

# Investor sentiment and short selling: do animal spirits sell stocks?

A Thesis

Presented to

The Faculty of the Department of Financial Economics

Radboud University

In Partial Fulfilment

Of the Requirements for the Degree of

Master of Science

By

Paul Reesink

July, 2020

## Abstract

This paper examines the relationship between investor sentiment and short selling. This is done by calculating daily investor sentiment for every stock in the S&P 500 between January 1990 and December 2019 and group the stocks into four portfolios based on their investor sentiment. The effect of relative overpricing of stocks on their short interest ratio is compared for the different sentiment portfolios. Main result is that the relationship between relative overpricing and the short interest ratio is strongly positive for the highest investor sentiment portfolio and becomes less positive when relative investor sentiment declines. Taking economic recession periods into account does not seem to change the relationship between investor sentiment and short selling, although the effect of the relative overpricing on the short interest ratio becomes less positive for all sentiment portfolios.

*Keywords:* Behavioural finance, firm-specific investor sentiment, short interest ratio, return anomalies

## Table of contents

	<b>Page</b>
Abstract	2
Table of contents	3
Chapter	
I.    Introduction	4
II.   Literature	5
III.  Research method	12
IV.  Results	17
V.   Conclusion and discussion	20
References	23

## I. Introduction

In classical finance, with rational investors and full information, deviations from fundamental value would be immediately exploited by arbitrageurs who bring back the value to their fundamentals. However, episodes like the Great Depression, Black Monday and the dotcom-bubble make it unlikely that stock prices are entirely based on fundamentals. There seem to be periods of systematic overpricing caused by overoptimistic investors who deviate the prices away from fundamental value. This theory of investor sentiment got increasing attention over the years. A key assumption in this framework is that short sellers are not able to exploit these overpricing because of short selling constraints. However, much of the investor sentiment literature focuses on the impact on future returns and take this assumption as given. In this paper I therefore want to zoom in to these short-selling constraints and especially if it becomes more difficult to short-sell as investor sentiment increases.

I investigate the effect of investor sentiment on short-selling in the S&P 500 between 1990-2019. To do this I rank the firms in portfolios based on the firm-specific investor sentiment indicator of Seok et al. (2019a) and divide them in four daily sentiment portfolios. I then calculate the relative overpricing of every stock based on the average of the nine groups of anomalies in Stambaugh et al. (2012). I calculate the relative overpricing for every month in the sample and then perform a panel regression with this relative overpricing and the short interest ratio for the different sentiment portfolios. If investors are purely rational one would expect that there is a positive correlation between the level of overpricing and the short interest ratio, as short-selling becomes more profitable the further stock prices deviate away from fundamental value, independent from the level of investor sentiment. However if investors are prone to behavioural biases and other short-selling constraints one would expect a different picture. In that case it would become more difficult to take an opposing view if overall sentiment becomes more optimistic. One would expect therefore that for the portfolio with the high investor sentiment stocks the correlation between investor sentiment and relative overpricing is negative while this is positive for the portfolio with the low investor sentiment stocks.

Contradicting the hypotheses I find a positive relationship between the relative overpricing of stocks and the short interest ratio in the highest investor sentiment portfolio. This means that if investor sentiment is relatively high short selling increases with the relative overpricing of a stock which is in line with the assumption of rational investors. Furthermore, for the lower sentiment portfolios this relationship is less positive. This means that if investor sentiment for a stock becomes relatively lower, people will less actively short-sell the stock if it gets relatively more overpriced. This is an interesting result. According to behavioural finance one would expect that if investor sentiment decreases the amount of sentiment-driven traders decreases which would increase short-selling because sentiment traders do not sell short (Barber and Odean, 2006). Also, following Barberis et al. (1997), high investor sentiment will increase noise traders their further expectations about stock prices which makes it more difficult for rational investors to short-sell. However, the results point to another reaction. They indicate that short-selling activity is the highest for high-sentiment stocks and decreases with sentiment although it stays positive for highest three sentiment portfolio.

A possible explanation would be that investor sentiment is measured in relative terms and not in absolute terms. This means that in times of low (high) overall investor sentiment the highest (lowest) sentiment portfolio still can have low (high) investor sentiment in absolute terms. This is important, because investor sentiment in absolute terms determines the amount of sentiment traders in the market and the strength of the representative heuristic. In bear markets the highest sentiment portfolio can still have also no sentiment traders in the market and rational expectations dominating the investment decisions of investors. In such a market it would not be surprising to see a positive effect of relative overpricing on short-selling activity even for the highest sentiment portfolio.

To test this explanation I made a distinction between periods where the US economy faced a recession and growth-years. One would expect that in periods of a recession the general investor sentiment is lower than in periods of economic growth and sentiment traders almost completely left the market which would increase the effect of relative overpricing on the short-selling activity. I observed the opposite. Although the panel regressions themselves give mostly insignificant results, the picture as a whole gives a clear trend with clearly lower coefficients in recession periods than in growth periods, with even a significant negative relationship between relative overpricing and the short interest ratio. A possible explanation here is the relative overpricing variable. Because this only gives a value relative to other stocks in the S&P 500 this could mean that even the highest relative overvalued stock is under-priced in terms of fundamental value, especially in bear markets. As rational investors can determine fundamental value they will close their short positions when the stocks approach fundamental value, which could explain a negative relationship between relative overpricing and the short interest ratio.

In chapter II I will explore the literature on investor sentiment and short-selling. I will discuss the main models of investor sentiment, how according to the literature investor sentiment affects results and then I will describe the main indicators that are constructed to capture investor sentiment. Next I explain the concept of short-selling and its connection to investor sentiment, derive my main hypotheses and explain the main anomalies that are important for the overpricing variable. Chapter III describes the methodology I used to come to the results. Chapter IV describes the main results of this paper. Chapter V concludes and provides a brief discussion of the results.

## II. Literature

The role of investor sentiment got increasing attention since the first formalization into a model by DeLong et al. (1990). They provided a framework with rational investors or arbitrageurs and noise or sentiment traders. Rational investors trade on fundamentals while noise traders trade on private signals who can deviate from fundamental value. Because noise traders have strong beliefs that their private signals can beat the market rational investors face noise trader risk if they take contrarian positions. When rational investors for example detect that stock prices are above fundamental value and take a short position, there can be another private signal who makes the noise trader even more bullish and deviates prices further from fundamental value. If the investment horizon of rational investors is shorter than that of noise traders they can lose money in this case.

After the introduction of the DeLong model in 1990 the investigation to investor sentiment increased rapidly and can be broadly distinguished in three categories: models of investor sentiment, the relationship between investor sentiment and stock returns, and indicators that capture investor sentiment.

Especially in the early years when investment sentiment was not part of the mainstream yet, models were built to provide a theoretical framework. This helped to explain the recent investor anomaly that investor sentiment was at that time. When the theoretical foundation was established researchers became more and more interested about the relationship between investor sentiment and future returns. After this causal relationship was established the focus in recent years lies more on building indicators to capture investor sentiment and therefore predict future returns. In the next sections I will discuss the relevant literature based on these three distinctions.

### *Models*

Daniel et al. (1998) propose a model with irrational traders that explains short-term momentum and long-term reversals. Key aspects here are overconfidence and the self-attribution bias. There are private and public signals. The public signals are available for everyone, while the private signals are only available for some investors. Those investors are overconfident about this private signals which leads to a stock price increase above fundamental value in case of bullish signals and a stock price decline below fundamental value in case of bearish signals. As the public signals comes and reveals the true information, self-attribution bias makes that the investor thinks that it was skill if his private signal was

right and it was bad luck in case the private signal was wrong which increases confidence and stock price even further. In the long run the public signal brings the stock price back to fundamental value which explains the long-run reversal.

Hong and Stein (1998) explain the price deviations from fundamental value with a slightly different model of bounded rational traders. In their model there are two types of investors: news watchers and momentum traders. News watchers trade solely on privately observed news about the stock fundamentals and not on past of current stock prices. When they get a bullish (bearish) signal about the company the stock price will rise (fall) but not directly to fundamental value which leads to an initial underreaction. Momentum traders on the other hand trade only on past prices. They can profit from the slow adaption to fundamental value by the news watchers. However, their signal will stretch too long so only the early momentum traders will profit while late momentum traders will buy when the stock is already at or above fundamental value. This explains the stock overreaction after the initial underreaction.

In Daniel et al. (2001) the authors elaborate further on the idea of overconfident investors. Because of overconfidence in the private signal there appears mispricing in the stock market. In their model the fundamental value/price ratio serves a proxy for mispricing, because of overconfidence the mispriced stocks are those with a low fundamental value relative to their price. Risk-averse arbitrageurs exploit this mispricing on the individual stock level because they can hedge the idiosyncratic risk, but at the industrial and market level mispricing persists because those risks cannot be arbitrated away.

After earlier papers that noise traders exist in the market, Wang (2000) showed that noise traders can also survive in the market. On the individual level they will go bankrupt sooner than rational traders because their fundamental risk is higher. As a group, however, they can survive in the market because they trade more aggressively and therefore their risk and return are higher than that of rational investors. This is especially the case for moderate bullish sentiment traders. In markets where it is difficult to value the stock, and so for every trader the fundamental risk is high, they even may dominate the market.

Cross et al. (2005) add to earlier work by building a behavioural model which no longer assumed that investors are fully rational. They therefore tested three behavioural assumptions added to the rational framework: investors have the tendency to herd, they react more strongly on new information in times of extreme sentiment and with the passage of time the temptation to modify their portfolio increases. They found that these three assumptions can lead to the observed fat-tails, volatility correlations and long-term volatility clusters observed in real data.

By analysing the impact of investor sentiment on merger activity Lamont and Stein (2006) distinguish between the aggregate stock market and the firm-specific level and state that on the individual stock level markets are highly efficient so investor sentiment has no impact while there is a relative large impact on the less efficient aggregate stock market.

Schleifer and Vishny (2010) provide a model which links investor sentiment to the deep recession of 2008. In their model banks have limited cash and with this money they can finance projects or buy securities. To generate more money to invest they can securitize their loans and sell them in the market. The prices for these securities are affected by investor sentiment and therefore high if sentiment is high while the underlying value stays the same. This makes it attractive for banks to sell all their loans if sentiment is high even with the knowledge that there is no buffer for bad times which has a destabilizing effect on the financial markets.

Stambaugh and Yu (2017) combine earlier work on rational risk factors with behavioural finance and built four factor model with a market factor, size factor and two mispricing factors based on 11 well-known anomalies. They find that lagged investor sentiment is correlated with the mispricing factors and

their size factor, constructed not to be affected by mispricing, has a size premium double that of Fama and French (1993).

### *Returns*

Neal and Wheatley (1998) tried to predict stock returns using investor sentiment. They use odd-lot ratios, net redemption and fund discounts as proxy for investor sentiment and found that the last two can predict the size premium. This is important because in the noise trader model the sentiment traders are mostly small individual investors and their participation ratio in small firms tend to be higher than in large firms so effects of investor sentiment would be mostly expected in the returns of small stocks. Because small individual investors are assumed to be the irrational investors Barber et al. (2009) investigate their trades specifically and in line with the herding assumption they find that trades of individual investors are highly correlated. Strongly bought stocks outperform strongly sold stocks in the next four weeks and then the picture reverses. This is especially true for stocks that are difficult to arbitrage, like stocks with high idiosyncratic risk. On the other hand strongly bought stocks by institutional investors underperform strongly sold stocks in the next weeks.

Lee et al. (2002) go even further and found that increased investor sentiment leads to higher current excess returns, lower volatility and higher future excess returns and is therefore a priced risk factor. They hypothesize that in the short-run the increased return is because of herding behaviour of sentiment traders while in the long-run the larger amount of noise traders in the market scares away rational investors so higher returns are needed for compensation. In later studies these short-term excess returns and long-term reversal are also revealed.

Brown and Cliff (2005) find that in the short-run high investor sentiment leads to overvaluation of stocks and therefore high excess returns while in the long-run the stock prices reverse to fundamental value and this leads to abnormal low returns.

By analysing valuation ratios Coackley and Fuertes (2006) find evidence consistent with the short-term overreaction and long-run reversal of earlier studies. They state that in bullish times prices overreact and this can persist for some time but as markets become bearish the stock prices deviate back to fundamental value.

Baker et al. (2012) look into the role of investor sentiment in a broader perspective. They find that global investor sentiment is a reliable contrarian predictor of future global market return and also in the cross-section between countries. Furthermore, they find indications that sentiment spreads between countries through international capital flows.

Moskowitz et al. (2012) investigate time-series momentum. They find that last year's stock return is a good predictor of the stock return next month due to positive auto-covariance. They link this to investor sentiment by assuming an initial underreaction and a later overreaction as the return reverses. Also on very short time frames this pattern is established.

Garcia (2013) finds that there is a correlation between investor sentiment and returns the day of publishing and this effect reverses the next four trading days. The correlation is especially strong during recessions.

In contrast to the studies mentioned above there are also results that do not show that investor sentiment is a return predictor. Das et al. (2005) investigated the predictability of future returns based on online discussions and found that discussion drives sentiment but that the discussion is highly influenced by past returns so that past returns drive sentiment and not the other way around. A similar conclusion was found by Wang et al. (2006) who examined the relationship between sentiment, returns and volatility. They find that returns and volatility affect sentiment and not the other way around. Investor sentiment can predict future volatility but this forecasting power is limited when past returns are included.

Stambaugh et al. (2015) look into the relationship of idiosyncratic risk (IVOL), investor sentiment and expected returns. They explain that stocks with higher IVOL have an higher arbitrage risk and therefore are more over- and under-priced than low IVOL stocks. Because of arbitrage asymmetry this effect is stronger for overvalued stocks which leads to an aggregate negative effect of IVOL on expected returns. The relationship between IVOL and expected returns is stronger in times of extreme investor sentiment because mispricing will be stronger at that times. Stambaugh et al. (2012) investigate mispricing during different times of investor sentiment. Using the 11 anomalies left with the Fama and French 3-factor model (3FF) they short the most overvalued assets of each anomaly and long the most undervalued ones. They find that profits are highest in times of high investor sentiment, because overpricing is then highest. Also they find that especially the short leg contributes to these profits, which is consistent with the view that because of arbitrage asymmetry there is a lot of overpricing but very little under-pricing. Because there is no under-pricing investor sentiment is also not correlated with profits in the long leg.

### *Indicators*

Soon after the first models of investor sentiment researchers began the search for variables that indicate the level of investor sentiment. Lee et al. (1991) explained the closed-end fund puzzle, shares in a closed-end mutual fund trade at a discount relative to the underlying assets, by a model of investor sentiment. They show that rational investors that participate in a closed-end fund want to be compensated for the (noise trader) risk that irrational investors become bearish about the fund and so they need to be compensated for that through a discount.

Another proxy was used by Eleswarapu and Reinganum (2004). Because the value of glamour stocks depends highly on (uncertain) future growth rates they are harder to value than other stocks and this may make them highly prone to investor sentiment. Eleswarapu and Reinganum (2004) therefore try to predict future stock market returns based on past glamour stocks returns as proxy for investor sentiment. They found that past glamour stocks returns are negatively correlated with future returns in the stock market and think this is because high glamour stocks returns are a signal of an overpriced stock market in general.

Baker and Stein (2004) explain that liquidity could be a sentiment indicator. Sentiment traders overreact to their private signals and therefore trade more than rational investors and provide liquidity to the market. Liquidity can therefore be a measure of the relative presence of sentiment traders in the market and because of short-sale constraints those sentiment traders always overreact and never underreact to private signals so their presence is a measure of market-overconfidence.

Da et al. (2015) construct an investor sentiment indicator based on pessimistic economic search trends by households (FEARS). They find that this FEARS indicator is negatively correlated to returns on the same day but that there is a reversal the days after and this effect is stronger for stocks that are hard to arbitrage. Furthermore, FEARS is correlated to volatility and a high FEARS number seems to trigger an outflow of equity of mutual funds into bonds.

Huang et al. (2015) use the six proxies of the investment sentiment indicator of Baker and Wurgler to build a new indicator where they use partial least squares instead of principal component analysis to filter relevant information for expected returns out of error and noise. They find that this aligned indicator negatively and significantly forecasts future dividend growth (a cash flow proxy) but does not forecast future dividend/price ratio (a discount rate proxy) so high investor sentiment leads to overoptimistic future cash flow growth and therefore low future returns.

The studies above all try to capture investor sentiment on the market level. In contrast to that Jian et al. (2019) and Seok et al. (2019a) used indicators to find investor sentiment on the firm level. Jian et al. (2019) built a management sentiment indicator based on the difference between the amount of positive and negative words in firm financial statements divided by the total number of words. This indicator is negatively correlated with returns the next month until next year. This is driven by managers



overoptimistic view of cash flow growth which results in negative earnings surprises the next year, and current overinvestment which results in low investment growth the next two years. Seok et al. (2019a) use a daily firm-specific sentiment indicator to capture the reactions on earnings announcements and find that the returns of high sentiment companies are higher following positive earnings announcements than for low sentiment companies which indicates that the latter underreact to positive news. This may be because investors only partially update their beliefs after announcements and for high sentiment companies optimistic announcements were expected, while for the latter pessimistic announcements were expected.

### *Short selling*

A key assumption in the investor sentiment theories is the existence of short-selling constraints. An early study of the effect on short-selling constraints on share prices was done by Miller (1977). In his theoretical model he assumed rational investors with divergence of opinion. In a world without short-selling the supply of a stock is fixed and the price of a stock with  $N$  shares is determined by the most optimistic investors who together want to hold  $N$  shares. However, with short-selling the investor who wants to sell short borrows a stock from an investor owning the stock and sells it to another investor while promising that he will return it in the future to the initial owner while in the meantime paying him all the dividends if they occur. A short-seller thus duplicates a stock and therefore with short-selling the supply of a certain stock is increased. With divergence of opinion and an increased stock supply the price of the stock will decrease. Without short-selling constraints, and therefore unlimited stock supply, the stock price will decrease to fundamental value. However, because short-sellers don't get their proceeds immediately, they need to hold a margin and the ability to short shares is not unlimited while the supply of stocks is fixed which means that stock prices still will be determined by a group of optimistic investors and be higher than the average opinion of all investors.

In the model of Diamond and Verrecchia (1987) the authors find, contrary to Miller, that short-selling constraints do not result in biased prices because investors anticipate on the short-selling constraints. They assume a market with rational investors where one group of informed traders have private information, while the other group of uninformed traders have only public information. With short-selling constraints pessimistic traders are left out of the market. This means that private bad information will not directly be incorporated into the stock prices so the speed of adjustment to private information will decrease. Because bad private information cannot enter the stock price, bad information will only be incorporated into the stock prices when it becomes public. At those announcements the swings will therefore be higher which results in larger excess returns. Because good news can enter the stock price at all moments the swings at announcements will mostly be corrections for bad news so the returns at announcements are skewed to the left.

Follow-up studies provided similar results regarding short-selling constraints and the efficiency of share prices. By using the rebate rate as proxy for short-selling difficulties Reed (2002) find that when short-selling is costly there is a reduction in informational efficiency, because stocks adjust slower to bad earnings announcements. Because negative private information is kept out of the market, stocks with short-selling constraints have larger absolute values and are more left skewed because when the negative information becomes public the reaction will be stronger (negative) with short-selling constraints.

Saffi and Sigurdsson (2011) find that low levels of lending supply lead to slower adjustment of prices to market shocks and therefore less price efficiency. Higher lending supply leads to less skewness and a lower frequency of extreme positive returns. Lower fees are associated with less downside risk and lower volatility.

Boehmer and Wu (2013) find empirical evidence that short-selling activity leads to prices that follow the random walk more closely. Furthermore monthly and annual price delays become smaller, post-earnings-announcement drift vanish for negative announcements, and there are less extreme price movements based on non-information.

Charoenrook and Daouk (2009) investigate the impact of short-selling restrictions on liquidity. They argue that the allowance of short-selling increases liquidity especially during bad times because with short-selling a whole group of investors, those with pessimistic beliefs that do not already own stocks of the company, can enter the market and cannot trade if short-selling is prohibited. This leads to less volatility, higher aggregate stock prices and lower expected returns as with short-selling better risk sharing is possible which lowers the required return and cost of capital. Beber and Pagano (2013) investigate this empirically by looking at the world-wide bans on short-selling during the financial crisis of 2007-2009 and found indeed a decrease in liquidity and slow-down of the price discovery process but only higher returns in the US. In another paper Grullon et al. (2015) confirm those findings for the US by showing that the release of short-sale constraints in the US lead to a drop in share prices and in equity issue and investment especially for small and overvalued firms. They hypothesize that this could be the case because short-selling decreases overvaluation or because a decline in share price makes managers reluctant to invest and this drives share prices further down so short-selling could be an instrument to manipulate share prices.

Massa et al. (2015) examine the effect of short-selling on stock prices and look directly into the effect on company earnings by investigating the disciplining effect of short-selling. They find that the amount of lendable shares, and therefore short-selling potential, have a strong negative impact on the amount of positive words and the height of the earnings in earnings report indicating that short-sellers have a strong disciplining effect on earnings management. The reasoning behind this is that managers don't manipulate reports if there is the threat that they will be punished by short-sellers through a sharp declining in stock prices if their manipulation is uncovered.

### *Hypothesis*

Besides rational investors who trade on fundamental value there are investors who think they can beat the market with other information than by determining fundamental value. Those are noise traders (Black, 1986). Noise traders can exist in the market, because there are limits to arbitrage. One of the limits to arbitrage is the risk that a stock might deviate even further from fundamental value ones the arbitrageur has taken his position (De Long et al., 1990). Because, especially in the case of short selling, this can result in the need for the arbitrageur to liquidate his position before it reverts to fundamental value and therefore suffers a loss. This noise trader risk means that if a noise trader, for whatever reason, is bullish and believes that the share price will increase above fundamental value it is of course possible that he will become even more bullish in the future (De Long et al., 1990).

Barberis et al. (1997) built a model for investor sentiment where the individual can face two modes: in mode one good or bad news comes as a surprise and the investor expects that the stock return is mean-reverting, but after a sequence of good or bad news the individual switches to mode two: here he expects that the sequence of good or bad news will continue. The psychological reasoning behind this is based on the representative heuristic first described by Kahneman and Tversky (1974). This means that one overestimate the probability of an event because it determines this probability directly from the immediate history of these events. If a short-seller is therefore himself a victim of the representative heuristic or he is aware of the phenomenon does not matter: it will form believes for the noise trader to expect higher returns and with enough noise traders in the market the stock price will increase independent on the thoughts of rational investors. The higher the investor sentiment (and so the sequence of good news) the stronger this heuristic will get with new good news and further deviation from fundamental value so the short seller will decrease his opposite position.

Yu and Yuan (2010) found that in times of high investor sentiment the mean-variance relation between variance and risk is far weaker than in times of low investor sentiment. This means that in times of low investor sentiment an increase of the variance of returns is accompanied with a strong increase in excess return, while in times of high sentiment this relationship is far weaker or even non-existent. Furthermore the negative relationship between return and volatility innovations is also weaker in times of high

investor sentiment than in times of low investor sentiment. The reasoning behind this is that in times of high investor sentiment, in line with Karlsson, Loewenstein and Seppi (2005), sentiment driven traders participate and trade more in the stock market. Furthermore, based on Barber and Odean (2006), the authors conclude that sentiment-driven traders do not sell short. Because sentiment-driven investors do not sell short they stay out of the market in times of low sentiment when the expected return is lower than the risk-free rate and participate heavily during high sentiment. This means that in high sentiment there are relatively few short-selling candidates. This makes companies who experience a period of higher investor sentiment relatively more difficult to short-sell than companies who experience a period of lower investor sentiment.

So investor sentiment can affect the amount of short-selling in different ways. The first way is that in high investor sentiment noise traders become more bullish and as prices deviate further from fundamental value the noise trader risk for the arbitrageur increases because his loss will become larger if he has to close his position before the prices come back to fundamental value. Because of representative heuristic companies at the high end of investor sentiment will be expected to continue with good news and those investors will be extremely bullish which makes it hard to short sell this companies and with further deviation from fundamental value the beliefs of investors will strengthen further which makes it even more difficult to short sell. In this case it does not matter if the arbitrageur himself believes the representative heuristic or not, because in both cases his noise trader risk will increase and this will decline his ability to short sell.

The other channel is the amount of short sellers in the market. As investor sentiment increases the amount of sentiment driven traders increase in the market. Assuming that those noise traders do not short-sell, the relative amount of short seller decreases with investor sentiment which makes it likely that the amount of short selling decreases with investor sentiment. My first hypothesis is therefore:

*Hypothesis I: the correlation between relative overpricing and the short interest ratio is negative for companies that experience high investor sentiment*

On the other hand on the low end of investor sentiment there are relatively less noise traders in the market and their beliefs are also relatively weak which decreases noise trader risk and makes it easier to short sell those companies. My second hypothesis is therefore:

*Hypothesis II: the correlation between relative overpricing and short interest ratio is positive for companies that experience low investor sentiment*

To measure relative overpricing one needs a set of stock characteristics that lead to predictable abnormal (low) returns in the future because investors put too much emphasis on these characteristics. This is exactly what Stambaugh et al. (2012) did:

- Financial distress: stocks with a higher failure probability are accompanied with higher risk so in the mean-variance framework this should be compensated with higher returns. Campbell et al. (2007), however, find that this companies earn lower instead of higher future returns;
- Net stock issues: as managers believe that their companies are overvalued they will issue shares to profit from this mispricing. In line with Ritter (1991) one would expect that stocks with high net stock issues underperform in future years;
- Total accruals: firms with higher total accruals earn lower subsequent returns probably because investors are too fixed on the accrual part in forming earning expectations (Sloan, 1996);
- Net operating assets: Hirshleifer et al. (2004) find that high relative net operating assets leads to low future returns, because investors focus too much on accounting profitability and forget cash;
- Momentum: high past returns in recent period forecasts high returns in next period (Jegadeesh and Titman, 1993) and this effect is stronger in times of high investor sentiment (Antonioni et al., 2010);

- Gross profit: firms with relative high gross profit/assets earn higher subsequent returns (Novy-Marx, 2010);
- Asset growth: relative high asset growth leads to low subsequent returns, probably because investors overreact to this growth in their business forecast (Cooper et al., 2008);
- Return on assets: firms with higher past return on assets earn higher future returns (Chen et al., 2010). This is especially the case for companies with high arbitrage costs which suggests that mispricing plays a key role (Wang and Yu, 2010);

### III. Research method

With the help of the NSM Library Team I got the daily closing price, volume and number of shares of every stock on the S&P 500 from January 1, 1990 until December 31, 2019 based on ISIN and date out of the Eikon database. I took 1990 as starting point because short interest data is only available from then on in the databases.

I followed the procedure of Seok et al. (2019a) to build daily sentiment indicators based on the relative strength index (RSI), the psychological line index (PLI), the adjusted turnover rate (ATR) and the logarithm of trading volume (LTV). To process the data I entered it into Stata16, sorted the stocks on ISIN and date and cleaned the dataset by removing duplicates and missing values.

I calculated the indicators based on Seok et al. (2019a and 2019b). In Seok et al. (2019a) the RSI is generated by:

$$(1.)RSI_{i,t} = \left( \frac{RS_{i,t}}{1 + RS_{i,t}} \right) * 100; RS_{i,t} = \frac{\sum_{k=0}^{13} \max(P_{i,t-k} - P_{i,t-k-1}, 0)}{\sum_{k=0}^{13} \max(P_{i,t-k-1} - P_{i,t-k}, 0)}; P_{i,t} \\ = \text{closing price of stock } i \text{ at time } t$$

Which means that the relative strength is the ratio of the total sum of upward movements in the stock's closing price over fourteen days divided by the sum of the downward movements and the RSI is the relative strength in percentage. The fourteen day period is common in the literature to calculate the RSI so I followed this procedure (Chen, Chong, and Duan, 2010).

To calculate the PLI Seok et al. (2019a) use the following formula:

$$(2.)PLI_{i,t} = \left( \sum_{k=0}^{11} \left[ \frac{\max(P_{i,t-k} - P_{i,t-k-1}, 0)}{(P_{i,t-k} - P_{i,t-k-1})} \right] / 12 \right) * 100$$

This means that the PLI counts the amount of upward movements in the closing stock price over a twelve day period and turns this into a percentage. The twelve day period is common in the literature so I followed this approach (Yang and Gao, 2014). In some cases the closing price is exactly the same as the day before so there is neither an upward nor a downward movement and the indicator gives an error. Because the PLI counts the number of upward movements and a flat line is not an upward movement I registered those cases as a zero.

The next indicator is the adjusted turnover rate given in Seok et al. (2019a) by:

$$(3.)ATR_{i,t} = \frac{V_{i,t}}{\text{number of shares outstanding}_{i,t}} * \frac{R_{i,t}}{|R_{i,t}|}; V_{i,t} \\ = \text{trading volume of stock } i \text{ at time } t; R_{i,t} = \left( \frac{P_{i,t}}{P_{i,t-1}} \right) - 1$$

The adjusted turnover rate is trading volume of a stock on a given day divided by the number of shares outstanding times the simple return that day divided by the absolute value of the same simple return. The adjusted turnover rate is therefore the proportion of stocks traded where the simple return divided

by his absolute value gives the sign of the trade. If the return is positive the value of the right part of the equation becomes one while it becomes minus one if the return is negative. In cases where the return is zero the formula gives an error so an adjustment had to be made. Because a positive return gives value one and a negative return get value minus one a zero return should get the value zero so automatically the adjusted turnover rate is also zero in these cases.

The last indicator is the LTV which is:

$$(4.)LTV_{i,t} = \ln (V_{i,t})$$

The logarithm of the volume of stock  $i$  at time  $t$ .

As done in Seok et al. (2019b) I regressed the four indicators separately to the market equity premium. Because my analysis is performed on the stocks of the American S&P 500 I used the overall S&P 500 return as the market return and the return on a three-month Treasury Bill as risk-free rate to calculate the equity premium. The indicators are based on daily stock prices so I also used the daily S&P 500 returns and transform the annualized Treasury Bill return in daily returns to get the daily equity premium. Then I regressed each of the four indicators to this daily equity premium:

$$(5.)Comp_{k,i,t} = a_0 + a_1 * Market_t + \epsilon_{k,i,t}$$

This way I obtained an alfa, beta and residual value for each component on every stock on a daily basis. The alfa and beta are the amount of the indicator driven by market fluctuations so these are the residuals that give a market-free approximation of each sentiment indicator for stock  $i$  at time  $t$ .

The next step was to combine the fluctuations in each sentiment indicator by getting a common factor of the four indicators for every stock  $i$  at time  $t$ . Because each residual measures different values and the principal component analysis measures the degree of variation of the different values I first needed to standardize each component so each indicator is measured in there z-value. With the standardized indicators I computed the common factor and generated this into a variable. Then I ranked the stocks on each date based on the common factor and separated this rank into four groups to get sentiment portfolios.

The idea here is that with portfolios with approximately 125 stocks the individual firm characteristics disappear and the analysis is truly based on sentiment, while with four portfolios the difference in sentiment between the upper and lower portfolio is still large enough to make an empirical distinction in sentiment.

In the work of Stambaugh et al. (2012) there are eleven main stock anomalies that predict overpricing based on nine categories. Because of data limitations and to prevent putting too much weight on some categories I have selected one anomaly out of every category so I have nine variables to capture overpricing: financial distress, net stock issues, total accruals, net operating assets, momentum, gross profitability premium, asset growth, return on assets, and investment-to-assets.

For the momentum variable I used the methodology of Carhart (1997) and calculated the cumulative return of every stock from month  $t-12$  until the end of  $t-2$  to get a ranking at month  $t$ . As with every anomaly I grouped the rankings into deciles of 50 stocks so the rankings are stable in the very short-run but are able to fluctuate over a longer period of time. Because the relative overpricing is in the end a simple average of the nine rankings this will prevent one anomaly to have a too strong on the relative overpricing variable. The momentum ranking is calculated for every month based on the formula:

$$(6.)MO_{i,t} = \sum_{k=2month}^{12month} \ln \left( \frac{P_{i,t-k}}{P_{i,t-1-k}} \right)$$

According to Ritter (1991) overvalued stocks will have higher net stock issues, because managers want to convert the overvaluation of their stock into cash. In line with Fama and French (2008) I calculated net stock issues as the annual logarithmic change in split adjusted shares outstanding:

$$(7.) NSI_{i,t} = \ln(CSHO_{i,t}) - \ln(CSHO_{i,t-1})$$

The last reporting year is the last fiscal year that ended at least four months before t. This means that, for example, the last fiscal year from January 1<sup>st</sup> 2008 until April 30<sup>th</sup> 2008 is 2006 and from the 1<sup>st</sup> of May until April 30<sup>th</sup> 2009 this is 2007. The reasoning behind this is that it will take four months in the new fiscal year for all the companies to update last year's company reports so only from then on investors can make decisions based on these figures (Stambaugh et al., 2012).

Total accruals is a negative predictor for future stock returns. Following Sloan (1996) I calculated accruals based on:

$$(8.) Ac_{i,t} = \frac{\Delta(\Delta ACT_{i,t} - \Delta CHE_{i,t} - \Delta DLC_{i,t} + \Delta LCT_{i,t} + \Delta TXP_{i,t}) - DP_{i,t}}{(AT_{i,t-2} + AT_{i,t-1})/2}$$

Where  $\Delta(\Delta ACT - \Delta CHE - \Delta DLC + \Delta LCT + \Delta TXP)$  is the annual change in noncash working capital, subtracted by depreciation and amortization expense, divided by the average total assets of the previous two fiscal years. Accruals is measured yearly and the most recent reporting year is the one that ends at least four months after the end of t-1. I ranked accruals in deciles where decile ten is the 50 stocks with the highest accruals and decile one are the 50 stocks with the lowest or missing accruals.

High net operating assets signals overvaluation of stocks (Hirshleifer et al., 2004). In line with their paper I calculated net operating assets based on the following formula:

$$(9.) NOA_{i,t} = \frac{(AT_{i,t} - CHE_{i,t}) - (AT_{i,t} - DLC_{i,t} - DLTT_{i,t} - CEQ_{i,t} - MIB_{i,t} - PSTK_{i,t})}{AT_{t-1}}$$

Where  $(AT - CHE)$  are operating assets and  $(AT - DLC - DLTT - CEQ - MIB - PSTK)$  are operating liabilities, divided by lagged total assets. Net operating assets is calculated yearly and the last reporting year is the one that ends at least four months before the end of t-1. The stocks with the highest net operating assets are in decile ten and the stocks with the lowest or missing net operating assets are in decile one.

Cooper et al. (2008) found that the growth of total assets can be a predictor of overvaluation. I calculated this variable as the growth of total assets over the last reported year which is the last full year that ends at least four months before the end of t-1. Again I ranked the stocks in deciles of 50 stocks where decile ten are the 50 stocks with the highest asset growth and decile one are the 50 stocks with the lowest or missing asset growth.

Titman et al. (2004) show that investment-to-assets is a negative predictor for future stock returns. Following their research I calculated investments-to-assets as the annual change in gross property, plant and equipment, plus annual change in inventory, divided by lagged total assets:

$$(10.) ITA_{i,t} = \frac{\Delta PPEGT_{i,t} + \Delta INVT_{i,t}}{AT_{t-1}}$$

The most recent reporting year is the one that ends at least four months before the end of t-1. Stocks are ranked in deciles with decile ten containing the 50 stocks with the highest investment-to-assets and decile one containing the 50 stocks with the lowest or missing investment-to-assets.

The O-score is a static measure by Ohlson (1980) to calculate the failure probability of a company. Firms with a high failure probability earn lower subsequent returns and are therefore relative overvalued (Campbell et al., 2008). To calculate the O-score I used Ohlson's (1980) formula:

$$(11.) O_{i,t} = -0,407 \ln(AT_{i,t}) + \frac{6,03(DLC_{i,t} + DLTT_{i,t})}{AT_{i,t}} - \frac{1,43(ACT_{i,t} - LCT_{i,t})}{AT_{i,t}} + \frac{0,076LCT_{i,t}}{ACT_{i,t}} - 1,72(LT_{i,t} > AT_{i,t})$$

Where  $AT$  equals total assets,  $DLC$  is current liabilities,  $DLTT$  is the amount of long-term debt,  $ACT$  is current assets,  $LCT$  is the portion of debt included in current liabilities and  $LT > AT$  is a dummy variable which equals one if total liabilities is larger than total assets and zero otherwise.

Last reporting year is the year that ends at least four months before the end of t-1. The stocks are grouped in deciles where decile ten are the 50 stocks with the highest failure probability and decile one are the 50 stocks with the lowest or missing O-score.

Companies with higher gross profit earn higher subsequent returns (Novy-Marx, 2013). Following his study I calculated gross profitability as total revenue minus cost of goods sold, scaled by total assets:

$$(12.) GPP_{i,t} = \frac{RVT_{i,t} - COGS_{i,t}}{AT_{i,t}}$$

The last reporting year is the year that ends at least four months before the end of t-1. The stocks are grouped in deciles where decile ten are the 50 stocks with the lowest gross profitability because those are the most overvalued, and decile one are the 50 stocks with the highest gross profitability.

Chen et al. (2010) found that firms with high return on assets earn higher subsequent return which indicates that return on assets captures undervaluation. In their work Chen et al. (2010) calculated return on assets based on income before extraordinary items in the current quarter divided by last quarter's total assets. Because of data limitations I only have yearly income before extraordinary items and total assets available so I calculated return on assets as:

$$(13.) ROA_{i,t} = \frac{IBQ_{i,t}}{ATQ_{t-1}}$$

Where the last reporting year is the one that ends at least four months before the end of t-1. I ranked the stocks in deciles with decile ten are the 50 stocks with the lowest return on assets because those are the most overvalued, and decile one are the 50 stocks with the highest return on assets.

In Stambaugh et al. (2012) the eleven anomalies are grouped in nine categories. Because of data limitations and because I did not want to put too much weight on some categories I picked one anomaly out of every category. I ranked the stocks for these nine anomalies from one to ten on every point in time where ten is the most overvalued stock in the anomaly category and one is the less overvalued stock in the anomaly category. I then calculated the simple average of every stock on every point in time to get the relative overvaluation of every stock over time.

The next step is to regress the short interest ratio to the overpricing level in each sentiment portfolio. I took the monthly short interest of every stock in my data sample from January 1990 until December 2019 out of FactSet. The data in FactSet is only available from March 1993 onward so I deleted the first three years with left 26 years and ten months of monthly data in my time-series.

After I merged the short interest data to the master dataset I first had to calculate the short interest ratio. Therefore I divided the short interest by average trading volume. Because the short interest is a monthly figure I calculated the average trading volume also over the last month.

Then I performed a factor panel regression where I created an interaction variable between the relative overvaluation level and the sentiment portfolio to measure the effect of investor sentiment on the effect of the relative overvaluation on the short interest ratio. I chose this method instead of a fixed effect regression because by default each stock is a different panel and for a fixed effect regression I should convert this to a setting where each sentiment portfolio is a different panel. In this scenario each sentiment portfolio should have a unique short interest ratio and overvalue level at every point in time which means that I should merge all the short interest ratio and overvalue levels of the stocks in every portfolio and therefore lose a lot of valuable data. Therefore I chose to perform a factor panel regression which measures the effect of the relative overvaluation level of a stock on its short interest ratio taking into account that it is panel data and measuring the interaction between the relative overvaluation level and investor sentiment. To account for serial correlation I included lags of the dependent and independent variable. To control for the general market fluctuations I added the monthly equity premium of the S&P 500 in the regression. I calculated the equity premium by calculating the monthly return of the S&P 500 for every month in the sample and subtract the monthly return of the three month Treasury bill.

Subsequently I performed tests to make sure that my analysis was not biased. First I ran a Granger causality test where I ran a factor panel regression with short interest ratio, the overvaluation level and the equity premium as independent variable, all three lagged one month. The one period lag was at first not significant, but when a two period lag was included this changed indicating a causal relationship over a longer period of time. Next step was to investigate if there is a causal relationship the other way around. Therefore I performed a regression with the relative overvaluation level as dependent variable and the overvaluation level, equity premium and short interest ratio as independent variables, all lagged one period. In this regression the short interest ratio coefficient was not significant both one period lagged as two period lagged so I conclude that there is no causality the other way around.

Then I tested for stationarity. I used the Fisher unit root test, because I have panel data which are unbalanced because not for every month is data available and the Fisher test takes this into account. Within the Fisher test I used the Phillips-Perron unit root because the short interest ratio is serial correlated and the Phillips-Perron test is robust for serial correlation. As I used three lags in the basic regression I added three lags to the Phillips-Perron test and removed cross-sectional means with the 'demean' option. The test strongly rejected the null hypothesis that the short interest ratio levels of all stocks are nonstationary.

Next I checked for multicollinearity in my dataset by performing a variance inflator factor test (VIF) in Stata. The VIF of all the overvaluation variables was above 15 while the VIF of other variables were around one. Taking into account this large difference in VIF and the rule of thumb that a VIF higher of ten indicates multicollinearity I deleted the lagged values of the overvaluation variable and this resulted in a drop of the remaining overvaluation variable close to one. Because eight of the nine overvaluation components are calculated yearly and the lags are monthly that makes sense.

Last steps were to test for auto-correlation and heteroskedasticity. To test for auto-correlation I could use the 'xtserial' option in Stata which showed that there is auto-correlation in my dataset. To check for heteroskedasticity I performed two feasible generalized least squares (FGLS) regressions. I chose this regression because this linear panel regression can account for heteroskedasticity. Therefore I performed one regression that accounted for heteroskedasticity and one regression that did not and then compared how much they fitted. Because the likelihood that they fit is significant there is heteroskedasticity in my dataset.

To account for the auto-correlation and heteroskedasticity in my dataset I again ran a FGLS regression, where I specified that there was heteroskedasticity and auto-correlation in my dataset. Because I have a large amount of stocks in my file and I relative long time period the auto-correlation for stocks could vary. Therefore I specified that the correlation coefficient could be unique for each panel. Data gaps



were eliminated in the calculation of this coefficient. Corrected for auto-correlation and heteroskedasticity the short interest ratio is significant until five lagged periods where in the last lag the sign flips and the equity premium until two lagged periods. For the highest three sentiment portfolios there is a strong positive effect of overvaluation on the short interest ratio. The effect of the lowest sentiment portfolio appears not to be significant.

## IV. Results

The main results are presented in figure I:

**Table I: panel regression short interest ratio**

The table reports the regression of the short interest ratio for every stock per month as the dependent variable and the interaction between the relative overvaluation of a stock and investor sentiment, the (lagged) equity premium of the S&P 500 and lags of the short interest ratio are independent variables. The interaction variable distinguishes the effect of the independent variables on the dependent variable for stocks with different levels of investor sentiment. *Senti* represents the *i*th investor sentiment portfolio with 1 as the top 25% daily ranked and 4 the bottom 25% sentiment stocks daily ranked. The sample period is from March 1993 through December 2019 (322 months).

VARIABLES	(1) Short interest ratio
Sent1 x relative overpricing	52.289*** (3.952)
Sent2 x relative overpricing	47.789*** (4.016)
Sent3 x relative overpricing	44.647*** (3.118)
Sent4 x relative overpricing	-2.222 (4.390)
Monthly equity premium	14.143*** (1.274)
<i>Lagged one month</i>	21.245*** (1.333)
<i>Lagged two months</i>	5.978*** (1.237)
Short interest ratio (lagged one month)	0.552*** (0.003)
<i>Lagged two months</i>	0.031*** (0.004)
<i>Lagged three months</i>	0.139*** (0.003)
<i>Lagged four months</i>	0.019*** (0.003)
<i>Lagged five months</i>	-0.000*** (0.000)
Constant	679.891*** (22.936)

Standard errors in parentheses  
 \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

The main result is that the coefficients of the interaction variable for the first three portfolios is strongly positive. This means that the relative overvaluation of a stock in the group of 75% stocks with the highest investor sentiment has a positive effect on the short interest ratio. For the 25% of stocks with the lowest investor sentiment there is no significant effect. Also the size of the coefficient drops with investor sentiment. Where the coefficient is 52,29 for the 25% stocks with the highest investor sentiment, it is 47,79 for the second group and 44,65 for the third group. This means that the relative overvaluation of stocks has a positive effect on the short interest ratio, but this effect becomes smaller when investor sentiment is declining.

In the first regression I found that three out of four of the interaction variables of relative overpricing and investor sentiment had a significant positive effect on the short interest ratio. To confirm the results I had to test if these interaction variables as a group also have a significant effect on the short interest ratio. The chi-squared test rejected the null hypothesis that the interaction variables as a group had no impact on the short interest ratio at the 1% significance level so I assume that there is an impact.

As mentioned in the literature section the hypotheses that I want to test are:

*Hypothesis I: the correlation between relative overpricing and the short interest ratio is negative for companies that experience high investor sentiment*

*Hypothesis II: the correlation between relative overpricing and short interest ratio is positive for companies that experience low investor sentiment*

For hypothesis I to be true the coefficient of the interaction of portfolio 1 with the relative overvaluation level should be negative and significant. However, the exact opposite is true. The relative overvaluation level seems to have a positive effect on the short interest ratio and therefore I can reject hypothesis I.

Hypothesis II states that relative overpricing has a negative effect on short interest ratio for the portfolios with the lowest investor sentiment. Because the results for portfolio 4 are not significant, the evidence for this hypothesis is not conclusive. However, there is a clear negative trend for the effect of overpricing on the short interest ratio from the highest investor sentiment portfolio to lower investor sentiment which is the opposite trend that I would expect if the hypotheses were true.

To test this negative trend I analysed:

$$(14.) \text{Senttrend}_{i,t} = (\text{Sent4} * \text{relative overpricing})_{i,t} - (\text{Sent1} * \text{relative overpricing})_{i,t}$$

Which resulted in a coefficient of -54,080 and significance at the 1% level. Therefore I concluded that the relationship between relative overpricing and the short interest ratio is negatively related to investor sentiment which confirms the violation of hypothesis II. Therefore I also reject hypothesis II.

### *Robustness*

To check for robustness I tested if the main results change in recession periods compared to non-recession periods. The idea behind this is that both my measures of investor sentiment and the level of overpricing are relative so only compares to other stocks in the panel and says nothing about their absolute levels. One would expect that in months of a recession this absolute values differ from periods with growth which could also affect the effect of overpricing on the short interest ratio. To test this difference I created a dummy with value zero in non-recession years and with value one in months where the U.S. economy hit a recession (NBER, 2020). Results are shown below:

*Table II: panel regression short interest ratio with recession dummy*

**Table I: panel regression short interest ratio**

The table reports the regression of the short interest ratio for every stock per month as the dependent variable and the interaction between the crisis dummy, relative overvaluation of a stock and investor sentiment, the (lagged) equity premium of the S&P 500 and lags of the short interest ratio are independent variables. The interaction variable distinguishes the effect of the independent variables on the dependent variable for stocks with different levels of investor sentiment in recession and non-recession periods.  $Sent_i$  represents the  $i$ th investor sentiment portfolio with 1 as the top 25% daily ranked and 4 the bottom 25% sentiment stocks daily ranked.  $Crisis_i$  is a dummy variable which equals 1 if the U.S. economy was in a recession according to the National Bureau of Economic Research, and 0 otherwise. The sample period is from March 1993 through December 2019 (322 months).

VARIABLES	(1) Short interest ratio
Crisis0 x Sent1 x relative overpricing	27.166*** (5.038)
Crisis0 x Sent2 x relative overpricing	27.322*** (5.030)
Crisis0 x Sent3 x relative overpricing	2.893 (5.074)
Crisis0 x Sent4 x relative overpricing	-20.304*** (5.360)
Crisis1 x Sent1 x relative overpricing	9.349 (7.830)
Crisis1 x Sent2 x relative overpricing	8.881 (7.768)
Crisis1 x Sent3 x relative overpricing	-6.704 (7.937)
Crisis1 x Sent4 x relative overpricing	-39.936*** (9.045)
Monthly equity premium	6.146*** (1.399)
<i>Lagged one month</i>	15.189*** (1.441)
<i>Lagged two months</i>	0.809 (1.366)
Short interest ratio (lagged one month)	0.498*** (0.003)
<i>Lagged two months</i>	0.141*** (0.003)
<i>Lagged three months</i>	-0.000 (0.000)
<i>Lagged four months</i>	0.000* (0.000)
Constant	1,184.301*** (27.398)

Standard errors in parentheses  
 \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

The main result is that the trend of the original panel regression remains intact. The portfolios with the highest sentiment capture a positive effect of relative overpricing on the short interest ratio both in non-crisis as crisis years and this effect is negative for stocks with the lowest investor sentiment. The coefficients are in general lower for stocks in crisis years and also the coefficients in non-crisis years are lower than in the original regression. The overpricing level in general has now a smaller impact on the short interest ratio and this effect is even significantly negative for stocks in the lowest portfolio quartile. The results, however, should be interpreted with causation because three out of four crisis portfolios show results that are not significant which can be explained by the relative short cumulative recession period.

Next I tested the following formulas:

$$(15.)CrisisHStrend_{i,t} = (Crisis1 * Sent1 * relative\ overpricing)_{i,t} - (Crisis0 * Sent1 * relative\ overpricing)_{i,t}$$

$$(16.)CrisisLStrend_{i,t} = (Crisis1 * Sent0 * relative\ overpricing)_{i,t} - (Crisis0 * Sent4 * relative\ overpricing)_{i,t}$$

Equation 15 subtracts the relative overpricing effect on the short interest ratio for the highest sentiment portfolio in non-recession periods from this interaction effect during recessions. The coefficient is -17,816 and is significant at the 1% level which shows that in recession periods the effect of relative overpricing on the short interest ratio in the highest sentiment portfolio is less positive compared to non-recession periods.

Equation 16 subtracts the relative overpricing effect on the short interest ratio for the lowest sentiment portfolio in non-recession periods from this interaction effect during recessions. The coefficient is -19,631 and is significant at the 5% level which shows that in recession periods the effect of relative overpricing portfolio on the short interest ratio in the lowest sentiment portfolio is less positive compared to non-recession periods.

Lastly I tested if the crisis dummy interaction variable as a group has a significant effect on the short interest ratio. The chi-squared test gives no significant at the 10% level so the null hypothesis that the interaction variable of recession, investor sentiment and relative overpricing as a whole has an impact on the short interest ratio cannot be rejected.

This means that taking periods of financial recessions into account diminishes the effect of the overvaluation on the short interest ratio. A possible explanation is that for most portfolios there was no significant effect of overvaluation on the short interest ratio in crisis periods which can be declared by the small sample of crisis periods. I therefore conclude that U.S. financial recessions are not a good indicator for the effect of relative overpricing on the short interest ratio.

## V. Conclusion and discussion

In this paper I have investigated the effect of investor sentiment on the short interest ratio. To research this subject I tried to capture the effect of the overpricing of stocks on the short interest ratio for groups of stocks that experience different amounts of investor sentiment. In this paper I defined investor sentiment as the systematic overpricing of stocks relative to their fundamental value caused by overoptimistic investors. Based on the method of Seok et al. (2019a and 2019b) I calculated the daily investor sentiment of every stock over the period 1990-2019. I divided the stocks into four groups with each 25% of the shares of the S&P 500 ranked on daily investor sentiment. Then I calculated the relative

overvaluation on stocks based on nine groups of anomalies found by Stambaugh and Yuan (2017) that capture deviations away from fundamental value. I ranked every stock in deciles for every anomaly and calculated the relative overvaluation per month by taking the simple average of the relative ranking for each anomaly. Next I compared the effect of the change in relative overvaluation on the change in short interest ratio for the different sentiment portfolios. Based on the representative heuristic (Barberis et al., 1997) and the evidence that sentiment-driven traders do not sell short (Barber and Odean, 2006) I expected to find a negative relationship between the relative overpricing of stocks and the short interest ratio for stocks in the highest sentiment portfolio and a positive relationship between relative overpricing of stocks and the short interest ratio for stocks in the lowest sentiment portfolio. Instead I found a positive relationship between relative overpricing and the short interest ratio in the portfolio with the highest investor sentiment and a (weak) negative relationship between relative overpricing and the short interest ratio for stocks in the lowest investor sentiment portfolio. What could explain this difference?

The short interest ratio is measured in absolute terms, while overpricing and investor sentiment are measured relative to other stocks of the S&P 500. What the results show is that for stocks with relative high investor sentiment the short interest ratio (in absolute terms) increases more when stocks get relatively more overpriced than for stocks with relative low investor sentiment. However, this says nothing about the absolute values of investor sentiment and therefore the amount of sentiment-driven traders in the market nor the pattern of overpricing that could trigger the representative bias.

In order to say something about the absolute values of investor sentiment I compared the results in years with economic growth to recession years, because I expect that the absolute values of investor sentiment in the latter case will be lower. Although mostly insignificant, the comparison shows that the effect of relative overpricing on the short interest ratio is less positive for all sentiment portfolios in years of a recession, which confirms the observed trend. The results could be biased because short interest data is only available starting 1993 and on a monthly basis. The amount of datapoints is therefore quite small, especially for recession periods. To overcome this problem one could use security lending as proxy for short selling which figures are present on a daily basis. However, both in FactSet as in Eikon this data was unavailable so further research should investigate this.

Another way to test the results is by using a different proxy for investor sentiment. The most prominent one is the investor sentiment indicator of Baker and Wurgler (2006). However, this indicator measures overall investor sentiment and I need it on the firm-specific level to distinguish between stocks. Another promising indicator is Google trends to keep track how often a certain stock is searched for (Bank et al., 2011). The problem here is that intensive use of Google is only a relative new phenomenon so the sample period would be very small. An interesting extension would be to use the manager sentiment index of Jiang et al. (2017) and see how it affects investor sentiment. Jiang et al. (2017) use the aggregated textual tone in corporate financial disclosures as indicator and find that manager sentiment negatively predicts stock returns, especially for companies that are difficult to value and hard to arbitrage. One would expect that the textual tone of financial disclosures also affect investors and it would be interesting to investigate the effect of manager sentiment on investor sentiment but this is beyond the scope of this paper.

A third robustness check is to use a different proxy for overpricing. The relative overpricing measure in this paper is based on anomalies with data that are mostly yearly available in FactSet and Eikon. This means that eight out of the nine anomalies change only yearly and this greatly diminishes the amount of unique datapoints in the sample and therefore the robustness of my research. Alternatives would be to find a database with monthly data, or to include more anomalies. Stambaugh et Yuan (2017) provide a list of 73 anomalies so if those are all included there will be more anomalies with monthly available data and so more datapoints in the research. Another solution would be to calculate the overpricing of stocks in absolute terms. Therefore one needs to calculate the enterprise value of each firm and subtract his market value. This would also enable analysing patterns in absolute overvaluation levels and therefore could shed light on the precise role of the representative bias in investor sentiment. However, the valuation of a company has a subjective part which makes a fair comparison difficult. Therefore further

research needs to establish clear and transparent rules to value companies so this subjective part is minimized.

This paper has tried to connect investor sentiment to short-selling behaviour. This relationship seems not as straightforward as in many work on investor sentiment is assumed. The results indicate rational behaviour of investors. Because this is contradicting modern behavioural finance theories and the concepts of overpricing and investor sentiment are not clear-cut, further research needs to dive further into this relationship.

## References

- Antoniou, C., Doukas, J.A., Subrahmanyam, A., 2010. Sentiment and momentum. Working paper, UCLA
- Bank, M., Larch, M., Peter, G., 2011. Google search volume and its influence on liquidity and returns of German stocks. *Financial Markets and Portfolio Management* 25
- Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. *Journal of Financial Economics* 49, 307–343
- Barber, B.M., Odean, T., 2006. Do noise traders move markets? Zurich meetings papers
- Barber, B.M., Odean, T., Zhu, N., 2012. Do retail traders move markets? *Review of Financial Studies* 22, 151-186
- Baker, M., Stein, J.C., 2004. Market liquidity as a sentiment indicator. *Journal of Financial Markets* 7, 271-299
- Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61, 1645–1680
- Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives* 21, 129-151
- Baker, M., Wurgler, J., Yuan, Y., 2012. Global, local, and contagious investor sentiment. *Journal of Financial Economics* 104, 272-287
- Beber, A., Pagano, M., 2013. Short-selling bans around the world: evidence from the 2007-09 crisis. *Journal of Finance* 68, 343-381
- Black, F., 1986. Noise. *Journal of Finance* 41, 529-43
- Boehmer, E., Wu, J., 2013. Short selling and the price discovery process. *Review of Financial Studies* 26, 287-322
- Brown, G.W., Cliff, M.T., 2005. Investor sentiment and asset valuation. *Journal of Business* 78, 405-440
- Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. *Journal of Finance* 63, 2899-2939
- Charoenrook, A., Daouk, H., 2009. A study of market-wide short-selling restrictions. Department of Applied Economics and Management, Working Paper 2009-21
- Chen, H., Chong, T.T.L., Duan, X., 2010. A principal-component approach to measuring investor sentiment. *Quantitative Finance* 10, 339-347
- Chen, L., Novy-Marx, R., Zhang, L., 2010. An alternative three factor model. Working paper
- Coakley, J., Fuertes, A-M., 2006. Valuation ratios and price deviations from fundamentals. *Journal of Banking & Finance* 30, 2325-2346
- Cooper, M.J., Gulen, H., Schill, M.J., 2008. Asset growth and the cross-section of stock returns. *Journal of Finance* 63, 1609-1652
- Cross, R., Grinfeld, M., Lamba, H., 2005. A threshold model of investor psychology. *Physica A - statistical Mechanics and its applications* 354, 463-478
- Da, Z., Engelberg, J., Gao, P., (2015). The sum of all FEARS investor sentiment and asset prices. *Review of Financial Studies* 28, 1-32
- Das, S., Martinez-Jerez, A., Tufano, P., 2005. eInformation: A clinical study of investor discussion and sentiment. *Financial Management* 34, 103-137
- Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and over-reaction, *Journal of Finance* 53, 921-965
- Daniel, K., Hirshleifer, D., Subrahmanyam, A., 2001. Overconfidence, arbitrage and equilibrium asset pricing. *Journal of Finance* 56, 1839-1886
- DeLong, B., Shleifer, A., Summers, L.H., Waldmann, R., 1990. Noise trader risk in financial market. *Journal of Political Economy* 90, 703–738
- Eleswarapu, V.R., Reinganum, M.R., 2004. The predictability of aggregate stock market returns: Evidence based on glamour stocks. *Journal of Business* 77, 275-294
- Fama, E., French, K., 2008. Dissecting anomalies, *Journal of Finance* 63, 1653-1678
- Garcia, D., 2013. Sentiment during recessions. *Journal of Finance* 68, 1267-1300
- Grullon, G., Michenaud, S., Weston, J.P., 2015. The real effects of short-selling constraints. *Review of Financial Studies* 28, 1737-1767

- Hirshleifer, D., Hou, K., Teoh, S.H., Zhang, Y., 2004. Do investors overvalue firms with bloated balance sheets. *Journal of Accounting and Economics* 38, 297-331
- Huang, D., Jiang, F., Tu, J., 2015. Investor sentiment aligned: a powerful predictor of stock returns. *Review of Financial Studies* 28, 791-837
- Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for market efficiency. *Journal of Finance* 48, 65-91
- Jiang, F., Lee, J., Martin, X., 2019. Manager sentiment and stock returns. *Journal of Financial Economics* 123, 126-149
- Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 46, 171-185
- Karlsson, N., Seppi, D.J., Loewenstein, G., 2005. The 'ostrich effect': Selective attention to information about investments. SSRN 772125
- Kelly, M., 1997. Do noise traders affect stock prices? *Journal of Money, Credit, and Banking* 29, 351-363
- Lamont, O.A., Stein, J.C., 2006. Investor sentiment and corporate finance: Micro and macro. Conference: 118<sup>th</sup> Annual Meeting of the American-Economic-Association
- Lee, C.M.C., Schleifer, A., Thaler, R.H., 1991. Investor sentiment and the closed-end fund puzzle. *Journal of Finance* 46, 75-109
- Lee, W.Y., Jiang, C.X., Indro, D.C., 2002. Stock market volatility, excess returns, and the role of investor sentiment. *Journal of Banking and Finance* 26, 2277-2299
- Massa, M., Zhang, B., Zhang, H., 2015. The invisible hand of short selling: does short selling discipline earnings management? *Review of Financial Studies* 28, 1701-1736
- Moskowitz, T.J., Ooi, Y.H., Pedersen, L.H., 2012. Time series momentum. *Journal of Financial Economics* 104, 228-250
- National Bureau of Economic Research, 2020. US business cycle expansions and contractions. Viewed on June 23<sup>th</sup>, 2020 from: <https://www.nber.org/cycles.html>
- Neal, R., Wheatley, S.M., 1998. Do measures of investor sentiment predict returns? *Journal of Financial and Quantitative Analysis* 33, 523-547
- Novy-Marx, R., 2010. The other side of value: Good growth and the gross profitability premium. Working paper, University of Chicago
- Reed, A.V., 2002. Costly short selling and stock price adjustment to earnings announcements. Citeseer, 1-44
- Ritter, J.R., 1991. The long-run performance of initial public offerings. *Journal of Finance* 46, 3-27
- Saffi, P.A., Sigurdsson, K., 2011. Price efficiency and short selling. *Review of Financial Studies* 24, 821-852
- Schleifer, A., Vishny, R.W., 2010. Unstable banking. *Journal of Financial Economics* 97, 306-318
- Seok, S.I., Cho, H., Ryu, D., 2019. Firm-specific investor sentiment and daily stock returns. *North American Journal of Economics and Finance* 50, UNSP 100857
- Sloan, R.G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review* 71, 289-315
- Stambaugh, R.F., Yu, J., Yuan, Y., 2012. The short of it: investor sentiment and anomalies, 2012. *Journal of Financial Economics* 104, 288-302
- Stambaugh, R.F., Yu, J., Yuan, Y., 2015. Arbitrage asymmetry and the idiosyncratic volatility puzzle. *Journal of Finance* 70, 1903-1948
- Stambaugh, R.F., Yuan, Y., 2017. Mispricing factors. *Review of Financial Studies* 30, 1270-1315
- Subrahmanyam, A., 2008. Behavioural finance: a review and synthesis. *European Financial Management* 14, 12-29
- Titman, S., Wei, K.C.J., Xie, F., 2004. Capital investments and stock returns. *Journal of Financial and Quantitative Analysis* 39, 677-700
- Wang, F.A., 2000. Overconfidence, investor sentiment, and evolution. *Journal of Financial Intermediation* 10, 138-170
- Wang, Y.H., Keswani, A., Taylor, S.J., 2006. The relationship between sentiment, returns and volatility. *International Journal of Forecasting* 22, 109-123
- Wang, H., Yu, J., 2010. Dissecting the profitability premium. Working paper, University of Minnesota



- Yang, C., Gao, B., 2014. The term structure of sentiment effect in stock index futures market. *North American Journal of Economics and Finance* 30, 171-182
- Yu, J., Yuan, Y., 2011. Investor sentiment and the mean–variance relation. *Journal of Financial Economics* 100, 367–381
- Zhou, G., 2018. Measuring investor sentiment. *Annual Review of Financial Economics*