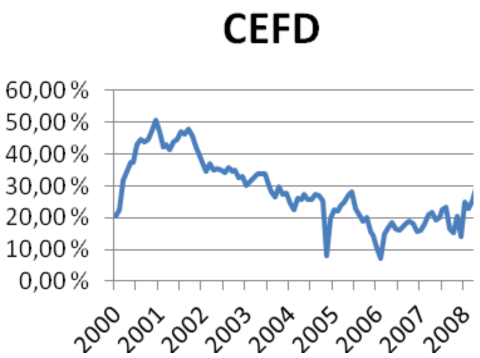
Table 2 Sentiment proxy correlations

Pearson correlations, significance levels, means and standard deviations for measures of investor sentiment during 2000-2008. Turnover is the ratio of turnover per share divided by the average turnover per share of previous 24 months. CEFD is the value weighted average discount on closed-end funds. VDAX is the average monthly implied volatility derived from DAX equity options. PC-ratio is the ratio of put trading volume divided by call trading volume of Finnish equity options traded at Eurex Exchange. DIVPREM is the log ratio of the value weighted average market-to-book ratios of dividend payers and nonpayers. Industrial Conf and Cons Conf are the industrial and consumer confidence indices gathered from DG ECFIN. ADV/DEC is the advance decline ratio which is just the number of advancing stocks divided by declining stocks. Equity Share is the proportion of equity issues divided by all issues. The bold figures denote statistical significance at the 5% level.

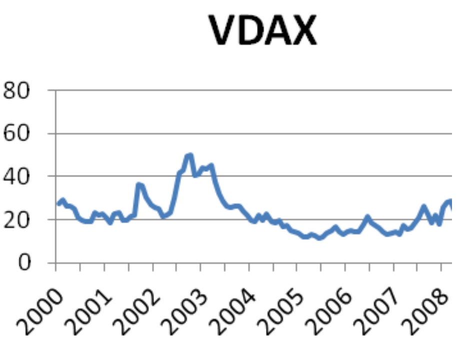
		Correlations								
		Turnover	CEFD	VDAX	PC-Ratio	DIVPREM	Ind Conf	Cons Conf	ADV/DEC	Equity Share
Turnover	Correlation	1								
	Sig.									
CEFD	Correlation	**-0.239**	1							
	Sig.	0.013								
VDAX	Correlation	**-0.339**	**0.456**	1						
	Sig.	<0.001	<0.001							
PC-Ratio	Correlation	-0.155	-0.178	-0.078	1					
	Sig.	0.110	0.065	0.422						
DIVPREM	Correlation	**-0.311**	**-0.557**	-0.029	0.052	1				
	Sig.	0.001	<0.001	0.764	0.592					
Ind Conf	Correlation	**0.606**	**-0.416**	**-0.620**	-0.075	-0.145	1			
	Sig.	<0.001	<0.001	<0.001	0.443	0.135				
Cons Conf	Correlation	**0.617**	**-0.270**	**-0.556**	-0.106	-0.177	**0.802**	1		
	Sig.	<0.001	0.005	<0.001	0.276	0.066	<0.001			
ADV/DEC	Correlation	-0.184	-0.136	-0.060	0.024	**0.275**	-0.106	-0.175	1	
	Sig.	0.057	0.161	0.539	0.802	0.004	0.277	0.069		
Equity Share	Correlation	0.056	0.184	0.164	-0.163	**-0.206**	-0.030	0.001	-0.047	1
	Sig.	0.563	0.056	0.090	0.091	0.033	0.757	0.995	0.627	

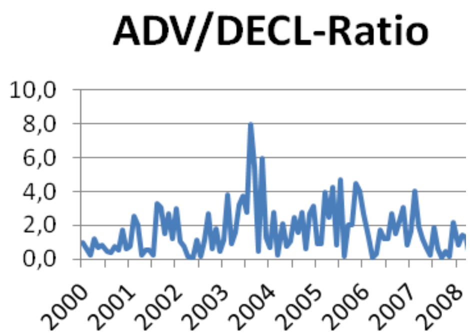

Turnover
3
2
1
0
2000 2001 2002 2003 2004 2005 2006 2007 2008

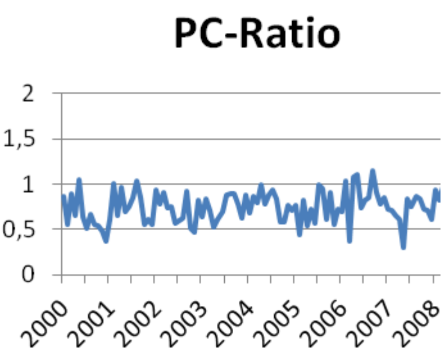

Consumer confidence
30,0
20,0
10,0
0,0
-10,0
2000 2001 2002 2003 2004 2005 2006 2007 2008

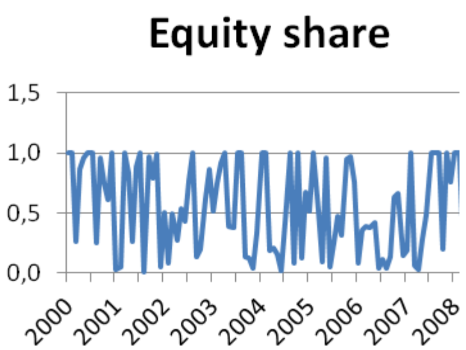

CEFD
60,00 %
50,00 %
40,00 %
30,00 %
20,00 %
10,00 %
0,00 %
2000 2001 2002 2003 2004 2005 2006 2007 2008


Industrial confidence
40,0
20,0
0,0
-20,0
-40,0
2000 2001 2002 2003 2004 2005 2006 2007 2008


VDAX
80
60
40
20
0
2000 2001 2002 2003 2004 2005 2006 2007 2008


ADV/DECL-Ratio
10,0
8,0
6,0
4,0
2,0
0,0
2000 2001 2002 2003 2004 2005 2006 2007 2008


PC-Ratio
2
1,5
1
0,5
0
2000 2001 2002 2003 2004 2005 2006 2007 2008


Equity share
1,5
1,0
0,5
0,0
2000 2001 2002 2003 2004 2005 2006 2007 2008

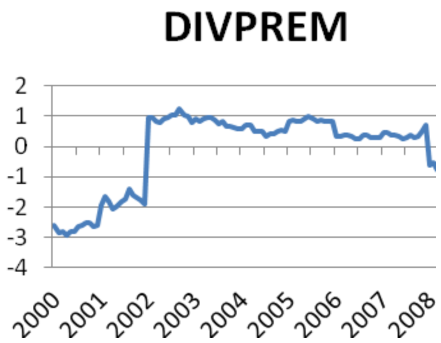


Figure 1 Investor sentiment proxies 2000-2008

The put/call-ratio, equity share and the advance/decline-ratio are not very highly correlated with the other sentiment proxies. Turnover, Industrial Confidence, Consumer Confidence, Closed-end fund discount and VDAX each have multiple significant correlations with each other and all in the expected directions. As aforementioned, the Closed-end discount and VDAX proxies are contrarian indicators and should have a negative relationship with other proxies. All of these proxies seem to flow in the expected direction of the sentiment although there are a few exceptions. The CEFD seems to be particularly high from mid 2000 until late 2001. A part of it may be due to an internal crisis that faced both Neomarkka and Norvestia in the beginning of the new millennium. Norvestia owned a part of Neomarkka and the CEO of Norvestia was also the vice president of the board of directors in Neomarkka. Neomarkka had paid substantially high rewards to these Swedish board members for consultancy services which amounted to over 10 % of the fixed costs of the company. Additional audits and additional general meetings of the shareholders were held and the CEO of Neomarkka and other Finnish board members quit their work relationships with Neomarkka. Board member Ralf Lehtonen resigned due to moral and ethical reasons (Neomarkka pörssitiedotteet 2000).

The VDAX index is clearly quite high in 2002-2003 compared to other sentiment indicators although the direction is the same. Advance/Decline-ratio and Equity Share do not seem to grasp the essence of sentiment very well in our time period and data set. They hardly correlate with other variables and the plots do not seem to indicate clear patterns of sentiment that are similar to other proxies. Equity Share might not be a good indicator in the Finnish market as the Finnish market is quite small and companies are not in need of capital every month and there can be months when no equity or no debt capital is raised in the whole market. The Put/Call-Ratio does not correlate with the other variables although the plot seems somewhat sensible. The Dividend Premium variable is a bit controversial. It does correlate with four other variables, but with only Turnover in the expected direction. On the base of the correlations and the plots, six variables are selected for further analysis: Turnover, CEFD, VDAX, Industrial confidence, Consumer confidence and Dividend premium. The first five proxies seem to act as expected and the Dividend premium is also chosen due to its highly significant correlations with Turnover and CEFD, although it should have a positive correlation with CEFD as both are expected to be contrarian sentiment indicators. The correlation of these proxies is -0.557. Principal Components Analysis will indicate which proxies should be included in the index.

Principal Components Analysis is a method that can be used to reduce a large number of variables into a smaller number of components and those variables can be grouped. The variables need to be correlated with each other and if no variable has a correlation of over 0.3 with another variable, PCA should not be used (Metsämuuronen 2008, 28). PCA is interested in the common variation of variables in order to create principal components. The goals of PCA are to condense the information in as few components as possible, to summarize patterns of correlations among observed variables, to provide an operational definition for an underlying process by using observed variables or to test a theory about the nature of an underlying process (Tabachnick & Fidell 2007, 608). If variables have a high correlation with each other and are regarded to measure a similar matter, they can be grouped into the same component. PCA is most suitable in cases where the researcher has a bunch of variables that are related to the same subject (Jokivuori & Hietala 2007, 93).

Mathematically PCA produces several linear combinations of observed variables, of which each combination is a principal component. In a way, the covariance or correlation matrix is broken in order to create new linear combinations. The main idea is to find those linear combinations that best explain the common variation of the variables. If two principal components can be extracted, the value of variable x can be formulated as following (Metsämuuronen 2008, 29):

$$x = a \times P_1 + b \times P_2 \quad , \tag{5}$$

where a and b are loadings (weights) and P_1 and P_2 are principal components. In practice, the first step is to calculate the zero mean data matrix which is used to create a correlation matrix denoted by **R**. The variance in the correlation matrix is condensed into eigenvalues. Eigenvalue is a number that depicts the principal components' ability to explain the common variation of the variables. The component with the largest eigenvalue has the most variance and components with smaller eigenvalues explain a smaller portion of the total variance. The calculation of eigenvalues and eigenvectors is usually left up to computers as the computations get increasingly difficult as the size of the matrix grows. The solution of an eigenproblem is solved by the following equation (Tabachnick & Fidell 2007, 616-617, 931):

$$(D - \lambda I)V = 0 \quad , \tag{6}$$

where **D** is the matrix, whose eigenvalues are sought, λ is the eigenvalue, **I** the identity matrix and *V* the eigenvector. The component loading matrix is created next, which is a matrix of correlations between components and variables. The first column of the matrix indicates correlations between the first component and each variable in turn (Tabachnick & Fidell 2007, 619). A loading indicates the correlation between an original variable and its component, and whether a variable belongs to a specific component or not. Loadings of 0.30 are considered significant and loadings of 0.50 are considered very significant when the sample size is over 50 (Hair, Anderson, Tatham 1987, 249). The goodness of a variable can be analyzed by its loadings or to be more precise, by the square of its loading, also called as communality. The closer the communality is to 1, the better the component can explain the variation of a chosen variable. If the variable's communality value is below 0.3, it is considered to be garbage and it usually isn't included in the component (Metsämuuronen 2008, 31).

The unrotated component loading matrix can be formed by the following equation:

$$A = V\sqrt{L} \quad , \tag{7}$$

where **A** is the component loading matrix, **V** is the eigenvector matrix and **L** is the eigenvalue matrix. Finally, the interest is directed towards the component scores, which are regression-like coefficients that can be used later on in regression analysis as dependent or independent variables. Component score coefficients are formed by the following way:

$$B = R^{-1}A \quad , \tag{8}$$

where *B* is the component score coefficients, $\mathbf{R}^{-1}$ is the inverse of the correlation matrix and **A** is the component loading matrix. This can be turned to component scores by:

$$F = ZB \quad , \tag{9}$$

where F denotes the component scores, Z standardized scores on variables and B component score coefficients (Tabachnick & Fidell 2007, 619, 623). Component scores have the advantage of representing a composite of all variables loading on the specific component and have a mean of zero and a standard deviation of 1 (Hair, Anderson, Tatham 1987, 260; Tabachnick & Fidell 2007, 650).

Usually when the researcher is not determined to produce just a single principal component, the components are rotated in order to increase interpretability in a way that the loadings of components concentrate on single components as highly as possible. But when the researcher has a clear vision of the phenomenon he or she is modeling, and an intuition of the number of components that is to be extracted, the researcher can force the variables to load on a predetermined number of components (Jokivuori & Hietala 2007, 94). In the case of this thesis, the variables are forced to a single component that hopefully accounts for an adequate amount of the common variation of the variables. Before a PCA is conducted, the researcher should test whether the correlation matrix is suitable for Principal Components Analysis. This can be done by the Bartlett's test of sphericity and by the Kaiser-Meyer-Olkin test of sampling adequacy. Bartlett's test investigates the hypothesis of whether the values of the correlation matrix are zero, in other words whether the correlation matrix is an identity matrix. If the value of the Bartlett's test is under 0.05, the correlation matrix is suitable for PCA and the variables are indeed correlated with each other. Kaiser-Meyer-Olkin measure tests whether the partial correlations among variables are small. It is calculated by dividing the correlation with the sum of the correlation and the partial correlation. The detailed formulas of Bartlett's and Kaiser's tests can be found from the appendix. If the KMO measure gives low values, it can be concluded that the bivariate correlations cannot be explained by other variables. The measure should be at least 0.6 in order for the correlation matrix to be used in PCA. The closer the value of KMO test is to 1, the better the suitability of the data for PCA (Metsämuuronen 2008, 32; Jokivuori & Hietala 2007, 96).

The limitations of PCA are somewhat different compared to for example regression analysis. In addition to the need of correlations among variables, any outliers should be fixed or removed from the sample. An outlier is a case with an extremely high or low value. The sample size should be large enough; there should be at least 5 times the amount of observations in relation to the amount of variables. Nonlinearity and multicollinearity are not problems in PCA (Metsämuuronen 2008, 28-29; Hair et al. 1987, 237).

Some preliminary tests are performed next, and it is assessed whether our sentiment proxies are suitable for PCA by performing Bartlett's test and KMO test and if all of our variables are suitable for PCA by analyzing the communalities of the variables. Table 3 provides information about the communalities of the sentiment proxies and Table 4 provides the communality values of the variables when only the variables that exhibit common variation are included.

Table 3 Communalities of the selected variables

This table presents the communality values of the selected variables when the extraction method is Principal Components Analysis. Turnover is the ratio of current turnover per share divided by the 2 year rolling average of turnover per share in the Helsinki Stock Exchange, CEFD is the closed end fund discount which is the value weighted average discount of closed end funds, VDAX is the DAX 30 implied options monthly average volatility, Dividend Premium is the log ratio of dividend payers and non-payers, Industrial Confidence and Consumer Confidence are monthly surveys that are controlled by DG ECFIN and conducted by Tilastokeskus.

	Turnover	CEFD	VDAX	DIVPREM	IND	CONS
Communality	0.553	0.286	0.579	0.010	0.838	0.766

The KMO test gives a value of 0.664 which indicates that our correlation matrix seems to be suitable for PCA, although the KMO test does give a relatively low value. The communalities table reveals that Dividend Premium variable indeed is garbage as was expected. Also CEFD seems to be under 0.3, but it needs to be assessed, how the extraction changes after Dividend Premium is removed from the correlation matrix.

Table 4 Communalities of the selected variables 2

This table presents the communality values of the selected variables when the extraction method is Principal Components Analysis. Turnover is the ratio of current turnover per share divided by the 2 year rolling average of turnover per share in the Helsinki Stock Exchange, CEFD is the closed end fund discount which is the value weighted average discount of closed end funds, VDAX is the DAX 30 implied options monthly average volatility, Industrial Confidence and Consumer Confidence are monthly surveys that are controlled by DG ECFIN and conducted by Tilastokeskus.

	Turnover	CEFD	VDAX	IND	CONS
Communality	0.536	0.316	0.587	0.832	0.755

KMO measure increases from 0.664 to 0.776, which means that this particular variable set is far more suitable for PCA than the previous one, where Dividend Premium was included. All the communalities are above 0.3, and thus all of the variables can be included in the sentiment index. As now we have determined the sentiment proxies that are used in the index, the outlier cases need to be corrected. Tabachnick & Fidell (2007, 73, 77) suggest that cases with standardized scores in excess of 3.29 are potential outliers. The outlier may be deleted or transformed. Deletion is not an option in our sample so the cases need to be transformed. An option is to change the score of the outlier so that it still remains the most extreme value, but the value is simply just one unit or standard deviation larger than the next biggest or smallest value. Just a few cases of our variables in the beginning of 2000 and in the end of 2008 contain extreme values which are modified in a way that the z-scores of those cases stay in the boundaries of 3.29.

Some of the sentiment proxies may reflect the same shift in sentiment sooner than others and a decision concerning timing is to be made. Generally proxies that involve firm supply decisions are further down the chain of events compared to proxies that are based on investor trading patterns or market variables (Baker & Wurgler 2007, 139). Both the current and lagged values of sentiment proxies need to be analyzed. A similar methodology than in the article of Baker & Wurgler (2006) is followed and a first stage index is constructed, which is the first principal component of the current and lagged proxies and gives us 12 loadings. Correlations are computed for the first stage index and each of the current and lagged proxies, and finally the SENT index is constructed from each of the proxies that indicate highest correlation with the first stage index. Each of the index components has been standardized and the SENT index has a mean of zero and unit variance.

The correlations with the first stage index and the current and lagged values indicate that only CEFD should be included in the index by its lagged value (-0.552 vs -0.584). The communality value of CEFD increased to 0.348, meaning that the lagged value indeed has more of the same characteristics as the other variables. The first principal component explains 61.276 percent of the total variance of the variables, which could be interpreted to be the common sentiment part of these variables. In comparison, the sentiment index of Baker & Wurgler (2006) only explained around 49 percent of the sample variance. However, this is not unexpected as they had a time period ranging from 1935 so not all the best proxies might have been available.

Our sentiment index is comprised of the loadings of the component score matrix and the index can be described as:

$$SENT_t = 0.25TURN_t + 0.299IND_t + 0.279CONS_t - 0.193CEFD_{t-1} - 0.244VDAX_t \quad , \tag{10}$$

where *TURN* is the turnover, *IND* industrial confidence, *CONS* consumer confidence, *CEFD* is the lagged closed-end fund discount and *VDAX* is the average implied volatility derived from DAX options. The correlation between the 12-term first stage index and the SENT index is 0.987, so little information is lost dropping the six terms. More detailed tables about the Principal Component Analysis results can be found from Appendix 2.

3.2.2 *Removing macroeconomic influences of the index*

As in the articles of Baker & Wurgler (2006), Lemmon & Portniaguina (2006) and Brown & Cliff (2004), we are also interested to determine a sentiment index that has been adjusted for macroeconomic foundations. The sentiment index can contain rational expectations based on risk factors that have been shown to predict future performance (Brown & Cliff 2005). Thus, there is an interest to form an index that suits the noise trader hypothesis and removes the common business cycle components from the index. The effects of macroeconomic influences are removed by regressing each of the sentiment proxies by a set of variables and then performing a Principal Components Analysis of the regressed residuals. Monthly data availability limits the choice of macroeconomic variables and following Baker & Wurgler (2006) and Lemmon & Portniaguina (2006), we choose inflation rate, industrial production growth, unemployment rate and short term rate (3 month euribor) as independent variables as the data is available on monthly basis. The data is collected from the website of the Bank of Finland and from Datastream. It is acknowledged that an important rational factor might be missing, but however the index should be a cleaner measure of irrational investor sentiment. Table 5 offers descriptive statistics of the macroeconomic variables

Table 5 descriptive statistics of the macroeconomic variables

This table offers descriptive statistics of the macroeconomic variables that are regressed on each of the sentiment proxies. AC denotes autocorrelation with lags of 1, 2 and 3 months. Autocorrelation coefficient values of over 0.2 signal significant autocorrelation.

	Inflation	Unemployment	Industial prod.	Short term rate
Mean	1.895	8.375	3.930	3.398
Std.Dev	1.255	1.128	5.577	1.035
Max.	4.700	11.200	25.100	5.110
Min.	-0.600	5.900	-14.500	2.030
Kurtosis	-0.761	-0.523	2.071	-1.413
Skewness	0.256	-0.497	0.193	0.126
AC (1)	0.956	0.763	0.644	0.985
AC (2)	0.918	0.546	0.497	0.956
AC (3)	0.878	0.452	0.464	0.924

It can be observed that the macroeconomic variables are quite autocorrelated and especially inflation and short term rate seem to be strongly autocorrelated. This is quite expected as macroeconomic measures fluctuate in cyclical patterns. Each of the variables has experienced higher and lower values during the time series and especially industrial production has had quite deviant extreme values in relation to its mean. The variables seem to be somewhat skewed and exhibit moderate kurtosis.

The main assumptions concerning regression analysis are that there should be an adequate number of observations, that the independent variables do not exhibit multicollinearity and that the residuals are normally distributed and that the disturbance terms should have the same dispersion. If there are problems with the disturbance terms regarding normality or dispersion, the reason can be found from the initial variables and their values. In these cases, transformations are made to the variables (Metsämuuronen 2008, 88-89). Each of the sentiment proxies is regressed by the following equation:

$$PROXY_t = c + iINF_t + uUNEMP_t + pINDP_t + rSHORT_t + \varepsilon_t \,, \qquad (11)$$

where *PROXY* indicates the sentiment proxy that is being regressed, *c* is a constant, *INF* denotes inflation, *UNEMP* the unemployment rate, *INDP* the industrial production growth and *SHORT* denotes the short term rate. After each of the proxies is regressed by the aforementioned independent variables, the same method as in the previous chapter is applied for the regression residuals. Screening continuous variables for normality is an important early step usually in multivariate analysis, even though normality of the variables may not always be required for analysis. However, the solution is usually better when the variables are normally distributed. In regression analysis, the residuals should exhibit a normal distribution, which can be enhanced by increasing the normality of the variables (Tabachnick & Fidell 2007, 79-81). Kolmogorov-Smirnov test is used to determine whether the independent and dependent variables are normally distributed. If the significance value is below 0.05, the variable is not normally distributed and if the significance value exceeds 0.05, the variable can be claimed to be normally distributed. The formula and the results from the Kolmogorov-Smirnov test can be found from Appendix 2.

Running the Kolmogorov-Smirnov test indicates that short term rate, unemployment and VDAX variables are not normally distributed. The descriptive statistics reveal that there may be issues with skewness and kurtosis with the variables. Several transformation techniques are suggested by Tabachnick and Fidell (2007, 86-89) and by Metsämuuronen (2008, 101-103). The most common methods for fixing variables are log transformation, square root transformation, inverse transformation and square transformation. The unemployment rate is log differenced in order to remove both the skewness and kurtosis. The short term rate variable is squared in order to take care of the kurtosis and a square root is taken from VDAX. After these transformations, all the variables follow a normal distribution at least in a greater extent and each exhibit a normal distribution according to the Kolmogorov-Smirnov test. The variables indicate significance levels of 0.248 for unemployment, 0.055 for short term rate and 0.124 for VDAX and thus the normality of the variables can be accepted at the 0.05 level.

All the sentiment proxies are regressed by the macroeconomic variables and then Principal Components Analysis is performed for the standardized residuals. This is done to ensure that nothing is lost in the process, although one could consider simply regressing the raw sentiment proxy with the macroeconomic factors. The results from the regressions show that each of the macroeconomic variables has some explanatory power. The most influential variable is the industrial production growth that has statistical significance in every sentiment proxy regression. The goodness of fit indicator, R_2, tells how well the sample regression line fits the data. In other words, it tells how much of the variation of the dependent variable can be explained by the model in percents (Gujarati 2003, 81; Metsämuuronen 2008, 96). The R_2 figures in the sentiment proxy regressions indicate that the macroeconomic conditions do have some impact on the proxies but a lot of the variation remains unexplained. The unexplained part is regarded as the excessive sentiment unwarranted by fundamentals (Lemmon & Portniaguina 2006).

Table 6 Sentiment proxies regressed by macroeconomic variables

This table presents the results of the sentiment proxy regressions on macroeconomic variables. Each regression is of the following form:

$$PROXY_t = c + iINF_t + uUNEMP_t + pINDP_t + rSHORT_t + \varepsilon_t$$

Where c is a constant, INF denotes inflation, UNEMP the change in unemployment, INDP the industrial production growth and SHORT denotes the short term rate. . *** denotes statistical significance at the 1 % level, ** at the 5 % level and * at the 10 % level and restrictions. t-statistics based on Newey-West standard errors are presented in parentheses. R_2 is the coefficient of determination.

	c	INF	UNEMP	INDP	SHORT	R_2
Turnover	1.120***	-0.036	0.064**	0.141***	0.011	0.206
	(23.25)	(-0.23)	(2.37)	(3.34)	(0.07)	
VDAX	4.770***	0.510**	0.025	-0.200**	-0.310	0.130
	(56.21)	(2.29)	(0.29)	(-2.26)	(-1.37)	
Consumer	14.080***	-2.620	-0.056	2.280***	2.694*	0.388
	(24.14)	(-1.65)	(-0.23)	(4.34)	(1.82)	
Industrial	3.730**	-2.540	-0.330	7.250***	2.150	0.383
	(2.17)	(-0.71)	(-0.35)	(4.99)	(0.64)	
CEFD	<0.001	-0.280	0.001	-0.240*	0.700	0.226
	(0.001)	(-0.66)	(0.001)	(-1.89)	(1.35)	

The regression variables do not exhibit severe collinearity and multicollinearity seems not to be an issue. The VIF values stay well below the critical values as the highest VIF value is 7.06 for the short term rate. Values of over 10 are usually considered to indicate a severe multicollinearity problem. In addition, multicollinearity usually causes the coefficients of determination to be very high, while t-statistics are low. This does not seem to be the case in our regressions (Gujarati 2003, 362)

The standardized residuals of these regressions are used in Principal Components Analysis to form the ADJSENT index. All of the following results can be viewed in more detail in the appendix. KMO test (0.768) and Bartlett's test (sig. 0.00) indicate that the variables are suitable for PCA. Each of the variables has a communality of over 0.3, which points out that all the macroeconomically adjusted variables are suitable for the index. The first principal component explains 56.014 percent of the common variation of the variables, which remains at a fairly high level. Removing the rational macroeconomic component from the variables decreases the common variation of the first principal component that indicates sentiment. This seems reasonable and expected as part of the common variation of the first principal component is due to macroeconomic factors. However, this is not in line with the findings of Baker & Wurgler (2006), who noticed that removing macroeconomic conditions from the index actually increased the common variation of the first principal component, which is somewhat surprising. In their study, the macroeconomically adjusted proxies were more correlated with each other than the raw proxies. If the raw variables were driven by common macroeconomic conditions, one would expect the opposite. The correlations between raw sentiment proxies are higher than the correlations of the macroeconomically adjusted proxies, which clearly indicate that macroeconomic conditions do have an impact in the sentiment proxies used in this thesis and data set.

The resulting sentiment index can be expressed as following:

$$ADJSENT_t = 0.256TURN_t^{\perp} - 0.266VDAX_t^{\perp} + 0.290CONS_t^{\perp} + 0.318IND_t^{\perp} - 0.187CEFD_{t-1}^{\perp}, \quad (12)$$

where ADJSENT is the macroeconomically adjusted or orthogonalized sentiment index, each proxy is marked as before with an addition of $\perp$ which indicates that the proxy has been regressed by the macroeconomic variables and is the residual of the regression. Each variable still enters with the expected sign.

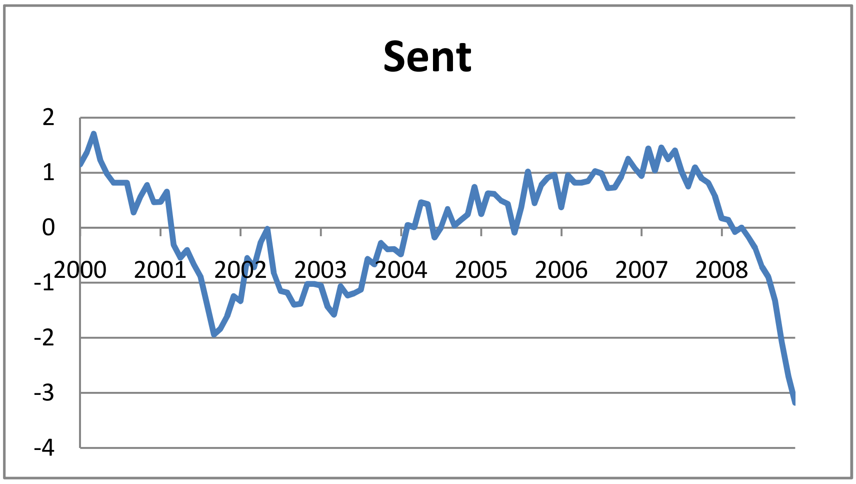
Table 7 provides descriptive statistics of the sentiment indices. The sentiment levels indices are accompanied by the sentiment changes indices. Baker and Wurgler (2007) used sentiment changes indices in their latest investor sentiment related article and investor sentiment levels indices in Baker and Wurgler (2006).

Table 7 Descriptive statistics of the sentiment indices

This table presents descriptive statistics of the sentiment levels and changes indices. SENT is the raw sentiment index and ADJSENT is the macroeconomically adjusted sentiment index. ΔSENT and ΔADJSENT are the sentiment changes indices of SENT and ADJSENT that are first differenced. AC is the autocorrelation with lags of 1, 2 and 3 months. Autocorrelation coefficient values of over 0.2 indicate significant autocorrelation.

	SENT	ADJSENT	ΔSENT	ΔADJSENT
Mean	0.000	0.000	-0.041	-0.023
Std.Dev	1.000	1.000	0.342	0.566
Max	1.728	2.189	0.691	1.437
Min	-3.228	-2.217	-1.159	-1.475
Kurt	3.253	2.526	3.653	2.834
Skew	-0.722	0.075	-0.325	-0.106
KS	1.234	0.665	0.613	0.723
	(0.095)	(0.768)	(0.847)	(0.673)
AC(1)	0.886	0.826	-0.054	-0.148
AC(2)	0.773	0.677	-0.033	-0.255
AC(3)	0.699	0.610	0.200	0.200

Sentiment changes indices are less autocorrelated and the Kolmogorov-Smirnov test statistics indicate that all of the indices are reasonably normally distributed. There is slight negative autocorrelations for the changes indices. It is not surprising that the negative autocorrelation is stronger for the orthogonalized changes index as overreaction is characteristic for irrational sentiment, which shows in the autocorrelations. The ΔSENT index is more normally distributed than SENT index probably due to smaller skewness. The descriptive statistics on the sentiment levels indices are based on 108 observations, whereas the sentiment changes indices are based on 107 observations. The next figures plot the SENT and ADJSENT indices during the time period of 2000-2008 and compare it to the fluctuations of the Finnish general stock market index.



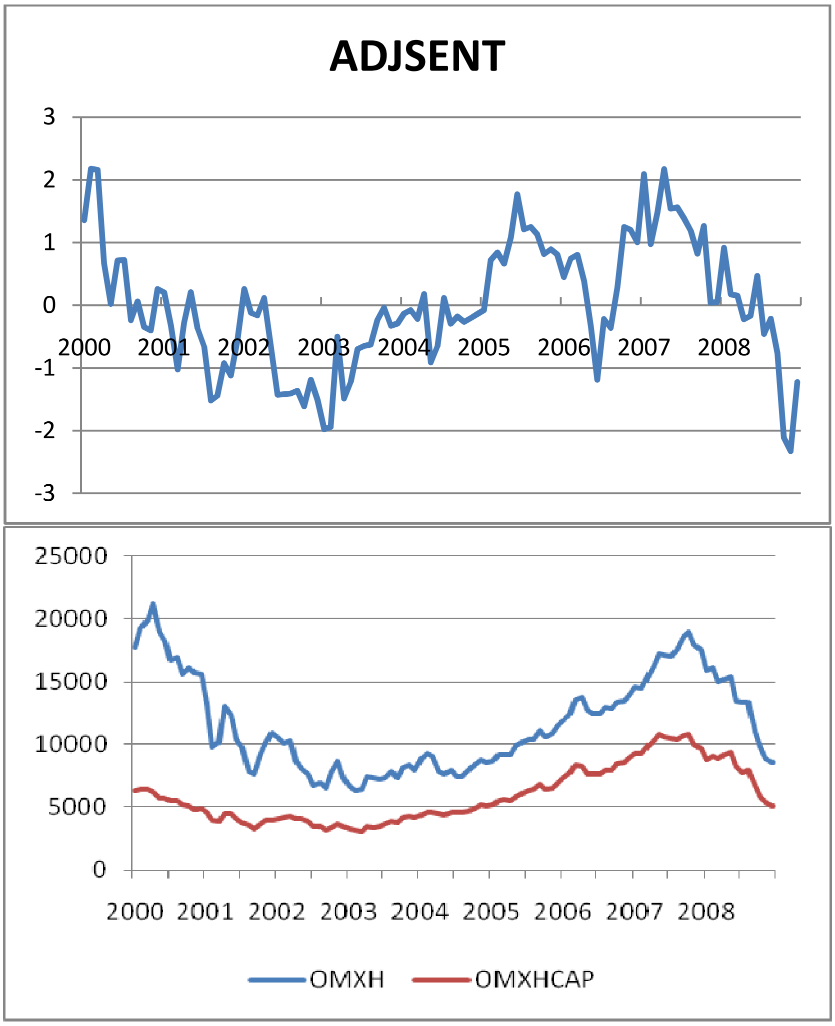


Figure 2 Sentiment and adjusted sentiment indices and the OMXH stock index

As can be seen from the figures, the SENT index is far more persistent than the orthogonalized index which fluctuates more aggressively. As the ADJSENT index is supposed to reflect the irrational component of investor sentiment, it is not surprising that the fluctuation is more aggressive. Both indices seem to line up quite well with the fluctuations of the general stock market indices although the ADJSENT index hits negative numbers in the beginning of 2006, which is due to high number in macroeconomic factors, especially industrial production growth. The sentiment indices seem to act as expected during the internet boom in the beginning of the new millennium and during the financial crisis in 2008. The sentiment proxies have a statistically significant correlation of 0.826 at the 1 percent level which means that even after controlling for macroeconomic conditions they exhibit similar characteristics.

3.2.3 Linear regression approach

Regression analyses are a set of statistical techniques that facilitate the assessment of the relationship between one dependent variable and several independent variables. The goal of regression is to produce regression coefficients for the independent variables that bring the Y values predicted from the regression equation as close as possible to the actual Y values obtained by measurement. The regression coefficients minimize the sum of the squared deviations between the predicted and obtained Y values, and optimize the correlation between the predicted and obtained Y values for the given data set (Tabachnick & Fidell 2007, 117-118). The ordinary least squares procedure is interested in minimizing the residual sum of squares and yields estimates of b_1 and b_2 that are unbiased and most efficient of their type, provided that certain conditions referred to Gauss-Markov conditions are satisfied. The least squares technique is by far the most popular regression analysis technique (Dougherty 2002, 52, 117):

The main interest in this thesis is directed towards studying the effects of investor sentiment on portfolio returns. This can be done by analyzing the explanatory power of the sentiment coefficient. In addition to finding out what the regression coefficients are, we are interested in determining whether the regression model is a good fit and whether the regression coefficients are significant. Whether the regression model is a good fit can be observed by the coefficient of determination R^2, which is the correlation of the observed and predicted values of Y squared. It tells how many percents the model is able to explain of the total variance of the dependent variable. The formula can be expressed in the following way (Metsämuuronen 2008, 96; Hanke & Wichern 2005, 277):

$$R^2 = \frac{\sum (\hat{Y} - \bar{Y})^2}{\sum (Y - \bar{Y})^2} \tag{13}$$

The adjusted R_2 is also used later on in regression equations. It enables the comparison of the goodness of fit between several equations that might have a different number of independent variables. If the R_2 was used, the increase in independent variables would inflate the R_2 even though the equation might not be a better fit. The adjusted R_2 accounts for the inclusion of additional variables (Gujarati 2003, 217). The formula can be found from Appendix 1.

The individual regression coefficients can be tested for significance by the t-test. A null hypothesis is formed which states that a particular regression coefficient does not have a linear influence on the dependent variable. If the t-statistic exceeds the chosen level of significance t-statistic value, the null hypothesis can be rejected and the coefficient can be claimed to have an influence on the dependent variable at the chosen level of significance. The t-statistic can be computed in the following way (Gujarati 2003, 249-252):

$$t = \frac{\hat{\beta}_i - \beta_i}{se(\hat{\beta}_i)} \quad , \tag{14}$$

where se stands for standard error and β_i is replaced by 0 when testing the null hypothesis. Another way of reporting the significance of the regression coefficients is through the use of p-values. It indicates the probability of obtaining the corresponding t-statistic as a matter of chance if the null hypothesis $H_0=\beta_i=0$ were true. The p-value gives the exact probability of committing a type I error which means that the coefficient is seen to have an influence on the dependent variable when it in reality did not have an influence (Dougherty 2002, 99-101). In this thesis, the t-values are presented and the statistical significance of the coefficients is expressed by a symbol *.

There are a number of assumptions that are made about the X_i variables and the error terms that are critical to the valid interpretation of the regression estimates when applying the method of least squares. First of all, the regression model needs to be linear in the parameters and X is assumed to be nonstochastic. There should be an adequate number of observations and the regression model should be correctly specified. There should not be perfect multicollinearity, which is perfect linear relationship between independent variables (Gujarati 2003, 65-75). In addition to these assumptions, the assumptions concerning the disturbance terms are also known as Gauss-Markov conditions. These conditions need to be satisfied if ordinary least squares regression analysis is to give the best possible results. The first condition is that the expected value of the disturbance term in any observation should be 0. The second condition is that the population variance of the disturbance term should be constant for all observations. The values of the disturbance term should be independent of each other and the disturbance terms should be distributed independently of the explanatory variables. Furthermore the normality of disturbance terms is usually assumed (Dougherty 2002, 76-79).

The most interest in the validity of the regression analysis focuses on the attributes of the disturbance terms or in other words, residuals. The residuals should be linear, normally distributed and homoscedastic, meaning that the variance of the residuals is the same for all predicted scores. The residuals should not show any systematic patterns also known as autocorrelation. If the Gauss-Markov conditions are not met, it will be possible to find out other estimators that are more efficient than the ordinary least squares estimator (Tabachnick & Fidell 2007, 125-127; Gujarati 2003, 65-75, 79-81).

When using time series data, the problem with autocorrelation is quite often present. From a forecasting perspective, autocorrelation is not of great concern. In a regression framework autocorrelation is usually fixed by modifying the standard regression model or more typically autocorrelation is handled by changing the nature of the error term. Strong autocorrelation can make two unrelated variables appear to be related and produce a significant regression. The examination of residuals will ordinarily reveal the problem as well as specific tests for autocorrelation like the Durbin-Watson test. Another matter that plagues usually regressions with time series data is heteroscedasticity, which occurs when the variance of the disturbance term is not constant (Hanke & Wichern 2005, 327-331, 346-347). As with the autocorrelation, the attendance of heteroscedasticity can obscure the statistical tests. It is quite likely that the standard errors will be underestimated, meaning that the t statistics are overestimated, and as a result, they give a misleading impression of the precision of the regression coefficients. Heteroscedasticity can be spotted from the residual plots or by running specific heteroscedasticity tests such as Glejser's test (Dougherty 2002, 223-228). In short under both heteroscedasticity and autocorrelation, the usual ordinary least squares estimators are no longer minimum variance among all linear unbiased estimators, and the usual tests of significance may not be valid (Gujarati 2003, 442).

Both autocorrelation and heteroscedasticity need to be removed from the equations and this can be done by using a procedure developed by Newey and West. It is an extension of White's heteroscedasticity consistent standard errors and is known as the heteroscedasticity and autocorrelation consistent standard errors or simply as Newey-West standard errors. By using the Newey-West method one is able to still use ordinary least square regression and the autocorrelation or heteroscedasticity will not obscure the results (Gujarati 2003, 484-485). The mathematics behind the Newey-West method is beyond the scope of this thesis and thus they are not presented.

In addition, it has been observed that most macroeconomic data are non-stationary and running regressions with such data could produce spurious results. In order for the regression to produce reliable results, the data should be stationary. A stationary series has a mean and the series has a tendency to return to that mean, whereas a non-stationary series wanders more widely. A stationary series has a constant variance and constant autocovariance for each lag. When working with economic time series data it has become very important to test for non-stationarity before proceeding with further estimations. The majority of financial and economic time series contain a single unit root. The augmented Dickey-Fuller test has become one of the most popular tests for non-stationarity due to its simplicity and its strong results (Kennedy 1998, 268-269, 274, 284). In many cases, stationarity can be obtained by simple differencing of the data. If taking a first difference produces a stationary process, the series is said to be integrated of order one and is denoted I(1). If the series needs to be differences d times the series is integrated of order d and is denoted I(d). The main objective of the Dickey-Fuller test is to examine the null hypothesis that $\phi=1$ in

$$y_t = \phi y_{t-1} + u_t \quad , \tag{15.1}$$

against the one-sided alternative $\phi<1$. Hence the null hypothesis is that the series contains a unit root, whereas H_1 is that the series is stationary. In practice, three types of equations are usually employed. The first equation does not include a constant term or a linear trend in the model, whereas the second equation includes a constant and the third includes both a constant and a deterministic linear trend (Brooks 2002, 378).

$$y_t = \phi y_{t-1} + \mu + u_t \tag{15.2}$$

$$y_t = \phi y_{t-1} + \mu + \lambda t + u_t \tag{15.3}$$

The augmented Dickey-Fuller test accounts for some forms of serial correlation and is more suitable when the time series exhibits autocorrelation. The formula can be expressed as following (Brooks 2002, 380):

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-1} + u_t \tag{16}$$

The lags of Δy_t now absorb any dynamic structure present in the dependent variable in order to ensure that u_t is not autocorrelated. However, a problem arises in determining the correct number of lags of the dependent variable. If monthly data was used, 12 lags could be utilized. Another choice could be to use an information criterion, such as the Schwarz information criterion or Akaike information criterion (Brooks 2002, 380).

The standard approach to test a hypothesis would be to construct a t-test, but under non-stationarity the statistic computed does not follow a standard t-distribution. A separate Dickey-Fuller distribution is utilized that has been generated by using Monte Carlo techniques (Harris 1995, 29). The Dickey-Fuller statistic is referred to as a τ-statistic. Table 8 presents the critical values for the DF-test with a sample size of 100.

Table 8 Dickey-Fuller critical values

This table presents the critical values for the Dickey-Fuller test. The critical values increase as additional variables are included in the model. These additional variables indicate a drift and a deterministic time trend.

Model	Level of significance		
	0.01	0.05	0.10
None	-2.60	-1.95	-1.61
Constant	-3.51	-2.89	-2.58
Constant and trend	-4.04	-3.45	-3.15

The bigger the τ-statistic value, the more likely it is that the variable does not have a unit root and is a stationary variable. The critical τ-statistic values are higher than in the ordinary t-test due to the inherent instability of the unit root process, the wider distribution of the t-ratios in the context of non-stationary data and the resulting uncertainty in inference (Brooks 2002, 379). The Dickey-Fuller test statistic formula can be found from Appendix 1. The unit root test results of the sentiment indices are presented in Table 9.

Table 9 Unit root test for the sentiment indices

This table presents the Augmented Dickey-Fuller statistics of the sentiment indices. τ-statistics for each of the sentiment indices are presented and the significance levels are reported in the parentheses. Schwarz information criterion is used to determine optimal lag length with a maximum lag length set to be 12.

Model	SENT	ADJSENT	ΔSENT	ΔADJSENT
None	-0.519	-3.076	-10.767	-10.257
	(0.489)	(0.024)	(0.000)	(0.000)
Constant	-0.482	-3.059	-10.291	-10.925
	(0.889)	(0.033)	(0.000)	(0.000)
Cons. and trend	-0.300	-3.160	-10.931	-10.237
	(0.989)	(0.098)	(0.000)	(0.000)

It can be observed from the unit root results that the raw sentiment index is non-stationary at all levels and the macroeconomically adjusted sentiment index is considered to be stationary at the 5 percent level without a deterministic trend. However, both indices exhibit stationary characteristics when first differenced. According to Brown and Cliff (2005) the raw sentiment index is likely to be non-stationary and if the macroecomically adjusted index remains to be non-stationary, the index has not been sufficiently cleaned of macroeconomic factors. Although the ADJSENT index is stationary at the 5 percent level, it could be wiser to use the first differenced sentiment indices as they are clearly more likely to be stationary.

In this thesis the most weight of the empirical tests are placed on different tests utilizing regression analysis. Regression analyses are run on several different portfolios and during extremely high and low sentiment periods as well as during the whole time period. Following Baker & Wurgler (2007) we run regressions of the form:

$$R_{pt} = c_p + d_p SENT_{t-1} + \beta_p R_{mt} + \varepsilon_{pt} \quad , \tag{17}$$

where R_p is the return of a portfolio, c is a constant, d and β are regression coefficients, *SENT* is the sentiment changes index and R_m is the value weighted market return. The returns are matched by the sentiment of the preceding month. When regressions are run with the orthogonalized sentiment proxy, SENT is simply replaced by ADJSENT. Regressions are run with returns of the following month and following 3 months. OMXH is added as a control variable because high volatility stocks are likely to have higher market betas which should be accounted for (Baker & Wurgler 2007).

In addition to these regressions, it is tested whether higher levels of positive or negative sentiment affect the returns more than during low levels of sentiment. This is done by following the method of Finter, Niessen and Ruenzi (2010) who studied the effects of sentiment in the German stock market. The sentiment index as an independent variable is replaced by two dummy variables. DS^H and DS^L where DS^H is given a value of 1 when the sentiment is over a standard deviation above the average and 0 otherwise and DS^L gets a value of 1 when the sentiment is more than 1 standard deviation below the average and 0 otherwise. The sentiment levels between -1 and 1 are considered to be in the “gray zone” when the sentiment is at moderate levels either on the positive or the negative side. Finter et al. used sentiment levels of over 1 and under -1 in a similar way in their article in order to divide the observations to the different dummy categories.

The regression equation can be expressed as following:

$$R_{pt} = c_p + d_1 DS^H{}_{t-1} + d_2 DS^L{}_{t-1} + \beta_p R_{mt} + \varepsilon_{pt} \quad , \tag{18}$$

where R_p is the return of the portfolio, c is a constant d_1, d_2 and β are regression coefficients, DS^H is the dummy variable for high sentiment and DS^L is the dummy variable for low sentiment and R_m is the return of the market portfolio OMXH.

4 EMPIRICAL RESULTS

4.1 Descriptive statistics and preliminary tests

This subchapter provides descriptive statistics of the return portfolios and looks for patterns and characteristics of the return portfolios under different sentiment levels. In addition, it is tested whether sentiment is related to the returns of the portfolios by testing the correlations of the current level of sentiment matched by the current returns of the portfolios. If sentiment is related to stock returns of the selected portfolios, the correlations should be statistically significant. It is also tested whether sentiment has a linear relationship with the current returns of stocks in general. In the later sections of this chapter tests are performed in order to assess whether sentiment can be used to explain future returns and if the sentiment index could be used as a trading strategy. These are tested by conducting several regression analyses.

Table 10 shows the average monthly return, minimum and maximum monthly returns of the time period and the monthly standard deviation of the return portfolios. The numbers are reported in percents.

Table 10 Descriptive statistics on portfolio returns

Descriptive statistics of the market portfolio, MSCI portfolios sorted by size and investment style, the equally weighted portfolio of all stocks and portfolios sorted by historical volatility. Lg equals large, Sm equals small, Gr. equals growth and Val. equals value companies. Autocorrelation is reported at lags of 1, 2 and 3 months. Autocorrelation coefficient values of over 0.2 signal significant autocorrelation. The returns are expressed in percentages.

	Mean	Std.Dev	Max	Min	Kurtosis	Skewness	AC (1)	AC (2)	AC (3)
OMXH	-0.71	8.48	25.66	-31.30	1.87	-0.47	0.22	-0.13	-0.05
Growth	-1.15	12.11	31.49	-43.10	1.88	-0.75	0.19	-0.16	-0.09
Value	-0.80	6.86	17.19	-17.59	-0.18	-0.12	0.13	-0.08	0.20
Small	0.34	5.60	11.56	-15.63	0.31	-0.69	0.30	0.17	0.22
Lg Gr.	-1.32	12.37	32.03	-44.59	1.94	-0.71	0.18	-0.17	-0.06
Lg Val.	-0.17	8.15	19.06	-32.54	1.52	-0.59	-0.06	-0.03	0.01
Sm Gr.	-0.85	8.23	12.66	-31.68	1.78	-1.16	0.37	0.19	0.09
Sm Val.	0.30	5.53	10.41	-21.11	1.77	-1.05	0.39	0.25	0.20
All ew	-0.36	5.65	11.46	-26.52	3.17	-1.05	0.25	0.19	0.12
Vol10	-2.45	10.55	34.90	-33.71	1.51	0.14	0.06	0.11	0.06
Vol20	-2.04	9.10	25.03	-28.96	1.34	-0.25	0.15	0.10	0.07
Vol40	0.09	5.42	9.02	-31.13	7.67	-1.78	0.18	0.19	0.07
Vol80	0.75	4.04	9.37	-19.68	5.91	-1.51	0.34	0.21	0.16
Vol90	0.75	4.17	9.46	-22.31	7.96	-1.69	0.29	0.20	0.09

The Table consists of MSCI benchmark indices, OMXH index, equal-weighted portfolio of all stocks (ALL EW) and of the five portfolios that have been formed by the previous historical volatility. Vol10 portfolio consists of the companies that are part of the largest volatility decile, Vol20 of the companies that are part of the largest quintile, Vol40 of the companies that are in the mid quintile (40-60) and Vol80 and Vol90 of the companies that are in the smallest quintile and decile. The results are based on 108 observations from 2000 until 2008.

On average, the returns of value companies have exceeded those of growth companies and the returns of small companies have topped those of larger companies. The most volatile companies have fared quite badly compared to the less volatile companies. The least volatile companies have experienced an average monthly return of 0.75 % during the time period. The less volatile companies are likely to have higher book-to-market ratios. The higher returns of both small stocks and stocks with high book-to-market ratios have been noticed in previous research (Fama & French 1996). The size effect has also been noticed in many research articles covering the Finnish stock market (Kallunki et al. 1997).

The minimum and maximum returns of companies tend to be large especially for the growth companies and for the volatile companies. This is not surprising as both the IT-bubble and the financial crisis took place in the time series. Growth stocks were those information technology companies that seemed to have endless profit potential and experienced a sharp rise and fall in the beginning of the new millennium. The two previously mentioned stock market crashes result for the low average monthly returns of the selected portfolios and indices.

Many of the portfolios seem to be positively skewed and exhibit moderate to strong kurtosis. Most of the return observations cluster around zero and hence it is not surprising that the portfolios are leptokurtic. Ideally, the variables should be as close to a normal distribution as possible. However, by performing variable transformations the interpretability of the results is weakened. Hence, no variable transformations are made. It is interesting to note that the returns of small companies are clearly more autocorrelated than those of larger companies. This is probably due to thinner trading.

Table 11 depicts the average monthly returns under different sentiment levels for three time periods: returns of the same month as the sentiment, returns of the following month in relation to sentiment and returns of the following three months in relation to sentiment. Sentiment is divided into four different levels. High, very high, low and very low where very high levels of sentiment denote levels of sentiment that are more than one standard deviation above the average and very low levels denote sentiment levels that are more than one standard deviation below the sentiment average. The sentiment levels are regarded high or low when the levels are positive or negative. The sentiment index has a mean of zero and a standard deviation of one.

Table 11 Average portfolio returns under different market scenarios

The returns of equal weighted portfolio, market portfolio, MSCI portfolios and portfolios sorted by volatility under different sentiment conditions. Sentiment is high when it is above the average and low when it is below average. Sentiment is very high or very low when it is over a standard deviation above or below the average. Sentiment is based on SENT index. C denotes the return of the current month in relation to the sentiment, 1m the return of the following month in relation to the sentiment and 3m the return of the following 3 months in relation to the sentiment. The returns are expressed in percentages.

	Sentiment					
	High			Low		
	c	1m	3m	c	1m	3m
ALL EW	0.03	0.12	-0.55	-0.83	-0.67	-2.19
OMXH	-0.10	0.01	-0.01	-1.46	-1.73	-6.24
Growth	-0.34	-0.04	-0.41	-2.25	-2.89	-9.71
Value	-0.53	-0.54	-1.70	-0.99	-1.08	-4.23
Small Cap	0.58	0.63	1.62	0.09	0.07	0.20
Large Gr.	-0.60	-0.35	-1.36	-2.30	-2.90	-9.75
Large Value	0.55	0.57	3.03	-0.92	-1.07	-3.71
Small Gr.	-0.65	-0.82	-2.55	-1.04	-0.75	-2.18
Small value	0.74	0.70	2.10	-0.21	-0.06	-0.46
Vol10	-2.35	-2.17	-6.07	-2.76	-2.84	-6.92
Vol20	-2.27	-2.07	-6.46	-1.96	-1.99	-6.43
Vol40	0.49	0.50	1.07	-0.62	-0.45	-1.04
Vol80	1.17	1.12	3.13	0.09	0.25	0.88
Vol90	1.16	1.13	3.45	0.06	0.24	1.12
	Very High			Very Low		
	c	1m	3m	c	1m	3m
ALL EW	1.72	0.49	-1.78	-2.30	-0.62	1.86
OMXH	2.96	2.81	3.17	-1.88	0.46	0.36
Growth	4.16	4.90	8.87	-2.10	0.91	-1.04
Value	0.27	0.03	-4.79	-2.22	-0.47	-1.11
Small Cap	1.48	1.00	-1.34	-1.20	0.59	4.55
Large Gr.	3.61	4.79	8.90	-1.99	1.01	-1.02
Large Value	-1.30	-1.12	0.14	-1.44	-0.58	-1.29
Small Gr.	2.37	0.76	-2.09	-2.28	0.30	5.39
Small value	0.31	0.33	-3.15	-1.84	-0.25	3.18
Vol10	-0.25	-2.01	-9.04	-4.21	-2.39	-1.46
Vol20	-0.59	-1.77	-8.49	-3.70	-1.80	-0.66
Vol40	1.09	1.35	-0.33	-2.44	-0.83	1.57
Vol80	0.89	0.77	1.06	-1.12	0.03	3.03
Vol90	0.80	0.19	0.37	1.07	-0.22	2.76

There are 62 monthly observations for the high sentiment and 46 monthly observations for the low sentiment. 15 of these observations are regarded very high and 23 to be very low. Some conclusions can be drawn even from these observations. When the sentiment is at moderate levels, the returns of value stocks seem to surpass the returns of growth stocks. By comparing the OMXH and the equal-weighted market portfolio it can be seen that the equal-weighted portfolio acts much more like what is proposed by the behavioral finance theory. This is not surprising as small stocks affect the returns of the equal-weighted portfolio more than in the value-weighted OMXH index.

The returns of the current and following month head to the same direction as the sentiment until they are reversed in the time period of the following three months. The returns of the following months are in negative relation with the sentiment during longer time intervals as the particular stocks are considered overvalued. The effects of sentiment can be observed by mispricing correction (Baker & Wurgler 2006).

Small stocks seem to act as predicted by the behavioral theory especially during levels of very high and very low sentiment. This is in line with the theory that small stocks are mainly owned by individuals who are considered to be more prone to sentiment, and during high levels of optimism or pessimism sentiment is more likely to affect the decisions of these individuals. What is interesting to note is that large companies and especially large growth companies exhibit quite intense momentum. When sentiment is very high, the growth and large growth companies yield monthly returns of around 4 percent and returns of almost 9 percent for the next three months. On the other hand, when sentiment is low, growth stocks and large stocks exhibit negative momentum. This may be due to the fact that foreign investors have a large ownership stake in Finnish companies. Almost 70 percent of the total market capitalization of Finnish stocks in 2001 was owned by foreigners and 32 percent if Nokia was excluded from the calculations. Foreign investors have been identified to practice a momentum trading strategy and execute large trades (Karhunen & Keloharju 2001; Grinblatt & Keloharju 2001).

The highly volatile stocks seem to earn a negative monthly return which does not seem to depend on the sentiment level. The most volatile stocks might contain companies that are highly distressed, unprofitable or thinly traded. It is likely that especially the Vol10 portfolio is comprised of at least some of these kinds of companies. However, the main interest of these portfolios is directed towards the low, mid and high quintile portfolios or Vol20, Vol40 and Vol80 portfolios. The least volatile stocks seem to be quite profitable regardless of sentiment level, although the profits are larger during high levels of sentiment and during very low levels of sentiment in the three months' time period.

In general, there is a fewer number of low sentiment time periods compared to high sentiment ones and the returns are more moderate during high levels of sentiment. That is the stock returns are higher in absolute values compared to the returns during high levels of sentiment. People are usually considered to be optimistic, but also loss averse and losses feel emotionally much worse than winnings of similar amount. The loss aversion of an investor depends on his prior investment performance. After prior gains the investor is less loss averse and after prior losses the investor becomes more loss averse. Prior winnings serve as a cushion for future setbacks and if the investor has recently experienced losses, he is more sensitive to additional setbacks. In addition, individual investors tend to realize their winnings too early and avoid selling their losses in order to minimize psychological burden (Barberis, Huang & Santos 2001). When the sentiment is low, the fear of losses may drive the decisions of investors as when the sentiment is high investors are less affected by similar gains. It can be seen from Table 11 that the returns of large stocks and especially large growth stocks have a rapid decline in value when the sentiment is low compared to smaller stocks. The effect during low sentiment might be explained by the notion that smaller stocks are more owned by individual investors who hold on to their losses for too long.

Next it is assessed whether sentiment is related to the returns of the portfolios. The returns and sentiment levels of the same time period are correlated in order to give a picture whether these two measures are interrelated. The correlations between both the SENT index and returns and the macroeconomically adjusted ADJSENT index and returns are expressed in table 12. Also the statistical significance of the correlations is shown at 1, 5, and 10 percent levels. The SENT index has a highly significant correlation of 0.253 with the equal-weighted portfolio of all stocks quoted at the Helsinki Stock Exchange and the correlation is even higher with the orthogonalized index. A positive correlation will exist between aggregate market returns and the sentiment index if the average stock is affected by sentiment (Baker & Wurgler 2007). There is not a significant correlation between the SENT index and the OMXH index. This can be interpreted that sentiment is more related with the returns of smaller companies as they have a bigger emphasis on the equal-weighted market portfolio.

Table 12 Correlations between portfolio returns

Correlations among portfolio returns and the raw sentiment index SENT and the orthogonalized ADJSENT index for the same time period. *** denotes statistical significance at the 1 % level, ** at the 5 % level and * at the 10 % level.

	SENT		ADJSENT	
	Coefficients	Significance	Coefficients	Significance
ALL EW	0.253	0.008***	0.349	0.000***
OMXH CAP	0.213	0.027**	0.283	0.003***
OMXH	0.158	0.105	0.215	0.026**
Growth	0.123	0.206	0.178	0.069*
Value	0.144	0.140	0.212	0.028**
Small Cap	0.185	0.057*	0.264	0.006***
Large Value	0.112	0.251	0.094	0.338
Large Growth	0.112	0.250	0.167	0.086*
Small Value	0.212	0.029**	0.266	0.006***
Small Growth	0.152	0.119	0.264	0.006***
Vol10	0.120	0.218	0.256	0.008***
Vol20	0.128	0.189	0.248	0.010***
Vol40	0.259	0.007***	0.338	0.000***
Vol80	0.254	0.008***	0.308	0.001***
Vol90	0.245	0.011**	0.279	0.004***

Small Cap and Small Value portfolios seem to be more correlated with the SENT index than other MSCI indices. It is surprising that the least volatile stocks seem to be more correlated with the SENT index than the most volatile stocks. Vol40, Vol80 and Vol90 all have a statistically significant correlation with the SENT index. The SENT index does not seem to have an impact on the returns of larger stocks or growth stocks in the same time period. However, the orthogonalized or in other words macroeconomically adjusted ADJSENT index seems to be more correlated with the MSCI index and volatility portfolio returns. Just the large value index does not exhibit a statistically significant correlation at any significance level. Large growth and growth indices also have smaller correlations with the ADJSENT index. The correlation with the equal-weighted index is quite high at 0.349. It is interesting to note that after controlling for macroeconomic foundations the sentiment index is more correlated with the return indices. Now the small growth index and the most volatile stock portfolios are all highly significantly correlated with the ADJSENT index. In comparison, the correlation between the equal-weighted portfolio and the sentiment index is smaller than in the tests conducted by Baker and Wurgler (2007) who found a correlation of 0.43.

In addition, a regression analysis is executed in order to determine whether sentiment has an impact on the average stock and whether sentiment affects stocks with different size and investment style dissimilarly. The returns of the equal weighted portfolio and MSCI portfolios are matched by the market return and investor sentiment change of the same month. The results show that sentiment is related to the current returns of stocks. The raw sentiment index SENT has more explanatory power than the orthogonalized sentiment index ADJSENT. This is not surprising as sentiment is partly formed by rational expectations about future that are reflected in the macro economy. It seems that smaller stocks and growth stocks are more influenced by sentiment. This finding is in line with the behavioral finance theory. Large value portfolio is the only portfolio that is not affected by changes in sentiment. It should be noted that sentiment changes of a single unit are extremely rare. The ADJSENT index experienced a change of at least a single unit three times and SENT index zero times during the time series. Table 13 summarizes the results.

Table 13 Sentiment and current portfolio returns

This table presents regressions results of portfolios regressed by a constant b_0, the sentiment changes indices SENT and ADJSENT and the market portfolio return OMXH.

$$R_{pt} = c_p + d_p SENT_t + \beta_p R_{mt} + \varepsilon_{pt}$$

Panel A presents the results when SENT is used as an independent variable and Panel presents the results when ADJSENT is applied as an independent variable. The first differenced sentiment levels indices or in other words, sentiment changes indices are used due to their stationarity and smaller autocorrelation. *** denotes statistical significance at the 1 % level, ** at the 5 % level and * at the 10 % level and restrictions. t-statistics based on Newey-West standard errors are presented in parentheses. Newey-West automatic bandwidth and lag length are used.

	Panel A				Panel B			
	c	d	β	Adj.R_2	c	d	β	Adj.R_2
ALL EW	0.001	0.043***	0.386***	0.463	-0.001	0.025***	0.387***	0.460
	(0.04)	(3.23)	(5.59)		(-0.25)	(3.26)	(5.40)	
Large value	0.002	0.003	0.432***	0.192	0.002	-0.007	0.444***	0.194
	(0.33)	(0.19)	(3.94)		(0.31)	(-0.66)	(3.93)	
Large growth	-0.004	-0.023**	1.400***	0.897	-0.004	-0.015*	1.400***	0.898
	(-1.19)	(-2.55)	(27.86)		(-1.00)	(-1.92)	(26.60)	
Small value	0.007	0.038**	0.300***	0.295	0.006	0.019**	0.305***	0.278
	(1.24)	(2.46)	(3.32)		(0.97)	(2.29)	(3.20)	
Small growth	-0.002	0.049***	0.646***	0.529	-0.003	0.024**	0.654***	0.515
	(-0.28)	(3.21)	(6.85)		(-0.47)	(2.29)	(6.75)	
Small cap	0.008*	0.032***	0.436***	0.514	0.007	0.016**	0.439***	0.504
	(1.77)	(2.84)	(6.73)		(1.47)	(2.23)	(6.63)	

The indices declined by more than a single unit six times and once respectively. Changes of at least 0.5 or -0.5 were more common. ADJSENT experienced an increase or decrease of at least half a unit 21 for both scenarios or 42 times in total and SENT increased by more than 0.5 five times and decreased by more than -0.5 nine times. Hence, it is wiser to interpret the results by assessing how much the portfolio return changes when the sentiment index changes by half a unit.

In conclusion, it seems that the average stock is indeed somewhat affected by investor sentiment. This finding is in line with Hypothesis 1. It is pleasing to find out that sentiment seems to affect small stocks and stocks with larger growth prospects more than larger stocks and value stocks. These findings suggest that sentiment might be a relevant measure. However, the biggest question is whether sentiment could be used to forecast returns. The next chapters are dedicated to finding whether previous changes in sentiment indices could be used to explain subsequent month returns of stocks.

4.2 Using sentiment to predict stock returns

This section is dedicated to studying whether investor sentiment and the irrational component of it have a statistical relationship with the future returns of different types of stocks and if relevant patterns of mispricing correction can be spotted. In other words, it is studied whether sentiment can be used to forecast returns. Ordinary least squares regression is used to determine the effect of sentiment on returns. The returns are logarithmized which makes them more normally distributed. The sentiment indices are reasonably close to a normal distribution. The issues of autocorrelation and heteroscedasticity are addressed by performing a Newey-West correction for the standard errors.

The sentiment regressions with a forecast horizon of one month are presented first, followed by the regressions with returns of the subsequent three months. As the sentiment changes indices are utilized, the results are based on 107 observations. In addition to the portfolios and indices that have already been introduced, it is tested whether a simple long-short portfolio delivers more significant results. LSVol portfolio is long on the most volatile quintile of stocks and short on the least volatile quintile. Baker & Wurgler (2006) used long-short portfolios in their study by giving equal weight for both the long and the short portfolio. They did not reinvest the gains of the short portfolio in the long portfolio. The portfolio is simply a difference portfolio of the most volatile and least volatile stocks. The results of the regression results are presented next for returns of the following month in table 14.

Although the constant c is reported, the main focus of interest is not in the precise predicted values of the regression equations, but in the effect of sentiment. β values reflect the beta of the portfolio or index compared to the market portfolio, although the risk free rate of interest has not been deducted. The main interest is directed towards the coefficient values and significance levels of d which indicate the effects of investor sentiment during the entire time period. The coefficient values demonstrate how much the monthly or quarterly returns increase or decrease when the sentiment increases by a single unit. A single unit corresponds to an increase of one standard deviation for the levels index (Metsämuuronen 2008, 91).

The adjusted R_2 provides information about the goodness of fit of the regression equation and the adjusted R_2 facilitates the comparison between regression equations as it accounts for the inclusion of additional independent variables. It can be seen from the results that in general the coefficients of determination are at a sufficient level indicating that a linear regression function with the chosen variables is suitable and is in some relation to the returns of the portfolios. The R_2 figures pretty much vary with the levels of β. It seems that sentiment does indeed have an influence in the portfolio returns of the subsequent month. The coefficient values can be interpreted as following: when the adjusted sentiment index changes or increases by a single unit, the return of e.g the small cap portfolio increases by 2.0 percent in the following month. Yet again it is important to notice that sentiment changes of a single unit are very rare. SENT index experienced a change of at least half a unit 14 times and ADJSENT 42 times during the whole time series. Even though the coefficients of SENT index are higher in comparison to the coefficients of ADJSENT, the changes of that magnitude are far more common for the ADJSENT index.

As we are most interested in the effects of irrational sentiment on stock returns and whether it could be used to forecast returns, most emphasis is placed on analyzing the statistical relationship between ADJSENT and stock returns. When ADJSENT increases by half a unit, the subsequent month returns of small growth, small value and small cap indices increase by 1.3, 1.05 and 0.9 percent respectively. In a similar scenario the returns of the volatility portfolios' Vol40, Vol80 and Vol90 increase by 1.0, 1.0 and 0.9 percents respectively. Each of these coefficients is statistically significant at the 1 percent level. Especially small stocks and least volatile stocks seem to be affected by sentiment and the explanatory power of the sentiment variable is the highest among these stocks.

Table 14 Regression results with a forecast horizon of 1 month

Regressions of portfolio returns on lagged sentiment and the market portfolio return OMXH.

$$R_{pt} = c_p + d_p SENT_{t-1} + \beta_p R_{mt} + \varepsilon_{pt}$$

The sample period includes monthly returns from 2000 until 2008 and the results are based on 107 observations. The monthly returns are matched by the sentiment of the previous month. SENT is the change in the raw sentiment index and ADJSENT is the change in the orthogonalized sentiment index. The portfolios' consist of MSCI indices, portfolios sorted by historical volatility and of the difference portfolio LSVol. LSVol is long on the most volatile quintile of stocks and short on the least volatile quintile.. The results of the regression using SENT as an independent variable are presented in panel A and panel B presents the results of regressions with ADJSENT as an independent variable. *** denotes statistical significance at the 1 % level, ** at the 5 % level and * at the 10 % level and restrictions. t-statistics based on Newey-West standard errors are presented in parentheses. Newey-West automatic bandwidth and lag length are used.

	Panel A				Panel B			
	c	d	β	Adj.R_2	c	d	β	Adj.R_2
Growth	-0.004	-0.031***	1.384***	0.927	-0.003	-0.011**	1.381***	0.923
	(-1.06)	(-3.19)	(34.15)		(-0.74)	(-2.01)	(31.79)	
Value	-0.002	0.031**	0.609***	0.608	-0.003	0.011	0.612***	0.592
	(-0.46)	(2.29)	(10.21)		(-0.67)	(1.27)	(9.53)	
Small	0.008*	0.029***	0.444***	0.496	0.007	0.018***	0.442***	0.500
	(1.72)	(2.81)	(7.77)		(1.52)	(3.34)	(7.04)	
Lg Gr.	-0.005	-0.031**	1.393***	0.899	-0.004	-0.010	1.389***	0.894
	(-1.38)	(-2.44)	(29.28)		(-1.09)	(-1.64)	(28.74)	
Lg Val.	0.007	0.042***	0.467***	0.319	0.006	0.020*	0.469***	0.303
	(1.39)	(2.63)	(4.44)		(1.12)	(1.83)	(4.31)	
Sm Gr.	-0.002	0.028*	0.668***	0.491	-0.003	0.026***	0.662***	0.507
	(-0.35)	(1.69)	(6.97)		(-0.43)	(3.06)	(6.64)	
Sm Val	0.007	0.033**	0.315***	0.275	0.007	0.021***	0.313***	0.280
	(1.22)	(2.60)	(3.64)		(1.04)	(2.89)	(3.39)	
Vol10	-0.020**	0.004	0.629***	0.243	-0.020	0.025**	0.616***	0.260
	(-2.42)	(0.23)	(5.63)		(-2.30)	(1.99)	(8.03)	
Vol20	-0.015	0.007	0.667***	0.376	-0.015**	0.014	0.660***	0.383
	(-2.15)	(0.54)	(11.20)		(-2.33)	(1.29)	(8.46)	
Vol40	0.005	0.035***	0.291***	0.251	0.003	0.020***	0.290***	0.248
	(1.02)	(3.73)	(3.73)		(0.82)	(3.04)	(3.56)	
Vol80	0.011***	0.041***	0.199***	0.293	0.010**	0.020***	0.200***	0.252
	(3.19)	(4.73)	(3.22)		(2.51)	(3.37)	(2.80)	
Vol90	0.011***	0.038***	0.188***	0.246	0.010**	0.019***	0.188***	0.223
	(2.98)	(4.27)	(2.91)		(2.53)	(3.05)	(2.72)	
LsVol	-0.026***	-0.031*	0.467***	0.254	-0.024***	-0.005	0.460***	0.237
	(-3.49)	(-1.78)	(7.23)		(-3.88)	(-0.44)	(5.33)	

Growth, Large Growth and LsVol portfolios have a negative relationship with sentiment. Smaller stocks, value stocks and less volatile stocks seem to have a positive relationship with sentiment. It was expected that sentiment prone stocks should be most affected by investor sentiment and that the relation between sentiment and future returns should have been negative for the speculative stocks. The effects of sentiment on growth stocks, less volatile stocks and value stocks was as anticipated but the relationship between sentiment and small stocks is not as predicted by the behavioral finance theory. In general, small capitalization stocks are considered to be more speculative. Thus, it was expected that there was a negative relationship between small stock returns and investor sentiment. However, it is pleasing to notice that especially the irrational component of sentiment has a good explanatory power among small stock return portfolios and the most volatile Vol10 portfolio.

Finter et al. (2010) found out that sentiment prone stocks fared better in the short term when sentiment was high and that the effect reversed in a longer time period. It is possible that investors who are mostly interested in small stocks gain courage to participate in the stock market only after the market has experienced a rise and when there is collective euphoria in the market. This might lead to a more delayed price pressure on these specific stocks. Lemmon and Portniaguina (2006) discovered that small stocks get relatively overvalued during high levels of sentiment. They observed a negative relationship between high levels of sentiment and stock returns of the following quarter. However, our equation implies how much the subsequent month return changes when sentiment changes by a single unit. An increase of one unit does not imply that sentiment is high per se, but rather that sentiment is gathering pace and might still be negative. It should be noted that sentiment did have a weaker or dissimilar influence on larger stocks and large growth stocks which more traded by foreign investors and institutional investors who are regarded to be smart money traders. These stocks are also more liquid and it is quite possible that the effects among small stocks are lagged.

At least part of the findings are in line with the findings of Baker and Wurgler (2006;2007). Least volatile stocks and value stocks have a positive relationship with sentiment and fare better than speculative stocks when sentiment has increased. In addition the LsVol portfolio has a negative relationship with sentiment. These effects are considered to be due to relative overvaluation of speculative stocks and relative undervaluation of safer, more bond-like stocks during collective market euphoria. Our findings suggest that sentiment does have a dissimilar impact among stocks but the effects are not as clear as proposed by the behavioral finance theory. It seems that the findings are in line with hypotheses 1 and 2. The effect of sentiment on future stock returns is assessed next in a tad longer time period of three months. It might be that the effects of sentiment in returns change in a longer time period. Table 15 summarizes the return regressions with returns of the subsequent quartile.

Table 15 Regression results with a forecast horizon of 3 months

Regressions of portfolio returns on lagged sentiment and the market portfolio return OMXH.

$$R_{pt} = c_p + d_p SENT_{t-1} + \beta_p R_{mt} + \varepsilon_{pt}$$

The sample period includes rolling quarterly returns from 2000 until 2008. The rolling quarterly returns are matched by the sentiment change of the previous month. SENT is the change in the raw sentiment index and ADJSENT is the change in the orthogonalized sentiment index. The portfolios' consist of MSCI indices, portfolios sorted by historical volatility and of the difference portfolio LSVol. LSVol is long on the most volatile quintile of stocks and short on the least volatile quintile. The results of the regression using SENT as an independent variable are presented in panel A and panel B presents the results of regressions with ADJSENT as an independent variable. *** denotes statistical significance at the 1 % level, ** at the 5 % level and * at the 10 % level and restrictions. t-statistics based on Newey-West standard errors are presented in parentheses. Newey-West automatic bandwidth and lag length are used.

	Panel A				Panel B			
	c	d	β	Adj.R_2	c	d	β	Adj.R_2
Growth	-0.010	-0.039**	1.328***	0.914	-0.009	- 0.017**	1.327***	0.914
	(-1.06)	(-1.98)	(26.66)		(-0.93)	(-2.10)	(24.74)	
Value	-0.007	0.049*	0.619***	0.611	-0.009	0.019	0.622***	0.602
	(-0.59)	(1.75)	(8.97)		(-0.69)	(1.45)	(8.36)	
Small	0.026*	0.045*	0.511***	0.493	0.025*	0.032***	0.511	0.499
	(1.81)	(1.80)	(5.68)		(1.88)	(2.88)	(5.63)	
Lg Gr.	-0.015	-0.034	1.336***	0.887	-0.014	- 0.020	1.335***	0.887
	(-1.37)	(-1.57)	(22.93)		(-1.60)	(-1.84)	(27.43)	
Lg Val.	0.016	0.057	0.461***	0.347	0.014	0.013	0.465***	0.328
	(1.12)	(1.63)	(5.27)		(0.96)	(0.85)	(5.00)	
Sm Gr.	0.001	0.031	0.801***	0.503	0.001	0.041**	0.798***	0.516
	(0.03)	(0.78)	(5.95)		(0.01)	(2.37)	(5.81)	
Sm Val	0.024	0.042***	0.410***	0.305	0.023	0.042***	0.410***	0.316
	(1.41)	(3.12)	(3.11)		(1.34)	(3.12)	(3.09)	
Vol10	-0.043*	-0.028	0.802***	0.461	-0.042*	0.026*	0.793***	0.464
	(-1.90)	(-0.64)	(9.53)		(-1.95)	(1.85)	(9.04)	
Vol20	-0.041*	-0.008	0.839***	0.548	-0.040*	0.021*	0.835***	0.552
	(-1.86)	(-0.21)	(11.69)		(-1.96)	(1.734)	(11.42)	
Vol40	0.013	0.038**	0.354***	0.293	0.013	0.034***	0.359***	0.312
	(1.11)	(2.08)	(3.01)		(1.09)	(2.83)	(3.09)	
Vol80	0.032***	0.049***	0.288***	0.309	0.031***	0.030***	0.288***	0.312
	(3.20)	(3.03)	(2.79)		(3.11)	(2.89)	(2.80)	
Vol90	0.034***	0.041**	0.280***	0.295	0.033***	0.030***	0.280***	0.307
	(3.23)	(2.50)	(2.87)		(3.24)	(2.98)	(2.91)	
LsVol	-0.072***	-0.056*	0.552***	0.367	-0.071***	-0.009	0.546***	0.351
	(-3.38)	(-1.70)	(4.62)		(-3.41)	(-0.71)	(4.35)	

The coefficients of determination are at pretty much same level as in the former equation. It is higher on all the volatility portfolios and value portfolios, which is positive as the regression equations are considered to be a better fit. It seems that sentiment has a quite similar impact on returns in a longer time period. Small stocks and less volatile stocks remain highly statistically significant and large growth stocks are in a negative relation with investor sentiment. The coefficients of SENT changes are generally higher than the coefficients of ADJSENT changes due to two factors: the ADJSENT index is more volatile and the average change of the index is higher than in the SENT index and because SENT index contains rational macroeconomic expectations about the future.

If ADJSENT increased by a half a standard deviation, the return of the subsequent quarter of small cap, small growth and small value portfolios increase by 1.6, 2.05 and 2.1 percent respectively. Some form of trend reversal can be spotted from the most volatile portfolios as the coefficients have turned to negative when SENT is used as an independent variable. The effects of investor sentiment seem to remain strong among small capitalization stocks and less volatile stocks. It seems that the ADJSENT index has more explanatory power on the small cap, small growth and the most volatile portfolios than the SENT index. These stocks are considered to be more sentiment prone and they are more affected by irrational sentiment. Small stocks and least volatile stocks seem to outperform larger stocks, more volatile stocks and growth stocks. Despite the relationship between small stocks and sentiment, the effects are as expected. Growth stocks are in a negative relationship with sentiment and less volatile stocks clearly outperform the most volatile portfolios. In addition sentiment does not have sizeable influence on large stocks and value stocks. The findings are somewhat in line with previous research. Finter et al. (2010) observed that sentiment prone stocks outperform in short term from one month to a quarter, Baker and Wurgler (2006;2007) found out that less volatile stocks should outperform most volatile stocks when sentiment is high.

However, it is possible that the returns of small companies and less volatile companies are related to some other known risk factors such as the abnormal profits of small stocks or stocks that have high book-to-market values. Kallunki et al. (1997) found out that the size anomaly is quite prevalent in the Finnish stock market. Evidently, this is an issue that needs to be scrutinized and accounted for.

4.3 Stock returns under extreme sentiment levels

In this subchapter it is assessed whether sentiment has a bigger impact on the returns of stocks when the sentiment is at extreme levels. As we have previously studied the effects of sentiment changes on returns, it is necessary to chart out the effects of extreme sentiment levels on returns. This is done by replacing the sentiment index from the equation by two dummy variables. The methodology is similar to the methodology applied in the article of Finter et al. (2010) who studied the effects of investor sentiment on the German stock market. The moderate levels of sentiment which range from -1 to 1 are used as reference levels for the dummy variables. The DS^H dummy is given values of 1 when the sentiment is over 1 and 0 otherwise and the DS^L is given values of 1 when the sentiment is lower than -1 and 0 otherwise. This way it is possible to assess whether sentiment needs to be sufficiently high or low to have an impact on returns and whether there is a difference between the effects of sentiment during low and high sentiment periods.

As the time period selected is quite short, the amount of observations for the extreme levels of sentiment is not particularly high. However, by using dummy variables, the whole time series can be utilized. It is possible to detect whether the average returns after extreme sentiment periods are statistically deviant from the returns after moderate sentiment levels. The SENT index hits sentiment levels of more than one standard deviation above the average 15 times and sentiment levels below 1 standard deviation 23 times. The ADJSENT index is more than 1 standard deviation on either positive or negative side 19 times for both levels of sentiment.

In addition, it is important to note that the dummy variables are formed on the basis of sentiment levels index. The SENT levels index has a unit root and is regarded as a non-stationary time series. As the dummy variables are derived from the levels index, the dummy variables might be non-stationary as well. The ADJSENT levels index could be considered to be stationary. The utmost interest is directed towards the effects of the ADJSENT index and the SENT index dummy variables are included in the tables as additional information. However, the statistical significance values of the SENT dummies should be treated with a higher degree of suspicion. Table 16 presents the results of the return regressions.

Table 16 Regression results during extreme sentiment levels

Regressions of portfolio returns on lagged sentiment dummies and the market portfolio return OMXH.

$$R_{pt} = c_p + d_1 DS^H{}_{t-1} + d_2 DS^L{}_{t-1} + \beta_p R_{mt} + \varepsilon_{pt}$$

The sample period includes rolling monthly returns from 2000 until 2008. The rolling monthly returns are matched by the sentiment of the previous month. DS^H gets values of 1 when sentiment is over 1 and 0 otherwise. DS^L gets values of 1 when sentiment is below -1 and 0 otherwise. Panel A presents results using dummies based on SENT and panel B ADJSENT dummies. *** denotes statistical significance at the 1 % level, ** at the 5 % level and * at the 10 % level and restrictions. t-statistics based on Newey-West standard errors are presented in parentheses. Newey-West automatic bandwidth and lag length are used.

	Panel A					Panel B				
	c	d_1	d_2	β	Adj.R_2	c	d_1	d_2	β	Adj.R_2
Growth	-0.006	0.010	0.008	1.368***	0.921	-0.008*	0.017*	0.015*	1.356***	0.924
	(-1.55)	(1.11)	(1.03)	(34.06)		(-1.96)	(1.78)	(1.87)	(33.42)	
Value	0.001	-0.126	-0.009	0.629***	0.589	0.006	-0.023*	-0.027***	0.648***	0.614
	(0.14)	(-0.92)	(-1.40)	(10.09)		(1.02)	(-1.92)	(-2.71)	(11.31)	
Small Cap	0.012**	-0.015	-0.014	0.466***	0.481	0.014**	-0.017	-0.019*	0.477***	0.489
	(2.49)	(-1.06)	(-1.31)	(7.46)		(2.37)	(-1.32)	(-1.89)	(7.67)	
Large Gr.	-0.009**	0.012	0.011	1.376***	0.894	-0.010**	0.017	0.017**	1.365***	0.896
	(-2.07)	(1.18)	(1.30)	(28.53)		(-2.29)	(1.26)	(2.11)	(27.89)	
Large Val.	0.006	-0.029	0.001	0.457***	0.203	0.019***	-0.045**	-0.048***	0.488***	0.261
	(1.01)	(-1.51)	(0.05)	(4.05)		(2.85)	(-2.52)	(-3.61)	(4.79)	
Small Gr.	0.002	-0.014	-0.012	0.689***	0.483	0.002	-0.010	-0.015	0.693***	0.483
	(0.20)	(-0.66)	(-0.73)	(7.10)		(0.19)	(-0.57)	(-0.85)	(7.19)	
Small Val.	0.014**	-0.020	-0.024	0.337***	0.264	0.016**	-0.025*	-0.028***	0.354***	0.280
	(2.53)	(-1.29)	(-1.57)	(3.79)		(2.31)	(-1.77)	(-2.65)	(3.96)	
Vol10	-0.017	-0.009	-0.006	0.641***	0.241	-0.020*	-0.002	0.001	0.637***	0.240
	(-1.44)	(-0.38)	(-0.33)	(6.50)		(-1.75)	(-0.14)	(0.06)	(6.59)	
Vol20	-0.011	-0.021	-0.006	0.683***	0.379	-0.013	-0.012	-0.001	0.680***	0.374
	(-1.16)	(-0.98)	(-0.42)	(10.92)		(-1.37)	(-0.83)	(-0.07)	(11.00)	
Vol40	0.008*	-0.005	-0.022	0.303***	0.221	0.007	-0.008	-0.015	0.312***	0.208
	(1.70)	(-0.46)	(-1.49)	(3.60)		(1.19)	(-0.65)	(-1.49)	(3.63)	
Vol80	0.014***	-0.010*	-0.019	0.213***	0.198	0.014***	-0.007	-0.019**	0.220***	0.192
	(3.41)	(-1.68)	(-1.43)	(2.90)		(2.70)	(-0.86)	(-2.54)	(2.95)	
Vol90	0.014***	-0.010	-0.019	0.200***	0.171	0.013***	-0.008	-0.016**	0.208***	0.159
	(3.38)	(-1.29)	(-1.25)	(2.86)		(2.72)	(-0.86)	(-2.48)	(2.82)	
LSVol	-0.025***	-0.010	0.013	0.468***	0.241	-0.027***	-0.003	0.016	0.457***	0.240
	(-2.75)	(-0.54)	(0.78)	(6.30)		(-3.04)	(-0.27)	(0.98)	(6.22)	

It can be observed from table 16 that the SENT index has a much smaller statistical impact on the returns of portfolios during extreme levels of sentiment compared to the ADJSENT index. The reason to this is probably that the changes of the SENT index are more conservative and the effect of sentiment may have already affected the returns. Another reason might be that during extreme sentiment levels, the irrational component of sentiment has a larger effect on investment decisions of people. During times of extreme optimism or pessimism the effect of mood is more amplified and macroeconomic data might be considered less valuable.

In general, the effects of SENT to the next month returns of portfolios do not deviate from the effects during moderate sentiment levels. However, when the orthogonalized sentiment index is utilized, the returns of value stocks, larger stocks and least volatile stocks are most affected by sentiment especially during periods of notable pessimism.

The returns of the subsequent month for growth portfolios are on the positive side when the previous month sentiment hits higher or lower levels. Value stocks and smaller stocks experience contrary results than growth stocks. The returns of smaller stocks and value stocks decline when sentiment is over a standard deviation positive or negative. Only the portfolios of the least volatile stocks seem to be more affected by sentiment during extreme levels. It is interesting to notice that small stocks and value stocks have a negative relationship with sentiment as they were in a positive relationship with sentiment changes. It might be due to mean reverting of returns when sentiment has reached its peak. When sentiment decreases, the effect on small cap stocks is negative and if the previous month's sentiment level is under -1, the returns are lower than during moderate levels of sentiment. This implies that pessimism has a more persistent effect on the returns of small cap stocks.

When the sentiment is over a standard deviation positive, the returns of Growth and Value portfolios yield a return that deviate by 1.7 and -2.3 percent from the returns during moderate sentiment levels. The returns of Value portfolios are smaller when sentiment is particularly high or low. The Large Value portfolio is significant at the 5 percent level when ADJSENT is positive and at the 1 percent level when ADJSENT is negative. The returns are 4.5 percent smaller during high levels of sentiment and 4.8 during low levels of sentiment than during moderate sentiment levels. Altogether, the relation with sentiment is similar with all the value portfolios. It could be that growth stocks experience momentum when the sentiment is high and thus the returns of the subsequent month are positive. During the whole time period, the value stocks, small stocks and least volatile stocks have yielded a better monthly average return than growth stocks but during strong levels of sentiment the returns of the following month are higher than those of these stocks. In addition, smaller stocks seem to earn lower returns.

It is possible that the more experienced investors “ride the wave” and keep investing in the larger growth companies as the more inexperienced investors realize their winnings too early or mimic the actions of the larger market participants. Many institutional investors and mutual funds are considered to utilize a momentum strategy and herd in and out of stocks. The momentum effect is considered to be largest among companies that are harder to value and companies that are more frequently traded (Chan, Hameed & Tong 2000; Glaser et al. 2003).

It seems that the stocks that are most traded by retail traders experience depreciations in value during the subsequent time periods. Kumar and Lee (2006) found that retail investors tend to overweight value stocks relative to growth stocks. It is possible that Finnish retail investors prefer large value stocks as they are considered safer and during extreme sentiment levels they might be lured in by the potential profits of the growth stocks and adjust their portfolios accordingly. Schmeling (2009) found that consumer confidence was related to the profits of value companies in Finland which could indicate that Finnish investors prefer value companies in general. The preference for value stocks and small stocks is probably smaller during high periods of sentiment. Karhunen and Keloharju (2001) found out that Finnish investors prefer smaller stocks compared to foreign investors. On the other hand, senior citizens have usually more investment capital and they prefer large and stable companies.

It is unclear why the subsequent month returns of the growth stocks are higher during low sentiment periods compared to moderate levels of sentiment if there was momentum involved. Is it due to the stocks being struck too low and that there is a slight positive rebound or are the retail investors slower to adjust their stock portfolios and the effects of sentiment take longer to impact the returns. Another interesting find is that value portfolios yield negative returns during extreme sentiment levels. The safer stocks should yield higher returns than speculative stocks when sentiment is high according to the behavioral finance framework. However, our previous findings did indicate that the returns of value stocks and least volatile stocks decreased when sentiment decreases and it seems that this effect is pronounced during periods of extremely low sentiment. This indicates that pessimism and negative investor sentiment has a clearly larger influence on the returns of these stocks.

Finter et al. (2010) found out that sentiment prone stocks deliver higher returns than sentiment insensitive stocks in the short run and that the effects of sentiment are especially significant during periods of very low sentiment. The results presented above are quite similar. Growth stocks and more volatile stocks yield higher returns than safer stocks during extreme levels of sentiment and sentiment seems to have more impact during times of high pessimism. This is consistent with the general tendency of investors to react more strongly in response to bad news than good news (Conrad, Cornell & Landsman 2002).

Next it is assessed whether the effects of extreme levels of sentiment are stronger in a longer time period of three months. Table 17 presents the results of these regressions. The coefficients of determination seem to be slightly higher or at the same level compared to the regression equation that utilized the sentiment levels indices. In general, sentiment does seem to affect the returns of MSCI portfolios in greater extent compared to the volatility portfolios. Unlike in the previous equation, also SENT does have statistical significance in several portfolio return regressions when the time period is three months.

It is interesting to note that when the time period is extended to three periods, low levels of sentiment have a weaker impact on returns than before. On the other hand, the returns of portfolios during high levels of sentiment seem to be much more affected by sentiment when the time period is extended to three months. The reason why sentiment seemed to have more impact on a shorter time period during low sentiment might be that as previously mentioned, fear of losing money is a more powerful feeling than the chance of prosperity. It is likely that sell decisions are made more rapidly and the selling pressure of stocks clears in a shorter time period. The effects of sentiment during low levels of sentiment wane in a longer time period.

The growth stocks except the Small Growth portfolio seem to earn higher future returns during extreme levels of sentiment and especially during high levels. The Small Growth portfolio earns a negative return when sentiment is high and a positive return when the sentiment is low. Although the coefficients are not statistically significant for this portfolio, it can be observed that the returns of larger growth stocks have a different pattern compared to the Small Growth portfolio. This can only be attributed to the size factor. Smaller, less volatile and value stocks earn negative returns during extreme levels of sentiment. Finter et al. (2010) found out that sentiment prone stocks outperform less speculative stocks when the time period was three months. This effect completely reversed in the longer term from six to twelve months following the sentiment. These results are partially in line with the findings of this thesis. Large growth stocks seem to outperform other stocks when the sentiment is high. However, small stocks and small growth stocks underperform in one month and three month periods relative to larger stocks. Previously we have presumed that these smaller stocks should be more sentiment sensitive stocks.

Table 17 Regression results during extreme sentiment levels II

Regressions of portfolio returns on lagged sentiment dummies and the market portfolio return OMXH.

$$R_{pt} = c_p + d_1 DS^H{}_{t-1} + d_2 DS^L{}_{t-1} + \beta_p R_{mt} + \varepsilon_{pt}$$

The sample period includes rolling quarterly returns from 2000 until 2008. The rolling quarterly returns are matched by the sentiment of the previous month. DS^H gets values of 1 when sentiment is over 1 and 0 otherwise. DS^L gets values of 1 when sentiment is below -1 and 0 otherwise. Panel A presents results using dummies based on SENT and panel B ADJSENT dummies. *** denotes statistical significance at the 1 % level, ** at the 5 % level and * at the 10 % level. t-statistics based on Newey-West standard errors are presented in parentheses. Newey-West automatic bandwidth and lag length are used.

	Panel A					Panel B				
	c	d_1	d_2	β	Adj.R_2	c	d_1	d_2	β	Adj.R_2
Growth	-0.002*	0.058***	0.012	1.309***	0.922	-0.024**	0.059***	0.024	1.294***	0.923
	(-1.93)	(2.75)	(0.64)	(25.62)		(-2.32)	(2.64)	(1.33)	(25.79)	
Value	0.004	-0.064*	-0.022	0.636***	0.608	0.009	-0.063*	-0.045	0.654***	0.616
	(0.28)	(-1.70)	(-0.89)	(9.11)		(0.72)	(-1.76)	(-1.84)	(9.00)	
Small Cap	0.035**	-0.057*	-0.008	0.532***	0.498	0.039***	-0.063**	-0.017	0.548***	0.507
	(2.54)	(-1.81)	(-0.40)	(5.90)		(2.63)	(-2.04)	(-0.68)	(6.18)	
Large Gr.	-0.029**	0.065***	0.021	1.311***	0.895	-0.031***	0.065**	0.029	1.295***	0.896
	(-2.55)	(2.64)	(1.09)	(22.20)		(-2.80)	(2.29)	(1.53)	(22.00)	
Large Val.	0.028	-0.051***	-0.028	0.487***	0.342	0.040**	-0.064***	-0.077*	0.516***	0.386
	(1.49)	(-2.64)	(-0.62)	(5.27)		(2.46)	(-3.12)	(-1.86)	(5.62)	
Small Gr.	-0.001	-0.045	0.028	0.805***	0.499	-0.003	-0.036	0.038	0.806***	0.500
	(-0.03)	(-1.01)	(0.83)	(5.45)		(-0.13)	(-0.83)	(0.91)	(5.75)	
Small Val.	0.039**	-0.078**	-0.028	0.438**	0.317	0.043**	-0.085**	-0.037	0.460***	0.331
	(2.43)	(-2.24)	(-0.93)	(3.49)		(2.61)	(-2.45)	(-1.30)	(3.68)	
Vol10	-0.040	-0.045	0.022	0.805***	0.464	-0.028	-0.063	-0.016	0.829***	0.468
	(-1.39)	(-1.03)	(0.51)	(8.36)		(-0.97)	(-1.51)	(-0.30)	(7.65)	
Vol20	-0.039	-0.050	0.023	0.842***	0.552	-0.027	-0.069*	-0.013	0.867***	0.557
	(-1.50)	(-1.19)	(0.67)	(10.27)		(-0.99)	(-1.70)	(-0.29)	(9.83)	
Vol40	0.024**	-0.041	-0.030	0.372***	0.293	0.022	-0.040	-0.016	0.378***	0.287
	(2.10)	(-1.50)	(-0.97)	(3.20)		(1.49)	(-1.50)	(-0.74)	(3.23)	
Vol80	0.044***	-0.050***	-0.033	0.311***	0.315	0.039***	-0.033*	-0.020	0.311***	0.285
	(4.69)	(-2.91)	(-1.20)	(3.03)		(3.40)	(-1.78)	(-1.23)	(3.01)	
Vol90	0.048***	-0.056***	-0.035	0.303***	0.323	0.043***	-0.036*	-0.025	0.305***	0.286
	(5.00)	(-3.22)	(-1.34)	(3.22)		(3.67)	(-1.74)	(-1.29)	(3.20)	
LSVol	-0.082***	-0.001	0.055	0.532***	0.363	-0.066**	-0.037	0.006	0.557***	0.351
	(-3.48)	(-0.15)	(1.22)	(4.71)		(-2.62)	(-0.99)	(0.15)	(4.58)	

The findings are partly parallel to the findings of Lemmon and Portniaguina (2006) for the value stocks. They observed that high levels of sentiment predict low future returns on value stocks but did not find any evidence that sentiment forecasts the returns on growth stocks. Sentiment has the largest statistical significance on the returns of Growth, Large Growth and Large Value portfolios, and when SENT is used as an independent variable, also on the least volatile stock portfolios. These results are quite surprising, as it was expected that larger stocks and less volatile stocks should not be very much influenced by investor sentiment. It is unclear whether this is due to wealthy, Finnish private investors preferring larger, less volatile stocks or due to the effect of large foreign investors. However, there is a possibility that sentiment reflects some other known stock market risk factors which will be scrutinized in more detail in the robustness tests chapter.

When sentiment is extremely high, the Growth portfolio yields a return that is 5.8 or 5.9 percent larger, depending on the sentiment index, than the return during moderate sentiment. Large Growth portfolio yields a return that is 6.5 percent larger than normally. These returns may be due to momentum. Cooper, Gutierrez and Hameed (2004) suggest that the momentum profits are larger when there has been positive past market returns as the previous successful trades increase the overconfidence of traders which leads to increased trading. But the returns during extremely low sentiment should not be that much larger unless the stock was not undervalued.

During moderate levels of sentiment, value stocks, small stocks and least volatile stocks yield the best returns and growth stocks and more volatile stocks yield negative returns. Nevertheless, when sentiment is extremely positive or negative according to these results one should refrain from investing in these safer stocks at least in the short term. Only the returns of small growth and to some extent Vol20 portfolio act in a way proposed by the behavioral finance theory and the findings of Baker and Wurgler (2006; 2007). This is that there should be a contrarian relation with the sentiment index and the returns of speculative stocks. When the sentiment is high, safe stocks should earn higher returns than speculative stocks and when sentiment is low the effect should be reversed. In their study, they noticed that during periods of high sentiment there was no size effect. The results presented above show that returns of small stocks are weakly affected by sentiment in the following month but are more strongly affected in the following three months. It also seems that the results regarding the returns of the volatility portfolios are quite mixed. No clear conclusions can be made from these regressions except that the effects of sentiment are not very similar to the findings in the American stock market.

Most companies yield negative returns when the sentiment has been extremely high or low with the exception of Growth and Large Growth portfolios. However, it is quite likely that this effect among large growth stocks reverses in a longer time frame. The findings are in line with hypothesis 2, but not clearly in line with hypothesis 3. It seems that sentiment affects different types of stocks in dissimilar manners, but it cannot be said that during high levels of sentiment, the more sentiment prone stocks would earn lower returns and during low levels of sentiment higher returns, which would reflect misvaluation. It is more likely that investors switch between stocks in relation to the prevailing sentiment. When sentiment is high, companies with the greatest growth prospects intrigue people more than more stable companies. The less speculative stocks probably do not reflect overvaluation but rather loss of public interest and disposition effect. The patterns among portfolio returns are not explicit enough to draw precise conclusions, but it can be stated that the effects of sentiment on the future returns of portfolios need to be more scrutinized.

4.4 Robustness tests

In this chapter it is investigated whether the effects of sentiment persist when additional known risk factors are included in the regressions as independent variables. De Long et al. (1990) suggest that noise traders create an additional noise trader risk, which should be accounted for by market participants. Baker and Wurgler (2007) propose that known asset pricing models do not grasp the effects of sentiment and that sentiment should be taken into consideration in asset pricing.

Fama and French (1993) suggest that the CAPM models’ beta is not a sufficient measure of systematic risk. Variables that have no special standing in asset pricing theory exhibit reliable explanatory power in the cross section of average returns. In addition to the excess market return, size and book-to-market factors have been shown to have explanatory power in determining asset returns. The size and book-to-market factors are related to economic fundamentals. Firms with low book-to-market ratios tend to have low earnings on assets and the low earnings are highly persistent. On the other hand, companies with low book-to-equity tend to earn high earnings. However, the earnings growth generally is mean reverted faster than investors expect. When the size and the book-to-market factors are used combined in an equation, they seem to grasp the effects of leverage and the P/E-anomaly.

The pricing error of the CAPM is three to five times larger than the pricing error of the three factor model. The three factor model explains largely the returns of portfolios, but it might not grasp the effect of short term momentum among stocks (Fama & French 1996).

The Fama-French factors are included in the previous regression equations as additional control variables to determine whether sentiment is an additional risk factor. These Fama-French factors are constructed by following the methodology that can be found from the website of French (2010). The average returns of value-weighted MSCI portfolios based on size and investment style are utilized. The size factor SMB is created by deducting the average return of all available MSCI small portfolios from the average return of all available MSCI large portfolios. There are five available small cap and five large cap portfolios that are used to create the average small cap and large cap portfolio returns. The book-to-market factor HML is created by deducting the average return of high book-to-market portfolios from low book-to-market portfolios. The portfolios utilized are yet again the MSCI portfolios that are sorted by investment style as growth and value portfolios. In other words, the average return of the four available MSCI value portfolios is deducted by the average return of the four available MSCI growth portfolios. The regression equations can be expressed in the following way:

$$R_{pt} = c_p + d_p SENT_{t-1} + \beta_p Rmrf_t + s_p SMB_t + h_p HML_t + \varepsilon_{pt} \tag{19}$$

$$R_{pt} = c_p + d_1 DS^H{}_{t-1} + d_2 DS^L{}_{t-1} + \beta_p Rmrf_t + s_p SMB_t + h_p HML_t + \varepsilon_{pt} \tag{20}$$

In the upper equation, the sentiment index is used as an independent variable and in the lower equation it is replaced by two dummy variables, DS^H and DS^L. R_p is the return of the portfolio, c is the constant, *SENT* is the sentiment index which is either the raw index or the orthogonalized one. *Rmrf* is the excess market return over the risk free rate that is the one month or three month euribor depending on the regression time frame. *SMB* is the small factor and *HML* is the book-to-market factor and ε is the error term. The robustness tests results are assessed next.

Table 18 Robustness test regression results

This table presents the results of regressions on portfolio returns of the following form:

$$R_{pt} = c_p + d_p SENT_{t-1} + \beta_p Rmrf_t + s_p SMB_t + h_p HML_t + \varepsilon_{pt}$$

The regression equation consists of portfolio return R_p, constant c, sentiment changes index SENT and the Fama-French factors Rmrf, SMB and HML.The sample period includes rolling monthly from 2000 until 2008. The rolling monthly returns are matched by the sentiment change of the previous month. Panel A presents results using the raw sentiment proxy SENT and Panel B presents results using the orthogonalized ADJSENT as an independent variable.*** denotes statistical significance at the 1 % level, ** at the 5 % level and * at the 10 % level. t-statistics based on Newey-West standard errors are presented in parentheses. Newey-West automatic bandwidth and lag length are used.

Panel A : SENT

	c	d	β	s	h	Adj.R_2
Growth	0.004***	-0.009*	1.016***	-0.371***	-0.391***	0.980
	(2.89)	(-1.89)	(22.89)	(-15.15)	(-6.87)	
Value	-0.006	0.008	0.975***	0.224***	0.528***	0.767
	(-1.55)	(0.93)	(21.43)	(2.88)	(7.47)	
Small cap	0.005***	0.004	0.864***	0.564***	0.329	0.917
	(5.32)	(0.88)	(36.64)	(14.98)	(7.64)	
Large Gr.	0.003	-0.006	0.971***	-0.452***	-0.425***	0.969
	(1.53)	(-1.01)	(20.90)	(-10.04)	(-7.05)	
Large Val.	-0.001	0.012	0.888***	-0.223***	1.039***	0.715
	(-0.11)	(1.28)	(12.28)	(-2.64)	(6.79)	
Small Gr.	-0.003	0.004	1.088***	0.829***	0.092	0.830
	(-0.99)	(0.51)	(22.16)	(12.33)	(0.84)	
Small value	0.002	0.002	0.835***	0.594***	0.499***	0.835
	(1.23)	(0.32)	(17.98)	(8.55)	(6.68)	
Vol40	0.001	0.011	0.688***	0.496***	0.343***	0.605
	(0.40)	(1.52)	(9.27)	(4.30)	(4.48)	
Vol80	0.008***	0.022***	0.511***	0.374***	0.285***	0.669
	(3.68)	(3.43)	(9.29)	(4.87)	(4.44)	
Vol90	0.008***	0.021**	0.459***	0.356***	0.218***	0.546
	(3.04)	(2.60)	(7.55)	(3.77)	(2.74)	
LsVol	-0.023***	-0.029	0.459***	0.243*	-0.235	0.291
	(-3.19)	(-1.42)	(3.23)	(1.79)	(-1.24)	

Panel B : ADJSENT

	c	d	β	s	h	Adj.R_2
Growth	0.005***	-0.001	1.003***	-0.376***	-0.405***	0.979
	(3.49)	(-0.288)	(22.11)	(-14.09)	(-7.29)	
Small cap	0.005***	0.006***	0.859***	0.558***	0.327	0.920
	(3.77)	(2.51)	(32.79)	(16.17)	(9.81)	
Large Val.	-0.001	0.013**	0.882***	-0.234***	1.043***	0.722
	(-0.18)	(2.22)	(11.79)	(-2.88)	(8.98)	
Small Gr.	-0.003	0.010*	1.074***	0.817***	0.086	0.834
	(-0.99)	(1.97)	(22.52)	(12.84)	(0.84)	
Small value	0.003	0.006*	0.825***	0.585***	0.493***	0.839
	(1.25)	(1.69)	(17.49)	(8.58)	(6.78)	
Vol10	-0.018**	0.014	0.892***	0.648***	-0.036	0.372
	(-2.37)	(1.16)	(5.97)	(4.37)	(-0.21)	
Vol40	0.001	0.009*	0.688***	0.491***	0.351***	0.609
	(0.30)	(1.92)	(9.39)	(4.33)	(4.50)	
Vol80	0.007***	0.011***	0.525***	0.374***	0.310***	0.659
	(3.17)	(2.78)	(9.05)	(4.95)	(4.39)	
Vol90	0.008***	0.011**	0.469***	0.354***	0.241***	0.540
	(2.85)	(2.48)	(7.01)	(3.86)	(3.04)	

After including the Fama-French factors, the adjusted R_2 increases in all of the equations meaning that these regression models indeed do a better job in explaining the future returns of these portfolios. The coefficient of determination is especially increased in the small cap and volatility portfolios. It is no surprise that the SMB factor has a particularly strong explanatory power among small stocks.

The explanatory power of sentiment on stock returns diminishes once the Fama-French factors are included in the model. Both the SMB and HML factors mostly grasp the effects of sentiment. According to the findings of Kallunki et al. (1997), the size anomaly and P/E-anomaly are quite strong in Finland. Similar conclusions can be made by applying the Fama-French factors in the time series as these are taken into consideration by the model.

When SENT is used as an independent variable the growth portfolio is statistically significant at the 10 percent level and the Vol80 and Vol90 portfolios are statistically significant at the 1 and 5 percent levels respectively. The return of the growth portfolio decreases by 0.45 percent when the previous month's SENT index changes by half a unit. The relationship is more powerful for the least volatile portfolios. A similar change in the SENT index is rewarded by a 1.1 percent increase in the Vol80 portfolio. The orthogonalized sentiment index ADJSENT remains statistically significant on the small portfolios, least volatile portfolios and on the Large Value portfolio. The coefficient is at least 1.0 for the Small Growth, Large Value, Vol80 and Vol90 portfolios. Although, the ADJSENT fluctuates more aggressively and changes of at least 0.5 units are more common, creating a profitable trading strategy based solely on sentiment is not likely to be possible. In a world of zero or close to zero transaction costs one could achieve occasional profits, but not consistently.

The observation that small capitalization stocks were statistically significant when the orthogonalized sentiment index was used indicates that small stocks are indeed more influenced by irrational sentiment. Sentiment remained to have a positive relationship between small cap stocks which does not indicate that small cap stocks are overvalued but rather that the effect of sentiment is lagged on the returns of these stocks. The effect of irrational sentiment on Large Value and the least volatile portfolios is either lagged or due to relative undervaluation as proposed by Baker and Wurgler (2006). It is possible that previous market rises lure private investors to participate in the market and these investors prefer safer stocks and smaller stocks. Table 19 presents the effects of sentiment on the subsequent quarter returns of stocks.

Table 19 Robustness test regression results on the subsequent quarter returns

This table presents the results of regressions on portfolio returns of the following form:

$$R_{pt} = c_p + d_p SENT_{t-1} + \beta_p Rmrf_t + s_p SMB_t + h_p HML_t + \varepsilon_{pt}$$

The regression equation consists of portfolio return R_p, constant c, sentiment changes index SENT and the Fama-French factors Rmrf, SMB and HML. The sample period includes rolling quarterly returns from 2000 until 2008. The rolling quarterly returns are matched by the sentiment of the previous month. Panel A presents results using the raw sentiment proxy SENT and Panel B presents results using the orthogonalized ADJSENT as an independent variable.*** denotes statistical significance at the 1 % level, ** at the 5 % level and * at the 10 % level. t-statistics based on Newey-West standard errors are presented in parentheses. Newey-West automatic bandwidth and lag length are used.

Panel A : SENT

	c	d	β	s	h	Adj.R_2
Growth	0.013***	-0.006	1.047***	-0.379***	-0.330***	0.982
	(3.51)	(-0.94)	(40.36)	(-19.30)	(-9.18)	
Value	-0.016	0.005	0.917***	0.222**	0.511***	0.807
	(-1.67)	(0.37)	(18.28)	(2.61)	(7.87)	
Small cap	0.016***	0.005	0.823***	0.584***	0.300	0.949
	(5.22)	(0.77)	(34.18)	(17.79)	(6.99)	
Small value	0.008	0.002	0.834***	0.586***	0.523***	0.864
	(1.46)	(0.14)	(17.36)	(6.16)	(4.68)	
Vol40	0.002	-0.002	0.657***	0.443***	0.359***	0.694
	(0.34)	(-0.19)	(7.85)	(4.87)	(3.97)	
Vol80	0.022***	0.009	0.521***	0.339***	0.300***	0.671
	(4.40)	(0.98)	(7.60)	(3.63)	(3.57)	
Vol90	0.025***	0.008	0.521***	0.339***	0.295***	0.671
	(4.09)	(0.98)	(7.60)	(3.63)	(3.57)	
LsVol	-0.063***	-0.037	0.436***	0.298	-0.458***	0.487
	(-3.31)	(-1.49)	(3.02)	(1.56)	(-2.85)	

Panel B : ADJSENT

	c	d	β	s	h	Adj.R_2
Growth	0.013***	0.007**	1.038***	-0.386***	-0.342***	0.982
	(3.74)	(2.02)	(41.42)	(-19.41)	(-9.75)	
Small cap	0.016***	-0.001	0.826***	0.585***	0.285***	0.949
	(5.53)	(-0.10)	(36.94)	(20.51)	(8.72)	
Large Gr.	0.010**	0.007	1.021***	-0.444***	-0.357***	0.969
	(2.12)	(1.41)	(41.40)	(-8.56)	(-7.33)	
Small Gr.	-0.004	0.005	1.056***	0.896***	-0.060	0.879
	(-0.52)	(0.52)	(17.25)	(14.13)	(-0.63)	
Small value	0.008	0.002	0.833***	0.585***	0.523***	0.864
	(1.44)	(0.38)	(15.69)	(6.40)	(4.93)	
Vol10	-0.039**	0.007	0.878***	0.610***	-0.241	0.622
	(-2.16)	(0.56)	(6.32)	(4.03)	(-1.64)	
Vol20	-0.039**	-0.003	0.975***	0.657***	-0.150	0.748
	(-2.34)	(-0.36)	(8.53)	(4.73)	(-1.27)	
Vol40	0.002	0.006	0.651***	0.438***	0.352***	0.695
	(0.41)	(0.73)	(7.91)	(5.00)	(4.27)	
Vol80	0.022***	0.006	0.555***	0.358***	0.338***	0.732
	(4.58)	(0.79)	(8.45)	(5.92)	(4.42)	

It seems that sentiment does not have an impact on returns in a longer time frame of three months after controlling for Fama-French factors. It is interesting to notice that many of the sentiment related return patterns reverse when the time frame is extended. The sentiment regression coefficients decrease in value which indicates that the return pattern is reversed. The coefficients were 0.006, 0.010 and 0.011 for Small cap, Small growth and Vol80 portfolios when the forecast horizon was 1 month. When the forecast horizon was extended to three months the coefficients were -0.001, 0.005 and 0.006. The return of the Growth portfolio has turned from negative to statistically significant at the 5 percent level with a coefficient value of 0.007, which means that the return of the growth portfolio increases by 0.7 percent in the following quarter when the ADJSENT has increased by a single unit. The difference in the sentiment related return patterns among stocks yet again indicates that sentiment does affect returns of stocks in different ways.

It can be concluded that large changes in sentiment do have an impact on the future returns of small stocks, least volatile stocks and the Large Value portfolio. This effect fades out when the forecast horizon is extended. The findings are partly in line with the findings of Baker and Wurgler (2006;2007). However, sentiment prone stocks did not exhibit patterns of clear overvaluation, but rather stocks that were considered to be more safe fared better than speculative stocks when the previous month's sentiment increased. Sentiment is usually related to previous good or bad news, returns or macro developments. Previous returns tend to drive sentiment which affects the subsequent returns (Schmeling 2009). Thus, it is likely that more speculative and volatile stocks have already experienced price changes and the sentiment is already compounded in the prices, but the safer and smaller stocks are slower to adjust due to private investors' slower participation on the market.

Even after controlling for Fama-French factors, the changes in sentiment did remain statistically significant in a few cases. In order to be even more robust, one should add a momentum factor in the model but it is likely that it would not significantly change the results as momentum has been observed to be more prevalent among stocks that are frequently traded (Glaser & Nöth 2003). Small stocks and less volatile stocks are usually not very highly traded. In addition, it is important to point out that the sample size was relatively small with 107 observations when the sentiment changes were utilized and 108 observations with sentiment levels data. Some of the portfolio return distributions were leptokurtic and might not have been normally distributed. Although it is sensible to treat the results with a healthy dose of criticism, we believe that the results provide a fairly good description of the phenomenon. The robustness of the extreme sentiment regressions is presented next in Table 20.

Table 20 Robustness test regression results during extreme sentiment levels

This table presents regression results of the following form:

$$R_{pt} = c_p + d_1 DS^H{}_{t-1} + d_2 DS^L{}_{t-1} + \beta_p Rmrf_t + s_p SMB_t + h_p HML_t + \varepsilon_{pt}$$

The regression equation consists of portfolio return $R_{p,}$ constant c, sentiment index dummies DS and the Fama-French factors Rmrf, SMB and HML. The sample period includes rolling monthly or quarterly returns from 2000 until 2008. The rolling monthly and quarterly returns are matched by the sentiment of the previous month. Panel A presents results using the raw sentiment proxy SENT and Panel B presents results using the orthogonalized ADJSENT as an independent variable. Panel numbers 1 and 2 signify the time period where 1 is monthly returns and 2 is quarterly returns. *** denotes statistical significance at the 1 % level, ** at the 5 % level and * at the 10 % level. t-statistics based on Newey-West standard errors are presented in parentheses. Newey-West automatic bandwidth and lag length are used.

Panel 1A :SENT

	c	d_1	d_2	β	s	h	Adj.R_2
Vol80	0.008***	-0.002	-0.009	0.538***	0.350***	0.357***	0.623
	(3.13)	(-0.31)	(-1.10)	(9.53)	(4.90)	(4.97)	
Vol90	0.009***	-0.003	-0.010	0.483***	0.322***	0.297***	0.494
	(2.92)	(-0.45)	(-0.99)	(7.54)	(3.76)	(3.97)	

Panel 1B : ADJSENT

	c	d_1	d_2	β	s	h	Adj.R_2
Growth	0.006***	0.001	-0.007*	0.996***	-0.370***	-0.431***	0.979
	(2.95)	(0.37)	(-1.70)	(22.69)	(-14.70)	(-7.73)	
Value	-0.004	-0.005	-0.002	0.987***	0.252***	0.501***	0.763
	(-1.07)	(-0.53)	(-0.37)	(20.92)	(3.12)	(6.75)	
Small Cap	0.005***	-0.003	-0.005	0.868***	0.548***	0.349***	0.915
	(2.99)	(-0.92)	(-0.16)	(32.57)	(15.79)	(9.62)	
Lg Gr.	0.006*	-0.001	-0.007*	0.954***	-0.442***	-0.475***	0.967
	(1.88)	(-0.10)	(-1.75)	(22.02)	(-9.42)	(-7.52)	
Lg Value	-0.002	-0.007	0.003	0.904***	-0.303***	1.159***	0.718
	(-0.44)	(-0.68)	(0.32)	(10.41)	(-2.71)	(8.15)	
Vol80	0.005	0.007	0.001	0.543***	0.360***	0.370***	0.619
	(1.40)	(1.25)	(0.03)	(8.79)	(4.60)	(4.49)	
Vol90	0.006	0.004	-0.001	0.487***	0.332*	0.304***	0.486
	(1.62)	(0.60)	(-0.16)	(6.96)	(3.53)	(3.58)	

Panel 2A : SENT 3 month return

	c	d_1	d_2	β	s	h	Adj.R_2
Growth	0.015***	0.008	-0.014	1.043***	-0.375***	-0.344***	0.982
	(2.97)	(1.09)	(-1.60)	(42.53)	(-18.56)	(-8.96)	
Value	-0.021**	-0.007	0.018	0.917***	0.234***	0.522***	0.785
	(-2.11)	(-0.22)	(1.10)	(20.44)	(2.95)	(6.76)	
Small Cap	0.011***	0.008	0.015*	0.832***	0.582***	0.308***	0.951
	(2.80)	(1.60)	(1.89)	(31.85)	(15.88)	(8.08)	
Lg Gr.	0.008	0.012	-0.005	1.027***	-0.427***	-0.349***	0.969
	(1.12)	(0.85)	(-0.47)	(35.81)	(-9.77)	(-6.82)	
Lg Val.	-0.007	0.007	0.041*	0.798***	-0.187*	0.905***	0.709
	(-0.56)	(0.44)	(1.78)	(9.39)	(-1.67)	(6.87)	
Vol80	0.023***	0.001	-0.007	0.558***	0.367***	0.334***	0.729
	(4.60)	(0.12)	(-0.45)	(8.31)	(4.50)	(4.07)	
Vol90	0.029***	-0.010	-0.014	0.519***	0.341***	0.277***	0.670
	(4.67)	(-0.98)	(-0.79)	(7.69)	(3.82)	(3.29)	

Panel 2B : ADJSENT 3 month returns

	c	d_1	d_2	β	s	h	Adj.R_2
Growth	0.013**	0.007	-0.007	1.043***	-0.376***	0.339***	0.982
	(2.43)	(0.85)	(-0.85)	(42.28)	(-17.90)	(8.03)	
Value	-0.020*	-0.001	-0.008	0.916***	0.237***	0.519***	0.782
	(-1.91)	(-0.05)	(-0.56)	(19.23)	(2.99)	(7.08)	
Small Cap	0.015***	-0.001	0.005	0.828***	0.580***	0.293***	0.949
	(3.05)	(-0.08)	(0.61)	(31.15)	(16.07)	(6.29)	
Lg Gr	0.007	0.009	-0.002	1.026***	-0.429***	-0.348***	0.969
	(0.98)	(0.57)	(-0.17)	(37.43)	(-9.63)	(-6.30)	
Lg Val	-0.003	0.004	0.028	0.795***	-0.191*	0.900***	0.700
	(-0.24)	(0.21)	(1.26)	(9.77)	(-1.80)	(6.42)	
Vol20	-0.016	-0.053*	-0.059**	0.946***	0.663***	-0.298**	0.761
	(-0.79)	(-1.68)	(-2.10)	(9.63)	(5.41)	(-2.39)	
Vol80	0.013*	0.024**	0.016	0.567***	0.368***	0.390***	0.738
	(1.72)	(2.22)	(1.32)	(8.09)	(4.66)	(4.59)	
Vol90	0.020**	0.014	0.005	0.528***	0.347***	0.322***	0.670
	(2.35)	(1.24)	(0.35)	(7.34)	(3.91)	(3.75)	

The results indicate that after controlling for the Fama-French factors, the explanatory power of the sentiment indices is greatly diminished. Hence, it appears that the small and market-to-book factors grasp same components as the sentiment. However, sentiment remains statistically significant at some cases. In general, the regression models are a much better fit once the Fama-French factors are included. The adjusted coefficients of determination increase substantially and some of them approach 1 closely. It is interesting to note that the signs of the returns may have turned from positive to negative or vice versa compared to the previous regressions.

Sentiment level is not a very relevant measure when the time period is one month. Only the growth and large growth experience statistical significance. When sentiment is extremely low the return of the next month is 0.7 percent lower than during moderate sentiment levels. When the time frame is extended to three months, small cap, large value, Vol20 and Vol80 portfolios are statistically significant. As the raw sentiment proxy is utilized as an independent variable and sentiment is extremely low, the subsequent three month returns of Small cap and Large Value are 1.5 and 4.1 percent larger than normally. When the orthogonalized sentiment proxy is used, the returns of Vol20 decline by 5.3 when the sentiment is high and 5.9 percent when the sentiment is low. Vol80 portfolio earns a return that is 2.4 percent higher than normally when sentiment is high. Small cap and Large Value are not statistically significant in this case. In summary, the effects of sentiment are not very powerful. Only the DS^H on Vol80 and DS^L on Vol20 are statistically significant at the 5 percent level. SMB and HML seem to be highly statistically significant in almost every equation. Especially HML does seem to have higher explanatory power when the dummy variables are utilized instead of the sentiment levels index.

The behavioral finance theory propose that sentiment prone stocks should fare better than bond like stocks in the future when sentiment is low and do worse when sentiment is high. These patterns are due to over and undervaluation and the habit of people to be swayed by their emotions. In addition, sentiment should affect the most and least volatile groups of stocks the most.

Although the effects might not be particularly strong, it is pleasing to notice that sentiment does have the biggest impact on the most and least volatile stocks. The effects of sentiment are as predicted during high levels of sentiment when the time period is three months. The least volatile quintile earns higher than usual returns and the returns are much higher compared to those of the most volatile quintile. However, the effects are not quite as expected when sentiment is low. Vol20 portfolio continues to earn lower returns than normally. On the other hand, the Large Value portfolio earns higher than usual returns in both three month regressions when sentiment is low and the same applies for the small cap stocks. The common denominator between these two might be the trades of retail investors. During low levels of sentiment, the stocks that are held by retail investors get undervalued and the future returns are higher. In addition, the effects of extreme sentiment levels on returns are similar as our previous findings. The least volatile stocks tend to earn higher returns than normally when sentiment is high or has increased substantially. Thus, it seems that the effects of sentiment on these stocks are a bit lagged as investors reallocate their wealth towards safer stocks.

It can be detected from the results that sentiment has more impact on returns during extremely low levels of sentiment. This can be attributed to investor psychology as losses feel much worse than winnings and bad news has been noticed to affect the decision making of investors more (Barberis et al. 2001; Conrad et al. 2002).

The findings in this thesis are interesting in many aspects. Sentiment is clearly related to the current returns of stocks and has explanatory power on the future returns of small capitalization stocks, least volatile stocks and on the Large Value portfolio even after controlling for Fama-French factors. When there is a strong change in investor sentiment or when the sentiment is at extreme levels, the returns of these aforementioned stocks are mostly affected. It is unclear whether there is a common denominator among the effects on these specific stocks and if that denominator is the trades of private investors.

Nevertheless, sentiment cannot be regarded as an entirely futile variable and some of the findings are at least partially in line with previous findings of Baker and Wurgler (2006; 2007). Generally safer stocks yielded a higher return when sentiment increased and large growth stocks were in a negative relationship with sentiment. However, the effects among small capitalization stocks were not as proposed by the behavioral finance theory. In addition, sentiment did seem to have the most impact on speculative stocks and less speculative stocks when sentiment was extreme levels and sentiment had different cross sectional effects among stocks. Hypotheses 1 and 2 could be confirmed. Hypothesis 3 could not be validated as the results were a bit mixed. The speculative stocks were not generally overvalued, although the Vol20 portfolio did yield a larger negative return during extreme sentiment. After additional risk factors were included and the macroeconomic foundations were taken into account no clear patterns of over-valuation could be detected. However, some signs of prior undervaluation could be observed in the case of the least volatile stocks during high sentiment and Small Cap and Large Value portfolios when the raw sentiment index was low. The findings were also partially in line with Finter et al. (2010) who found out that sentiment was mostly related to undervaluation of stocks and that the effects were mainly driven by notable pessimism.

In conclusion, although sentiment seems to be related to both current and future returns of especially specific stocks, it can hardly be used in order to create a systematically profitable trading strategy. However, gaining a better insight of the behavior of market participants and the market mechanisms is achievable as the effects of sentiment are far from a resolved issue. Thus, more research on the matter is required. The next chapter summarizes and concludes the thesis and the research and presents new research topics.

5 SUMMARY AND CONCLUSIONS

This thesis has tried to provide insight to the behavioral element of market participants and search for answers regarding whether investor sentiment has an impact on stock prices in the small stock market of Finland. The traditional finance does not recognize investor sentiment as a factor that would influence stock returns. Market participants in general are seen to be rational and even if not everyone is rational, their trades will offset each other or the rational traders will take advantage of the situation and the stock price returns to its correct level. One cannot gain abnormal returns without taking additional risk.

Behavioral finance suggests that many mental and psychological factors affect the decision making of a person. Behavioral finance provides a larger social science perspective by incorporating sociology and psychology in finance. The two main characteristics that enable the effects of investor sentiment according to the behavioral framework are noise traders and limits of arbitrage. Noise traders are uninformed investors that base their decision mainly on irrelevant information and are prone to be swayed by their emotions and are most affected by investor sentiment. They are usually recognized as retail investors, but also institutional and larger investors have been claimed to act in an irrational manner at times. It is also expected that the trades of noise traders can be highly correlated as people have been proven to make similar types of biased decisions.

De Long et. al (1990) suggest that noise trading creates an additional risk element which should be accounted for and securities that are subject to greater noise trader risk tend to exhibit greater volatility and mean reversion. Limits of arbitrage mean that rational investors are unable and unwilling to take extensive positions against noise traders due to noise trader risk. One of the greatest risks for a rational arbitrageur is that the noise traders' beliefs will not revert to the mean for a long time and might become even more extreme. The investment horizon of rational arbitrageurs is usually quite short as the funds they are operating with are rarely entirely their own. Short selling can also incur additional psychological and sociological costs or it might even be restricted or forbidden in a country.

Sentiment drives the relative demand for speculative investments, thus causing cross-sectional effects among stocks despite arbitrage forces. The characteristic that makes some stocks more speculative than others is due to the difficulty and subjectivity of determining their true intrinsic values. Speculative stocks are more difficult to value and to arbitrage because of the subjective attributions of the stocks as there is greater uncertainty of their future. Uncertainty means that the effect of many psychological factors such as overconfidence, representativeness and conservatism is more emphasized.

Psychological and sociological factors are considered to be the cause of the correlated irrational decisions among people. Investors are usually overconfident about their skills and use simplified heuristics in order to make decisions. Emotions have been recognized to be relevant in decision making which might lead to overreacting from time to time. People rarely make investment decisions in a vacuum and the media and other people around us influence our decisions. Herding is a pivotal element in the propagation of a stock market bubble and investor sentiment and herding are closely interrelated. Sentiment extremities can cause bubbles and crashes and investor sentiment can even impact the behavior of corporations.

In order to test whether investor sentiment has an impact on the Finnish stock market and if the effects were as claimed by the behavioral finance framework, an investor sentiment index was composed by Principal Components Analysis. The investor sentiment index was comprised of five sentiment proxies that were: turnover, closed end fund discount, VDAX, consumer confidence and industrial confidence. In addition another index was formed that accounted for macroeconomic factors in order to create a measure that reflected the mood of the market that could not be warranted by macroeconomic fundamentals. As the raw sentiment index appeared to have a unit root, the indices were first differenced in order to make the series stationary.

Sentiment did seem to be in relation with the current returns of stocks. Especially the orthogonalized sentiment index was highly statistically correlated with portfolio returns. The regression analysis indicated that sentiment was clearly related to the current returns of various portfolios. Smaller stocks and growth stocks were most influenced by changes in investor sentiment. The SENT index had more explanatory power due to macroeconomic conditions. Although it was assessed whether sentiment is related to the current returns of portfolios sorted by size and investment style or to the returns of portfolios formed by historical volatility, the main interest was to discover whether there was sentiment related mispricing correction. The behavioral theory posited that during high levels of sentiment, the prices of speculative stocks would be overvalued and vice versa during low levels of sentiment. This information could be investigated by assessing the relationship between sentiment and future returns of stocks.

Changes in investor sentiment did have a quite high explanatory power on especially small stocks and least volatile stocks. Sentiment had a highly statistically significant relationship with the returns of these specific stocks. The irrational component of sentiment, the ADJSENT index had a clearly larger influence on the returns of small stocks. This indicates that smaller stocks are indeed more influenced by irrational investors. Growth, Large Growth and LsVol portfolios had a negative relationship with sentiment. Whereas, smaller stocks, value stocks and less volatile stocks had a positive relationship with sentiment. When the forecast horizon was extended to three months, the relationship between sentiment and returns stayed rather similar. No clear return reversal patterns were detected even though the effect of sentiment seemed to wear a bit down. After controlling for additional known stock market risk factors which are known as the Fama-French factors, the explanatory power of sentiment was significantly diminished. However, the effects of sentiment on the least volatile portfolios and small capitalization stocks remained highly statistically significant. A single unit change in irrational investor sentiment increased the subsequent month return of small growth portfolio by 1 percent and the portfolio of least volatile quintile of stocks by 1.1 percent. It is necessary to remember that sentiment changes of this magnitude are quite rare and a change of half a unit is a lot more frequent. Nevertheless, these findings indicate that sentiment is a relevant measure. When the time frame was extended to three months, sentiment had a weak influence on the returns of stocks after controlling for Fama-French factors.

It was also assessed whether sentiment had a larger impact during extreme levels of sentiment. It was evident that sentiment indeed did have a bigger impact on larger variety of stocks. The large growth stocks seemed to outperform other stocks when sentiment hit extreme levels and value stocks, small stocks and less volatile stocks seemed to underperform. The irrational component of sentiment had a bigger impact during extreme levels of sentiment, which is not surprising, as when uncertainty increases the effects of market psychology increases. Sentiment had more impact when sentiment was extremely low during the one month time period and more impact on stocks when the time period was extended to three months. This indicates that the effects of sentiment are more rapid when there is high pessimism and fear in the market.

When the Fama-French factors were included in the equations the effects of sentiment weakened and it appeared that the small and book-to-market factors grasp same components as the sentiment. Extreme levels of sentiment seemed to have more impact during low levels of sentiment. The latter is in line with the notion that people react to bad news more powerfully than to good news and that the fear of losses might drive the trades. There has also been evidence that people are far more upset of losses than pleased by winnings of equivalent amounts (Conrad et al. 2002; Shiller 2003).

The Small Cap and Large Value portfolio yielded 1.5 and 4.1 percent larger returns than normally during the next three months when sentiment was extremely low when the raw sentiment index was used. The least volatile quintile portfolio Vol80 experienced an increase in returns and the portfolio consisting of most volatile stocks depreciated in value when sentiment was extremely high. These findings were partially in line with the behavioral finance theory and the findings in the U.S stock market. However, when sentiment was low the returns of the most volatile portfolio did not yield higher returns than normally, but lower returns which was not expected. Sentiment had the biggest impact on the least and most volatile portfolios and the effects were dissimilar among different sorts of stocks.

The hypotheses 1 and 2 could be confirmed but the hypotheses 3 could not be clearly confirmed. It can be summarized that sentiment did have an influence on especially small stocks and least volatile stocks when the sentiment change was sufficiently forceful or when the sentiment prevailing sentiment was extreme. The effects of sentiment on stock returns related to prior undervaluation of the aforementioned stocks or delayed demand of these stocks. Sentiment prone stocks did not exhibit patterns of overvaluation, which is in line with the findings of Finter et al. (2010). In addition, sentiment had a larger impact on stocks when the prevailing sentiment was extremely pessimistic. The findings are parallel to both behavioral and traditional views of finance. Sentiment did have an impact on returns of stocks even after accounting for macroeconomic conditions and other stock market risk factors. However, the relation to future returns was quite weak and generating a profitable investment strategy based on sentiment seems quite difficult.

More research into the matter is necessary. It could be investigated whether sentiment has a larger impact on returns when the time period is extended or reduced to be weeks. Another interesting research topic could be how sentiment affects the average earnings announcement returns of different sorts of stocks. It could be tested if sentiment is related to false earnings expectations around earnings announcements and if there was mispricing correction among stocks when sentiment is particularly high or low.

6 REFERENCES

Ackert, L. – Church, K. – Deaves, R. (2003) Emotion and financial markets. *Federal Reserve Bank of Atlanta Economic Review*, Q2/2003, 33-41.

Akerlof, G. – Shiller, R. (2009) How animal spirits destabilize economies. *Mckinsey Quarterly*, No. 3, 126-135.

Allen, W. – Evans, D. (2005) Bidding and overconfidence in experimental financial markets. *Journal of Behavioral Finance*, Vol. 6, No. 3, 108-120.

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

Baker, M. –Wurgler, J. (2002) Market timing and capital structure. *The Journal of Finance*, Vol. 57, No. 1, 1-32.

Baker, M. – Wurgler, J. (2004) A catering theory of dividends. *The Journal of Finance*,Vol. 59, No. 3, 1125-1165.

Baker, M. – Wurgler, J. (2006) Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance*, Vol. 61, No. 4, 1645-1680.

Baker, M. –Wurgler, J. (2007) Investor sentiment in the stock market. *The Journal of Economic Perspectives*, Vol. 21, No. 2, 129-151.

Bandopadhyaya, A. (2006) *Measures of Investor Sentiment: Who wins the horse race?* Financial services forum, University of Massachusetts.

Barber, B. – Odean, T. (2001) Boys will be boys: Gender, Overconfidence and Common stock investment. *The Quarterly Journal of Economics,* February 2001, 261-293.

Barber, B. – Odean, T. (2006) All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. Forthcoming in: *The Review of Financial Studies.*

Barber, B. – Odean, T. – Zhu, N. (2009) Systematic Noise. *Journal of Financial Markets* Vol. 12, 547-569.

Barberis, N. – Shleifer, A – Vishny, R. (1998) A model of investor sentiment. *Journal of Financial Economic,*Vol. 49, 307-343.

Barberis, N. – Huang, M. – Santos, T. (2001) Prospect Theory and Asset Prices. *Quarterly Journal of Economics*, Vol. 116, No.1, 1-53.

Bazerman, M. (2006) *Judgment in managerial decision making*. John Wiley & Sons Inc, Hoboken.

Berglund, T. – Wahlroos, B. – Örnmark, A. (1983) The Weak-form Efficiency of theFinnish and Scandinavian Stock Exchanges. *Scandinavian Journal of Economics*, Vol. 85, No. 4, 521-530.

Bergman, N. – Roychowdhury, S. (2008) Investor sentiment, expectations and corporate disclosure. *Journal of Accounting Research*, Vol. 46, No. 5, 1057-1083.

Biais, B. – Weber, M.(2008) Hindsight bias, risk perception and investment performance. To appear in: *Management Science.*

Bitmead, A. – Durand, R. – Ng, H. (2004) Bubblelepsy: The Behavioral offspring of the internet stock phenomenon. *The Journal of Behavioral Finance*, Vol. 5, No. 3, 154-169.

Black, F. (1986) Noise. *The Journal of Finance*, Vol. 41, No. 3, 529-543.

Booth, G. – Kallunki, J. – Martikainen, T. (1999) Earnings news and the behavior of large and small traders in the Finnish stock market. *Applied Economic Letters* , Vol. 6, 771-774.

Brooks, C. (2002) *Introductory econometrics for finance*. Cambridge University Press, Great Britain.

Brown, G. – Cliff, M. (2004) Investor sentiment and the near term stock market. *Journal of empirical finance,* No. 11, 1-27.

Brown, G. – Cliff, M. (2005) Investor sentiment and asset valuation. *Journal of Business*, Vol. 78, No. 2, 405-440.

Brunnermeier, M. – Nagel, S. (2004) Hedge funds and the technology bubble. *Journal of Finance,*Vol. 59, No. 5, 2013-2040.

Canbas, S. (2009) Investor sentiment and stock returns: Evidence from Turkey. *Emerging Markets Finance and Trade,* Vol. 45, No. 4, 36-52.

Caparrelli, F. – D'Arcangelis, A. – Cassuto, A. (2004) Herding in the Italian stock market: A case of behavioral finance. *The Journal of Behavioral Finance*, Vol. 5, No. 4, 222-230.

Chan, K. – Hameed, A. – Tong, W. (2000) Profitability of momentum strategies in the international equity markets. *Journal of Financial and Quantitative Analysis*, Vol. 35, No. 2, 153-172.

Conrad, J. – Cornell, B. – Landsman, W. (2002) When is Bad News Really Bad News? *The Journal of Finance*, Vol. 57, No. 6, 2507-2532.

Cooper, M. – Gutierrez, R. – Hameed,A. (2004) Market states and momentum. *Journal of Finance*, Vol. 59, 1345-1365.

Daniel, K. – Hirshleifer, D. – Subrahmanyam, A. (1998) Investor Psychology and Security Market Under-and Over-Reactions. *Journal of Finance*, Vol. 53, No. 6, 1839-1885.

Davis, N. (2003) *The Triumph of Contrarian Investing*. McGraw-Hill, USA.

De Jong, A. – Rosenthal, L. – Van Dijk, M. (2008) The risk and return of arbitrage in dual-listed companies. Forthcoming in *Review of Finance.*

De Long, J. – Shleifer, A. – Summers, L. – Waldman, R. (1990) Noise trader risk in financial markets. *The Journal of Political Economy*, Vol. 98, No. 4, 703-738.

Deutsche Börse (2005) VDAX New presentation, <http://deutscheboerse.com/dbag/dispatch/de/binary/gdb_content_pool/imported_files/public_files/10_downloads/50_informations_services/30_Indices_Index_Licensing/21_guidelines/40_volatility_indices/VDAXNEW_Presentation.pdf>, retrieved from the Internet 20.06.2010.

Dittrich, D. – Güth, W – Maciejovsky, B. (2005) Overconfidence in investment decisions: An experimental approach. *The European Journal of Finance*, Vol. 11, No. 6, 471-491.

Dougherty, C. (2002) *Introduction to Econometrics.* Oxford University Press, New York.

Drakos, K. (2009) Terrorism activity, investor sentiment and stock returns. *Review of financial economics*, Vol. 19, No. 3, 128-135.

EC DGFin (2007) The joint harmonized EU programme of business and consumer surveys.

Edmans, A. – Garcia, D. – Norli, O. (2007) Sports Sentiment and Stock Returns. *The Journal of Finance*, Vol. 62, No. 4, 1967-1998.

Ekholm, A. (2002) Essays on stock market reactions to new information. Doctoral dissertation, Svenska Handelshögskolan, Helsinki.

Eviews (2007) *Eviews 6 user guide.*

Fama, E. (1970) Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, Vol. 25, No. 2, 383-417.

Fama, E. – French, K. (1996) Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, Vol.51, No. 1, 55-84.

Fama,E. – French, K. (1993) Common risk factors in the returns on stocks and bonds *Journal of Financial Economics*, Vol. 33, 3-56.

Ferguson, N. (2008) *Rahan nousu*. Hakapaino, Helsinki.

Finter, P. – Niessen, A. – Ruenzi, S. (2010) *The Impact of Investor Sentiment on the German Stock Market*. Centre for Financial Research Working paper, University of Cologne.

Forbes, W. (2009) *Behavioural finance*. John Wiley & Sons, Cornwall.

Gemmill, G. – Thomas, D. (2002) Noise Trading, costly Arbitrage, and asset prices: Evidence from closed-end funds. *The Journal of Finance*, Vol. 57, No. 6, 2571- 2594.

Glaser, M. – Nöth, M. – Weber, M. (2003) *Behavioral finance*. In: Blackwell Handbook of Judgement and Decision Making, eds. Koehler, D. – Harvey, N. John Wiley&Sons inc., Hoboken.

Gombola, M. –Marciukaityte, D. (2007) Managerial overoptimism and the choice between debt and equity financing. *The Journal of Behavioral Finance*, Vol. 8, No. 4, 225-235.

Greene, W. (2008) *Econometric analysis*. Pearson Prentice Hall, New Jersey.

Grinblatt, M. –Keloharju, M. (2001) What makes investors trade? *The Journal of Finance*, Vol. 56, No. 2, 589-616.

Grinblatt, M. – Keloharju, M. (2009) Sensation Seeking, overconfidence and trading activity *The Journal of Finance*, Vol. 64, No. 2, 549-578.

Grinblatt, M. – Titman, S. – Wermers, R. (1995) Momentum investment strategies, portfolio performance and herding: a study of mutual fund behavior. *The American Economic Review*, Vol. 85, No. 5, 1088-1105.

Gujarati, D. (2003) *Basic Econometrics*. McGraw-Hill, New York

Hair, J. – Anderson, R. – Tatham, R. (1987) *Multivariate Data Analysis*. McMillan Publishing Company, New York.

Hanke, J. – Wichern, D. (2005) *Business forecasting*. Pearson Prentice Hall, New Jersey

Harris, R. (1995) *Cointegration analysis in econometric modelling*. Prentice Hall, Essex

Harvey, N. (2007) Use of heuristics: Insights from forecasting research. *Thinking & Reasoning*, Vol. 13, No. 1, 5 – 24.

Hassett, K (2002) *Bubbleology: The amazing science of stock market bubbles*. Crown Publishing Group, Westminster.

Heikkilä, Toni (1997) *Information content of cash dividend and earnings announcements: Evidence from the Finnish stock market*. Licenciate's Thesis, Turku School of Economics, Turku.

Hirshleifer, D. (2001) Investor psychology and asset pricing. *The Journal of Finance*, Vol. 56, No. 4, 1533-1597.

Hirshleifer, D. – Shumway, T. (2003) Good Day Sunshine: Stock Returns and the weather. *The Journal of Finance*, Vol. 58, No. 3, 1009-1032.

Hwang, S. – Salmon, R. (2004) Market stress and herding. *Journal of Empirical Finance*, Vol. 11, 585-616.

Imam, S. – Barker, R. –Clubb, C. (2008) The use of valuation models by UK investment analysts. *European Accounting Review*, Vol. 17, No. 3, 503-535.

Investopedia (2010) VIX – CBOE Volatility Index definition <http://www.investopedia.com/terms/v/vix.asp>, retrieved 02.09.2010.

Jokivuori, P. – Hietala, R. (2007) *Määrällisiä tarinoita: Monimuuttujamenetelmien käyttö ja tulkinta*. WSOY, Helsinki.

Kahneman, D. – Riepe, M. (1998) Aspects of investor psychology. *The Journal of Portfolio Management*, Summer 1998.

Kallunki, J. – Martikainen, M. – Martikainen, T. – Yli-Olli, P. (1997) The Finnish stock market: A survey of some empirical evidence and its practical relevance. *Liiketaloudellinen aikakauskirja*, Vol. 4, 474-495.

Kamstra, M. – Kramer, L. – Levi, M. (2003) Winterblues: A SAD stock market cycle. *The American Economic Review*, Vol. 93, No. 1, 324-343.

Kaplanski, G – Levy, H. (2010) Sentiment and stock prices: The case of aviation disasters. *The Journal of Financial Economics*, Vol. 95, 174–201.

Karhunen, J. – Keloharju, M. (2001) Shareownership in Finland. *Liiketaloudellinen aikakauskirja*, Vol. 2, 188-226.

Kelly, M. (1997) Do noise traders influence stock prices. *Journal of Money, Credit and Banking*, Vol. 29, No. 3, 351-363.

Kennedy, P. (1999) *A guide to econometrics*. Blackwell publishers, Massachusettes.

Knif, J. – Löflund, A. (1997) The pricing of Finnish stocks: A survey of some empirical evidence. *Liiketaloudellinen aikakauskirja*, Vol. 4, 496-515.

Korhonen, A. (1977) *Stock prices, information and the efficiency of the Finnish stock market : empirical tests.* Doctoral dissertation, Helsinki School of Economics, Helsinki.

Kumar, A. – Lee, C. (2006) Retail investor sentiment and return comovements. *The Journal of Finance*, Vol. 61, No. 5, 2451-2486.

Lamont, O. – Stein, J. (2006) Investor sentiment and corporate finance: Micro and macro. *The American Economic Review*, Vol. 96, No. 2, 147-151.

Lee, C. – Shleifer, A. – Thaler, R. (1991) Investor sentiment and the closed-end fund puzzle. *Journal of Finance*, Vol. 46, 75-109.

Lemmon, M. –Portniaguina, E. (2006) Consumer confidence and asset prices: Some empirical evidence. *The Review of Financial Studies*, Vol. 19, No. 4, 1499-1529.

Lin, C. – Rahman, H. – Yung, K. (2009) Investor sentiment and REIT returns. *The journal of real estate finance and economics*, Vol. 39, 450-471.

Ljunqvist, A. – Nanda,V. – Singh, R. (2003) *Hot markets, investor sentiment and IPO pricing*. NYU Stern School of Business Working paper, New York.

Lucey, B. – Dowling, M. (2005) The role of feelings in investor decision-making. *The Journal of Economic Surveys*, Vol. 19, No. 2, 211-237.

Malkiel, B. (2003) The efficient market hypothesis and its critics. *The Journal of Economic Perspectives*, Vol. 17, No. 1, 59-82.

Malkiel, B. (2005) Reflections on the efficient market hypothesis: 30 years later. *The Financial Review* Vol. 40, No. 1, 1-9.

Malkiel, B. (2007) *Sattumankauppaa Wall Streetillä*. Talentum, Helsinki.

Marsden, A. – Veeraraghavan, M. – Ye, M. (2008) Heuristics of representativeness, anchoring and adjustment, and leniency: Impact on earnings' forecasts by Australian analysts. *Quarterly Journal of Finance and Accounting*, Vol. 47, No. 2, 83-102.

Marttila, J. (2001) *Järki ja tunteet osakemarkkinoilla.* Edita Oyj, Helsinki.

Mehra, R. – Sah, R. (2002) Mood fluctuations, projection bias and volatility of equity Prices.*Journal of Economic Dynamics and Control*, Vol. 26, No. 5, 869-887.

Meklari (2010) Pörssi onkin tiedemiesten taivas, No. 1, 26-29.

Metsämuuronen, J. (2008) *Monimuuttujamenetelmien perusteet*. Gummerus, Jyväskylä.

Michaely, R. – Womack, K. (1999) Conflict of interest and the credibility of underwriter analyst recommendations. *The Review of Financial Studies*, Vol. 12, No. 4, 653-686.

Mishkin, F. (1998) *The economics of money, banking and financial markets*. Addison Wesley, Massachusetts.

MSCI Barra (2007) Index construction objectives and methodology for the MSCI global investable market value and growth indices.

MSCI Barra (2008) Index construction objectives, guiding principles and methodology for the MSCI global investable market indices.

MSCI Barra (2010) Index calculation methodology for the MSCI equity indices.

Neal, R. – Wheatley, S. (1998) Do measures of investor sentiment predict returns? *The Journal of Financial and Quantitative Analysis*, Vol. 33, No. 4, 523 -547.

Neomarkka pörssitiedotteet (2000) Webpage of Neomarkka <http://www.neomarkka.fi/tiedotteet/release.html>, retrieved 30.6.2010.

Nesvetailova, A. (2007) *Fragile finance*. Palgrave Macmillan, London.

Nikkinen, J. – Rothovius, T. – Sahlström, P. (2005) *Arvopaperisijoittaminen*. WSOY, Helsinki.

Nikkinen, J. – Sahlström, P. – Takko, K. – Äijö, J. (2009) Turn-of-the-month intra month anomalies and U.S macroeconomic news announcements on the thinly traded Finnish stock market. *International Journal of Economics and Finance*, Vol. 1, No. 2, 3-12.

Nikkinen, J. – Vähämaa, S. (2010). Terrorism and stock market sentiment. *Financial Review*, Vol. 45, 263-275.

Nofsinger, J. (2005) Social mood and financial economics. *Journal of Behavioral Finance*, Vol. 6, No. 3, 144-160.

Nofsinger, J. (2008) *The psychology of investing*. Pearson Prentice Hall, New Jersey.

Nyberg, P. – Vaihekoski, M. (2010) A new value-weighted total return index for the Finnish stock market. *Research in International Business and Finance*, Vol. 24, No. 3, 267-283.

Paterson, R. (2002) Buy on the rumor: Anticipatory affect and investor behavior. *Journal of Psychology and Financial Markets*, Vol. 4, No. 2, 218-226.

Pörssisäätiö (2007) Development of the Finnish stock market <http://www.porssisaatio.fi/en/articles/development-of-the-finnish-stock-market>, retrieved 20.4.2010.

Read, C. (2009) *The Fear Factor*. Palgrave McMillan, London.

Saario, S. (2007) *Saarion sijoituskirja: miten sijoitan pörssiosakkeisiin*. WSOY, Helsinki.

Schadéwitz, H. – Blevins, D. (1998) A preliminary analysis of the semi-strong efficiency of the Helsinki stock exchange. *International Journal of Management*, Vol. 15, No. 2.

Scharfstein, D. – Stein, J. (1990) Herd behavior and investment. *The American Economic Review,* Vol. 80, No. 3, 465-479.

Schmeling, M. (2007) Institutional and individual sentiment: Smart money and noise trader risk. *International Journal of Forecasting*, No. 23, 127-145.

Schmeling, M. (2009) Investor sentiment and stock returns: Some international Evidence. *Journal of empirical finance* No. 16, 394-408.

Schwarz, N. (2010) Feelings-as-information theory. To appear in Handbook of theories of social psychology, eds. Van Lange, P. – Kruglanski, A. – Higgins,T., Sage.

Sherman, S. – Cialdini, R – Schwartzman, D. – Reynolds, K. (1985) Imagining can heighten or lower the perceived likelihood of contracting a disease: The mediating effect of ease of imaginery. *Personality and social psychology bulletin*, Vol.11, 118-127.

Sheung, H. – Lu, Y. – Wei, Y. (2010) Causalities between sentiment indicators and stock market returns under different market scenarios. *The International Journal of Business and Finance Research*, Vol. 4, No. 1, 159-171.

Shiller, R. (2003) From efficient markets theory to behavioral finance. *The Journal of Economic Perspectives*, Vol. 17, No. 1, 83-104.

Shiller, R. (2005) *Irrational exuberance.* Princeton University Press, New Jersey.

Shiller, R. (2009) *The Subprime Solution*. Princeton University Press, New Jersey.

Shleifer, A. (2000) *Inefficient markets: an introduction to behavioral finance.* Oxford university press, Oxford.

Shleifer, A. – Summers, L. (1990) The noise trader approach to finance. *The Journal of Economic Perspectives*, Vol. 4, No. 2, 19-33.

Shleifer, A. – Vishny, R. (1997) The limits of arbitrage. *The Journal of Finance*, Vol. 52, No. 1, 35-55.

Shleifer, A. – Vishny, R. (2003) Stock market driven acquisitions. *Journal of Financial Economics*, Vol. 70, 295-311.

Simon, D. – Wiggins III, R. (2001) SP Futures returns and contrary sentiment Indicators. *The Journal of Futures Markets*, Vol. 21, No. 5, 447–462.

Slovic, P. – Finucane, M. – Peters, E. – MacGregor, D. (2004) Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk and rationality. *Risk Analysis*, Vol. 24, No. 2, 1-12.

Sornette, D. (2003) *Why stock markets crash*. Princeton University Press, New Jersey.

Tabachnick, B. – Fidell, L. (2007) *Using multivariate statistics*. Pearson, Boston.

Tay, N. (2009) Social Network Characteristics and The Evolution of Investor Sentiment. In : *Agent-based approaches in economic and social complex systems V*, eds. Terano, T. – Kita, H. – Takahashi, S. – Teguchi, H. ,Springer, Japan.

Tetlock, P. (2007) Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, Vol. 62, No. 3, 1139-1168.

Tokic, D. (2007) Noise trading and market efficiency: the role of passive investors. *Institutional Investor*, June 2007, Vol. 41, 37-44.

Trueman, B. (1988) A theory of noise trading in securities markets. *The Journal of Finance*, Vol. 43, No. 1, 83-95.

Tversky, A. – Kahneman, D. (1974) Judgment under uncertainty: Heuristics and biases. *Science*, Vol. 185, No. 4157, 1124-1131.

Vaihekoski, M. (1997) The Finnish stock market : Recent trends and important events. *Liiketaloudellinen aikakauskirja*, Vol. 4, 526-543.

Vaihekoski, M. (2004) *Rahoitusalan sovellukset ja Excel*. WSOY, Helsinki

Virtanen, I. – Yli-Olli, P. (1987) Forecasting stock market prices in a thin security Market. *Omega International Journal of Management*, Vol. 15, No. 2, 145-155.

Wang, F. (2001) Overconfidence, investor sentiment, and evolution. *Journal of Financial Intermediation*, Vol. 10, 138-170.

Wang, Y. – Keswani, A. – Taylor, S. (2006) The relationship between sentiment, returns and volatility. *International Journal of Forecasting*, Vol. 22, 109-123.

Womack, K. (1996) Do brokerage analysts’ recommendations have investment value? *The Journal of Finance*, Vol.59, No. 1, 251-278.

Wong, S – Lievano, R. (2009) Investor sentiment as intervention. *Academy of Accounting and Financial Studies Journal*, Vol. 13, No. 4, 55-66.

Wärneryd, K. (2001) *Stock market psychology*. Edward Elgar Publishing, Cheltenham.

Zhou, R. – Lai, R. (2009) Herding and information based trading. *The Journal of Empirical Finance*, Vol. 16, 388-393.

APPENDIX 1 : FORMULAS AND EQUATIONS

Equations:

1.1 Coefficient of determination and adjusted coefficient of determination

$$R_2 = \frac{\Sigma(\hat{Y}_i - \overline{Y})^2}{\Sigma(Y_i - \overline{Y})^2}$$

$$\overline{R}_2 = 1 - (1 - R_2)(\frac{N-1}{N-k-1})$$

R_2 is the square of the correlation coefficient between Y and its estimate. The regression sum of squares is divided by the total sum of squares. The adjusted R_2 accounts for the inclusion of explanatory variables *k* (Tabachnick&Fidell 2007, 153; Gujarati 2003, 84).

1.2 Kaiser-Mayer-Olkin test of sampling adequacy

$$KMO = \frac{\sum r^2}{\sum r^2 + \sum a^2}$$

Kaiser-Mayer-Olkin test is conducted to assess whether a Principal Components Analysis can be performed on a set of variables. It indicates the common variation among the variables. A value of over 0.6 is usually desirable. r denotes correlation and a partial correlation (<http://www.helsinki.fi/atk/tilasto/Spssjatko/faktori/faktori.html>).

1.3 Kolmogorov- Smirnov test

H_o= The distribution of the variable is not deviant from normal distribution
H_1= The distribution of the variable deviates from normal distribution

$$D = \max\left|F_0(x_i) - S_N(x_i)\right|$$

$F_0(x_i)$ denotes cumulative frequencies and $S_N(x_i)$ denotes the corresponding values of the normal distribution. D is compared to chart values for a given significance level (usually 5 percent level) in order to determine whether the variable follows a normal distribution. However, statistical programs give a specific p-value. If the value is close to zero the null hypothesis will be rejected and the variable is not seen to follow a normal distribution (<http://www.uku.fi/~mauranen/bis/bis7_doc.htm>)

1.4 Dickey-Fuller test stat

$$TS = \frac{\hat{\phi}}{se\,\hat{\phi}},$$

where the OLS estimator of ϕ is divided by the standard error of the estimate of ϕ. The hypotheses are as following:

$H_0 = \phi = 0$

$H_1 = \phi = 1$

The test distribution does not follow a standard t-test distribution, but rather a Dickey-Fuller distribution (Ben Vogelvang 2005, Econometrics Theory and Applications, 285, Pearson Prentice Hall, Essex).

APPENDIX 2: ADDITIONAL RESULTS

Table A1 Correlations between the 1st stage index and the sentiment proxies

This table presents the correlations between the first stage index of SENT and the sentiment proxies and their lagged values. The proxies that are most correlated with the first stage index are selected to form the SENT index. The index is formed by Principal Components Analysis.

	Turn	CEFD	VDAX	IND	CONS	LagTurn	LagCEFD	LagVDAX	LagIND	LagCONS
1st stage	0,729	-0,552	-0,75	0,907	0,854	0,697	-0,584	-0,741	0,898	0,824

Table A2 Principal Components Analysis results on SENT index

This table summarizes the results of the Principal Component Analysis on the chosen sentiment proxies: Industrial confidence, Consumer confidence, Turnover, VDAX and lagged Closed end fund discount. The communalities, component loadings and component score coefficients, the total amount of variance explained by the first principal component and the results from KMO and Bartlett's test are included in the table.

	Communalities	Loadings	Scores
Industrial	0.840	0.916	0.299
Consumer	0.732	0.856	0.279
Turnover	0.585	0.765	0.250
VDAX	0.559	-0.747	-0.244
CEFD	0.348	-0.590	-0.193
Variance explained		61.276	
KMO and Bartlett		0.788	0.000

Table A3 Kolmogorov-Smirnov test on macroeconomic variables

This table presents the results of the Kolmogorov-Smirnov test before variable transformations and after. The short term variable is squared and the unemployment variable is log differenced. If the significance value is below 0.05 it indicates that the variable is likely not to be normally distributed.

	Short term	unemployment	inflation	Ind.production
Mean	3.40	8.36	1.89	3.84
Std.Dev	1.04	1.12	1.26	5.53
Kolm. Z	1.497	2.017	1.072	0.786
Sig.	0.023	0.001	0.200	0.568
	After transformations			
	Short term	unemployment	inflation	Ind.production
Mean	12.62	-0.005	1.89	3.84
Std.Dev	7.24	0.118	1.26	5.53
Kolm. Z	1.339	1.021	1.072	0.786
Sig.	0.055	0.248	0.200	0.568

Table A4 Kolmogorov-Smirnov test on sentiment proxies

This table presents the results of the Kolmogorov-Smirnov test before variable transformations and after. A square root is taken from the VDAX variable. If the significance value is below 0.05 it indicates that the variable is likely not to be normally distributed.

	TURN	CEFD	VDAX	CONS	IND
Mean	1.12	0.29	23.61	14.08	3.73
Std.Dev	0.34	0.10	9.71	4.37	12.09
Kolm. Z	0.675	0.725	1.565	0.570	1.125
Sig.	0.752	0.670	0.015	0.901	0.159
	After transformations				
	TURN	CEFD	VDAX	CONS	IND
Mean	1.12	0.29	4.77	14.08	3.73
Std.Dev	0.34	0.10	0.92	4.37	12.09
Kolm. Z	0.675	0.725	1.179	0.570	1.125
Sig.	0.752	0.670	0.124	0.901	0.159

Table A5 Correlations of raw and orthogonalized sentiment components

This table correlates the raw sentiment proxies and orthogonalized sentiment proxies. The orthogonalized proxies have been regressed by industrial production growth, inflation, short term rate and unemployment rate change.

Raw sentiment proxy correlations

	Turnover	VDAX	Industrial	Consumer	CEFD
Turnover	1				
VDAX	-0.378	1			
Industrial	0.633	-0.612	1		
Consumer	0.630	-0.510	0.797	1	
CEFD	-0.299	0.434	-0.452	-0.291	1

Orthogonalized sentiment proxy correlations

	Turnover Res	VDAX Res	Industrial Res	Consumer Res	CEFD Res
Turnover Res	1				
VDAX Res	-0.322	1			
Industrial Res	0.551	-0.602	1		
Consumer Res	0.567	-0.446	0.694	1	
CEFD Res	-0.227	0.424	-0.436	-0.289	1

Table A6 Principal Components Analysis results on ADJSENT index

This table summarizes the results of the Principal Component Analysis on the residuals of chosen sentiment proxies: Industrial confidence, Consumer confidence, Turnover, VDAX and lagged Closed end fund discount. The communalities, component loadings and component score coefficients, the total amount of variance explained by the first principal component and the results from KMO and Bartlett's test are included in the table.

	Communalities	Loadings	Scores
Industrial	0.794	0.891	0.318
Consumer	0.661	0.813	0.290
Turnover	0.515	0.718	0.256
VDAX	0.557	-0.746	-0.266
CEFD	0.312	-0.523	-0.187
Variance explained	56.014		
KMO and Bartlett	0.768	0.000	