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

Forecasting power of investor sentiment: A sector analysis

Author: A. Herceg
Student number: 429195
Thesis supervisor: Dr. Jan Lemmen
Co-reader: Dr. Sjoerd van den Hauwe

Finish date: [August, 2017]

PREFACE AND ACKNOWLEDGEMENTS

This master thesis is a crown of my Financial Economics degree at Erasmus University. Since this is the first time that I have conducted and written a proper research paper, the entire thesis process was a big challenge. However, now that I am writing this preface, it means I am just a few more steps away from the submission. I would like to say that I feel extremely proud of myself and thankful for all the support and help I have received during the entire period. Therefore, I wish to thank a couple of people without whom this master thesis period would not have come to an end.
Firstly, I am tremendously grateful to my supervisor, dr. Jan Lemmen for his help and guidance during this period. His comments and suggestions on how to improve my writing, what is still missing in the analysis and things that I have overlooked made this thesis possible.

Special thanks to my boyfriend, who was always there when I needed any sort of help or dose of motivation, but also for helping me proofread the entire paper.

Last but not least, I am enormously thankful to my parents and family, who have provided me with the opportunity to study at Erasmus University and supported me in all means. Without them, this thesis nor master's degree would not be possible.

Overall, it was a process of "learning by doing".

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

ABSTRACT

By conducting this research, we aim to shed light on the forecasting power of investor sentiment. The focus is to explore sectors within the U.S. and their relationship with investor sentiment. Four measures of investor sentiment are applied, of which two are market based and two are survey based. The finding suggests that the in the out-of-sample period, sectors that contain relatively young and extreme growth firms are more influenced by investor sentiment. Especially over the 3- and 10-month forecast horizons, investor sentiment has a strong forecasting power. Comparing our out-of-sample model with the Random walk model, we note that in almost all cases the Random walk model outperforms.

Keywords:
Investor sentiment, sector returns, forecasting, Granger causality, Random Walk

JEL Classification:
G17, G41

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 Introduction

What is the power of investor sentiment? Does investor sentiment have forecasting power on stock returns? The same question has been puzzling researchers and financial practitioners for many years. Historical and recent market events, e.g. the dot.com bubble in 2000 and the global financial crisis in 2008 pointed out that the traditional view towards finance; where markets are efficient and investors are rational, failed to explain such events. Even though traditional asset pricing theories leave no role for investor sentiment, researchers led by theoretical assumptions and real-life events took different perspective on examining financial markets and stock returns. Significant contribution in the field of behavioral finance was documented by De Long et al. (1990) who show significant influence of investor sentiment on equilibrium prices. Their findings contributed to various attempts in examining the importance of the sentiment-return relationship. Furthermore, Baker and Wurgler (2006) find a strong response towards investor sentiment for stocks that are young, hard to value and therefore hard to arbitrage. Their findings indicate that individual and uninformed investors are dominant among such group of stocks. Different results were found by Brown and Cliff (2004, 2005) who argued that investor sentiment is a better predictor of stock returns over longer horizons, e.g. 1-year or longer. Given that the rational and experienced investors can correct for the mispricing in the short run, in the long-run the mispricing persists and it is not possible to exploit it. As well, the investor sentiment grows over time and its effect becomes more pronounced with time. Additionally, Brown and Cliff (2004) also observe that sentiment is caused by returns, rather than vice versa. The same conclusion about the direction of causality was reached by Salhin et al. (2016).

Existing literature reports mixing results especially on the effects of investor sentiment. Theoretically speaking, sentiment has been defined as an overall feeling towards the market given that investors sometimes feel either optimistic or pessimistic. Such a psychological state can lead to irrational investments. Assuming that all humans are exposed to feelings and biases, theoretically we can conclude that sentiment has predictive power. On the other hand, empirical evidence provides inconsistency in the results, mainly due to the fact that investor sentiment is an unobservable variable and thus hard to measure (Brown and Cliff, 2004; Lemmon and Portniaguina, 2006).

More recently, Baker and Wurgler (2006, 2007) stated the obvious: investor sentiment has significant influence on stock returns, which has been theoretically and empirically documented. Firstly, it has been well documented that investor sentiment has a significant impact on stock prices, but it is yet unclear what is the best measure of investor sentiment. Thus, in this paper we attempt to explore if different investor sentiment measures have stronger forecasting power on sector returns. We therefore look at market based and survey based investor sentiment measures. Among market based measures, the most attention was given to Baker and Wurgler's sentiment index (2006, 2007). Researchers believe that an index based on market variables is able to define investor sentiment and capture its effects. Lee at al. (1991) analyzed closed-end funds and changes in investor sentiment, concluding that any fluctuation in discounts of closed-end funds are a consequence of changes in investor sentiment. Stambaugh, Yu and Yuan (2012) used the investor sentiment index to examine asset pricing anomalies, while Lemmon and Portniaguina (2005) looked at what effect investor sentiment has on value and growth stocks. In all of their studies, sentiment, no matter how it was measured, had some effect on price movements. On the other hand, some authors point out at several drawbacks of market based measures. Particularly, Da et al. (2015) stated that market variables included in the construction of investor sentiment "have a disadvantage of being the equilibrium outcome of many economic forces other than investor sentiment". To account for spurious results from just one investor sentiment measure, we contribute to the literature by including additional sentiment measures.

Secondly, survey based measures gained in popularity as they became highly used in researches concerning investor sentiment. Among many measures, the most commonly used survey based measure is the Consumer Confidence Index constructed on consumer's responses from survey questions. The questions are formed to capture consumer's current and future expectations of economic performance. The Consumer Confidence Index has been used as a proxy for investor sentiment in various sentiment-return studies. They often document strong correlations between stock market returns and the consumer confidence index (Schmeling, 2009). Schmeling (2009) examined how consumer confidence predicts stock returns in 18 industrialized countries. He reports negative relationship between the investor sentiment, measured by the consumer confidence and stock returns. Moreover, Lemmon and Portniaguina (2006) used consumer confidence as a proxy for investor sentiment to analyze the influence of sentiment on stock market returns. They indicate that investor sentiment (based on consumer confidence) has a forecasting power when it comes to predicting returns. Given the current findings between sentiment-return relationship, both above-mentioned market and survey based sentiment measure will be examined. The aim of this paper is to get a clearer picture of which investor sentiment measures is better at forecasting sector stock returns.

Most of the literature examining investor sentiment and stock returns consider aggregate market or stock price movements in particular countries. Only few have looked at how particular industry or sector reacts to investor sentiment. Salhin et al. (2016) examine the impact of managerial sentiment on sector return in the U.K.. To carry out their research, they used consumer and business confidence measures as proxies for customer and managerial sentiment respectively. Their findings suggest that managerial sentiment significantly influences sector returns. Interestingly, an especially prone sentiment-return relationship is observed within manufacturing firms. However, there is no supporting evidence that consumer sentiment predicts or influences returns. Chen et al. (2014) analyse how local and global market sentiment influences industry returns in 11 Asian countries. While Kadilli (2015) examined how financial firm's stock returns can be predicted by investor sentiment in developed countries. Overall, the majority of the literature examines the predictive power of investor sentiment in the aggregate market, whereas sectors and industries are left unexplored. Assuming that different sectors consist of different types of stocks, these different types of stocks will attract different groups of investors who might or might not be fully informed and rational. Sectors based on more growth stocks might demonstrate more sensitivity towards investor sentiment and vice versa.

To assess which investor sentiment measure has the strongest forecasting power, we split our sample into an in-sample and out-of-sample period. The out-of-sample period creates forecasts for 1-, 3- and 10-month periods horizons for each of the four investor sentiment measures. Evaluating the forecasts, our results suggest that investor sentiment performs better over the longer forecast horizons, namely 3- and 10-months ahead. Particularly interesting results appear for sectors such as the IT, Industry and Consumer Staples, whose constituents are relatively young, unstable and speculative firms. Such results fall in line with the literature, especially with Baker and Wurgler's (2006, 2007) that young and growth firms are more prone to investor sentiment than the more mature and stable firms. In addition, to evaluate the validity of our results, we compare the results from the forecasting period with the Random walk model. The Random walk outperforms our forecasting models for almost all sectors.

Our contribution to the existing academic literature is extensive. Firstly, we employ both, market based and survey based investor sentiment measures to capture the strongest forecasting power. Most of the existing literature focuses on only one investor sentiment measure and does not take such a broad perspective. Secondly, the focus of this research is on the sectoral stock return rather than on the aggregate market. Most literature examining the sentiment-return relationship is concentrated on aggregate markets or on specific geographical regions. Thirdly, we apply an unbiased forecasting procedure using the Bayesian Information Criterion which has not been applied in this area of research.

The remaining of this paper is organized as follows. Chapter 2 covers the most important literature related to traditional finance and asset pricing. Here we discuss potential market anomalies based on which the behavioral finance theory emerged. Next, Chapter 3 covers the data description, summary statistics and some preliminary results of our analysis. In Chapter 4 we outline the main methodology used to carry out the research, which is split into an in-sample and out-of-sample period. Chapter 5 reports main results and findings, while Chapter 6 concludes.

CHAPTER 2 Literature review

The motivation for this research comes from the lack of empirical evidence which examines the influence of investor's sentiment on sectoral stock returns. Prevailing literature that explores the relationship between investor sentiment and stock returns focuses on the aggregate markets. Throughout the literature, researchers agree that investor sentiment has an influence on the market. However, its quantifiable impact is still unclear. To shed light on this topic, we seek to investigate what impact investor sentiment has on different sectors across the U.S.. First, we introduce and briefly touch upon the traditional finance theories, its beliefs and emergence of the Efficient Market Hypothesis (EMH hereafter). By going over the main concepts and findings, we identify market anomalies and possible implications to EMH and traditional finance beliefs. The inability to explain market anomalies during the financial crisis, triggered researchers to consider human/psychological factors of market participants as a potential explanation. This field is called behavioral finance. Here, we will discuss the evolution of behavioral finance, significance of investor sentiment, as well as its predicting power.

2.1 Central pillars of financial markets

How are asset prices determined? The same question has puzzled researchers for two centuries. Many scholars in the field of financial economics have searched for valid theoretical and empirical evidence. It is a fundamental question around which the finance field has been developing. For the purpose of this research, it is crucial to understand the evolution of financial theory, thus we start by introducing the building blocks of financial markets. Specifically, understanding the traditional finance approach towards risk and return allows for better grasping the concepts and beliefs of the behavioral finance view on asset pricing. Whereas behavioral finance is a relatively new field in financial economics, traditional finance beginnings date back in the past. Notable contributions to the modern economics were made by Adam Smith already in the 18th century, who is also known as the "father of modern economics". Smith (1766) set economic origins towards efficiency of financial markets. In short, he presented a theory based on market trends in consumption and production which states that markets, by nature, are inclined to become efficient. His theory of an invisible hand indicated that an "invisible hand" is used as a guide, that leads market forces of supply and demand towards the most optimal and efficient level.

Firstly, we will look at the theory on efficient market and rationality. As already mentioned, the evidence of efficient markets and stocks prices appeared in the finance literature in the 18th century (Smith). A breakthrough in the research on efficient markets was done by Samuelson (1965) who observed a random fluctuation in the market when all available information is present on the stock market. During the same year, famous U.S. economist Fama (1965) contributed to finance literature by interpreting the meaning of efficient market; "a market in which prices fully reflect available information". Followed by efficient markets explanation, Fama (1965b) conducted an empirical research to test if all available information is incorporated into asset prices. Furthermore, he showed that there are no trends among stock prices as they follow a Random walk. This suggests that prices reflect all available fundamental information that appear and that any predictability of returns should be impossible. Following the research by Fama and Samuelson, Roberts (1967) introduced the term "Efficient Market Hypothesis" (EMH hereafter), around which new empirical findings emerged related to the field of financial economics. Apart from forming EMH, Roberts introduced different forms of market efficiency based on available information i.e. the weak form which contains historical price information; and the strong form including all public and private information. The third form of market efficiency, namely semi-strong form, was added by Fama (1970) and it should contain all publicly available information. While the all three forms of market efficiency are present in the finance literature, the strong form is considered to be unrealistic since it should incorporate private information. Such private information falls under the "insider information" which is considered illegal. Therefore, the past and on-going research is focused mainly on the weak and semi-strong form of market efficiency, while the weak form drew special attention. By definition, the weak form incorporates only historical price information and prices today are independent of the prices tomorrow or two days ahead. Thus, the weak form of market efficiency is closely related to the Random walk hypothesis.

Having mentioned that Fama outlined in his findings that stock prices in an efficient market should follow the Random walk, researches have associated EMH with the theory of a Random walk. The theory of Random walk in a stock market was presented in early 1863 by economist Jules Augustin Frederic Regnault and a few decades later by Pearson (1905). Ever since the term was present in empirical finance research and it is used as a benchmark for forecasting models. Firstly, the idea behind the Random walk is, as its name indicates, random movements in stock prices that do not appear as a consequence of past stock behavior. Rather, today's information will be incorporated into today's price change, neglecting historical changes and potential predicting power. Since the idea of EMH is closely linked to the Random walk hypothesis, around the 1990s, new evidence shed light on investor's and market rationality, questioning the random walk of stock prices. Testing the true randomness of stock prices is beyond the objective of this research, however grasping the concept of the Random walk as a benchmark model is necessary.

Random behavior of stock prices intrigued researchers for decades. Findings in this field of research triggered the evolution of behavioral finance. Before introducing the theories and models of behavioral finance, we will first discuss the Random walk of stock prices. A recent study done by Borges (2010) looked at stock market indices of six developed European countries between the period 1993 until 2007. The aim of the study was to test the weak form of market efficiency, thus the Random walk hypothesis. Several different tests have been employed to carry out the analysis, among which, the serial correlation test and Augmented Dickey Fuller (ADF hereafter). Firstly, the serial correlation test is used to detect any relationship in the time series data over a certain period. Essentially, the test checks if stock prices, over the period, show signs of serial correlation. If there is a positive serial correlation between the past and future stock returns, it indicates that stock returns are not random but are part of some sequence. Thus, serial correlation test directly assesses if a time series data follow the Random walk. In the research carried out by Borges (2010), the serial correlation test indicates presence of a random walk in some countries. Interestingly, results for Greece, Spain and Portugal did not support the Random walk hypothesis which can be due to the fact that in those countries a high level of speculation was present. Furthermore, there is less market integrity which consequently reflects a poor financial market's performance. Another commonly used testing procedure used to check for the presence of a unit root is the Augmented Dickey-Fuller (ADF hereafter) test. The ADF test is an extended version of the original Dickey-Fuller test and incorporates lagged changes that should account for any potential serial correlation. To estimate the coefficient, a regression model such as the Ordinary Least Square (OLS hereafter) is conducted. Based on the obtained coefficient from the equation, we can either reject or accept the null hypothesis. The null hypothesis assesses if the time series data contains a unit-root, such that failure to reject the null hypothesis indicates that the time series data contains a unit-root and thus appears random. Borges (2010) conducted ADF tests on both daily and monthly data for all the six countries in the data set and found that the null hypothesis cannot be rejected in any country nor for any type of returns (daily and monthly). Such results fall in line with the Random walk hypothesis. Although, the ADF test does not reject the null hypothesis on all levels, such results should be taken with caution. Meaning, the ADF test are normally supported and compared with other popular unit root tests, such as the Kwiatkowski-Phillips-Schmidt-Shin test (hereafter KPSS) and Phillips-Perron test (hereafter PP), to see if each method arrives at the same conclusion The ADF and PP test the null hypothesis that the times series data is non-stationarity. On the other hand, the KPSS test checks the null hypothesis of stationarity in the data. In our research, we complement the ADF test with the two popular models, namely KPSS and PP test which should indicate similar results.

Apart from Borges (2010), Gan et al. (2005) focused their research on testing the market efficiency in the U.S., Japan, New Zeeland and Australia after a liberalization period that had influenced the economic situation in these countries. Similar to Borges (2010), they conducted the unit root test to check the time series data for non-stationary. In their research, apart from using the ADF test, they included the PP test which can or cannot confirm the results from the ADF test. By using additional unit root tests this allows for more precise and accurate results, especially if both tests confirm the same results. Thus, Gan et al.'s (2005) research confirmed the presence of weak form of market efficiency in all countries from the sample by conducting the ADF and PP test. Their findings indicate that markets are efficient and past stock returns do not influence future stock returns, nor is the market predictable, which falls in line with the Random walk hypothesis. Lee et al. (2000) used unit root and variance ratio tests to examine the presence of the Random walk on French futures and options market. Evidence of the Random walk was found in both markets, which falls in line with the finding in Borges (2010) and Gan et al. (2005).

Moreover, not all papers testing the movements in stock prices reached the same conclusion.
Lo and Mackinlay (1988) apply the variance ratio test to examine the variance of the returns over different intervals in the data set. If the variance of the q period difference is the equivalent to q times the variance of the one period difference, then the data follows a Random walk. By testing the weekly stock returns, they observed outperformance of the Random walk versus their model. Meaning, their full and sub-sample of stock prices do not follow the Random walk

but appear to have a predictable component in their movements. Later in the research, Lo and Mackinlay's variance ratio test has received a lot of attention among researchers for testing the Random walk of stock prices. However, most of the papers do not find supporting evidence for the weak form of market efficiency nor for the Random walk of stock prices.

In addition, DeBondt and Thaler (1990) took a different approach in examining the predictability of stock returns. Already in 1990, they assumed that departure from EMH comes from irrationality and human biases. In their research, they assessed stocks with high and low long-term past returns, where they aim to capture the so called "long-term reversal" of stock returns. This means, that stocks which in the past had experienced long periods of high returns, in the future will consequently have lower returns; and vice versa. Such movements in stock returns come as a consequence of investor's overreaction. Based on DeBondt and Thaler's experiment, people tend to invest based on their feelings and perceptions towards the market. Their findings suggest that stock movements exhibit a predictable component and thus fail to follow the Random walk.

With the EMH grows in popularity, the number of studies in this field expanded, among which many questioned the validity and true efficiency of the theory. Among many researchers at that time, it is worth mentioning Grossman and Stiglitz's (1980) empirical research in which they point out that cost of information can offset return on investment. Investors would not be inclined to invest, leading to inefficiencies in the market. It was Shiller (1981) who was one of the first to oppose validity of efficient market by introducing factors other than fundamentals that could influence stock prices. In his research, he questioned investor's rationality on future dividends and stock prices. His findings indicate that stock prices fluctuation, in relation to discounted future dividend has a stronger effect motivated by actions of irrational investor that are not eliminated. Following Shiller's research, Black (1986) contributed to the finance literature by introducing the term "noise" associated with investors and trading. In his research, he argues that stock price fluctuation is a result of an unexpected noise in the market that as a consequence affects investor's return.

In summary, there are evident indications that traditional theories cannot explain certain price movements and market anomalies. Those discoveries and divergence among researchers gave rise to another school of thought. Namely, behavioral finance that examines an investor's behavior and psychological patterns, questioning the overall rationality in the market. It is important for this research to introduce the concept of efficient market and the classical approach towards risk and return.

2.1.1 Asset pricing and challenges

Among the traditional asset pricing model, the most famous one is the Capital Asset Pricing Model (CAPM hereafter) introduced by Sharpe (1964), Lintner (1965) and Black (1972). The CAPM nor any other traditional pricing model incorporates the role of investor sentiment in the prediction. In short, the idea of CAPM is to explain the relationship between risk and required rate of return, such that the expected future return is influenced only by systematic risk which cannot be diversified away. Because of the assumption that markets are efficient and investors are rational, idiosyncratic risk which is a firm specific risk, can easily be diversified and is not included in the model. The equation (1) explains the CAPM setting:

$$R_{jt} - R_{rft} = \alpha_j + \beta_j \left(R_m - R_{rf}\right) + \varepsilon_{jt} \tag{1}$$

Where $R_{jt} - R_{rft}$ represents the excess return of an asset j at time t, $\left(R_m - R_f\right)$ is the market risk premium and ε_{jt} is a noise term for asset j at time t. In this case, β_j represents a market or systematic risk. Since the rationale behind the CAPM lays in efficient market portfolios, where investors are seen as risk averse and any exposure to higher risk should be reflected in the return. In equation (1), such change in stock returns is justified by factor β_j. Overall, various studies have identified drawbacks of the CAPM. Among many, Roll and Ross (1980) point out on the existence of other risk factors other than market risk.

Thus, Fama and French introduced a new so called Fama-French Three Factor Model (hereafter 3F model) which builds upon the already existing CAPM. The idea behind the 3F model lays in adding two additional factors that should explain asset returns, namely size and value factors. Size factors in the 3F model should capture the size effect, which refers to better performance of small cap stocks than the large cap stocks. To capture the effect, Fama and French used the portfolio of small and big stocks to calculate the average return by differencing the average returns of the three small and three big portfolios (so called Small Minus Big or SMB factor).

On the other side, value factor should capture the value effect which corresponds to outperformance of the value stocks over the growth stocks. By constructing the two portfolios, to get a value factor one should take the difference between the average of the two value and two growth portfolios (so called High Minus Low, HML factor). Although the idea behind the 3F is to better explain the asset returns by including additional factors, the model received a lot of criticism. Evidence in the literature cannot agree if those size and value factors are true risk factors, or if they are a consequence of a psychological phenomenon or irrationality. Moreover, it is still questionable what is the correct way to construct portfolios based on the size and value. Even though testing the 3F model is beyond the scope of this research, the model provides potential behavior implications that will be discussed later in this paper. Fama and French (1996), Jegadeesh and Titman (1993) and Carhart (1997) among many, extended the CAPM framework and included other factors next to β to account for additional risk which could potentially explain the excess return of an asset j at time t. By adding additional factors, researchers suggest that the existing models did not perform as expected. But instead, extending the already existing models, by including different factors that should better explain the excess returns, signals the need for a change in the traditional assumptions on the asset pricing.

However, none of the models took investor sentiment into consideration. The biggest trigger for a growing amount of contradictions towards EMH and CAPM were historical events that did not fit into the traditional framework and expectations. Abnormal deviations on the stock market such as the bubble in the 1970's, the Black Monday on 1987, the Dot.com bubble in the year 2000 and recent global financial crisis in 2008 are concrete evidence of the model failure that could not be explained by the fundamentals. In particular, behavioral finance tries to explain such market anomalies by taking a psychological perspective on investor's beliefs and expectations. For instance, in the recent study by Ho and Hung (2009) they constructed a conditional CAPM based on the three-factor model and importantly, included the investor sentiment proxy that is incorporated in the beta. Moreover, they compared their conditional CAPM with the traditional CAPM of Sharpe, Lintner and Black to determine which model can better explains fluctuation in future stock returns. Interestingly, they observe better performance of their newly constructed conditional CAPM in describing the stock prices when investor sentiment is included in the model.

In summary, grasping the traditional finance and asset pricing models, helps with the understanding of the financial markets mechanism and their approach towards risk and return. Even though the literature reports many contradictions and limitations of the CAPM, and some of the researchers call it a “half true” model, it is still present in the financial literature. It was those shortcomings of the CAPM and EMH that were driving the further research that later on emerged into behavioral finance. Based on the observable market anomalies that could not be explained by the existing asset pricing models, it became evident that investors are exposed to other risk other than just market risk. Such findings have lead research into more human factors that could influence the decisions and trading activities.

2.2 Behavioral Finance development

Market anomalies discovered in the EMH led researchers towards a new view on the predictability of asset return. Historical finance literature does not tell us much about behavioral finance, but only after the 1980’s the new era of beliefs emerged, questioning the assumptions of the EMH. Prior literature supporting the EMH states that in the process of asset pricing, noise traders are not taken into account, since rational arbitrageurs eliminate the mispricing and prices will reflect the fundamental value (Fama 1965). This indicates that asset prices are not affected by investor’s irrationality and a role of noise traders does not exist in the EMH. However, research by De Long, Shleifer, Summers and Waldmann (1990, hereafter, DSSW) and Lee, Shleifer and Thaler (1991) suggests that noise traders have significant influence on asset prices and cause mispricing in the market, uncovering a possible extension to EMH and classical finance. It is evident that the rationale behind the behavioral finance lays in the inefficient markets due to investor sentiment and the noise, leading to biases and irrational trading. In other words, behavioral finance takes a psychological perspective on understanding investor’s decisions and determination of asset prices. To do so, behavioral finance theory introduces human biases and sentiment as a potential explanation for market anomalies identified by earlier traditional theories.

Baker and Nofsinger (2010) state that the “father” of the behavioral finance theory is R. Thaler, whose discoveries shaped the theory of behavioral finance. Later, Thaler, together with Barberis (2003) proposed the concept of behavioral finance to be split into two pillars; limits to arbitrage and psychology. Limits to arbitrage, on the one hand, in theory hold, but Thaler and Barberis, among others argue that, in reality, arbitrage strategies are costly and risky. Due to the uncertainty, rational investors might not be able to exploit the mispricing. On the contrary, psychological elements such as heuristics and biases, could affect investors judgement of the market and his or her investing behavior. That is why the behavioral finance theory is focusing on beliefs and preferences of market participants. Investor's beliefs, are identified in behavior finance as biases among which the most common are over and under confidence, optimism or pessimism, and representativeness. Among the psychological elements, the most important factor is investor sentiment. By the definition, it is an overall feeling of an investor towards the market which can be influenced by feelings, expectations or noisy information.

Regarding preferences, the prospect theory defined by Kahneman and Tversky (1979) explains investor decision making when there the risk is involved in the process. In short, prospect theory explains what value gains and losses have, for example stocks that pay high returns versus stocks on which an investor is losing money. The theory is illustrated by the value function, which represents steep and convex line in the loss area. Such steep slope indicates that losses are perceived as more painful for an investor and it leaves a stronger impact than gains. Gains on the other hand, are represented by a concave value function, which implies investors are risk seeking. Meaning, gains are followed by positive feelings and enthusiasm, which allure one to invest even more.

2.2.1 Investor Sentiment

Investor sentiment by definition, is a feeling or expectation of an investor based on which a decision is made in a financial market. Baker and Wurgler (2006) construct a simple but precise definition of investor sentiment as "the propensity to speculate". Their definition of sentiment was a driving force for many researches, who looked at stock prices and market behavior, particularly at trading strategies. Precisely, Shefrin (2008) proposes to look at investor sentiment from two perspectives, either as excessive optimism or pessimism about stock returns. Therefore, the relationship between investor sentiment and future stock returns should be negative. Higher sentiment encourages investor overpricing and irrational thinking, consequently leading to lower future returns and vice versa for lower sentiment. Baker and Wurgler (2006) and Lemmon and Pornuaguina (2006) emphasize that such a significant negative relationship between sentiment and returns is especially pronounced among small and growth firms, since it is difficult to arbitrage away any mispricing in a “hard to determine price” field. Kadilli (2015) examines financial companies in 20 developed countries from January 1999 to August 2011 by using panel regime-switching models and market wide measures of investor sentiment. He documents a negative, but insignificant impact of investor sentiment on future stock returns during normal times. However, during crisis times, he reports strongly positive and significant effects of investor sentiment towards future returns.

Having identified sentiment as a feeling or perception of an investor towards the market, it is important to comprehend how investor sentiment can help financial practitioners, but also have an impact on the asset prices. There are different directions when it comes to examining the role and effect of investor sentiment. However, the most plausible explanations come from a psychological perspective on the sentiment. This view on sentiment is based on feelings and experience of the decision maker, such that the decision maker creates less accurate forecasts for speculative firms, due to his high sentiment (Hribar and McInnis, 2012). Meaning, if an investor creates a forecast about young and uncertain firm, he might be overly optimistic given that his sentiment level is high. His or her predictability will solely rely on his past performance. Especially if in the past, he experienced more gains than losses. As well, firms with such characteristics attract unexperienced and unprofessional investors, whose sentiment is more likely to influence his or her decisions to invest. Interestingly, researchers agree that it is very difficult to measure such unobservable and qualitative factors which are not easily quantified (Baker and Wurgler, 2006, 2007). Despite the difficulty to directly measure investor sentiment, theoretical assumptions strongly support the existence of investor sentiment and its effect on the stock prices. Thus, we will outline the various existing investor sentiment measures in the next section.

2.2.2 Investor sentiment measures

Finance literature provides sufficient evidence which document the existing relationship between investor sentiment and stock returns (Baker and Wurgler, 2006; Baker and Wurgler, 2007; Da et al.,2015). The long running debate continued in the direction of defining the most prominent measure of sentiment, yielding numerous measures which are highly correlated among one another (Brown and Cliff, 2004). Thus, the choice of sentiment indicators does not represent the main concern and it is a matter of context of the research. Among many, sentiment measures can be grouped into two categories: market based and survey based measures.

Firstly, we will begin the discussion of survey based measures that have been widely used in empirical research in the past and recently due to the large horizon span and proven successful predictability. Consumer Confidence Index (hereafter referred as CCI) is a survey specifically made for targeting a consumer's belief about the overall economy, household spending, increase in personal consumption, etc. (Ludvigson, 2004, Qui and Welch, 2006; Bthia and Bredin, 2013). It is believed and empirically examined that changes and fluctuations in the market can have an influence on how consumers form their opinions about current and future performance of an economy. CCI dates back in the 1977, and is constructed on a monthly basis from the surveys which are based on 5000 households in the U.S.. CCI is issued by the Conference Board, however a similar index which is also widely used in the research is conducted by the University of Michigan, namely the University of Michigan Consumer Sentiment Index (hereafter MS). The Consumer confidence index constructed by the Michigan University is conducted on a monthly basis and is determined by the consumer's responses on five questions about their expected financial situation and economic performance in the upcoming year. Fisher and Statman (2003) and Lemmon and Portniaguina (2006) included consumer confidence as a potential investor sentiment proxy, arguing that consumers represent a large group of individual investors who have expectations and attitude towards a future economic state. Another popular source of Consumer confidence is the OECD database, which will be used along with the MS sentiment measure to account for consumer confidence. Such measures of investor sentiment are not only reflecting the financial markets and investing activities, but also take a broader perspective and look at individual consumption as well.

Several researchers have used CCI in explaining stock market development. On the one hand, Jansen and Nahuis (2003) examined the short-run relationship between CCIs and stock returns by using short-term EU data on 11 countries. They document positive correlation between sentiment and stock returns in most of the countries. However, they note a one-way causality from stock returns to consumer confidence in the short run. Keeping in mind that stock returns are extremely difficult to predict, their findings on the causality are not surprising. Moreover, Charoenrook (2005) uses the MS index as a measure of investor sentiment to test if investor sentiment has an influence on stock returns. He observes that changes in the MS index and excess market returns are negatively related in the short run (referring to one-month window) and in the long run (one-year window). On the other hand, Schmeling (2009) used the CCI as a proxy for investor sentiment to examine 18 industrialized countries and finds a negative relationship between sentiment and stock returns on average. His findings suggest that investing on higher sentiment levels consequently results in lower stock returns, since higher sentiment levels are followed by periods of over optimism and behavioral biases. As well, such results suggest that investor sentiment has a predictable component that translates into irrational behavior of an investor. Recently, Ferrer et al. (2016) in their analysis used CCI as a proxy for investor sentiment and came across interesting results. The CCI is an irrelevant indicator used as a proxy for investor sentiment, because their findings revealed inability to forecast stock returns. However, in their analysis, they have constrained their data sample on the EU countries where presence of IT firms is significantly smaller than in the U.S.. As well, they narrowed down their sample period to the post doc.com bubble, which is the period after crisis where, firstly the consumers became more risk averse while investing and more prone towards biases as they have experiences severe losses. Secondly, the market was still in the process of recovery and as consumers are not considered to be experienced investors, their confidence levels were still affected by the crisis.

The American Association of Individual Investors (hereafter AAII) conducts a survey that is commonly used and cited in the financial literature aiming to find out how market participants feel about the stock market over the next six months. There are three possible answers; bullish, bearish and neutral, where on a weekly basis individuals provide an answer that translates into bullish, bearish or neutral expectation on future stock market returns. Such information can be used to construct a measure of investor sentiment by taking the fraction of bullish investors (Fisherman and Statman, 2000; Brown and Cliff (2004). Similarly to the AAII sentiment index, is the Investors Intelligence (hereafter II) survey which aims to capture the expectations of institutional investors. In the II survey, the survey is conducted with market professionals, such as investment advisors and newsletter writers on a weekly basis, whereby they are asked to provide their expectation of the future stock market based on the same question as in the AAII survey. Importantly, those two surveys should take two different approaches as the II survey is based on the experienced and professional investors, whereas the AAII survey comprises of all participants in the financial market. Such participants do not need to be professionals and thus can be less reliable. One could expect that the II survey used as an investor sentiment proxy should indicate no correlation between sentiment and movements in stock prices, since professional investors should not be affected by their sentiment. Even though in this research we use survey based investor sentiment measures, because of the data availability we could not make use of the AAII and II surveys. It would be an interesting area for future research, to compare these two surveys and report if the professionals do or do not demonstrate investing patterns based on their sentiment.

To test if consumer confidence affects the stock returns, Fisher and Statman (2003) used two measures of consumer confidence, namely the Conference Board and the Michigan University confidence measures. They report a strong and positive correlation of 0.54 among the Conference Board and the Michigan University confidence measures. This positive correlation of the two survey based sentiment measures indicates that the two measures move in the same direction. Even though the confidence measures evaluate the same group of consumers, the formation of the questions is not the same. In our analysis, we also find a strong positive relationship between the consumer confidence proxies, the Michigan University (MS) and the OECD measure of 0.77. On the other hand, by comparing the AAII and II sentiment index, which measure the investor's sentiment, with the consumer confidence that find they report that the consumer confidence grows as the investor's confidence grows. Overall, they find the evidence that low consumer confidence is caused by the fall in the stock returns, rather than vice versa. Meaning, a significant drop in the consumer confidence should not raise a concern for investors, because the falling confidence is not influencing movements in stock returns.

Potential limitations of survey based sentiment measures are the volume of the response rate, error likelihood in data processing and human biases. During some periods, the response rate in surveys tends to be high, while in other periods it might be low, leading to unreliable information for that period. Another potential problem might be biases when it comes to revealing the actual expectations towards the future market. Some people have tendency not to tell the true story if they are not compensated for such information. Having that in mind, we used two survey based measures of sentiment, namely MS and OECD, as well as two market based measures, BW (Baker and Wurgler) sentiment indicator and S^{PLS} (updated version of the

Baker and Wurgler sentiment index by partial least square method). In contrast to survey based measures, market based measures are not constrained in data availability and provide general overview of the economic performance (Naik and Padhi, 2016).

Secondly, the most commonly used market based sentiment measure has been constructed by Baker and Wurgler (2006), referred as to 'BW index'. Because it is not possible to directly measure and observe investors behavior and sentiment towards the market, Baker and Wurgler (2006) constructed a sentiment index based on principal component analysis. They extracted six components e.g. equity shares, dividend premium, closed-end fund discount rate, share turnover, number of IPOs and first days returns on IPOs. Potential criticism to BW sentiment index points at the high correlation of the six sentiment variables with the business environment from which they are extracted. Therefore, one could assume that BW investor sentiment does not represent just investor sentiment, but also forces that are driving the business environment and thus the index might be misleading (Da et al., 2010; Chu et al., 2015). Furthermore, Baker and Wurgler realized the problem of possible high correlation. In order to create more accurate sentiment measure, they removed non-fundamental components from fundamental components, by regressing each variable on business specific variables. The new altered BW index is called the orthogonal BW index.

A relatively new investor sentiment index has been constructed by Huang et al. (2015) which is built upon the same six variables of Baker and Wurgler (BW) sentiment index. They use a partial least square (hereafter S^{PLS}) method to extract relevant and correct information about the future stock market contained in the six proxy variables from the incorrect and noisy information. Those six variables are: equity shares, dividend premium, closed-end fund discount rate, share turnover, number of IPOs and first days returns on IPOs. In their empirical analysis, they followed a common approach of using a linear predictive regression procedure for predicting stock returns. Evidence suggests that the updated version of the Baker and Wurgler sentiment index, namely S^{PLS} index, is better able to predict aggregate monthly stock returns, which is contradictory to Baker and Wurgler (2007) and Baker et al. (2012).

Qui and Welch (2006) and Da et al. (2015) draw attention to a potential disadvantage of using market based sentiment measure. Namely, to construct such a measure, the components used in the calculation are not only disclosing the investor's environment, but are encompassing the entire market and business conditions. To account for a potential drawback in this particular investor sentiment measure, in this study, we will employ market and survey based measures of investor sentiment. Market based measures include real indicators straight from the market where investors are actively and directly involved. On the contrary, survey based sentiment measures rely on individual's responses and feelings about the market, which could be incomplete and not true. Thus, BW and S^{PLS} index should forecast sector returns better than MS and OECD.

2.2.3 Predictive power of investor sentiment

There is a vast literature on investor sentiment and its effect on the overall stock market. Among which, recent focus has been on its predictive power in explaining future returns. When testing the predictive power of investor sentiment, many researchers look at shifts in investor sentiment. Those shifts in investor sentiment are of particular importance because they indicate when investors are trading on noise and are so called noise traders. If this is the case, and investors are prone to noisy information, then causality will run from investor sentiment to stock returns. Brown and Cliff (2004) tested the causality effect between sentiment and returns and report strong causality from returns to investor sentiment. While on the other hand, Wang et al. (2006) examined several measures of investor sentiment and reports that none of the sentiment measures Granger causes returns, rather, there is a strong causal relationship from returns to sentiment measures. Interestingly, their results imply that either high or low stock returns will have an evolving effect on the investor's sentiment. Thus, indicating that returns might have stronger predictive power than it was expected and investor sentiment only develops over time. In addition, Salhin et al. (2016) examined the sentiment-return relationship between the sectors within the U.K.. In terms of causality, they find that stock returns of Financials and Manufacturing sectors Granger cause the sector sentiment, rather than vice versa. Given that the constituent firms of those sectors are mature, stable and attract professional investors, the results are not surprising. Rather, it shows that among certain sectors, investor sentiment has no predicting power as the rational and experienced investors prevail, leaving no role for sentiment.

Earlier papers by Neal and Wheatley (1998) looked at several proxies for investor sentiment based on closed-end funds and mutual funds. They concluded that equity returns can be predicted by investor sentiment. The same conclusion was found in the empirical research by Simon and Wiggins (2001) and Baker and Wurgler (2006). This empirical research points out at the importance of the sentiment-return relationship and concludes that investor sentiment can predict future returns.

On the other hand, Fisher and Statman (2000) categorized investors and their sentiment into three groups: large investors which are Wall Street strategists, small/ individual investors and investor newsletter writers classified as institutional investors. Their findings indicate that Wall Street strategist's sentiment together with small and individual investor sentiment are negatively correlated with future stock returns. They also indicated that causality does not only run from sentiment to returns, but instead there is a solid evidence that causality runs in both directions. Interestingly, their results did not find significant evidence that the Wall Street strategists and institutional investors show difference in investing patters from the individual/ unprofessional investors. As Wall Street strategists are considered to be professionals and informed investors, they should demonstrate rational behavior and invest on their knowledge, rather than on their sentiment. A similar finding was discussed in the work of Brown and Cliff (2004) who researched the relationship between investor sentiment and equity returns. However, their conclusion indicated that returns have an important role in sentiment determination.

Interestingly, a majority of the literature studying the role of investor sentiment and its predictability are examining the aggregate market, where only few focus their research to specific regions, e.g emerging markets. Hence, our paper contributes to the rather scarce literature on examining sectors. Furthermore, Chen et al. (2012) examined influences of local and global sentiment on industry returns in 11 Asian countries by using monthly data from 1996 to 2010. Results show differences in local and global sentiment. Firstly, as reported by the relationship between local sentiment and return, industry's returns such as Materials, Consumer Services, Telecommunications and Utilities are influenced by higher local sentiment. On the other hand, global sentiment shows negative (positive) effect on stock returns when the market is pessimistic (optimistic) for industries Financials, Health, Oil and Gas and Industrials.

Overall, industries which are classified as young and speculative show a stronger sentiment-return relation affected by local sentiment.

Schmeling (2009) examined the sentiment-return relationship on 18 industrialized countries and reports that on average, investor sentiment negatively predicts returns across countries. In addition, he further investigates in which countries the investor sentiment has a stronger impact on returns and concludes that in countries where regulatory institutions are less organized and where herd-like investment behavior is present, the investor sentiment-return relationship is more evident. Herd-like behavior is a psychological situation, where investors and market participants do not follow their instincts, but invest or trade in the direction of the crowds. Such a situation is especially evident in the emerging and developing countries where it lacks market integrity. Herd-like effect offers a potential behavioral explanation why irrational investors are present in the market and why are they driving the mispricing. Kaplanski and Levy (2010) examined how the worldwide aviation disasters between 1950 and 2007 influence market sentiment and thus overall industries returns. Interestingly, they report that after an aviation disaster, less stable industries, with a majority of younger and growth firms, are more exposed to market sentiment than more stable and mature industries. The strongest influence has been documented among IT industry, whereas the lowest among Utilities. This conclusion confirms previous finding from Baker and Wurgler (2006) that uncertain and young stocks are more prone to investor sentiment.

By employing regime switching model on cross-selection of stock returns, Chung et al. (2012) distinguish between expansion and recession state to examine the predictive effect of investor sentiment. Despite their methodology not directly being relevant to our research, their findings are important for understanding the predictive power of investor sentiment. They find that investor sentiment has predictive power only in expansion states (low volatility). Meaning, expansion states demonstrates the period of higher levels of investor sentiment, that is followed by lower returns for specific types of stock e.g. small, growth, non-earning and non-dividend paying stock. Similarly, the vice versa follows for recession states. Interestingly, from their findings, we can observe that specific characteristics of stocks can determine to what extent investor sentiment affects them. For example, Baker and Wurgler (2007 and 2012) suggest that stocks that are more mature, easy to value and easy to arbitrage show less sensitivity towards sentiment and can have positive future return-sentiment relationship.

Table 1: Summary of main literature

The table summarizes the main literature related to our paper which is based on examining investor sentiment and stock return relationship.

AUTHOR	PERIOD	PURPOSE	METHOD	SAMPLE	FINDINGS
Brown and Cliff (2005)	Monthly data 1963-2000	Is investor sentiment able to predict stock returns over 1 to 3 years?	Regression	U.S.	Survey based investor sentiment affects asset valuation and is negatively related with future returns
Baker and Wurgler (2006)	Monthly data 1962-2001	How investor sentiment affects cross-section of stock returns	Regression	U.S. common stocks	Small, young and unprofitable stocks earn lower subsequent returns when sentiment is high
Lemmon and Portniaguina (2006)	Monthly data 1977- 2002	Does consumer confidence predict returns as explained by behavior finance	Regression	U.S.	During periods of high consumer confidence, investors overvalue small stocks versus larger stocks
Kaplansky and Levy (2010)	Daily data 1950- 2007	How aviation disasters affect stock returns?	Regression	Worldwide data on 228 aviation disasters	Less stable industries, smaller and riskier firms are more exposed to investor sentiment
Chen et al. (2012)	Monthly data 1996-2010	How industries are affected by local and global market sentiment	Predictive regression	11 Asian countries	Overall, sentiment affects future industry returns, however differently for local and global sentiment.
Huang et al. (2015)	Monthly data July 1965-December 2010	New sentiment index that will better predict aggregate stock returns	Regression	U.S. aggregate market	New investor sentiment has greater predicting power over aggregate stock market
Salhin et al. (2016)	Monthly data 1985-2014	Analyze relationship between managerial and consumer sentiment and sector returns	Regression	U.K.	Significant impact of managerial sentiment towards sector return, especially manufacturing firms

Kadilli (2015)	Monthly data 1999- 2011	Can investor sentiment predict stock returns of financial firms	Panel model with regime switches	20 developed countries worldwide	Positive and strongly significant effect of investor sentiment on returns only during crisis times.

CHAPTER 3 Data

This section is split into two parts. Firstly, we describe in depth how we obtained the data used to carry out our research. Secondly, we show some preliminary results such as summary statistics, correlation table and some graphs displaying time series plots of investor sentiment over the sample period.

3.1 Data Selection

In our study, we use a dataset of 10 sectors within the U.S.. The sample period consists of monthly data from January 1984 until December 2014. We restrict our sample period to December 2014 because of the data availability of the market based sentiment measures. Most of the sentiment data is not available for public for free and without market based sentiment measure this thesis would be inconsistent and incomplete. This is only a limitation to our research. The purpose is to capture the effect of both market based and survey based sentiment measures on sectoral stock returns. This time span captures several U.S. recession periods. The events are as follows: Oil price shock in early 1990, the dot.com bubble in early 2000 and U.S. housing bubble or so called the Great Financial Crisis from 2008 until mid-2009.

Several types of data have been used in our analysis, such as four different measure of investor sentiment, sector returns and S&P 500 returns. Data was retrieved from several sources such as market data obtained from the The Center for Research in Security Prices (CRSP), Consumer Confidence Index by the University of Michigan obtained from the FRED database and OECD index from the OECD database, while BW index and S^{PLS} were retrieved from J. Wurgler and professor Zhou database respectively.[1]

3.1.1 Sector returns

We retrieve our sector returns from S&P 500 index constituents that have been sorted to the corresponding sector by employing Global Industry Classification System (hereafter 'GICS'). Each firm in the index has a corresponding GICS assigned by Standard and Poor and MSCI. After sorting the firms by corresponding sector, we were left with 10 sectors: Energy (34 firms), Materials (25 firms), Industrials (67 firms), Consumer Discretionary (85 firms), Consumer

[1] Available at http://apps.olin.wustl.edu/faculty/zhou/#useful_links; http://people.stern.nyu.edu/jwurgler/

Staples (36 firms), Health Care (61 firms), Financials (66 firms), Information Technology (68 firms), Utilities (28 firms) and Telecommunication Services (4 firms).[2] We extracted closing prices from CRSP database and compute log returns to normalize the data. S&P 500 index is included in our analysis to assess the overall market performance and we refer to it as a benchmark. Outliers in the sample are not removed, because our sample period includes several market turmoil's where we expect extreme values for those periods.

3.1.2 Investor sentiment proxy

In this study, as a measure of investor sentiment we use data from market based and survey based measures as sentiment proxies. Ideally, we want to examine if specific investor sentiment proxy has stronger impact on forecasting sector returns over the others. Baker and Wurgler (2007) pointed out that among researchers, many are questioning the reliability and consistency of survey data and such measures. In line with this way of thinking, we choose to use both types of sentiment measures to account for any inconsistency in survey data. Firstly, we distinguish between two market and two survey based sentiment measures. For the market based sentiment measure, Baker and Wurgler (hereafter BW) and updated Baker and Wurgler sentiment index (by applying the partial least square approach) (hereafter S^{PLS}) sentiment index is employed. The choice of BW sentiment index is standard in the finance literature and widely used as it captures economic cycles and movements in the market. As mentioned before in section 2.2.2, it is composed of six investor sentiment constituents from which the first principal component is taken. Importantly, they noted a high correlation between raw sentiment variables and business cycles. Hence, before taking the first principal components, they regressed each of the variables on several business cycles proxies to get cleaner variables for constructing the orthogonal BW sentiment index. Therefore, we use the orthogonal $BW^{\perp}$ sentiment index.

Next, as the second market based sentiment measure, the S^{PLS} index is considered. S^{PLS} sentiment index is constructed on the idea of Baker and Wurgler (2006, 2007), adopting the same six variables as in BW index (Huang et al., 2015). However, they noticed that BW components include approximation errors that cannot account for stock movements and its predictability. Using the partial least square method, Huang et al. (2015) were able to separate

[2] Real Estate sector is excluded from our analysis since only in September, 2016 it was acknowledged as a new sector in GISC.

information contained in the six components on: information relevant for stock return predictability and noisy information. Since the S^{PLS} sentiment measure is relatively new in the literature and has not yet been widely used, it will be used to carry out our research. According to Huang et al. (2015) and Sun et al. (2016), the S^{PLS} sentiment index has higher predictive power and is able to predict aggregate stock market returns, while the BW index failed to do so.

Secondly, survey based sentiment measures as mentioned in Section 2.2.2 are commonly used in combination with market based measures. In this research we choose two measures; the University of Michigan Consumer Sentiment Index (hereafter MS) and Confidence Index developed by the OECD (hereafter OECD), obtained from the FRED and the OECD website respectively. CCIs are being considered as a classical measure of consumer's feeling and perception of the market in economics and finance, indicating optimism towards the current and future economic performance. Both surveys are based on U.S. household's response about current consumption and future expectation of the economy. Respondents provide an answer which is later turned into an index. The higher the level of the index, the more optimistic respondents are about the future economic state.

Based on the previous analysis by Lemmon and Portniaguina (2006) from looking at time series analysis, they document forecasting power of MS proxy on stocks that are held by individual investor and are small in nature. Similarly, Barsky and Sims (2012) emphasize on the importance of CCI as a predictor of future economic performance. They note that CCI contains information which can predict future economic events since it gathers the information from consumers that provide reliable information about their confidence levels. Thus, in our research, as part of the survey based measures we use two common surveys of CCI, namely MS and OECD indices.

3.2 Sample statistics and preliminary results

Table 2 presents all variables previously discussed. The statistics are presented for in-sample period from January 1984 until December 2004 including monthly percentage sector returns and monthly investor sentiment indices. Firstly, we report information regarding mean, median, standard deviation, skewness, kurtosis and autocorrelation levels. From Table 2. it can be observed the following: high returns for Financials, Health, and Consumer Discretionary sectors amount to 0.395, 0.412 and 0.366 percent respectively. However, they do not outperform the S&P 500 index with a mean monthly return of 0.912 percent. In terms of volatility, which is measured by the standard deviation, the S&P 500 has the highest standard deviation. The volatility is also large for IT and Telecom compared to the other sectors. The sector returns generally report negative skewness, suggesting a larger negative tail. Materials and Telecom have a small positive skewness of 0.006 and 0.077 respectively. This suggests that their distribution is more symmetric and has slightly more positive returns. Each of the return indices report large positive kurtosis, indicating that the majority of the distribution is centered around the mean. The Industrials report a kurtosis of 13.255 and it is the largest of the returns. The autocorrelations of the returns for all orders appear to differ from zero, suggesting the presence of autocorrelation effects. Finally, the investor sentiment indices show distinct differences. The differenced MS index is the most volatile of the four indices, and also displays the lowest autocorrelations. The other three indices, BW, OECD and S^{PLS}, report lower volatilities, and also significantly strong autocorrelations for all orders. The sentiment indices are displayed in Figure 1. The MS index appears to be the smoothest of the curves, showing the least variation compared to the other three. However, taking the difference appears to result in the largest variation. Finally, the investor sentiments display typical features with large changes occurring during periods of economic contraction.

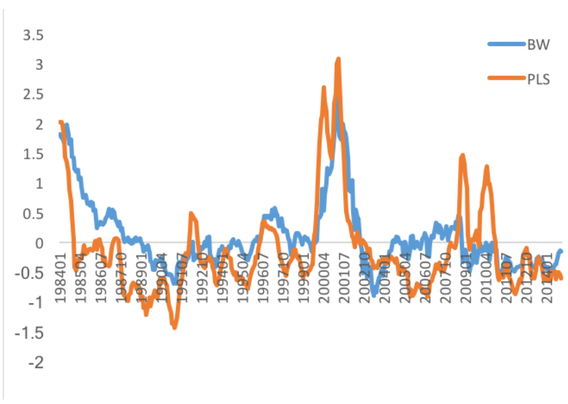
Table 2: Summary statistics

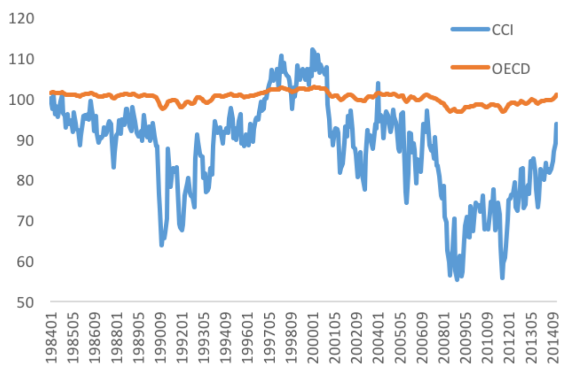
This table shows sample statistics of the sector returns, S&P 500 index and four investor sentiment indices applied in the sample period from January 1984 until December 2003. Each row represents one of the ten sectors ENE (Energy), FIN (Financials), HEA(Health), IND (Industrials), IT (Information Technology), MAT (Materials), TEL (Telecom), UTL (Utilities), CD (Consumer Discretionary), CS (Consumer Staples) and SP500 (S&P500 index). Four investor sentiment indices are Baker and Wurgler (BW), The Michigan University Consumer Sentiment Index (MS), OECD Consumer Confidence Index (OECD) and the updated version of Baker and Wurgler sentiment index S^{PLS} by Huang et al. (2015). Sector returns are monthly percentage log returns. The standard sample statistics are shown along with skewness and kurtosis. Notation ρ_k denotes k^{th} order autocorrelation, meaning ρ_1 is correlation between observations that are one period lagged and is called autocorrelation of order one.

	Mean	Median	SD	Skewness	Kurtosis	ρ_1	ρ_2	ρ_3	ρ_4
ENE	0.321	0.188	2.138	-0.437	7.405	-0.105	0.029	-0.040	-0.102
FIN	0.395	0.697	2.652	-0.684	5.152	0.027	-0.020	-0.075	-0.084
HEA	0.412	0.756	2.276	-1.039	7.033	-0.046	0.019	-0.009	-0.147
IND	0.301	0.527	2.389	-1.589	13.255	-0.073	-0.061	-0.023	-0.086
IT	0.333	0.616	3.604	-0.903	5.694	-0.052	0.003	0.110	-0.069
MAT	0.262	0.336	2.414	0.006	3.985	-0.121	0.019	-0.065	-0.132
TEL	0.100	0.334	2.712	0.077	5.067	0.004	0.014	0.074	-0.047
UTL	0.181	0.271	2.010	-0.370	3.505	0.019	-0.055	0.060	0.053
CD	0.366	0.587	2.168	-0.406	3.967	0.044	-0.088	0.008	-0.113
CS	0.314	0.364	1.925	-0.220	3.634	0.031	-0.015	-0.150	-0.126
SP500	0.912	1.135	4.513	-0.743	5.462	-0.008	-0.046	-0.018	-0.105
ENE	0.231	0.041	0.669	1.182	3.965	0.976	0.952	0.927	0.898
FIN	0.015	-0.250	3.573	0.444	5.665	-0.060	-0.002	-0.016	-0.062
HEA	0.000	0.002	0.209	-0.197	5.107	0.788	0.363	-0.023	-0.254
IND	-0.074	-0.274	0.847	1.647	5.666	0.980	0.938	0.879	0.809

Figure 1:Time series plot of investor sentiment measures

Time series plots of investor sentiment over the sample period January 1984 to December 2014. Left, the sentiment indexes BW and $S^{PLS}$are displayed. On the right, MS and OECD are displayed, note that these two series are not adjusted for non-stationarity.





Furthermore, Figure 1 represents the time series of all four investor sentiment proxies over the full sample period from January 1984 until December 2014. The sample period covers several market turmoil's e.g. 1987 Black Monday, the dot.com bubble in 2000 and the Great Financial crisis in 2008. All four proxies try to capture the movements in investor sentiment during the sample period, however there is an obvious difference among them. For instance, BW and S^{PLS}strongly react to market events, which can be seen by the increasing sentiment, especially the spike around the dot.com bubble in 2000. MS proxy noticeable falls after the Great Fiinancial crisis indicating pessimism around consumers about future economic performance. Overall, OECD proxy shows signs of mean reversion, where BW, $S^{PLS}$and MS appear to be more volatile over the entire sample period.

Next, Table 3. reports correlation coefficients between sentiment measures and sectoral returns, taking the lag lengths of 0, 1, 2 and 3. By taking the lagged sentiment, ideally, we would like to see if sentiment at time t-1, t-2 or t-3 is related with sector returns at time t. For example, lagged sentiment correlation of 2nd order means that sector returns today are compared with investor sentiment measure two months ago. The correlation coefficients between sentiment measures and sector returns show mixed results. Firstly, statistically significant correlation coefficient is observed in IT, Telecom, Consumer Discretionary and S&P 500. Though, the most significant correlations are observed for the first order lags. The largest and statistically significant correlations are observed between 1 period lagged MS index and Industrials with correlation coefficient of 0.329, while the second largest coefficient was between 2nd period lagged OECD index and IT sector of 0.268. Interestingly, correlation across all sectors and survey based measures, MS and OECD is positive, while with both market based measures BW and S^{PLS} negative. However, this relationship is not significant for all sectors, but IT and Telecom sector. For example, sectors like Energy and Utilities, are not at all correlated with any sentiment measure, which could be explained by the fact that these two sectors are mainly composed of mature and "safer" firms. Such companies are less prone to investor sentiment since they are profitable, dividend-paying firms with lower propensity to speculate. Thus, in sectors where lot of firms experience extreme growth and are relatively volatile, such as IT, Telecom and Consumer Discretionary, we can expect that sentiment will have higher influence on returns which is in line with theoretical predictions.

Table 3: Correlation Matrix

This table represents correlation coefficients between sector returns, S&P 500 index and investor sentiment indicators for the sample period from January 1984 until December 2014. BW and S^{PLS} are market based sentiment indicator, where BW is Baker and Wurgler (2006, 2007) sentiment index, S^{PLS} is Huang et al. (2015) updated version of the Baker and Wurgler sentiment index based on the partial least square method, MS is the Michigan University sentiment measure based on the consumer confidence and OECD is another consumer confidence survey based sentiment indicator from the OECD database. *: significance at 1% level.

Lagged correlation		ENE	FIN	HEA	IND	IT	MAT	TEL	UTL	CD	CS	SP500
0^{th} order	BW	-0.071	-0.094	-0.085	-0.119	-0.256*	-0.052	-0.193*	-0.109	-0.144	-0.032	-0.207*
	MS	0.033	0.128	0.079	0.100	0.117	0.046	0.046	0.104	0.108	0.066	0.110
	OECD	0.136	0.219*	0.103	0.253*	0.227*	0.189*	0.105	0.101	0.268*	0.134	0.253*
	S^{PLS}	-0.040	-0.099	-0.106	-0.107	-0.224*	-0.100	-0.215*	-0.008	-0.151	-0.117	-0.199*
1^{st} order	BW	-0.064	-0.095	-0.063	-0.114	-0.224*	-0.081	-0.179*	-0.101	-0.134*	0.000	-0.187*
	MS	0.081	0.228*	0.212*	0.329**	0.216*	0.224*	0.179*	0.079	0.342*	0.208*	0.309*
	OECD	0.092	0.234*	0.153	0.345	0.318*	0.220*	0.137	0.061	0.368*	0.195*	0.336*
	S^{PLS}	-0.042	-0.092	-0.104	-0.098	-0.209*	-0.082	-0.214*	-0.007	-0.123	-0.110	-0.184*
2^{nd} order	BW	-0.023	-0.048	-0.003	-0.085	-0.237*	-0.046	-0.178*	-0.013	-0.095	0.031	-0.161*
	MS	0.027	0.017	-0.057	0.123	0.239*	0.046	0.008	-0.072	0.091	0.007	0.116
	OECD	0.028	0.095	0.060	0.219*	0.268*	0.107	0.061	-0.032	0.246*	0.117	0.224*
	S^{PLS}	-0.043	-0.103	-0.099	-0.091	-0.199*	-0.070	-0.206*	-0.003	-0.107	-0.095	-0.173*
3^{rd} order	BW	-0.008	-0.032	0.022	-0.039	-0.209*	-0.034	-0.170*	0.016	-0.080	0.045	-0.129
	MS	0.005	-0.091	-0.016	-0.082	-0.098	-0.052	-0.071	-0.022	-0.007	-0.020	-0.058
	OECD	-0.028	-0.051	0.013	0.016	0.083	-0.024	0.006	-0.071	0.065	0.038	0.052
	S^{PLS}	-0.045	-0.085	-0.060	-0.089	-0.190*	-0.006	-0.197*	0.016	-0.097	-0.077	-0.157*

Table 4: Unit root test

Table 4 represents unit root test for the in-sample period from January 1984 until December 2003 for all variables used in this study. Three testing procedures are applied, namely the Augmented Dickey Fuller (ADF) test, Phillips-Perron (PP) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. The ADF is modelled with a constant and excluding a drift. The maximum number of lags used by the test is determined by the Bayesian Information Criterion (BIC). The null hypothesis of the ADF test is non-stationarity. For the PP test, a similar approach to the ADF test is used with a null hypothesis of non-stationarity. For the KPSS test, the null hypothesis of the test is stationarity. For the ADF and PP test, if the p-value is less than the 1 percent significance level, the value is capped at 1 percent. The KPSS test on the other hand restricts the p-value to 10 percent for all cases that exceed it.

	Level			Difference		
	ADF	PP	KPSS	ADF	PP	KPSS
ENE	-15.893	-14.652	0.034	-27.634	-26.725	0.002
(p-value)	*(0.001)*	*(0.001)*	*(0.100)*	*(0.001)*	*(0.001)*	*(0.100)*
FIN	-14.726	-12.625	0.058	-25.931	-24.528	0.002
(p-value)	*(0.001)*	*(0.001)*	*(0.100)*	*(0.001)*	*(0.001)*	*(0.100)*
HEA	-14.421	-14.105	0.114	-25.790	-25.016	0.002
(p-value)	*(0.001)*	*(0.001)*	*(0.100)*	*(0.001)*	*(0.001)*	*(0.100)*
IND	-16.353	-17.523	0.049	-26.936	-27.025	0.002
(p-value)	*(0.001)*	*(0.001)*	*(0.100)*	*(0.001)*	*(0.001)*	*(0.100)*
IT	-15.301	-15.021	0.148	-26.265	-25.881	0.002
(p-value)	*(0.001)*	*(0.001)*	*(0.048)*	*(0.001)*	*(0.001)*	*(0.100)*
MAT	-13.832	-12.177	0.035	-23.297	-22.982	0.003
(p-value)	*(0.001)*	*(0.001)*	*(0.100)*	*(0.001)*	*(0.001)*	*(0.100)*
TEL	-12.374	-10.481	0.141	-21.706	-20.996	0.004
(p-value)	*(0.001)*	*(0.001)*	*(0.059)*	*(0.001)*	*(0.001)*	*(0.100)*
UTL	-15.072	-16.288	0.042	-25.516	-25.827	0.002
(p-value)	*(0.001)*	*(0.001)*	*(0.100)*	*(0.001)*	*(0.001)*	*(0.100)*
CD	-11.598	-11.029	0.063	-19.716	-19.207	0.003
(p-value)	*(0.001)*	*(0.001)*	*(0.100)*	*(0.001)*	*(0.001)*	*(0.100)*
CS	-11.770	-10.498	0.055	-20.975	-20.024	0.003
(p-value)	*(0.001)*	*(0.001)*	*(0.100)*	*(0.001)*	*(0.001)*	*(0.100)*
SP500	-14.968	-15.614	0.076	-26.108	-26.821	0.002
(p-value)	*(0.001)*	*(0.001)*	*(0.100)*	*(0.001)*	*(0.001)*	*(0.100)*

Table 4 continued

BW	-2.611	-2.259	2.393	-15.316	-15.172	0.133
(p-value)	*(0.009)*	*(0.009)*	*(0.010)*	*(0.001)*	*(0.001)*	*(0.074)*
MS	-0.223	-0.104	2.015	-16.073	-15.136	0.030
(p-value)	*(0.569)*	*(0.569)*	*(0.010)*	*(0.001)*	*(0.001)*	*(0.100)*
OECD	-0.007	-0.015	2.147	-5.243	-4.901	0.109
(p-value)	*(0.649)*	*(0.649)*	*(0.010)*	*(0.001)*	*(0.001)*	*(0.100)*
$\mathbf{S^{PLS}}$	-2.463	-1.977	1.568	-8.672	-6.931	0.199
(p-value)	*(0.014)*	*(0.014)*	*(0.010)*	*(0.001)*	*(0.001)*	*(0.016)*

The stationarity tests for the data used in this paper is displayed in Table 4. The original data is evaluated for non-stationarity using ADF, PP and KPSS tests. For the return series, there is no significant signs of non-stationarity at the 1 percent significance level. However, for investor sentiment, MS and OECD show significant non-stationarity for the ADF and PP tests. For caution, we take the difference of these series and recomputed the tests, and find that the MS and OECD series are now stationary. The differenced series are used in the remainder of the paper.

Chapter 4 Methodology

The following section introduces the methodology used to obtain the results. It consists of two parts, an in-sample and out-of-sample period. Firstly, the methodology starts off with the Granger causality test, then it proceeds with the explanation of the in-sample period and the model setup. In-sample model is used for estimating and selecting the optimal model and its parameters. Secondly, we move on to the out-of-sample model, which is based on the optimal model selected from the in-sample model. To evaluate the out-of-sample performance, we explain the evaluation techniques.

4.1 In-sample period

In Section 3.2, sample statistics and some preliminary results have been reported, although they do not provide conclusive evidence that sentiment directly influences returns. There are some indications of correlation between sector returns and lagged investor sentiment measures, which will be examined further by looking at the Granger causality test. To examine if investor sentiment has a predicting power, the full sample is split in two periods: in-sample and out-of-sample period. In-sample period ranges from January 1984 until December 2003. The starting date is determined by the availability of the sector return data, whereas the end date is chosen to be end of 2003 for two reasons. Firstly, the in-sample period captures several market downturns, including the dot.com bubble. That way our in-sample period is highly dynamic thus incorporating more information in the model. Moreover, the in-sample period is 20 years long, which is a sufficiently large sample for running the least squares model. In-sample period is used for computing the optimal model, which will be applied throughout the out-of-sample period. The reason we opt for using the optimal model, instead of continuously re-computing the optimal model is to save on computation time.

4.1.1 Granger causality test

The aim of this study is to empirically test if investor sentiment can predict returns. In Section 3.2 we have done some preliminary tests, which showed potential evidence of correlation between sector returns and investor sentiment measures. Therefore, we start our analysis by firstly looking at the Granger causality test. The Granger causality test, as its name already says, uses time series data to determine if there is a causal relationship between variable x and y (Granger, 1969). We employ the Granger causality test to investigate if investor sentiment measures can forecast sector returns and/or vice versa. First, we evaluate the correlations between investor sentiment and sector returns to get a general understanding of the relationship. Thereafter, the Granger Causality test is applied to un-root the causation of the series (Granger, 1988). As we are attempting to explain and forecast returns, it is important to remain cautious with these econometric techniques. Sector returns are notoriously difficult to estimate, and therefore economic theory should also be used to justify economic setups. The analysis is started off by examining the causality between the series. The following equation shows Granger causality:

$$X_t = \sum_{j=1}^{m} \beta_j y_{t-1} + \varepsilon_t \qquad (2)$$

Where X_t is sector monthly return at time t, Y_{t-1} is a sentiment proxy one period in the past and ε_t is an error term. The Granger Causality model shows if a dependent variable Y can be better predicted by X while using its historical value alone (Wooldridge, 13th edition). Moreover, the test is performed only for in-sample period to help with the model selection procedure. Performing the Granger causality test on the out-of-sample period would introduce a bias to the forecasting, as we explicitly assume we do observe the data in that period.

4.1.2 In-sample model setup

To perform the in-sample analysis, we estimate an autoregressive model with lagged terms for the sector returns and investor sentiment. Using this procedure is standard when selecting lag length. The process is automatic and therefore does not require each regression to be manually examined using correlogram. By doing so, we are able to capture the past information. Huang et al. (2014) used similar lagged regression to examine at what lag length the investor sentiment has an effect on different industries. The model is setup as follows:

$$y_t = \alpha_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \beta_4 y_{t-4} + \theta_1 SE_{t-1} + \theta_2 SE_{t-2} + \theta_3 SE_{t-3} + \theta_4 SE_{t-4} + \varepsilon_t \quad (3)$$

where y_t denotes the sector returns, β_p are the coefficients of the p ordered lagged sector returns y_{t-p} , SE denotes the investor sentiment index, θ_q are the coefficients of the q ordered lagged investor sentiment returns SE_{t-q}. The model above is estimated using the standard least squares approach.

The optimal lag lengths p and q need to also be estimated. We constrain p and q such that they cannot be larger than 4. Therefore, there are a total of 16 different model combinations with different lag lengths that can be computed. To decide which model is optimal, we perform the following algorithm. First, we estimate a particular model. Using the likelihood of the returns, we can compute the Bayesian Information Criterion (BIC) shown below:

$$BIC = \ln(n)\, k - 2\ln(L) \quad (4)$$

where L is the value of the likelihood function, x is the observed data, n is the sample size, and k is the number of parameters. The model with a lower BIC is preferred. The BIC has a penalty term for models with larger number of parameters. This gives it an advantage over the Akaike Information criterion.

The BIC is therefore computed for each of the model combinations. The final model is then selected and can be used for forecasting. Note that the selected model is also corrected for serial correlation and heteroskedasticity using Newey West (1987) standard errors. Each of the models are not tested for these violations individually, therefore we assume their presence and control for heteroskedasticity and serial correlation in each model.

4.1.3 In-sample model estimation summary example

In this Section, we outline the in-sample model used for performing out- of sample forecasts. The model is estimated over the sample period from January 1984 to December 2003. The steps of the algorithm are outlined using Energy sector returns and BW investor sentiment data.

Step 1: The model setup of Equation (3) has a maximum of 4 lag lengths for returns (p) and investor sentiment (q). To determine the optimal lag length in the estimation sample, we first estimate each of the models using Newey West (1987) standard errors. We save the coefficients for each of the 16 models.

Step 2: For each of the models, the BIC is computed and reported in the table below.

Table 5: BIC for selecting optimal lag length with energy returns and BW sentiment

For each of lag length p and q for the model in Equation (3), the BIC is reported, where p and q represent lag lengths for returns and investor sentiment respectively. The maximum lag length of p and q are both set equal to 4.

p/q	1	2	3	4
1	952.35	945.61	945.69	946.94
2	947.34	950.11	950.20	951.44
3	948.35	951.24	954.70	955.94
4	946.85	948.82	952.65	957.13

For example, the model with p equal to 1, and q equal to 2 has a BIC of 945.61 and can be represented as:

$$y_t = \alpha_0 + \beta_1 y_{t-1} + \theta_1 BW_{t-1} + \theta_2 BW_{t-2} \quad (5)$$

Step 3: Select the optimal model by choosing the lowest BIC for p and q. In our case, this is p equal to 1 and q = 2 for the Energy sector and BW measure of investor sentiment. The optimal coefficients can also be found reported in Table 8.

Step 4: The steps 1 through to 3 are also repeated for the remaining 9 sector returns and 3 measures of investor sentiment.

4.2. Out-of-sample period

Next to the in-sample period, which is used to estimate the optimal model for validation, the second part of the methodology includes the out-of-sample period. The period ranges from January 2004 to December 2014. The choice for the starting point of the out-of-sample period corresponds to the post dot.com bubble period after which the markets and investor sentiment stabilizes. Following January 2004, the market is shortly stable, however, also experiences heavy fluctuations during the financial crisis. This combination of stable and dynamic activity provides a good sample for evaluating forecast performance in these market conditions. The model selection criteria discussed in the previous section is performed for each of the sector returns and investor sentiment measures. This gives a total of 40 optimal models for the 10 sector return indices and 4 measures of investor sentiment. The forecasting methodology is discussed below, after which the methodology used to evaluate the forecasts is outlined.

4.2.1 Out-of-sample model setup

As mentioned, to perform the forecasts, we use the previous model in Equation (3). For sake of illustration, we use a simple model with p and q both equal to 1. We create forecasts for 1 month, 3- and 10-months ahead using the following recursive relationships:

$$\hat{y}_{t+1} = \alpha_0 + \beta_1 y_t + \theta_1 SE_t$$

$$\hat{y}_{t+2} = \alpha_0 + \beta_1(\alpha_0 + \beta_1 y_t + \theta_1 SE_t) + \theta_1 SE_{t+1} \quad (6)$$

Substituting allows us to make forecasts multiple periods ahead. When forecasting, the sentiment is forecasted using the same approach for selecting lag lengths as in 4.1.2. We adopt a standard regression model from the finance literature that is used in forecasting analyses, such as in the Huang et al. (2014). The following model in Equation (7) is estimated using this procedure with up to 4 lags for sentiment:

$$SE_t = \alpha_0 + \theta_1 SE_{t-1} + \theta_2 SE_{t-2} + \theta_3 SE_{t-3} + \theta_4 SE_{t-4} + \varepsilon_t \quad (7)$$

Each of the forecasts is performed using a 1-month rolling window, where the initial window is set as in-sample period. After each forecast is made, the window is moved again and the forecast is repeated (see Section 4.2.3 for detailed explanations).

4.2.2 Forecast evaluation techniques

The resulting forecasts y_t for the sector returns can be compared to the actual returns from those periods. By doing so, we are able to see how close the investor sentiment indices predict the actual return series. Two standard methods are applied to evaluate the return forecasts, namely root mean square error (hereafter RMSE) and mean absolute error (hereafter MAE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y} - y_{t)})^2}{n}}$$

$$MAE = \sqrt{\frac{\sum_{i=1}^{n}|\hat{y} - y_t|}{n}} \quad (8)$$

where $\hat{y}_t$ denotes the forecasted returns, y_t denotes the actual returns and n denotes the sample size. The squared and absolute error approach differs from taking the conventional average. The reason is that when the average is taken of the errors, the magnitude is difficult to compare. Therefore, squaring or taking the absolute gives equal weight to positive and negative values. Different models can be compared using this estimation error, and indicate the best forecasting power.

The forecasted models can also be compared to the Random Walk model. This is done by comparing the squared forecasted errors of the two models. First, we define the forecasted returns $\hat{y}_{t+1}$ as the investor sentiment model Equation (6). Next, we define the AR(1) Random walk model as:

$$\ddot{y}_{t+1} = \alpha_0 + \beta_1 \ddot{y}_t. \qquad (9)$$

The errors for each of the sentiment model are computed by: $\hat{e}_{t+1} = \hat{y}_{t+1} - y_{t+1}$. The errors for each of the random walk models are given by: $\ddot{e}_{t+1} = \ddot{y}_{t+1} - y_{t+1}$. Finally, the two model errors are compared computing the following differential:

$$d_t = \hat{e}_{t+1}^2 - \ddot{e}_{t+1}^2. \qquad (10)$$

Each of the differentials is computed for each of the forecasts, and then evaluated whether on average they differ significantly from zeros. The Diebold-Mariano test statistic is used to do so. Under the null hypothesis, $E(d_t) = 0$ suggesting equal forecast performance. Under the alternative, $E(d_t) \neq 0$ suggesting superior forecast performance for one of the models depending on the sign. If the z-statistic statistic of the Diebold-Mariano test is positive, the Random walk model is superior, while if it is negative, the investor sentiment model is superior.

4.2.3 Out-of-sample model summary example

In the following section, we outline the steps used for performing out- of-sample forecasts using the optimal model derived in Section 4.1.2. The out- of sample period ranges from January 2004 until December 2014. The first step involves using the in-sample period to generate the model for investor sentiment, which is required for out-of-sample forecasting. Thereafter, we introduce the rolling window and outline the forecasting procedure. For illustration purposes, we are using Energy sector returns and BW investor sentiment data.

Step 1: The in-sample period ranges from January 1984 to December 2003. The model setup of Equation (7) has a maximum of 4 lag lengths for investor sentiment. To determine the optimal lag length in the estimation sample, we first estimate each of the models using Newey West (1987) standard errors. We save the coefficients for each of the 4 models.

Step 2: For each of the BW models, the BIC is computed and reported in the table below.

Table 6: BIC optimal lag length for BW investor sentiment

The BIC of Equation (6) is reported below for different lag lengths. The maximum lag length is set to 4.

Lag length	1	2	3	4
BIC	-231.38	-226.42	-220.15	-216.38

For example, the model with a lag length of 1 for sentiment has a BIC equal to -231.38 and can be represented as:

$$BW_t = \alpha_0 + \theta_1 BW_{t-1} \quad (9)$$

Step 3: Select the optimal model by choosing the lowest BIC. In our case, this is equal to 1. The optimal coefficients can also be found reported in Table 8.

Step 4: Now that we have the optimal model for returns as outlined in Section 4.1.3, and also the optimal model for investor sentiment, we can begin with out-of-sample forecasting. The length of the rolling window is from January 1984 to December 2003. This sample will be used to generate the first forecast.

Step 5: We generate forecasts for returns up to 10 months ahead using the models (5) and (9).

$$\begin{aligned}
\hat{y}_{t+1} &= \alpha_0 + \beta_1 y_t + \theta_1 BW_t + \theta_2 BW_{t-1} \\
\hat{y}_{t+2} &= \alpha_0 + \beta_1 \hat{y}_{t+1} + \theta_1 BW_{t+1} + \theta_2 BW_t \\
&= \alpha_0 + \beta_1 \hat{y}_{t+1} + \theta_1 (\alpha_0 + \theta_1 BW_t) + \theta_2 BW_t
\end{aligned}$$

The procedure above is repeated recursively to generate forecasts up to 10-months ahead. The forecasts for returns $\hat{y}_{t+1}$ 1 month, $\hat{y}_{t+3}$ 3 months and $\hat{y}_{t+10}$ 10 months are then saved.

Step 6: The saved forecasts in the previous step are then used to compute the forecast errors. The formula for the one month ahead is:

$$e_{t+1} = \hat{y}_{t+1} - y_{t+1} \quad (10)$$

This is also repeated for the 3- and 10-month horizons and the value is saved.

Step 7: The rolling window is now shifted one month, such that the period ranges from February 1984 to January 2004. Step 4 through to 6 is repeated until the end of the rolling window reaches the end of the out- of sample period.

Step 8: When each of the forecasts are completed in the previous steps and the errors are accumulated, we can then compute the RMSE and MAE using equation (8).

CHAPTER 5 Results

The following chapter presents the main findings of the research. First, we will examine the model setup using the Granger causality test and discuss the coefficients of the optimal models derived. Thereafter, the optimal models will be used to make forecasts which can be evaluated using several procedures. We will evaluate the forecast errors using plots as well as examining the statistics. The findings will allow us to draw conclusions on the performance of investor sentiment across the different sectors.

5.1 Model setup and analysis

The first step of investigating the influence of investor sentiment on returns is the Granger causality test. Thereafter, the models are estimated using the optimal lag length estimation scheme and the coefficients are reported.

5.1.1 The Granger causality test

The influence of investor sentiment on sector returns is computed across each of the sectors including the S&P 500. The statistics, along with the corresponding p-values are reported in Table 7.

The Granger causality test provides us with mixed results. As the p-values are large, we opt for a larger significance level of 10% to make comparisons easier. For the market based investor sentiment measures, BW and S^{PLS}, similar results are found. In panel A, BW has a significant effect on Energy and Industrial sector. As for Utilities and S&P500, there is a significant evidence that the sector returns cause BW. For the IT sector, causality appears to run in both directions. As for the S^{PLS} measure, causality only significantly runs in the direction of Telecom sector, whereas for the IT sector returns causality runs in the opposite direction.

Table 7: Granger Causality test

Below are results of the Granger causality test between sector returns and investor sentiment measures for in-sample period from January 1984 until December 2003. BW, CCI, OECD and S^{PLS} are sentiment indices. Null hypothesis is rejected at 10% significance level. If null hypothesis is rejected, we will accept alternative one. The Panel A shows market based sentiment measures, where BW stands for Baker and Wurgler (2006, 2007) index and S^{PLS} stands for the updated Baker and Wurgler sentiment index by the Partial Least Square method by Huang et al. (2015). Panel B represents survey based measures, namely OECD which is consumer confidence index obtained from the OECD website and the MS which is consumer confidence index constructed by the Michigan University.

Test 1: H_0: Sentiment is not Granger causal on return
Test 2: H_0: Returns is Granger causal on sentiment

Panel A:

	BW			
	Test 1		Test 2	
	F-stat	p-value	F-stat	p-value
ENE	2.197	*(0.045)*	0.945	*(0.464)*
FIN	0.666	*(0.677)*	1.570	*(0.157)*
HEA	0.738	*(0.619)*	1.524	*(0.172)*
IND	2.052	*(0.060)*	1.694	*(0.124)*
IT	2.926	*(0.009)*	1.887	*(0.085)*
MAT	0.867	*(0.521)*	1.086	*(0.374)*
TEL	1.151	*(0.337)*	1.191	*(0.315)*
UTL	0.873	*(0.515)*	3.373	*(0.003)*
CD	1.730	*(0.119)*	1.528	*(0.174)*
CS	0.446	*(0.847)*	1.360	*(0.235)*
SP500	1.533	*(0.168)*	1.885	*(0.084)*

	S^{PLS}			
	Test 1		Test 2	
	F-stat	p-value	F-stat	p-value
ENE	0.915	*(0.485)*	0.886	*(0.506)*
FIN	0.877	*(0.512)*	1.396	*(0.217)*
HEA	0.770	*(0.594)*	1.283	*(0.267)*
IND	0.428	*(0.860)*	1.204	*(0.305)*
IT	1.331	*(0.245)*	1.913	*(0.080)*
MAT	1.339	*(0.244)*	0.419	*(0.865)*
TEL	2.129	*(0.054)*	0.338	*(0.916)*
UTL	1.031	*(0.406)*	0.866	*(0.520)*
CD	1.329	*(0.248)*	0.657	*(0.684)*
CS	1.260	*(0.280)*	0.734	*(0.623)*
SP500	1.016	*(0.416)*	1.408	*(0.212)*

Panel B:

	OECD			
	Test 1		Test 2	
	F-stat	p-value	F-stat	p-value
ENE	1.177	*(0.320)*	0.363	*(0.901)*
FIN	2.778	*(0.013)*	1.074	*(0.379)*
HEA	0.750	*(0.610)*	1.432	*(0.204)*
IND	3.320	*(0.004)*	2.138	*(0.050)*
IT	2.365	*(0.031)*	2.211	*(0.044)*
MAT	1.472	*(0.192)*	0.566	*(0.757)*
TEL	0.577	*(0.748)*	2.033	*(0.065)*
UTL	0.658	*(0.684)*	0.381	*(0.891)*
CD	1.492	*(0.185)*	1.117	*(0.356)*
CS	0.515	*(0.796)*	0.855	*(0.530)*
SP500	2.947	*(0.009)*	1.403	*(0.214)*

	MS			
	Test 1		Test 2	
	F-stat	p-value	F-stat	p-value
ENE	1.464	*(0.192)*	0.739	*(0.619)*
FIN	1.230	*(0.292)*	3.103	*(0.006)*
HEA	1.050	*(0.394)*	2.378	*(0.031)*
IND	0.918	*(0.483)*	6.187	*(0.000)*
IT	1.648	*(0.136)*	4.973	*(0.000)*
MAT	1.004	*(0.425)*	1.617	*(0.147)*
TEL	0.701	*(0.649)*	1.665	*(0.134)*
UTL	0.311	*(0.931)*	0.542	*(0.776)*
CD	0.572	*(0.752)*	3.593	*(0.002)*
CS	0.332	*(0.919)*	1.200	*(0.310)*
SP500	1.277	*(0.269)*	4.612	*(0.000)*

Next, we look at the MS and OECD investor sentiment measures in the panel B of the Table 7. For the MS, causality only significantly runs in the direction of sector returns, specifically, Financials, Health, Industrial, IT, Consumer Discretionary and S&P500. Therefore, the MS measure is not found to Granger-cause sector returns. On the other hand, OECD has contradictory results. For Financial sector and S&P500, causality runs in the direction of sentiment to sector returns. Notably, for Industrials and IT sector, causality runs in both directions. Such findings are consistent with finding of Lemmon and Portniaguina (2006) who document forecasting power of investor sentiment on small and hard to value stocks. Kaplanski and Levy (2010) also document that market sentiment has the largest effect among young and fast growing sectors, such as the IT sector.

The Granger causality results at first lead to question of the influence of investor sentiment on returns, particularly for the MS index. Similar results are found by Brown and Cliff (2004) and Wang et al. (2006). However, we should keep in mind that stock returns are notoriously difficult to forecast, and therefore causality tests should be treated with some caution.

5.1.2 Coefficient evaluation

The optimal lag lengths of the sector return and investor sentiment are selected using the BIC criterion. For each industry, the model specifications are displayed in Table 8.

Table 8: Estimation coefficients

The Table 8 covers the in-sample period from January 1984 until December 2003. Estimated coefficients for each of the sectors use autoregressive terms for returns and investor sentiment. Each of the table uses the four measures of investor sentiment namely Baker and Wurgler sentiment index BW in panel A, updated Baker and Wurgler sentiment index by the partial least square method by Huang et al. (2015) S^{PLS} in panel B, the Michigan University confidence index MS in panel C and the OECD confidence index in the panel D. The coefficients are listed along with t-statistics and p-values. Significance level is determined as follows: when p-value is less than 0.01 we add ***, p value between 0.01 and 0.05 ** and p-value from 0.05- 0.1 is *. Symbol * represents the level of significance. $\beta_1$represents coeffiecients for sector returns at lag lenght 1, 2, 3 and 4, where $\theta_1$is the investor sentiment coefficient at lag length 1, 2, 3 and 4. Coefficients refer to the Equation (3).

Panel A: BW investor sentiment measure

Sector	α_0	β_1	β_2	β_3	β_4	θ_1	θ_2	θ_3	θ_4	R^2	BIC
ENE	0.344	-0.1				1.301	1.303			0.049	945
t-stat	2.341**	1.646*				1.409*	1.341*				
FIN	0.401	0.028				-2.017	0.708	2.375	-1.127	0.027	1158.878
t-stat	2.1855**	0.427				-1.715*	0.424	1.434	-0.951		
HEA	0.463	-0.078				-3.020	1.428	1.460		0.055	919.157
t-stat	2.863**	-1.119				-2.866**	0.960	1.370			
IND	0.342	-0.074				-0.970	-1.612	3.548	-1.042	0.045	1105.108
t-stat	2.096**	-1.154				-0.924	-1.086	2.395**	-0.976		
IT	0.494	-0.082	-0.034	0.072	-0.088	-1.193				0.054	1169.458
t-stat	1.934*	-1.196	-0.496	1.062	-1.294	-2.676***					
MAT	0.294	-0.110				-1.466	0.687	2.631	-2.110	0.085	720.882
t-stat	1.497	-1.394				-1.280	0.420	1.622	-1.839*		
TEL	0.194	-0.034				-0.100	-0.675			0.046	761.360
t-stat	0.878	-0.428				-0.077	-0.521				
UTL	0.194	0.017				-3.030	3.025			0.055	1017.679
t-stat	1.439	0.274				-3.480***	3.441***				
CD	0.394	0.028				-2.020	1.197	2.540	-2.112	0.077	689.339
t-stat	2.198**	0.357				-1.961*	0.818	1.735*	-2.040*		
CS	0.375	0.011	-0.014	-0.147	-0.122	0.026				0.043	658.527
t-stat	2.254**	0.142	-0.173	-1.837*	-1.502	0.110					
SP500	1.039	-0.012				-1.866	-1.604	6.446	-3.612	0.040	1407.776
t-stat	3.305***	-0.182				-0.937	-0.571	2.301**	-1.798*		

Panel B: S^{PLS} investor sentiment measure

Sector	α_0	β_1	β_2	β_3	β_4	θ_1	θ_2	θ_3	θ_4	R2	BIC
ENE	0.363	-0.111	0.014	-0.049	-0.113	-0.113				0.055	946.701
t-stat	2.402**	-1.625	0.199	-0.729	-1.684*	-0.650					
FIN	0.431	0.017	-0.025	-0.076	-0.083	-0.175				0.022	1160.146
t-stat	2.399**	0.266	-0.386	-1.172	-1.282	-0.855					
HEA	0.391	-0.047				0.107	-3.034	4.214	-1.498	0.042	923.090
t-stat	2.403**	-0.664				0.096	-1.589	2.192**	-1.321		
IND	0.361	-0.090	-0.085	-0.046	-0.101	-0.172				0.032	1108.262
t-stat	2.264**	-1.389	-1.315	-0.715	-1.566	-0.932					
IT	0.254	-0.068	-0.021	0.078	-0.085	-0.650				0.042	1172.183
t-stat	1.022	-0.995	-0.305	1.138	-1.229	-2.089**					
MAT	0.353	-0.142	-0.028	-0.089	-0.161	-0.323				0.087	720.422
t-stat	1.813*	-1.789*	-0.355	-1.115	-2.042**	-1.459					
TEL	0.121	-0.043	-0.035	0.028	-0.092	-0.753				0.073	758.537
t-stat	0.561	-0.539	-0.436	0.354	-1.159	-2.842***					
UTL	0.169	0.017	-0.056			-0.079				0.011	1028.540
t-stat	1.290	0.271	-0.867			-0.513					
CD	0.453	0.028	-0.132	0.000	-0.150	-0.391				0.067	691.071
t-stat	2.480**	0.357	-1.655*	0.004	-1.867*	-1.896*					
CS	0.405	-0.003	-0.022	-0.151	-0.124	-0.256				0.056	656.456
t-stat	2.511**	-0.036	-0.274	-1.898*	-1.542	-1.448					
SP500	1.073	-0.036	-0.078	-0.045	-0.132	-0.872				0.046	1406.307
t-stat	3.430***	-0.563	-1.214	-0.701	-2.042**	-2.441**					

Panel C: MS investor sentiment measure

Sector	α_0	β_1	β_2	β_3	β_4	θ_1	θ_2	θ_3	θ_4	R2	BIC
ENE	0.383	-0.114	0.007	-0.055	-0.111	0.051				0.060	945.523
t-stat	2.556**	-1.677*	0.103	-0.808	-1.653*	1.267					
FIN	0.395	-0.033				0.178	0.056	-0.086	0.002	0.076	1146.773
t-stat	2.339**	-0.490				3.630***	1.140	-1.802*	0.035		
HEA	0.425	-0.065				0.139	-0.025	-0.065		0.069	916.203
t-stat	2.695***	-0.931				3.234***	-0.565	-1.549			
IND	0.340	-0.171				0.236	0.112	-0.070	-0.051	0.162	1074.254
t-stat	2.356**	-2.653***				5.682***	2.568**	-1.705*	-1.249		
IT	0.365	-0.143				0.279	0.258	-0.072	-0.039	0.144	1148.087
t-stat	1.578	-2.087**				4.276***	3.867***	-1.092	-0.622		
MAT	0.308	-0.145	-0.027	-0.079	-0.156	0.137				0.118	715.282
t-stat	1.620	-1.861*	-0.340	-1.021	-2.024**	2.725***					
TEL	0.061	-0.005				0.128				0.034	762.460
t-stat	0.285	-0.058				2.269**					
UTL	0.177	0.012	-0.055			0.030				0.013	1028.150
t-stat	1.359	0.187	-0.858			0.808					
CD	0.327	-0.037				0.207	0.069			0.131	677.458
t-stat	1.980**	-0.465				4.621***	1.497				
CS	0.359	0.001	-0.007	-0.145	-0.100	0.100				0.078	652.795
t-stat	2.264**	0.007	-0.094	-1.838*	-1.252	2.419**					
SP500	0.977	-0.097				0.387	0.177	-0.129	-0.061	0.118	1387.819
t-stat	3.441***	-1.465				4.767***	2.123*	-1.613	-0.773		

Panel D: OECD investor sentiment measure

Sector	α_0	β_1	β_2	β_3	β_4	θ_1	θ_2	θ_3	θ_4	R^2	BIC
ENE	0.337	-0.124				2.464	-1.801			0.063	942.390
t-stat	2.354**	-1.851*				2.288**	-1.687*				
FIN	0.410	-0.054				5.427	0.083	-4.021	2.051	0.126	1133.465
t-stat	2.501**	-0.842				3.149***	0.028	-1.355	1.194		
HEA	0.421	-0.056				3.771	-2.596			0.057	917.421
t-stat	2.664***	-0.816				3.267***	-2.248**				
IND	0.343	-0.174				3.828	3.368	-5.514	2.018	0.178	1069.635
t-stat	2.402**	-2.720***				2.530**	1.276	-2.124**	1.342		
IT	0.367	-0.138				2.477	10.499	-12.166	5.515	0.170	1141.637
t-stat	1.610	-2.073**				1.055	2.575**	-2.989***	2.346**		
MAT	0.298	-0.177	-0.051	-0.089	-0.145	2.684				0.128	713.551
t-stat	1.574	-2.250**	-0.649	-1.157	-1.902*	3.042***					
TEL	0.005	-0.008				2.181	2.210	-6.458	5.099	0.067	759.544
t-stat	0.024	-0.097				1.010	0.586	-1.656*	2.219**		
UTL	0.168	0.009				1.990	-1.607			0.023	1025.647
t-stat	1.300	0.141				1.978**	-1.605				
CD	0.357	-0.048	-0.134	0.029	-0.077	3.790				0.170	673.090
t-stat	2.103**	-0.618	-1.792*	0.393	-1.022	4.809***					
CS	0.354	-0.011	-0.014	-0.136	-0.113	1.568				0.072	653.776
t-stat	2.226**	-0.144	-0.182	-1.714*	-1.417	2.199**					
SP500	0.983	-0.096				7.993	4.725	-10.567	5.321	0.154	1377.754
t-stat	3.538***	-1.492				2.767***	0.940	-2.126**	1.846*		

In Table 8, it can be noted that considerable numbers of lag length (q>1) are needed with the exception of the IT and Consumer Staples sector. For example, for the Financial sector, it can be observed that the optimal number of lags for the returns is 1, and investor sentiment is 4. This may suggest that the investor sentiment is including superior information multiple periods in the past. In terms of R-squared, the best performing models with BW as independent variable are Materials and Consumer Discretionary sector of 0.085 and 0.077 respectively. The poorest performing model is the Financial sector with an R-squared of 0.027.

On the contrary, the other market based measure, S^{PLS} reported in the panel B of the Table 8, finds evidence strongly in favor of returns. Apart from the Health sector, each of the sector returns have 4 lags for the return data and only 1 lag for the investor sentiment. This may indicate weak information coming from investor sentiment as the returns are more informative in the optimal model. Further, we find R-squared that falls into a similar range as the previous BW table.

Mixed results are observed for panel C in the Table 8. Energy, Materials and Consumer Staples sectors results show 4 lags for the return series and only 1 lag for the investor sentiment with R-squares in the range of 0.01 and 0.118. On the other hand, for Financial, Health, and Industrial sector results suggest the contrary with R-squares ranging from 0.118 to 0.162. The difference is very large, suggesting superior in-sample fit of the investor sentiment models. Finally, OECD reports similar results to MS, with the highest R-squares corresponding to the models with multiple lags for the investor sentiment measures.

5.2 Forecast performance analysis

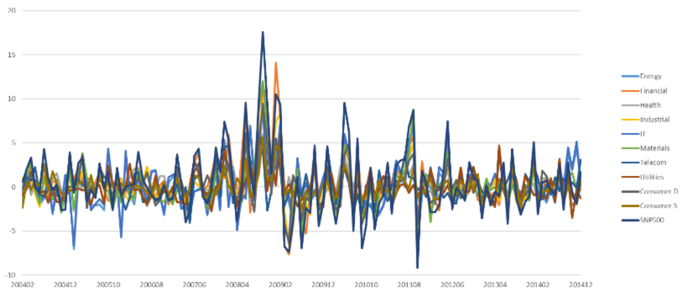
The optimal models developed in the previous section along with the models for investor sentiment reported in Appendix B are used to make forecasts 1-, 3- and 10-months ahead using a rolling window. The forecasts are then compared to the actual returns allowing us to compute forecast errors. These error plots are evaluated in the following section along with the corresponding statistics.

5.2.1 Forecast error plots

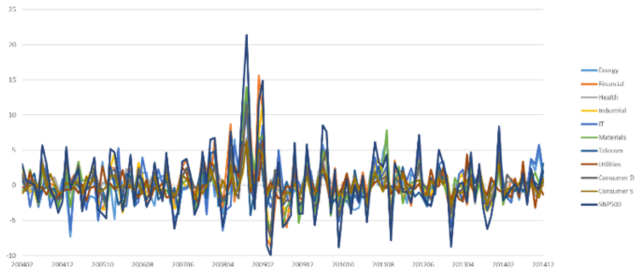
The errors for each of the investor sentiment models up to 1-month ahead are plotted in Figure 3. Generally, spikes are seen in the errors during periods of financial turmoil. Particularly, in 2008 and 2009, the largest errors occur. Visually, it could be noticed that during the 2008 crisis, the forecast error for MS is larger than 20%, while it is between 16 and 19% for the other investor sentiment measures. It is to be expected that the largest errors would occur during the financial crisis, especially among unprofessional individuals. Meaning, the consumer confidence level (MS) indicates that consumers in the U.S. during the financial crisis were more affected by the crisis and thus, less accurate in formulating their expectations of the market performance. This could be due to the asymmetric information among uninformed individuals when making an investment decision. A similar result is observed for the 3-month and 10-month ahead forecasts in Appendix A.

Figure 3: Out-of-sample error plot

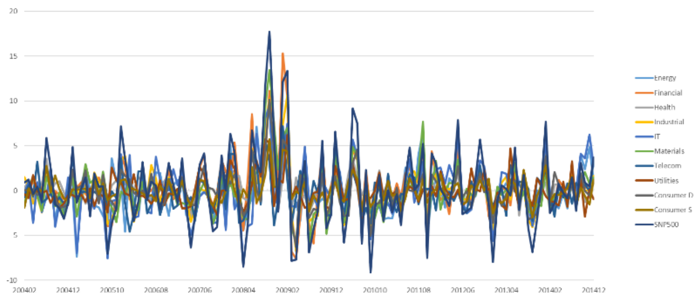
1-month forecast horizons is shown in graphs a, b, c and d for BW, MS, OECD and S^{PLS} respectively. The forecast period spans over the January 2004 until December 2014. The graph plots the errors which are calculated as the difference between the forecasted and actual returns for each of the sectors.



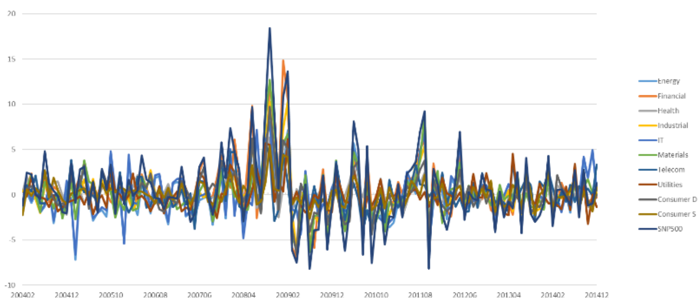
(a) BW



(b) CCI



(c) OECD



(d) S^{PLS}

5.2.2 Forecast evaluation statistics

Further, we observe that errors differ only slightly per model. To extensively examine the errors, the RMSE and MAE are being computed to draw a fair comparison. Table 9 reports the forecast statistics. Generally, the RMSE and MAE lead to the same conclusion, and for this reason, we stick to the RMSE when carrying out the evaluation.

Table 9: Forecast errors

In this table RMSE and MAE are reported, where RMSE and MAE are root mean square error and absolute mean error respectively. Numbers describe the forecast performance of the models, a smaller RMSE and MAE suggest a better predictability. Each of the performance measures are computed for 1-, 3- and 10-month horizon. Out-of-sample forecasting period spans over the January 2004 until December 2014. Panel A reports the Baker and Wurgler sentiment index, BW, panel B reports S^{PLS} which is the updated version of the Baker and Wurgler sentiment index, panel C reports the Michigan University sentiment index, MS, and panel D reports the OECD sentiment index.

	Forecast Horizon									
Panel A: BW sentiment measure										
ENE				**FIN**				**HEA**		
	1m	3m	10m		1m	3m	10m		1m	3m	10m
RMSE	6.751	6.768	7.022		9.240	9.554	10.011		2.737	2.676	2.729
MAE	1.951	1.960	2.009		2.0226	2.062	2.1254		1.232	1.227	1.234
IND				**IT**				**MAT**			
RMSE	5.540	5.435	5.681		7.063	7.088	7.069		6.827	6.601	6.844
MAE	1.614	1.618	1.663		1.981	1.990	1.992		1.833	1.828	1.857
TEL				**UTL**				**CD**			
RMSE	3.671	3.700	3.849		2.945	2.971	3.059		4.814	4.913	5.119
MAE	1.489	1.499	1.538		1.278	1.285	1.304		1.632	1.653	1.695
CS				**SP500**							
RMSE	1.832	1.810	1.745		16.507	16.675	17.372				
MAE	1.013	1.006	1.030		2.959	2.991	3.082				

Table 9 continued**:**

Panel B: S^{PLS} sentiment measure

ENE				**FIN**			**HEA**		
	1m	3m	10m	1m	3m	10m	1m	3m	10m
RMSE	7.003	7.043	6.998	9.699	9.958	9.709	2.634	2.612	2.656
MAE	1.999	2.013	2.008	2.005	2.044	2.056	1.194	1.184	1.189
IND				**IT**			**MAT**		
RMSE	5.914	5.749	5.645	6.785	6.881	6.841	7.583	7.160	6.804
MAE	1.619	1.617	1.635	1.976	1.986	1.984	1.948	1.905	1.872
TEL				**UTL**			**CD**		
RMSE	3.756	3.825	3.881	2.769	2.793	2.878	5.098	5.299	5.068
MAE	1.501	1.514	1.536	1.210	1.227	1.248	1.694	1.720	1.702
CS				**SP500**					
RMSE	1.767	1.749	1.677	17.539	17.502	16.843			
MAE	1.000	0.995	1.009	3.007	3.020	3.029			

Panel C: MS sentiment measure

ENE				**FIN**			**HEA**		
	1m	3m	10m	1m	3m	10m	1m	3m	10m
RMSE	7.050	7.032	7.006	10.466	10.459	10.516	3.111	3.079	3.110
MAE	2.015	2.025	2.057	2.160	2.186	2.192	1.328	1.322	1.327
IND				**IT**			**MAT**		
RMSE	8.119	7.621	7.572	9.572	9.342	9.478	8.045	7.676	7.173
MAE	2.077	2.065	2.047	2.406	2.399	2.430	1.996	1.986	1.967
TEL				**UTL**			**CD**		
RMSE	4.358	4.409	4.461	2.843	2.934	2.999	5.885	5.914	5.879
MAE	1.604	1.624	1.650	1.230	1.258	1.281	1.825	1.839	1.818
CS				**SP500**					
RMSE	2.101	2.119	2.063	22.320	21.829	21.963			
MAE	1.075	1.083	1.109	3.539	3.520	3.536			

Table 9 continued

Panel D: OECD sentiment measure

ENE					**FIN**				**HEA**		
		1m	3m	10m		1m	3m	10m	1m	3m	10m
RMSE		6.871	6.813	7.055		9.717	9.677	9.698	2.682	2.669	2.717
MAE		1.991	2.003	2.075		2.074	2.114	2.118	1.243	1.248	1.262
IND					**IT**				**MAT**		
RMSE		6.974	6.552	6.511		9.564	9.345	9.342	7.835	7.278	6.923
MAE		1.937	1.907	1.880		2.401	2.377	2.385	1.975	1.952	1.960
TEL					**UTL**				**CD**		
RMSE		4.703	4.760	4.737		2.971	3.051	3.117	4.907	5.233	5.049
MAE		1.681	1.701	1.702		1.294	1.315	1.337	1.630	1.709	1.676
CS					**SP500**						
RMSE	1.894	1.898	1.911			20.026	19.772	19.887			
MAE	1.040	1.047	1.093			3.307	3.317	3.329			

BW sample suggests that out of the ten sector return models, Consumer Staples sector has a smaller error for the 10-month horizon and thus performs better over longer forecast horizons. For the remaining models, we notice that shorter term forecast horizons perform better. The largest error corresponds to the S&P500 at all forecast horizons. While Utilities and Health sectors have the smallest RMSE over all forecast horizons, indicating superior performance.

For the OECD sentiment measure, we find slightly more outperformance at longer (3- and 10-months ahead) forecasts horizons. Of the ten sectors, 4 outperform at forecast horizons larger than 1-month. The Consumer Staples sector reports the smallest errors, while once again the S&P500 has the largest errors. Finally, for the MS and S^{PLS} measures, more than half of the sectors have lower errors at larger forecast horizons. The Consumer Staples sector has the lowest corresponding errors, while the S&P500 has the largest.

To conclude, it can be observed that many models perform better over longer horizons, 3 or 10 months. A possible explanation to this may be the inclusion of the investor sentiment measures. We observe that for the OECD sentiment measures, Industry and IT sector outperform over longer forecast horizons. Closely examining the coefficients of these models, we observe that investor sentiment had four lags included in the model. This could be a possible explanation to the superior forecast performance over longer horizons. Furthermore, given that those are the sectors containing extreme growth and relatively young firms, the results fall in line with the conclusion of Baker and Wurgler (2006, 2007), Baker et al. (2012) and Fink et al. (2010). Such firms are more sensitive to investor sentiment and thus investors are more likely to speculate. Also, the S&P500 outperforms over longer horizons and has four lags for investor sentiment. This suggests that investor sentiment has a delayed influence on future stock returns. Interestingly, we note that S&P 500 has smaller errors over the longer horizons for the MS and OECD investor sentiment proxies, which does not fall in line with Fisher and Statman (2003) results. They argue that the consumer confidence measures do not have the forecasting power over the S&P 500, because firms that fall under the S&P 500 are easy to value and mature firms, only attracting a rational and experienced investor.

5.3 Random Walk Forecast Evaluation

Assessing the Random walk comparison results, we see the MS model significantly underperforms the Random Walk model at the 1-month ahead horizon. This can be observed by the significant positive Diebold-Mariano statistics, indicating outperformance of the Random walk model. As for the BW and MS at the 1-month horizon, there are several cases were the Random walk outperforms the out-of-sample model. For longer forecast horizons, the results tend to vary. Outperformance of the out-of-the sample model versus the Random walk model can be seen for the IT sector at the 10-month horizon when examining the BW sentiment measure. This result would suggest that the IT sector consists of firms that are more affected by the investor sentiment than the others. For the rest of the sectors we observe negative Diebold-Mariano statistics which would suggest outperformance of the out-of-the sample models, however, it is not statistically significant. Generally, we can conclude that forecast performance between the two models is similar. The results suggest that our models, in almost all cases, does not perform better than the random forecasts. This is not surprising, as the Random walk models are notoriously difficult to outperform as observed in the literature.

Table 10: Diebold-Mariano z-statistic of the Random Walk forecasts

This table contains the Diebold-Mariano statistics of the differential between the squared errors of the BW, MS, OECD and S^{PLS} models with the Random Walk Model. The difference is computed across each of the ten sectors and forecast horizons of 1, 3 and 10-months. Under the null hypothesis, there is no significant difference in forecast performance. Under the alternative hypothesis, a positive z-statistic indicates an outperformance of the Random walk. A negative statistic indicates the outperformance of the investor sentiment model. Significance level is determined as follows: when p-value is less than 0.01 we add ***, p value between 0.01 and 0.05 ** and p-value from 0.05- 0.1 is *. Symbol * represents the level of significance.

	BW			MS		
	1-month	3-month	10-month	1-month	3-month	10-month
ENE	1.696*	-1.053	-0.780	1.483*	1.225	-0.359
FIN	0.514	-0.363	-0.468	1.943**	1.133	3.093***
HEA	0.973	0.196	-0.120	2.851***	2.486***	1.806**
IND	0.407	-0.663	-0.593	2.016**	1.189	0.840
IT	1.935**	1.420*	-1.599*	3.384***	2.596***	1.621*
MAT	0.698	-0.286	-0.334	1.972**	1.901**	0.712
TEL	0.836	-0.212	-0.755	2.446***	1.587*	1.544*
UTL	0.794	-0.953	-0.909	1.699**	1.476*	1.481*
CD	0.257	-0.157	-0.373	1.583*	1.381*	1.158
CS	0.528	0.358	-0.672	2.152**	1.985**	2.222**
SP500	0.786	0.137	-0.245	2.239**	1.627*	1.572*

	OECD			S^{PLS}		
	1-month	3-month	10-month	1-month	3-month	10-month
ENE	1.638*	-0.970	-0.626	0.285	0.166	-1.051
FIN	2.598***	1.756**	1.330*	1.791**	1.360*	1.180
HEA	0.726	-0.300	-0.719	0.777	-0.200	-0.472
IND	2.025**	0.564	0.224	1.371*	0.734	-0.325
IT	1.487*	0.820	0.546	2.378***	2.461***	1.545*
MAT	1.825**	1.572*	-0.006	1.697**	1.107	-0.865
TEL	0.109	-0.148*	-0.361	2.043**	2.613***	1.308*
UTL	0.744	-0.470	-0.571	1.247	1.146	0.571
CD	0.557	1.163	0.043	1.713*	1.914**	-0.393
CS	2.396***	2.572***	1.888**	0.690	0.378	-1.509*
SP500	2.406***	1.697**	0.980	1.735**	1.263	-0.334

5.4 Robustness check

Robustness checks are used to evaluate the models in the research process. However, in our paper we do not perform direct robustness check because of two reasons. Firstly, as a measure of investor sentiment we utilized in total four different proxies: two market based and two survey based investor sentiment measures. By employing four of them, we are able to assess which investor sentiment proxy is better able to predict the stock returns. This wide range does not include any selective bias. Secondly, the setup of the models uses the Bayesian information criterion to select the optimal lag length for each of the models. Therefore, we do not constrain the number of parameters in the model and the selection procedure remains objective. To sum up, given the varied data set and objective model selection criteria, no biases are introduced in the forecasting procedure.

CHAPTER 6 Conclusion

In this study, we assessed the forecasting power of investor sentiment on stock returns. The focus was on the sectoral stock returns within the U.S. that are classified according to GICS. The sample was split into an in-sample period, which was used to estimate the model parameters. The out-of-sample period applied a rolling window to continuously forecast the returns. The models were successfully estimated in the in-sample period using the Bayesian Information Criterion.

By examining different sectors, we observe similar results as have already been found in the literature. Firstly, sectors such as IT, Industries and Consumer Staples fall under the relatively volatile and unstable environments due to its constituent firms. For this reason, these sectors have a closer relationship between the investor sentiment and the stock returns. This is particularly the case for survey based investor sentiment measures, MS and OECD, where we observe stronger predicting power. Secondly, the results indicate stronger predictability over the longer forecast horizons, precisely 3- and 10-months. The superior forecasting performance over the 3- and 10-month horizons is an interesting relationship which has not been documented in the literature. Although the literature provides evidence on the predictability of the investor sentiment, it was not clear over the which horizon its effect is mostly pronounced. Brown and Cliff (2005) among others, document strong forecasting power of the investor sentiment over the longer horizons, e.g. 12- to 24-months. However, we focus on shorter horizons; 1-, 3- and 10-month forecast horizons and conclude that the investor sentiment demonstrates the forecasting power across the sectors.

Each of the forecasted return models are compared with a Random walk model. In almost all cases, the Random walk model performs equally well, or even better than the forecasted model. This would suggest that the returns are random, and this has been observed extensively throughout the research. From a practical perspective, asset managers should opt for the Random walk model when forecasting returns. However, this does not deem our results irrelevant. Investor sentiment influences sectors in different manners, particularly when forecasting 3- and 10-month horizons and this what we have been able to clearly highlight. Yet, it is still to define how to measure the exact amount of the sentiment present among the investors.

Possible limitation to our study is that we used the estimation sample to estimate the coefficients, instead of continuously re-estimating the model at each forecast step. This could be an area of future research. Moreover, as we have already discussed, the investor sentiment is extremely difficult to measure and thus how to quantify its effects is still unknown. Therefore, finding a sector related sentiment measure that captures sector specific trading patterns and stock movements would be a potential area for further research.

REFERENCES

Antoniou, C, Doukas, J.A. and Subrahmanyam, A. (2013). Cognitive dissonance, sentiment, and momentum. *Journal of Financial and Quantitative Analysis* 48 (1), 245-275.

Bachelier, L. (2011). *Louis Bachelier's theory of speculation: the origins of modern finance*. Princeton University Press.

Baker, M., Wurgler, J., and Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of Financial Economics*, *104 (2),* 272-287.

Baker, H. K., and Nofsinger, J. R. (Eds.). (2010). *Behavioral finance: investors, corporations, and markets*. John Wiley & Sons.

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

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, *61*(4), 1645-1680.

Barberis, N., and Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, 1(part B*)*, 1053-1128.

Bathia, D., and Bredin, D. (2013). An examination of investor sentiment effect on G7 stock market returns. *The European Journal of Finance*, *19*(9), 909-937.

Black, F. (1986). Noise. *The Journal of Finance*, *41*(3), 528-543.

Borges, M. R. (2010). Efficient market hypothesis in European stock markets. *The European Journal of Finance*, *16*(7), 711-726.

Brown, G. W., and Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of Empirical Finance*, *11(1),* 1 –27.

Brown, G. W., and Cliff, M. T. (2005). Investor sentiment and asset valuation. *Journal of Business* 78(2), 405–40.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57-82.

Charoenrook, A. (2005). Does sentiment matter. Unpublished working paper. *Vanderbilt University*.

Chen, N. F., Roll, R., and Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business*, *59(3),* 383-403.

Da, Z., Engelberg, J. and Gao, P. (2015). The sum of all fears: Investor sentiment and asset prices. *Review of Financial Studies, 28(1),* 1–32.

DeBondt, W. F., and Thaler, R. H. (1995). Financial decision-making in markets and firms: A behavioral perspective. *Handbooks in Operations Research and Management Science*, *9 (13)*, 385-410.

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

Fama, E. F., and French, K. R. (1996). The CAPM is wanted, dead or alive. *The Journal of Finance*, *51(5),* 1947-1958.

Fama, E.F., and French, K.R. (1993), Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33(1),* 3-56.

Fama, E. F. (1970). Multiperiod consumption-investment decisions. *The American Economic Review*, *60(1),* 163-174.

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

Ferrer, E., Salaber, J., and Zalewska, A. (2016). Consumer confidence indices and stock markets' meltdowns. *The European Journal of Finance*, *22(3),* 195–220.

Fink, J., Fink, K. E., Grullon, G., and Weston, J. P. (2010). What drove the increase in idiosyncratic volatility during the internet boom?. *Journal of Financial and Quantitative Analysis*, 45(5), 1253-1278.

Fisher, K. L., & Statman, M. (2003). Consumer confidence and stock returns. *The Journal of Portfolio Management*, *30*(1), 115-127.

Fisher, K. L., and Statman, M. (2000). Investor sentiment and stock returns. *Financial Analysts Journal*, *56(2),* 16–23.

Gan, C., Lee, M., Hwa, A. Y., and Zhang, J. (2005). Revisiting Share Market Efficiency: Evidence from the New Zealand Australia, US and Japan Stock Indices. *American Journal of Applied Sciences*, 2(5), 996- 1002.

Grossman, S. J., and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, *70*(3), 393-408.

Hill, R. C., Griffiths, W. E., and Lim, G. C. (2008). *Principles of Econometrics,* Hoboken, Wiley.

Ho, C., and Hung, C. H. (2009). Investor sentiment as conditioning information in asset pricing. *Journal of Banking & Finance*, *33*(5), 892-903.

Hribar, P., and McInnis, J. (2012). Investor sentiment and analysts' earnings forecast errors. *Management Science*, *58*(2), 293-307.

Huang, D., Jiang, F., Tu, J. and Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns, *Review of Financial Studies* 28, 791–837.

Huang, C., Yang, X., Yang, X., and Sheng, H. (2014). An empirical study of the effect of investor sentiment on returns of different industries. *Mathematical Problems in Engineering*, *2014(1),* 1-11

Jansen, W. J., and Nahuis, N. J. (2003). The stock market and consumer confidence: European evidence. *Economics Letters*, *79*(1), 89-98.

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

Kadilli, A. (2015). Predictability of stock returns of financial companies and the role of investor sentiment: A multi-country analysis. *Journal of Financial Stability, 21(C),* 26-45.

Kahneman, D., and A. Tversky. (1974). Judgment under uncertainty: Heuristics and biases. Science, *185(4157),* 1124–1131.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, *47(2),* 263-291.

Kaplanski, G., and Levy, H. (2010). Sentiment and stock prices: The case of aviation disasters. *Journal of Financial Economics*, 95(2), 174-201.

Lee, C., Shleifer, A. and Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle. *Journal of Finance, 46(1),* 75–109.

Lee, W. Y., Jiang, C. X., and Indro, D. C. (2002). Stock market volatility, excess returns, and the role of investor sentiment. *Journal of Banking and Finance*, *26(12),* 2277– 2299.

Lemmon, M., and Portniaguina, E. (2006). Consumer confidence and asset prices: Some empirical evidence. *The Review of Financial Studies*, *19*(4), 1499-1529.

Lo, A. W., and MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *The Review of Financial Studies*, *1*(1), 41-66.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, *47(1),*13-37.

Ludvigson, S. C. (2004). Consumer confidence and consumer spending. *The Journal of Economic Perspectives*, *18*(2), 29-50.

Ludvigson, S. C., and Ng, S. (2007). The empirical risk–return relation: A factor analysis approach. *Journal of Financial Economics*, *83*(1), 171-222.

Naik, P. K., and Padhi, P. (2016). Investor sentiment, stock market returns and volatility: evidence from National Stock Exchange of India. *International Journal of Management Practice*, *9*(3), 213-237.

Neal, R., and S. Wheatley. (1998). Do measures of investor sentiment predict stock returns? *Journal of Financial and Quantitative Analysis* 33(4), 523–547.

Newey W. and West K. (1987). A simple, positive semi-definite, hetroskedasticity and autocorrelation consistent covariance matrix, *Econometrica, 55(3),* 703-708.

Pearson, K. (1905). The problem of the random walk. *Nature*, 72(1865), 294.

Qiu, L., and Welch, I. (2006). Investor Sentiment Measures, *Brown University.*

Regnault, J. (1863). *Calcul des chances et philosophie de la bourse*. Mallet-Bachelier.

Salhin, A., Sherif, M., and Jones, E. (2016). Managerial sentiment, consumer confidence and sector returns. *International Review of Financial Analysis*, 47 (2016), 24-38.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *IMR; Industrial Management Review (pre-1986)*, *6*(2), 41.

Schmeling, M. (2009). Investor Sentiment and Stock Returns: Some International Evidence. *Journal of Empirical Finance 16(3),* 394–408.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 425-442.

Shiller, R. J. (1981). The use of volatility measures in assessing market efficiency. *The Journal of Finance*, *36*(2), 291-304.

Shefrin, H. (2008). *A behavioral approach to asset pricing*. Academic Press.

Simon, D. P., and Wiggins, R. A. (2001). S&P futures returns and contrary sentiment indicators. *Journal of Futures Markets*, *21*(5), 447-462.

Smith, A. (1937). The wealth of nations [1776].

Solt, M. E., and Statman, M. (1988). How useful is the sentiment index? *Financial Analysts Journal*, *44(5),* 45– 55.

Stambaugh, R. F., Yu, J., and Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, *104*(2), 288-302.

Yu, J., and Y. Yuan. (2011). Investor sentiment and the mean-variance relation*. Journal of Financial Economics 100(2)*, 367–381.

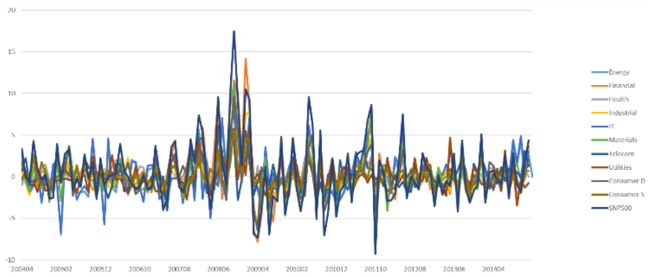
Wang, C. (2001). Investor sentiment and return predictability in agricultural futures markets. *Journal of Futures Markets*, *21(10),* 929–952.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Nelson Education.

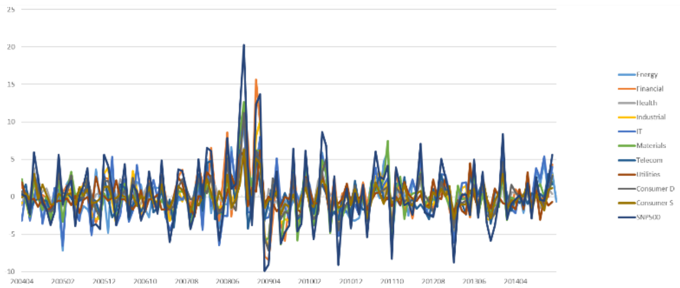
APPENDIX A

Figure 4: 3-month ahead forecasts

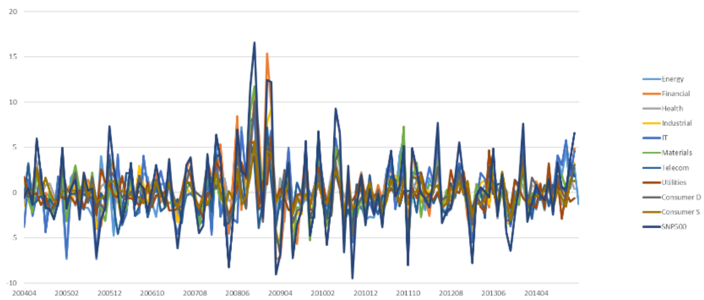
3-month forecast horizons is shown in graphs a), b), c) and d) for BW, MS, OECD and S^{PLS} respectively. The forecast period spans over January 2004 until December 2014. The graph plots the errors which is calculated as the difference between the forecasted and actual returns for each of the sectors



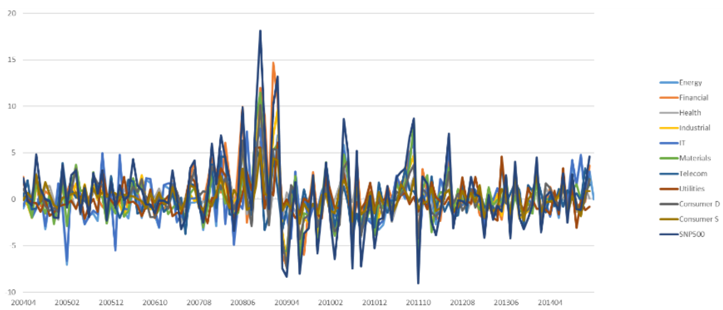
(a) BW



(b) CCI



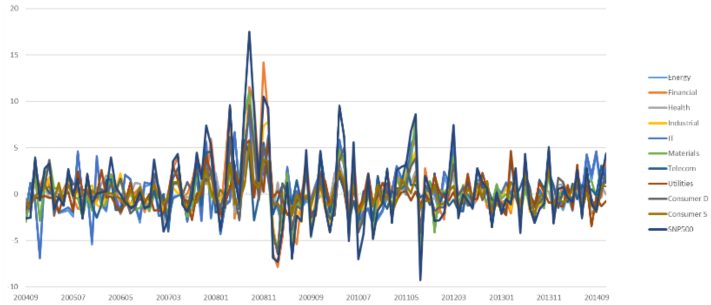
(c) OECD



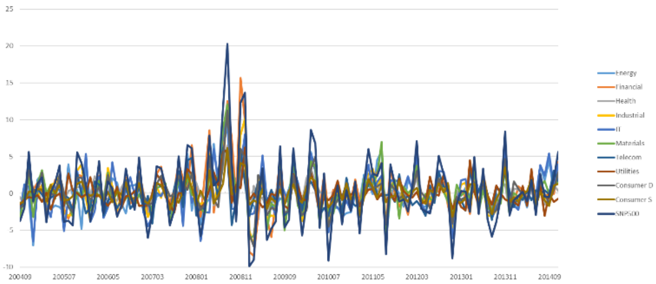
(d) S^{PLS}

Figure 5: 10-month ahead forecasts

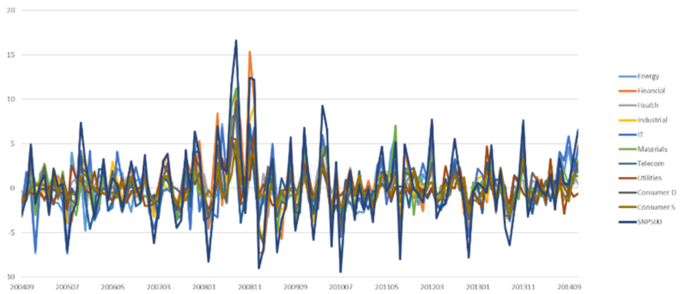
10-month forecast horizons is shown in graphs a, b, c and d for BW, MS, OECD and S^{PLS} respectively. The forecast period spans over January 2004 until December 2014. The graph plots the errors which is calculated as the difference between the forecasted and actual returns for each of the sectors



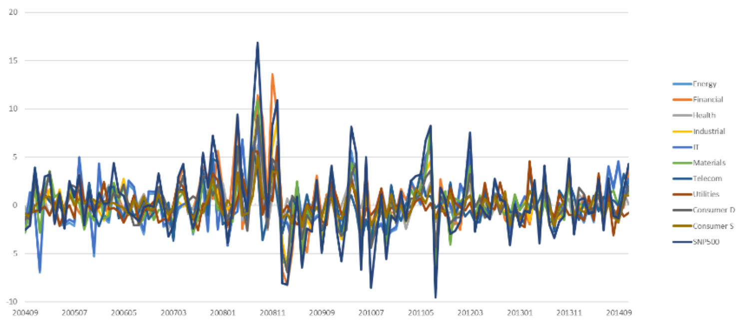
(a) BW



(b) CCI



(c) OECD



S^{PLS}

Appendix B

Table 11: Correlation coefficients

The table 11 reports correlation coefficients between the four investor sentiment for the sample period from January 1984 until December 2014. BW and S^{PLS} are market based sentiment indicator, where BW is the Baker and Wurgler (2006, 2007) sentiment index, S^{PLS} is Huang et al. (2015) updated version of the Baker and Wurgler sentiment index based on the partial least square method, MS is the Michigan University sentiment measure based on consumer confidence and OECD is another consumer confidence survey based sentiment indicator from OECD database.

	BW	CCI	OECD	S^{PLS}
BW	1	-0.160	-0.279	0.806
CCI	-0.160	1	0.722	-0.137
OECD	-0.279	0.722	1	-0.281
S^{PLS}	0.806	-0.137	-0.281	1

Table 12: Estimation coefficients Investor Sentiment

The table covers the estimation period from January 1984 until December 2003. Estimated coefficients for each Investor sentiment uses up to four lags. The table reports four measures of investor sentiment namely BW, PLS, MS and OECD. The coefficients are listed along with t statistics and p -values. Significance level is determined as follows: when the p-value is less than 0.01 we add ***, p value between 0.01 and 0.05 ** and p- value from 0.05- 0.1 is *. Symbol * represents the level of significance.

Investor sentiment	α_0	θ_1	θ_2	θ_3	θ_4	R^2	BIC
BW	0.012	0.985				0.994	-231.379
t-stat	*1.249*	*69.556*					
CCI	-0.054	-0.061	-0.022	-0.0296	-0.063	-1.914	1287.619
t-stat	*-0.238*	*-0.941*	*-0.352*	*-0.455*	*-0.988*		
OECD	-0.000	1.313	-0.670			0.997	-426.665
t-stat	*-0.095*	*27.385****	*-14.034****				
PLS	0.000	1.408	-0.241	-0.200		0.995	-233.852
t-stat	*0.087*	*22.167****	*-2.192****	*-3.110****			