Handelshøyskolen BI

GRA 19703 Master Thesis

Thesis Master of Science 100% - B

Predefinert informasjon

Startdato:	09-01-2023 09:00 CET	**Termin:**	202310
Sluttdato:	03-07-2023 12:00 CEST	**Vurderingsform:**	Norsk 6-trinns skala (A-F)
Eksamensform:	T		
Flowkode:	202310||11378||IN00||B||T		
Intern sensor:	(Anonymisert)		

Deltaker

Navn:	Alexander Andersen og Per Harboe Hesjedal

Informasjon fra deltaker

Tittel *:	Short Interest as a Predictor of Stock Return: An Empirical Analysis
Navn på veileder *:	Nataliya Gerasimova

Inneholder besvarelsen konfidensielt materiale?:	Nei	**Kan besvarelsen offentliggjøres?:**	Ja

Gruppe

Gruppenavn:	(Anonymisert)
Gruppenummer:	3
Andre medlemmer i gruppen:	

Master Thesis

Short Interest as a Predictor of Stock Return: An Empirical Analysis

Supervisor:

Nataliya Gerasimova

Hand-in Date:

03.07.2023

Campus:

BI Bergen

Examination code and name:

GRA 19703 Master Thesis

Programme:

Master of Science in Business with Major in Finance

Abstract

This thesis presents an empirical analysis of the portfolio returns generated by short interest investment strategies and evaluates the performance using the Fama-French five-factor model. We compare the results obtained from this model with the Sharpe ratio of the strategy. Our findings indicate that high levels of short interest are predictive of abnormal returns. The absence of short interest corresponds to no abnormal return and in the least conservative strategy, negative abnormal returns are observed. The Fama-French five-factor model provides only marginal explanatory power beyond that offered by the three-factor model. This suggests that additional factors should be considered in future research on the topic to better understand the drivers of portfolio returns in the context of short interest investment strategies.

Acknowledgments

This thesis is written as a part of our Master of Science in Business with Major in Finance at BI Norwegian Business School.

We extend our gratitude to our supervisor, Nataliya Gerasimova, whose guidance has been instrumental in shaping this thesis. Her insightful suggestions and constructive feedback have driven us to enhance our work, both our academic growth and the quality of this thesis.

In addition, we are grateful for the support of our friends and family throughout the process of writing this thesis. Their encouragement has provided strength and support.

Contents

1.0 Introduction and motivation

Many researchers and practitioners are attempting to explain and predict stock market returns. The pioneering work of Jensen, Black, and Scholes (1972) brings attention to the significant relationship between systematic risk and expected return. Since then, several models have emerged, including the Fama and French (1993) three-factor model, followed later by the Fama and French (2015) five-factor model. These models aim to provide a more comprehensive explanation for the variations observed in stock market returns. Despite the efforts of numerous researchers, none of the models have achieved a fundamental status, as we continue to discover investment strategies that outperform the predictions of these models.

Short interest has experienced a substantial increase in the United States over the past few decades This has resulted in a rise in interest from both professionals in the financial industry and researchers in academia (Chung & Wang, 2021) (see figure 1.1). Boehmer et al. (2008) report that short selling accounts for about 13% of overall NYSE trading volume and that this percentage is similar across market capitalization groups. Boehmer et al. (2009) show that this fraction increases to more than 40% by 2007. These percentages are much higher than the monthly short interest data would suggest. The reason is that short positions are, on average, much shorter-lived than long positions. The growing interest in short selling signifies the presence of informed investors who can generate profits by betting against stocks. Therefore, it becomes increasingly relevant to utilize public information regarding short sellers' opinions on stocks.

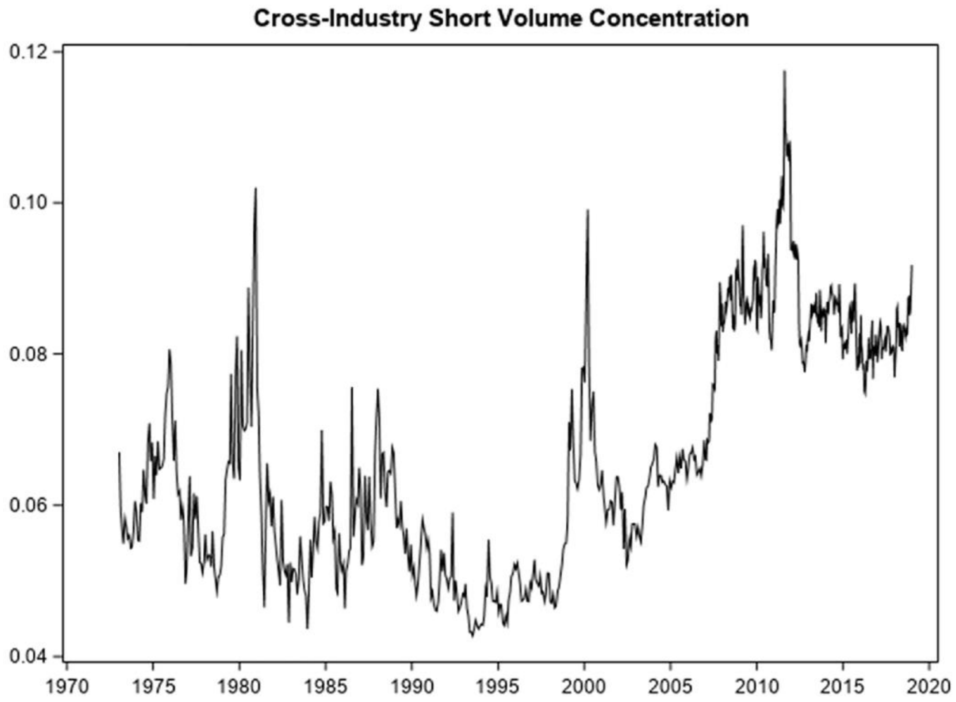

Figure 1.1 The Herfindahl-Hirschman Index (HHI) is computed by summing the squared percentages of the total cross-sectional short sales volume attributed to each FF48 industry. (Chung & Wang, 2021, p.2)

Short interest indicates the number of short selling positions in individual stocks as documented in publicly available records. For the US stock market, this information is accessible through various sources, including New York Stock Exchange (NYSE) data. Short interest data provides insight into the market sentiment surrounding individual stocks, as it reveals the extent to which investors expect stock prices to decline. An assumption of a short interest trading strategy is that a higher short interest indicates that a stock is likely to underperform in the future. This assumption is based on the idea that short selling is driven by informed investors who establish short positions in stocks they believe to be overvalued. Consequently, a high short interest may signal an overvalued stock that is likely to underperform compared to the market (Engelberg et al., 2012).

Numerous researchers examine the occurrence of abnormal negative returns in stocks characterized by elevated levels of short interest (Diamond & Verrecchia, 1987; Desai et al., 2002; Boehmer et al., 2008; Asquith et al., 2005; Mohamad et al., 2013) Conversely, alternative investigations focus on the performance of stocks with low levels of short interest (Boehmer et al., 2010). In cases where short sale constraints are not restrictive, the absence of short interest can potentially arise from short sellers purposefully avoiding such stocks due to their private valuations being equal to or surpassing the prevailing market price.

The presence of these anomalies serves as a motivating factor for our investigation into the development of an investment strategy encompassing long, short, and long/short positions. In this thesis, our primary focus is to investigate the potential of utilizing publicly available short selling information in an investment strategy to predict abnormal returns. Specifically, we aim to examine whether aggregated levels of short interest can serve as a reliable predictor of abnormal returns. Thus, our research question can be formulated as follows:

Is the publicly disclosed short interest ratio a reliable predictor of abnormal returns for United States (U.S.) stocks listed on the New York Stock Exchange (NYSE)?

To accomplish this, we analyze data on short interest ratios and monthly returns from U.S. stocks listed at NYSE spanning from March 1993 to February 2023. In order to ensure the availability of ex-ante trading signals, we sort stocks based on their short interest levels from the previous month, creating portfolios accordingly. Our investment strategy involves taking long positions in portfolios with the lowest short interest ratios and short positions in portfolios with the highest ratios. Additionally, we implement a long-short investment strategy that combines both approaches. To evaluate the performance of the portfolios, we regress their excess returns on the Fama-French five-factor model, aiming to identify any abnormal returns that cannot be explained by the model. This study expands upon prior research by utilizing an extensive and up-to-date data sample, while also incorporating additional factors to enhance the understanding of portfolio returns in the context of a short interest investment strategy.

Our findings indicate positive abnormal returns from both the short and long-short strategies, which cannot be explained by the Fama-French five-factor model. Despite the short-only strategy showing negative average returns and a slightly negative Sharpe ratio, it still generates positive abnormal returns. Conversely, the intercept analysis suggests that the returns from the long-only strategies are already priced into the model, despite observing positive average returns and the highest Sharpe ratio. Furthermore, the inclusion of the investment and profitability factors does not significantly enhance the explanation of excess portfolio returns. Specifically, the investment factor does not exhibit a significant coefficient in the regression results.

This paper is structured as follows: Section two provides an overview of the relevant literature pertaining to the topic, followed by a description of the data employed in the study. Section three outlines the methodology employed to develop and evaluate our investment strategy, while section four presents and discusses the results obtained. In section five, we summarize the key findings of our study and draw conclusions based on our analysis. Furthermore, we propose avenues for future research to extend and enhance the understanding of the topic at hand.

2.0 Literature review

The rise of short sales and short sellers has led to an increase in research focused on this subject. By examining the most significant research on the topic, this section aims to provide readers with insights and knowledge derived from similar content to ours. First, we present findings from research on the relationship between stock market return and short selling, and how various kinds of short selling information like short interest ratio can be used to explain stock market return. Last, we examine research that explores the impact of short sale constraints on asset prices, market efficiency, and market quality.

2.1 Short interest

Multiple studies show a correlation between high levels of short selling activity and poor abnormal returns in the stock market. The following studies show that even when controlling for various combinations of factor models investors can get an abnormal return by taking advantage of this anomaly. According to Boehmer et al. (2010), short interests are commonly viewed as a measure of investor heterogeneity. If this is true, a stock that can easily be shorted but is not targeted by short sellers indicates that market participants agree that it is not overvalued. It implies that short sellers do not have exclusive negative information about the stock.

Diamond and Verrecchia (1987) argue that investors who short sell are more motivated due to the high transaction costs of short selling, and therefore make more informed investment decisions. According to their research, an unforeseen increase in the short interest has a negative correlation with the stock price. Boehmer et al. (2008) find that highly shorted stocks underperformed light shorted stocks. Consistent with the findings of Desai et al. (2002) in the Nasdaq Market, they find that stocks with high short interest deliver negative returns after controlling for the market, book-to-market, size, and momentum factors. Their research reveals that stocks with high short interest display a persistent negative abnormal performance lasting 12 months.

Asquith et al. (2005) utilize short interest ratio data for NYSE-Amex stocks spanning from 1980 to 2002 and for Nasdaq stocks from June 1988 to December 2002. Their findings indicate that portfolios composed of stocks with high short interest exhibit underperformance compared to the market, as evidenced by the intercepts derived from four-factor time-series regression models. They categorize their portfolios based on institutional ownership; this reveals a consistent relationship with returns in equally weighted portfolios. Lower institutional ownership is associated with greater negative abnormal returns. However, value-weighted portfolios of high short interest stocks do not exhibit statistically significant evidence of underperformance or a consistent relationship with institutional ownership and subsequent returns. In contrast to the findings of Desai et al. (2002), their research reveals that high short interest ratio portfolios, which exclude stocks as soon as the short interest ratio falls below a predetermined threshold, experience more pronounced negative abnormal returns compared to portfolios that retain a firm for an extended inclusion period, such as 12 months, during the eight years following their sample period.

In contrast to earlier research focusing on the negative abnormal returns associated with stocks with high short interest, In their research, Boehmer et al. (2010) examine the predictive value of low short interest on future returns. Using data from Nasdaq, Amex, and NYSE covering the period from 1988 to 2005, they find that portfolios comprising stocks with low short interest exhibit significant and economically meaningful positive abnormal returns. These positive returns often exceed the negative returns observed in portfolios comprised of heavily shorted stocks. Consequently, the study highlights the slow assimilation of publicly available positive information regarding low short interest into market prices, thus raising concerns about overall market efficiency. Moreover, their results challenge prevailing theories regarding the impact of short

sale constraints.

Boehmer and Wu (2012) explore the role of short selling in the price discovery process, using a large dataset of daily short selling activity from the US market. Their findings suggest that short selling significantly contributes to price efficiency by accelerating the integration of negative information into stock prices. Further, their research indicates that short sellers tend to be more informed than buyers, leading to improved price discovery and a more efficient market, particularly during periods of high market uncertainty and increased stock-specific news flow. More recently, Rapach et al. (2016) demonstrate that short interest stands out as a potentially dominant predictor of aggregate stock returns. It outperforms various popular return predictors, both within the sample and out of sample. Their findings support the notion that short sellers possess valuable insights, allowing them to anticipate forthcoming aggregate cash flows and related market returns.

The presented studies above are based on US data and may not be representable for the international market. For instance, Boehmer et al. (2021) conduct a global study, examining stock market return data from 38 countries between 2006 and 2014. They find that the utilization ratio and days-to-cover ratio are the strongest predictors of stock market returns. Similarly, Bonne et al. (2017) conduct a global analysis of short interest information and find it to be a robust factor in explaining the abnormal stock return, both independently and within a multivariate framework.

Mohamad et al. (2013) research paper is centered on the UK market. Their findings show a negative correlation between short selling and stock returns. Their research suggests that an increase in short selling activity often precedes significant negative abnormal stock returns, implying that short selling can serve as an indicator of overvalued stocks. They also note a positive link between short selling activity, stock volatility, and trading volume, hinting at the potential role of short selling in improving market liquidity and price efficiency. However, they find that this pattern does not hold when it comes to short sales around the ex-dividend date, but persists during the 2008 financial crisis, mirroring the findings of Diether et al. (2008) in the US context.

2.2 Short sale constraints

Our extensive analysis period covers several periods that have experienced economic downturns and short sales constraints. Regulatory measures often place constraints on short sales. These limitations are typically the result of strategies deployed by regulators or exchanges to maintain stability and temper volatility in the market, especially during periods characterized by market stress or crisis. Several studies demonstrate how these constraints and their interplay with other factors can shape market dynamics. Saffi and Sigurdsson (2011), suggest that short sale constraints can enhance market stability, limit price manipulation by informed traders, and decrease the likelihood of price crashes. While other studies, such as that of Boehmer et al. (2013), point to potential negative effects, including inflated stock prices and volatility, diminished liquidity, and market inefficiencies due to the hindered ability of informed traders to correct mispricing.

Alternative studies propose that additional factors correlate with expected returns. Duffie et al. (2002) propose a model that predicts a link between expected returns and lending fees, in contrast to the somewhat contradictory literature regarding the relationship between returns and the current level of short interest. Boehmer et al. (2006) investigate the impact of the interaction between investor opinion dispersion and short sale constraints on stock valuation effects. Their findings indicate that stocks are not consistently overvalued when either of these conditions is not met.

This thesis distinguishes itself from previous research, such as Boehmer (2010) and Asquith et al. (2005), by adopting the Fama-French Five-Factor model instead of the four-factor model of Fama and French (1993) and Carhart (1997). By incorporating two additional factors into the analysis of the portfolio return of the short interest investment strategy, we gain new insights into the factors that drive variations in the strategy's return.

Our study utilizes a comprehensive dataset spanning the past 30 years, encompassing the period from March 1993 to February 2023. This extended dataset includes recent years that have not been previously analyzed, allowing us to examine the performance of short interest ratio as a predictor of stock return in more recent periods. Through this approach, we aim to contribute to the extensive body of literature on this topic.

3.0 Data

3.1 Data collection

In this thesis, we utilize a data sample obtained specifically from the New York Stock Exchange (NYSE). The dataset consists of the publicly disclosed information on short interest for a range of commonly traded U.S. stocks. We focus on stocks listed on the NYSE that are based in the U.S., excluding any non-U.S. stocks. U.S. firms are required to report their short interest positions in all equity securities two times a month, as mandated by FINRA (FINRA, 2023). The data is extracted from the Bloomberg terminal and includes essential variables such as the date, stock symbol, short interest, short interest rate, and the most recent recorded price for each equity.

We use data at a monthly frequency to ensure an understanding of the trends and patterns in the market. By choosing a monthly interval, we strike a balance between capturing shorter-term fluctuations without being overwhelmed by the noise of daily data. Our analysis covers a substantial period, ranging from March 1993 to February 2023, spanning a total of 30 years. This period includes various market phases, including periods of growth, recession, and recovery, thus providing a comprehensive backdrop for our study.

To ensure the quality of the data and mitigate potential errors, we filter the dataset to exclude missing observations in either short interest, short interest ratio, or last price. Once the dataset is streamlined to only include complete and dynamic price information, we proceed to calculate the log return based on the last recorded price for each stock. This process yields a total of 384,478 data points, which are distributed among 1551 different stocks.

In the context of the data required for factor models, our primary resource is the well-established and comprehensive database curated by Kenneth R. French (French, 2023). We specifically focus on extracting the US Fama-French three- and five-factor data. The Fama-French three-factor model incorporates the size factor, book-to-market, factor, and excess return on the market, while the five-factor model extends this by additionally considering the profitability and investment factor.

The alignment of the period for the stock, short interest data, and factor model data is crucial. This alignment ensures that we are comparing like with like and avoids the pitfalls of asynchronous data comparisons. Consequently, our analysis presents a synchronized and coherent view of market developments over these 30 years.

Table 3.1: Descriptive Statistics

	Observations	Mean	SD
Log return	384478	0,0031	0,1083
SIR	384478	0,0473	0,0625
MktRF	360	0,0069	0,0448
SMB	360	0,0015	0,0306
HML	360	0,0015	0,0329
RMW	360	0,0035	0,0273
CMA	360	0,0025	0,0218
RF	360	0,0018	0,0017

The table presents descriptive statistics for the log-return, short interest ratio and the Fama-French five-factor model in the period from 1993 to 2023.

3.2 Limitations

Our research may be subject to survivorship bias due to the inability to include delisted stocks where trading has ceased, which stems from limitations in extracting the dataset. As a result, our analysis predominantly focuses on the stocks of companies that have survived and continue to trade, while neglecting those that have failed or been delisted. This can create a distorted understanding of the performance and risks associated with investing in stocks, as the historical analysis becomes biased in favor of successful companies. Survivorship bias can lead to erroneous conclusions when examining historical market trends or attempting to discern patterns in the data. Consequently, this may give rise to deceptive beliefs regarding the true effectiveness of certain factors or investment strategies, as our analysis excludes failed companies (Elton et al., 1996; Carpenter & Lynch, 1999).

4.0 Methodology and Hypothesis

4.1 Portfolio construction

We utilize the short interest ratio obtained from the Bloomberg terminal to create long, short, and long/short portfolios. This methodology takes inspiration from the approach detailed in Fama and French's (2015) research, which outlines the creation of factor portfolios. The long portion of the portfolio comprises stocks with the lowest short interest ratio, while the short portion consists of stocks with the highest short interest ratio. This implies that stocks with a low short interest ratio are expected to yield higher returns, whereas stocks with a higher short interest ratio are anticipated to yield lower, or even negative, returns.

The construction of the different portfolios is inspired by Boehmer et al. (2010) with modifications to fit our data sample. First, we implement a univariate portfolio, sorting based solely on the short interest ratio. To ensure our trading signals, represented by this ratio, are ex-ante, we apply a one-month lag. Furthermore, our trading strategy involves sorting monthly return data into portfolios according to the previous month's short interest levels. We use three different sorting methods to generate 10, 50, and 80 portfolios, each constructed with equal weights according to the formula:

$$r_{Pt} = \sum_{i}^{N} r_i * \frac{1}{N} \quad (4.1)$$

Our investment strategy entails taking a long position in portfolios characterized by the lowest short interest ratios, specifically targeting the 10^{th}, 2^{nd}, and 1.25^{th} percentile portfolios. Additionally, we implement short positions in portfolios with the highest short interest ratios, focusing on the 90^{th}, 98^{th}, and 98.75^{th} percentile portfolios. Asquith et al. (2005) noted that portfolios composed of heavily shorted stocks offered higher returns when rebalanced regularly, particularly when short interest levels dipped below a defined threshold, as opposed to holding these portfolios for a predetermined period. Therefore, in line with these findings, we adhere to a monthly rebalancing cycle for our positions. At the commencement of each month, we rebalance and reconstruct our portfolio, to align it with the newly updated short interest levels from the preceding month.

4.1.1 Portfolio weighting

Recent literature has investigated the performance differences in equally weighted and value-weighted portfolios. Findings suggest that equal weights outperformed value weights on a general basis (DeMiguel et al., 2009; Malladi & Fabozzi, 2016; Plyakha et al. 2015). In addition, Asquith et al. (2005) and Boehmer et al. (2010) find that equally weighted portfolios deliver higher abnormal returns when short interest ratio is considered the main sorting factor.

Boehmer et al. (2010) present several compelling arguments in favor of employing equal weights in portfolio construction based on short interest. Firstly, they argue that value weighting fails to capture the average investor's net short position, which is zero for all stocks.. Unlike short sellers, the average investor does not hold short positions due to the absence of net debt in the global wealth portfolio. Second, while the supply of shares on the long side remains relatively stable in the short term, the short interest experiences significant fluctuations on a monthly basis, particularly for individual stocks. These fluctuations highlight the dynamic nature of short interest and the need to consider it in a portfolio construction framework.

Another important consideration put forth by Boehmer et al. (2010) is the potential bias introduced by a value-weighted portfolio when dealing with heavily shorted stocks that consistently underperform. In such cases, a value-weighted approach would assign reduced weights to these stocks each month, potentially obscuring the success of short sellers. By utilizing equal weights in the portfolio construction process, the performance of short sellers in heavily shorted stocks can be more accurately represented and evaluated. This approach ensures that the impact of short interest on portfolio performance is properly captured and avoids potential distortions caused by varying weights assigned to individual stocks over time.

Taking into account these considerations, we have made the informed decision to exclusively utilize equally weighted portfolios for our analysis. This approach aims to avoid the potential discrepancies associated with value-weighting and more accurately reflect the activities of short sellers.

4.2 Factor models

In the upcoming section, we present the methodology and theory underlying three key models: CAPM, the Fama-French three-factor model, and the Fama-French five-factor model. These models progressively build upon each other in an attempt to explain stock market returns. Specifically, the three-factor model extends the CAPM framework, while the five-factor model expands upon the three-factor model. In our research, we aim to investigate whether the short interest factor can predict stock market returns that cannot be explained by the five-factor model and if the two added factors provide an increase in the explanatory power of the model.

4.2.1 CAPM

Markowitz's groundbreaking work on portfolio selection laid the foundation for modern portfolio theory. In the 1960s, William Sharpe (1964) and John Litner (1965), as cited by Fama and French (2004), developed the capital asset pricing model (CAPM). This model establishes a connection between systematic risk and asset returns and is represented by equation 2.1. CAPM has been instrumental in understanding the relationship between risk and reward in investment portfolios.

$$E(r_i) = r_f + \beta_i\left(E(r_m) - r_f\right) \tag{4.2}$$

Where r_f is the risk-free rate, $E(r_i)$ is the expected return on an asset, $E(r_m)$ is the expected return on the market portfolio and β_i is the beta of asset i. The Capital Asset Pricing Model (CAPM) is based on the assumption that the market risk premium and a security's sensitivity to the market are the sole factors determining its expected return. In 1981, Banz (1981) proposed that CAPM is flawed in its specification and fails to consider an important additional factor: the "Size" factor.

Banz's findings revealed that smaller firms exhibited a higher average risk-adjusted return compared to larger firms.

4.2.2 Fama-French three-factor model

Fama and French (1992) conducted an evaluation that included the size factor proposed by Banz (1981), as well as Basu's (1983) earnings-price ratio (E/P) and Bhandari's (1988) findings on the correlation between leverage and expected return in models considering size (ME) and β. They also consider the research by Stattman (1980) and Rosenberg et al. (1985), which reveals a relationship between the average return on stocks and a firm's book-to-market ratio. Their findings suggest that size (ME) and book-to-market ratio provide explanatory power of average stock return on NYSE, Amex, and NASDAQ in the period of 1963-1990 and that β (the slope coefficient of stock return on the market) alone or in combination with other variable has little explanatory power on average return.

Fama and French (1993) extended the investigation conducted in Fama and French (1992) and discovered that the intercept from a three-factor regression incorporating book-to-market ratio, size, and the excess market return is nearly zero. This suggests that the risk proxies (book-to-market ratio and size) and the market factor account for a significant portion of the variation in average stock market returns. Their findings expand the CAPM framework by incorporating the book-to-market ratio and size factor to explain the cross-section of average stock returns. The Fama-French three-factor model is formulated as follows:

$$E(r_i) = r_f + \beta_1\left(E(r_m) - r_f\right) + \beta_2(SMB) + \beta_2(HML) + \varepsilon \qquad (4.3)$$

In equation (4.3) two components are added to equation (4.2), the return on a diversified portfolio of small stocks minus the return on a portfolio of large stocks (SMB), and the difference between the returns on diversified portfolios of high and low book-to-market ratio stocks (HML) (Fama & French, 2015).

4.2.3 Fama-French five-factor model

According to Fama and French (2015), a five-factor model that captures investment patterns in average stock return, value, profitability, and size perform better than the previously described Fama-French three-factor model. According to studies by Titman et al. (2004) and Novy-Marx (2013), the three-factor model (4.3) fails to account for much of the variation in average returns related to profitability and investment. To address this issue, Fama and French (2014) include profitability and investment factors in their model. The Fama-French five-factor model is given by:

$$E(r_i) = r_f + \beta_1\left(E(r_m) - r_f\right) + \beta_2(SMB) + \beta_3(HML) + \beta_4(RMW) \quad (4.4)$$

$$+\beta_5(CMA) + \varepsilon$$

Equation (4.4) introduces two additional factors to equation (4.3). RMW represents the return differential between portfolios consisting of stocks with strong and weak profitability, while CMA represents the return differential between portfolios comprising companies with low and high investment levels, commonly known as conservative and aggressive firms (Fama & French, 2014).

4.3 Performance measures

To assess the performance of the portfolio, we calculate conventional performance metrics, including the standard, mean of return, and Sharpe Ratio (SR). Additionally, our study enhances existing research conducted by Boehmer et al. (2010) and Asquith et al. (2005) by incorporating new factors into the asset pricing model, specifically utilizing the five-factor models. Through a regression analysis, we examine the intercept to assess whether the investment strategy has generated abnormal returns. This methodology enables us to build upon the previous findings and investigate the potential for the strategy to yield excess return beyond what can be explained by the model.

4.3.1 Regression

The Fama-French three-factor regression is given by the following formula:

$$R_{Pt} - Rf_t = \alpha_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + e_{it} \qquad (4.5)$$

Where R-squared (R_{Pt}) is the excess return of the portfolio we create based on the short interest ratio. In addition to employing the three-factor model, we further extend our analysis by regressing our results on the Fama-French five-factor model. This allows us to assess whether the model can explain a greater portion of the variance in the portfolio's returns and provides us with a deeper understanding of the factor exposure. The Fama-French five-factor regression is formulated as follows:

$$R_{Pt} - Rf_t = \alpha_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t \qquad (4.6)$$

$$+c_i CMA_t + e_{it}$$

When interpreting the result from the time series regression above our focus lies primarily on these key elements: the coefficients, the model's α_i, t-statistics, and R^2 and Adjusted R^2. The Intercept α_i is of particular interest, as it allows us to investigate if the portfolio generates an average monthly abnormal performance. The coefficients (b_i, s_i, h_i, r_i and c_i) from our regression represent the factor loading and give information about the portfolio return exposure to other well-known factors. We test if the intercept and the coefficient are different from zero by implementing a t-test. Further, R^2 represents the overall fit of the regression model. If R^2 is close to 1, there is an indication that the variables in the model explain most of the variation in the portfolio return.

4.3.2 Other Performance Measure

In addition to presenting the intercept, coefficients, and R^2 values obtained from our regression analysis, we complement our evaluation by incorporating a risk-adjusted return measure. We utilize the Sharpe ratio (SR) to assess the performance of our portfolios in a risk-adjusted manner. By considering the ratio of excess return to risk, the Sharpe ratio provides valuable insights into the risk-adjusted performance of our investment strategy.

The Sharpe ratio is a commonly employed metric for evaluating risk-adjusted returns. It incorporates risk into the calculation of a portfolio's total return, and its formula is as follows:

$$SR = \frac{R_P}{\sigma_P * \sqrt{12}} \tag{4.7}$$

In this formula, R_P represents the average portfolio return and σ_P denotes the standard deviation of the portfolio return, often interpreted as the total risk or volatility. This measure indicates the total return per unit of risk taken.

This measure provides valuable insights into the risk and return trade-off of our portfolios, allowing us to evaluate their performance in a more comprehensive manner. By incorporating these risk-adjusted return measures, we aim to provide a thorough and nuanced assessment of the portfolios constructed based on the short interest ratio.

4.4 Hypothesis

Our thesis aims to supplement and expand the current literature on short interest in the US stock market, utilizing an updated dataset sourced from the NYSE, which has not been previously examined. We seek to enhance the analysis by incorporating additional factors that have not been explored by prior researchers, thereby scrutinizing the performance of the short interest investment strategy in greater depth. We intend to examine the success of short sellers in identifying overvalued stocks for short selling, as well as their ability to identify and avoid undervalued stocks. Moreover, we want to test if short sellers are sophisticated traders that possess private information. To do so we test whether the Fama-French model misses some of the variations in average return related to shorting information. In addition, we explore the impact of incorporating additional factors into the model, to determine their ability to explain a greater proportion of the portfolio return.

If an asset-pricing model provides a complete explanation of expected returns, the regression of excess returns from any portfolio on the factor returns of the model will yield an intercept that is not significantly distinguishable from zero. When Fama and French developed the three-factor model and later the five-factor model, they stated that the models explained all the variance in average stock return and that the intercept of the regression model, α, is zero (Fama & French, 1993; Fama & French, 2015)

Following the estimation of the time-series regressions for the Fama-French model, we evaluate the performance by examining the estimated intercepts. Specifically, our research entails regressing the long, short, and long/short portfolios on the Fama-French Five-factor model to determine whether the intercepts significantly deviate from zero. Building upon the previous information presented, we have formulated the following hypothesis:

$$H_0: \alpha_i = 0$$

$$H_A: \alpha_i \neq 0$$

Where alpha (α_i) is the intercept in our regressions. We reject the null hypothesis that $\alpha_i = 0$ if the intercept has a significant deviation from zero at a confidence level of 95%. This implies that the investment strategy generates abnormal returns that are not accounted for or explained by the underlying model. We aim to assess the effectiveness of the five-factor model and determine whether its intercept differs statistically from the intercept derived from the three-factor model. To achieve this, we employ a t-test to compare the intercepts of the two regression models.

In addition to evaluating the performance of portfolio returns through intercept analysis, we incorporate the widely recognized Sharpe ratio, introduced by Sharpe (1966), as a key performance metric. Jobson and Korkie (1980) further emphasize the applicability of the Sharpe ratio in estimating efficient portfolios based on the principles of Modern Portfolio Theory and its value in comparing the risk-adjusted performance across various portfolios. Thus, to comprehensively assess portfolio performance while considering the tradeoff between risk and reward, we report both the intercept and the Sharpe ratio.

5.0 Results

Our study focuses on examining whether an investment strategy solely based on short selling information can generate abnormal returns. Specifically, we aim to quantify the extent to which the abnormal returns are associated with the short interest factor investment strategy and determine the portion that can be explained by the Fama-French five-factor. We seek to expand upon previous research by introducing additional factors that have not been previously considered in the regression analysis of portfolio returns from the short interest investment strategy. In addition, we endeavor to compare the explanatory power of the Fama-French five-factor model with the three-factor model, exploring the extent to which the new model provides new insights into portfolio returns.

Given this framework, our primary focus lies on the intercept obtained from the regression analysis. To evaluate the statistical significance of our findings, we establish a significance level of 5%, allowing us to assess whether we can reject the null hypothesis that the intercept is not significantly different from zero.

Our analysis starts with presenting the empirical findings derived from our regression analysis, where we applied the Fama-French five-factor model to the different percentile portfolios. We then examine the disparities between the regression outcomes obtained using the five-factor model versus the three-factor model. Lastly, we provide an overview of the supplementary performance measures utilized in our study.

5.1 Fama-French five-factor regression

The results of the Fama-French five-factor regression analysis are displayed in Table 5.1. The intercepts of the regression model hold significant importance as they indicate whether a portion of the portfolio returns can be attributed to factors other than those proposed by the Fama-French three-factor model. Moreover, we are interested in examining the coefficients to assess the exposure to the factors, and the R-squared value to determine the extent to which the Fama-French factors account for the variation in portfolio returns.

Table 5.1 shows that all intercepts associated with the three distinct short-only strategies display statistically significant deviations from zero at a significance level of 5%. Importantly, these intercepts exhibit positive values, underscoring their statistical significance and suggesting the presence of additional factors contributing to the portfolio returns beyond the ones considered in the Fama-French five-factor model. This finding aligns with the research conducted by Diamond and Verrecchia (1987), Asquith et al. (2005), Boehmer et al. (2008), and Rapach et al. (2016), which identified a significant negative intercept from regression on a long portfolio consisting of stocks exhibiting the highest short interest ratio.

Conversely, for the long-only regression involving the 2nd and 1.25th percentile portfolios, the intercepts are not significantly different from zero, indicating a lack of abnormal returns. Only the 10th percentile portfolio exhibits a statistically significant intercept, which is negative and economically small, with a value of -0.264%. This suggests that the portfolio's performance yields abnormal negative returns after accounting for the five Fama-French factors. In terms of the long-short strategy, similar to the short-only strategy, all intercepts demonstrate statistically significant differences from zero. Moreover, these intercepts are positively valued, suggesting the presence of additional factors contributing to the portfolio returns within the long-short strategy.

In summary, our intercept analysis reveals that both the short-only strategy and the long-short strategy generate abnormal returns that cannot be fully explained by the Fama-French three-factor model. These findings are consistent with a similar study conducted by Boehmer et al. (2010) which also observes the presence of abnormal returns beyond the model's explanatory power. Interestingly, Boehmer et al. (2010) find that the absence of short interest predicts positive abnormal returns. In our study, when implementing a strategy of going long on percentile portfolios with the lowest short interest, we find that two portfolios do not demonstrate a significant deviation from the expected returns, and one even exhibited a negative intercept. These findings shed light on the nuanced variations in the intercepts across different strategies and underscore the importance of considering additional factors to fully capture the complexities of portfolio returns.

Proceeding to analyze the slope coefficient in Table 5.1. First and foremost, all the coefficients for RM-RF, HML, and SMB exhibit significant deviations from zero, providing strong evidence to reject the null hypothesis at a significance level of 5%. This implies that these factors play a substantive role in influencing portfolio returns. For the short-only strategy, it is noteworthy that all the significant coefficients exhibit negative values. Specifically, the slope coefficient associated with the excess market return approaches approximately –1,1 across all the short portfolios. This suggests a strong inverse relationship between the portfolio and excess market returns for the short positions. The exposure to SMB for the short positions is negative and ranges from -0,6247 to -0,8692 indicating that the underlying stocks in the portfolio are exposed to small capitalization stocks. Examining the Profitability (RMW) factor, we observe that all the short portfolios exhibit a negative exposure within a range of -0.2808 to -0.2280. This indicates a substantial negative loading on the profitability premium, albeit with a relatively low level of sensitivity.

In contrast, the long-only strategy showcases positive coefficients for the first three factors. The coefficient for the excess market return is approximately 0.5, suggesting a positive association with the market. However, this coefficient also indicates a relatively low level of market risk. This finding aligns with the research conducted by Boehmer et al. (2010), who discovered a similar nature and magnitude of the market coefficient. It suggests a consistent exposure to the market, even after incorporating two additional factors and considering a more recent period. Furthermore, the SMB and HML factors demonstrate coefficients of approximately 0.2 which indicate a positive but rather low exposure to value and small capitalization stocks. Regarding the RMW factor within the long-only portfolios, both the 10th percentile and the 1.25th percentile portfolios show statistically significant positive factor loadings at a 5% significance level, suggesting a slightly favorable association with the portfolio's returns.

Within the long-short strategy, all significant coefficients display negative values. Notably, the market exposure coefficient for the 95th percentile strategy reaches -0.5618, highlighting a substantial negative impact of the market factor on portfolio returns. The coefficient associated with the exposure to the SMB factor for the 90th-percentile strategy demonstrates the least negative value of -0.2906 for the long-short strategy. This finding implies a noteworthy inverse relationship between the SMB factor and the portfolio returns. Furthermore, the only significant coefficient for RMW is observed for the 10th-90th percentile portfolio, revealing a relatively low but negative exposure to the profitability factor.

Another interesting finding that emerges from our analysis is that none of the slope coefficients associated with the investment (CMA) factor exhibit statistical significance at the 5% level. This implies that the variation in portfolio returns cannot be attributed to the difference in returns between firms adopting conventional investment practices and those pursuing more aggressive investment strategies. Therefore, the lack of significant coefficients suggests that the investment factor does not significantly contribute to explaining the variance in portfolio returns.

Alongside the intercept and slope coefficients, Table 5.1 also presents the R-squared values obtained from the regression analysis. Notably, the 90th-percentile short-only strategy exhibits the highest R-squared, indicating that the Fama-French five-factor model accounts for approximately 89% of the variance in portfolio returns.

Furthermore, a discernible pattern emerges when considering the various strategies. The strategies with a larger number of stocks display higher R-squared values. Additionally, as the shorting method becomes more conservative, involving the exclusion of more stocks based on the levels of short interest, the R-squared decreases. This pattern suggests that the explanatory power of the Fama-French three-factor model strengthens when applied to portfolios with more stocks and when employing less conservative shorting techniques.

In summary, this analysis provides valuable insights into the dynamics between the factors and portfolio strategies. The statistical significance of several factors and discernible positive/negative relationships increase our understanding of how these factors influence portfolio returns. Additionally, the analysis enhances our knowledge of the interplay between the factors and portfolio performance, as indicated by the R-square value.

Table 5.1: Regression outputs Fama-French five-factor model

Portfolios # Stocks	Intercept	RM-RF	HML	SMB	RMW	CMA	R^2
Short 90th #155 Stocks	0,0059***	-1,1156***	-0,5543***	-0,6247***	-0,2808***	0,0654	0,8892
Short 98th #31 Stocks	0,0071***	-1,1002***	-0,5142***	-0,7695***	-0,2280***	0,0803	0,7473
Short 98,75th #19 Stocks	0,0067***	-1,0600***	-0,5611***	-0,8692***	-0,2900***	0,1606	0,7041
Long 10th #155 Stocks	-0,0026**	0,5589***	0,1043**	0,2534***	0,1062**	-0,0351	0,5258
Long 2nd #31 Stocks	-0,0002	0,5331***	0,1251**	0,2630***	0,1207	0,1207	0,4126
Long 1,25th #19 Stocks	0,0004	0,5287***	0,1588***	0,2662***	0,2066***	-0,0102	0,3668
Long – Short 10th – 90th #310 Stocks	0,0051***	-0,5581***	-0,4472***	-0,3749***	-0,1767***	0,0303	0,5222
Long – Short 2nd - 98th #62 Stocks	0,0088***	-0,5685***	-0,3862***	-0,5101***	-0,1093	0,0323	0,3458
Long – Short 1,25th – 98,75th #39 Stocks	0,0090***	-0,5327***	-0,3994***	-0,6066***	0,0854	-0,1503	0,3239

Table 5.1 displays the results from the Fama-French five-factor regression on the monthly portfolio return. The different factors in the table are described in section 4.1.3. The portfolios Short 90^{th}, Short 98^{th} and Short $98{,}7^{th}$ include stocks that are in the portfolio with the highest short interest ratio (SIR) in month t−1, respectively. The portfolios Long 10^{th}, Long 2^{nd}, and Long $1{,}25^{th}$ include stocks that are in the portfolio with the lowest SIR in month t−1. The average number of stocks in the portfolios is shown under "# stocks." R^2 represents R-squared and shows how well the data fits the regression.

5.2 Comparison of models

In this section, we compare the findings from the three-factor model to the previously presented five-factor model. The purpose of this analysis is to assess the extent to which the two additional factors enhance the model and explain the variation in return observed in the portfolios. In Tabel 5.2, we report findings only for the findings using the 10^{th} and 90^{th} percentile for the Fama-French three-factor regression the other strategies are reported in the appendix.

Upon examining the results of the Fama-French three-factor regressions, a similar pattern emerges in the intercepts as observed in the five-factor regressions. Notably, when analyzing the short and long-short portfolio, all intercepts are found to be positive and statistically significant at a significant level of 5%. Conversely, for the long-only strategy, the intercepts are not significantly different from zero at the specified significance level of 5%. This indicates that the long-only strategy intercepts do not provide strong evidence of additional contribution to portfolio returns beyond those encompassed by the Fama-French three-factor model. In the regression with the five-factor model the intercept from the 10^{th} percentile portfolio was negative and significant. This implies that after accounting for the two new factors additional portfolio return is explained by the model. To assess the statistical disparity in the intercept between the two distinct models, we performed a t-test. The outcomes of this t-test can be found in Table A.1 in the appendix, revealing that none of the intercepts from the various strategies exhibited significant differences from each other at a significance level of 5%.

Continuing our analysis, we shift our focus to the slope coefficients outlined in Table 5.2. All coefficients in the three-factor model are significantly different from zero. Furthermore, we observe that these coefficients align closely in terms of both magnitude and nature with the findings observed in the preceding section. This consistency implies that the portfolio retains a consistent exposure to the underlying factors, even in the absence of the two factors that were removed from the model. This highlights the importance of investigating the explanatory power of the two models.

In Table 5.2, in addition to the intercept and slope coefficients, we also present the R-squared values derived from the Fama-French three-factor regressions. Notably, these values demonstrate a slight decrease compared to those in Table 5.1, with an average reduction of approximately 1% in the regression's explanatory power upon the removal of the two additional factors. However, it is crucial to note that R-squared always increases when variables are added to the equation, making it insufficient to solely rely on R-squared for gauging the added explanatory power of the model. To address this concern, we have reported the adjusted R-squared values for both strategies in Table 5.3.

Table 5.2: Regression outputs Fama-French three-factor model

Portfolios # Stocks	Intercept	RM-RF	SMB	HML	R^2
Short 90th #155 Stocks	0,0047***	-1,0934***	-0,4919***	-0,7048***	0,8726
Long 10th #155 Stocks	-0,0023	0,5517***	0,2044***	0,16068***	0,5197
Long – Short 10th – 90th #310 Stocks	0,0043***	-0,5429***	-0,2906***	-0,5435***	0,5094

Table 5.2 displays the results from the Fama-French three-factor regression on the monthly portfolio return. The different factors in the table are described in section 4.1.2. The portfolio Short 90^{th} includes stocks that are in the portfolio with the highest short interest ratio (SIR) in month t−1, respectively. The portfolios Long 10^{th} include stocks that are in the portfolio with the lowest SIR in month t−1. The average number of stocks in the portfolios is shown under "# stocks." R^2 represents R-squared and shows how well the data fits the regression model. Regression coefficients are reported with p-values, * stands for significance at 10% level, ** - at 5% level, *** -at 1% level.

Table 5.3 reveals that the inclusion of two additional factors only marginally improves the model's explanatory power. Notably, the regression analysis for the short-only 90th percentile portfolio exhibits the most substantial increase in explanatory power, with a modest improvement of 1.61%. These findings suggest that the added factors have limited influence in capturing additional variation in the portfolio return. Furthermore, these results provide valuable insights into the portfolio's relationship with the Profitability and Investment factors within the Fama-French model.

The inclusion of the two new factors in our analysis of portfolio returns has yielded limited to no additional explanatory power. Prior studies have examined the relationship between the short interest factor strategy's return and the Fama-French three-factor model, with the subsequent incorporation of the momentum factor. These investigations, conducted by Desai et al. (2002), Asquith et al. (2005), and Boehmer et al. (2010), have demonstrated a significant correlation between all factors and portfolio returns. However, the specific impact of the momentum factor on augmenting the explanatory power of the regression model was not thoroughly examined in these studies. Consequently, the degree to which the momentum factor enhances the model remains uncertain.

Table 5.3: Adjusted R-square

Portfolios	FF3	FF5
Short 90th	0,8715	0,8876
Long 10th	0,5156	0,5191
Long – Short 10th – 90th	0,5052	0,5154

Table 5.3 displays the adjusted R-square for the Fama-French three-factor and five-factor regression.

5.3 Other performance measures

In addition to the regression analysis, we assessed the risk-adjusted performance of the strategies. We have employed various metrics such as average return, standard deviation, and Sharpe ratio to evaluate their performance. By considering alternative performance metrics, we aim to gain a comprehensive understanding of how the portfolios perform, beyond the scope of the Fama-French factor utilized in our analysis.

The average monthly return from the different strategies is presented in Table 5.4. For the long-only strategy, we observe negative returns for the three different percentile strategies. The long-only strategy exhibits only a positive average monthly return and the 1,25th percentile strategy is the one with the highest value averaging at 0,7271% monthly return. Similarly, the long-short strategy delivers a positive average return ranging from 0,1254% in the least conservative portfolio sorting strategy to 0,5750% in the most conservative strategy.

In the Sharpe ratios presented in Table 5.3, we observe a similar pattern to the average monthly returns across different strategies. As the portfolio sorting becomes more conservative, the Sharpe ratio increases for all strategies. The long-short strategy stands out with the highest observed Sharpe ratio at 0.6031.

Table 5.4: Sharpe Ratio

Portfolios	Return	SR
Short 90th	-0,0026 (-0,0602)	-0,1518
Short 98th	-0,0012 (-0,0675)	-0,0610
Short 98,75th	-0,0015 (0,0981)	-0,0758
Long 10th	0,0039 (-0,0371)	0,3630
Long 2nd	0,0062 (-0,0404)	0,5343
Long 1,25th	0,0073 (0,0418)	0,6031
Long – Short 10th – 90th	0,0013 (-0,0430)	0,1010
Long – Short 2nd - 98th	0,0050 (-0,0565)	0,3092
Long – Short 1,25th – 98,75th	0,0058 (-0,0602)	0,3306

Table 5.4 displays the average monthly return from the portfolios. In the parenthesis are the standard deviation reported. SR reports the Sharpe ratio for each portfolio.

However, these findings contradict the results of the regression analysis. According to the analysis, the short strategy had the highest alpha, followed by the long-short strategy, while the long-only strategy had the lowest alpha, which was negative. One possible explanation for these contradictory findings is that the long-only strategy interacts differently with the Fama-French factors. Table 5.1 shows that all significant slope coefficients for the long strategy are positive. Additionally, we have provided the cumulative returns of all strategies and the

Fama-French factors in the appendix. The graph for the long-only strategy outperforms the long-short and short-only strategies in terms of cumulative returns. However, there is a noticeable correlation between the returns and the first three factors, particularly the market factor. This correlation can explain the absence of abnormal returns from the strategy as the market can explain a significant portion.

6.0 Conclusion

The primary objective of this thesis is to investigate if short sellers are informed traders and assess the potential for generating abnormal returns by exploiting short selling information. We employ a systematic approach to achieve this by sorting stocks listed on the NYSE into portfolios based on their short interest ratios from the preceding month. Moreover, our investment strategy involves taking short positions in portfolios exhibiting the highest percentile of short interest and taking long positions in portfolios characterized by the lowest percentile of short interest. By implementing this approach, we aim to capitalize on the potential opportunities presented by stocks with high short interest ratios and those with low short interest ratios, respectively.

Our findings reveal that both the short and long-short strategies yield positive abnormal returns that cannot be accounted for by the Fama-French five-factor model. It is noteworthy that even though the short-only strategy exhibits an average negative return and a slightly negative Sharpe ratio, it still generates positive abnormal returns. Contradictory, even though we find positive averages return and the highest Sharpe ratio in the long only strategies, the findings from the intercept analysis indicate that the return is already priced in the model.

Moreover, our findings indicate that the inclusion of the investment and profitability factor does not significantly improve the explanation of excess portfolio returns. In particular, the investment factor exhibits no significant coefficient in the regression results. This highlights the potential for future research to explore alternative factors that may better capture the variations in portfolio returns. Extending the scope of investigation in this manner might yield a more robust understanding and offer more insights into the variability of portfolio return from the short interest investment strategy.

7.0 References

Asquith, P., Pathak, P. A., & Ritter, J. R. (2005). Short interest, institutional ownership, and stock returns. *Journal of Financial Economics*, *78*(2), 243–276. https://doi.org/10.1016/j.jfineco.2005.01.001

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3–18. https://doi.org/10.1016/0304-405x(81)90018-0

Basu, S. (1983). The relationship between earnings' yield, market value and return for NYSE common stocks. *Journal of Financial Economics*, *12*(1), 129–156. https://doi.org/10.1016/0304-405x(83)90031-4

Bhandari, L. C. (1988). Debt/equity ratio and expected common stock returns: Empirical evidence. *The Journal of Finance*, *43*(2), 507–528. https://doi.org/10.1111/j.1540-6261.1988.tb03952.x

Boehme, R. D., Danielsen, B. R., & Sorescu, S. M. (2006). Short-sale constraints, differences of opinion, and overvaluation. *Journal of Financial and Quantitative Analysis*, *41*(2), 455–487. https://doi.org/10.1017/s0022109000002143

Boehmer, E., & Wu, J. J. (2012). Short selling and the Price Discovery Process. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.972620

Boehmer, E., Huszar, Z. R., & Jordan, B. D. (2010). The good news in short interest. *Journal of Financial Economics*, *96*(1), 80–97. https://doi.org/10.1016/j.jfineco.2009.12.002

Boehmer, E., Huszár, Z. R., Wang, Y., Zhang, X., & Zhang, X. (2021). Can shorts predict returns? A global perspective. *The Review of Financial Studies*, *35*(5), 2428–2463. https://doi.org/10.1093/rfs/hhab079

Boehmer, E., Jones, C. M., & Zhang, X. (2008). Which shorts are informed? *The Journal of Finance*, *63*(2), 491–527. https://doi.org/10.1111/j.1540-6261.2008.01324.x

Boehmer, E., Jones, C. M., & Zhang, X. (2013). Shackling short sellers: The 2008 shorting ban. *Review of Financial Studies*, *26*(6), 1363–1400. https://doi.org/10.1093/rfs/hht017

Bonne, G., Jain, V., Melas, D. (2017) Is There a Short Interest Factor? A Global and Multi-dimensional Analysis of Short Interest and the Cross Section of Stock Returns. MSCI. https://www.msci.com/documents/10199/4585b0fa-8415-4fb2-ad25-b21b11c06910

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Carpenter, J. N., & Lynch, A. W. (1999). Survivorship bias and attrition effects in measures of Performance Persistence. *Journal of Financial Economics*, *54*(3), 337–374. https://doi.org/10.1016/s0304-405x(99)00040-9

Chung, C. Y., Liu, C., & Wang, K. (2021). The big picture: The industry effect of short interest. *International Review of Financial Analysis, 76*, 101760. https://doi.org/10.1016/j.irfa.2021.101760

DeMiguel, V., Garlappi, L., & Uppal, R. (2007). Optimal versus naive diversification: How inefficient is the 1/*n* portfolio strategy? *Review of Financial Studies*, *22*(5), 1915–1953. https://doi.org/10.1093/rfs/hhm075

Desai, H., Ramesh, K., Thiagarajan, S. R., & Balachandran, B. V. (2002). An investigation of the informational role of short interest in the nasdaq market. *The Journal of Finance*, *57*(5), 2263–2287. https://doi.org/10.1111/0022-1082.00495

Diamond, D. W., & Verrecchia, R. E. (1987). Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Economics*, *18*(2), 277–311. https://doi.org/10.1016/0304-405x(87)90042-0

Diether, K. B., Lee, K.-H., & Werner, I. M. (2008). Short-sale strategies and return predictability. *Review of Financial Studies*, *22*(2), 575–607. https://doi.org/10.1093/rfs/hhn047

Duffie, D., Gârleanu, N., & Pedersen, L. H. (2002). Securities lending, shorting, and pricing. *Journal of Financial Economics*, *66*(2-3), 307–339. https://doi.org/10.1016/s0304-405x(02)00226-x

Elton, E. J., Gruber, M. J., & Blake, C. R. (1996). Survivor bias and Mutual Fund Performance. *Review of Financial Studies*, *9*(4), 1097–1120. https://doi.org/10.1093/rfs/9.4.1097

Engelberg, J. E., Reed, A. V., & Ringgenberg, M. C. (2012). How are shorts informed? *Journal of Financial Economics*, 105(2), 260–278. https://doi.org/10.1016/j.jfineco.2012.03.001

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, *47*(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and Bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/10.1016/0304-405x(93)90023-5

Fama, E. F., & French, K. R. (2004). The Capital Asset Pricing Model: Theory and evidence. *Journal of Economic Perspectives*, *18*(3), 25–46. https://doi.org/10.1257/0895330042162430

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

FINRA, (2023, May 05). *Short Interest Reporting.* https://www.finra.org/filing-reporting/regulatory-filing-systems/short-interest

French, K. R., (2023). Kenneth R. French - Data library. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.htmlhttps://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Jensen, M. C., Black, F., & Scholes, M. S. (1972). The capital asset pricing model: Some empirical tests.

Jobson, J. D., & Korkie, B. (1980). Estimation for Markowitz Efficient Portfolios. *Journal of the American Statistical Association*, 75(371), 544–554. https://doi.org/10.1080/01621459.1980.10477507

Malladi, R., & Fabozzi, F. J. (2016). Equal-weighted strategy: Why it outperforms value-weighted strategies? theory and evidence. *Journal of Asset Management*, *18*(3), 188–208. https://doi.org/10.1057/s41260-016-0033-4

Mohamad, A., Jaafar, A., Hodgkinson, L., & Wells, J. (2013). Short selling and stock returns: Evidence from the UK. *The British Accounting Review*, *45*(2), 125–137. https://doi.org/10.1016/j.bar.2013.03.001

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, *108*(1), 1–28. https://doi.org/10.1016/j.jfineco.2013.01.003

Plyakha, Y., Uppal, R., & Vilkov, G. (2015). Why do equal-weighted portfolios outperform value-weighted portfolios? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2724535

Rapach, D. E., Ringgenberg, M. C., & Zhou, G. (2016). Short interest and aggregate stock returns. *Journal of Financial Economics*, 121(1), 46–65. https://doi.org/10.1016/j.jfineco.2016.03.004

Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, *11*(3), 9–16. https://doi.org/10.3905/jpm.1985.409007

Saffi, P. A., & Sigurdsson, K. (2011). Price efficiency and short selling. *Review of Financial Studies*, *24*(3), 821–852. https://doi.org/10.1093/rfs/hhq124

Sharpe, W. F. (1966). Mutual Fund Performance. *The Journal of Business* (Chicago, Ill.), 39(1), 119–138. https://doi.org/10.1086/294846

Stattman, D. (1980). Book values and stock returns. The Chicago MBA: *A Journal of Selected Papers*, 4(1), 25–45.

Titman, S., Wei, K. C., & Xie, F. (2004). Capital Investments and Stock Returns. *Journal of Financial and Quantitative Analysis*, *39*(4), 677–700. https://doi.org/10.1017/s0022109000003173

Appendix

Figure A.1: Cumulative return of long-short 10th – 90th percentile portfolio with Fama-French five-factors

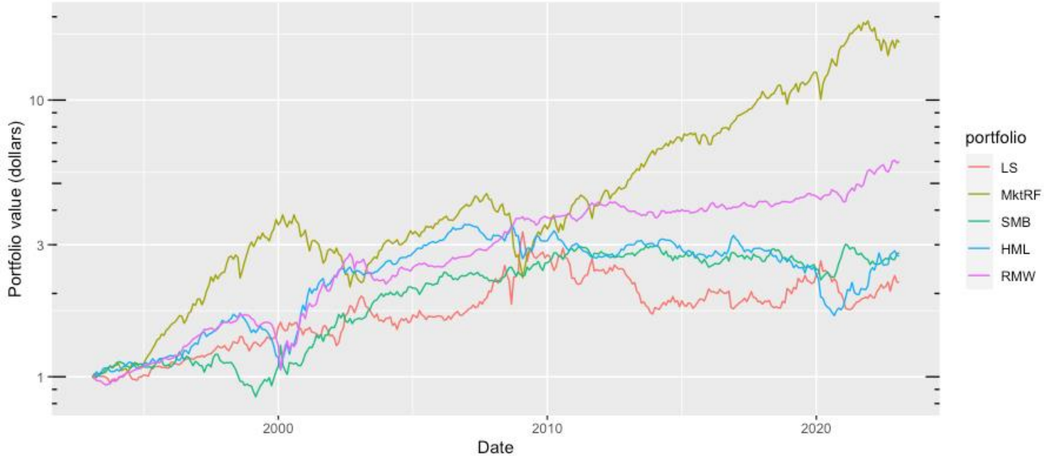

Figure A.2: Cumulative return of long 10th percentile portfolio with Fama-French five-factors

Figure A.3: Cumulative return of short 90th percentile portfolio with Fama-French five-factors

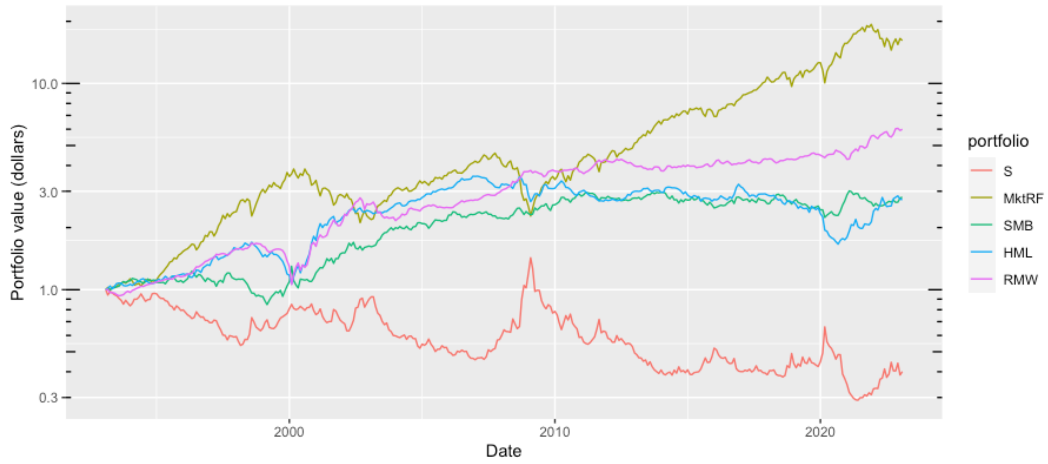

Figure A.4: Cumulative return of long-short 2nd-98th percentile portfolio with Fama-French five-factors

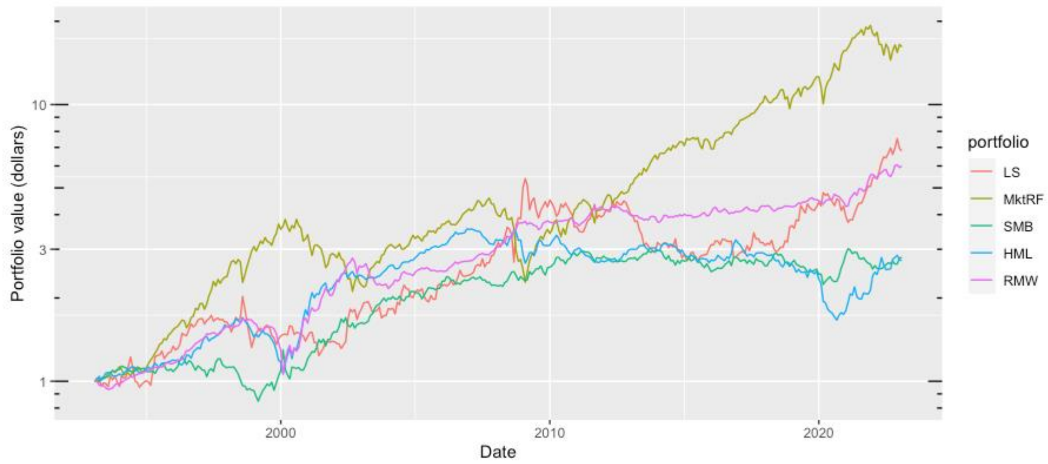

Figure A.5: Cumulative return of long 2nd percentile portfolio with Fama-French five-factors

Figure A.6: Cumulative return of short 98th percentile portfolio with Fama-French five-factors

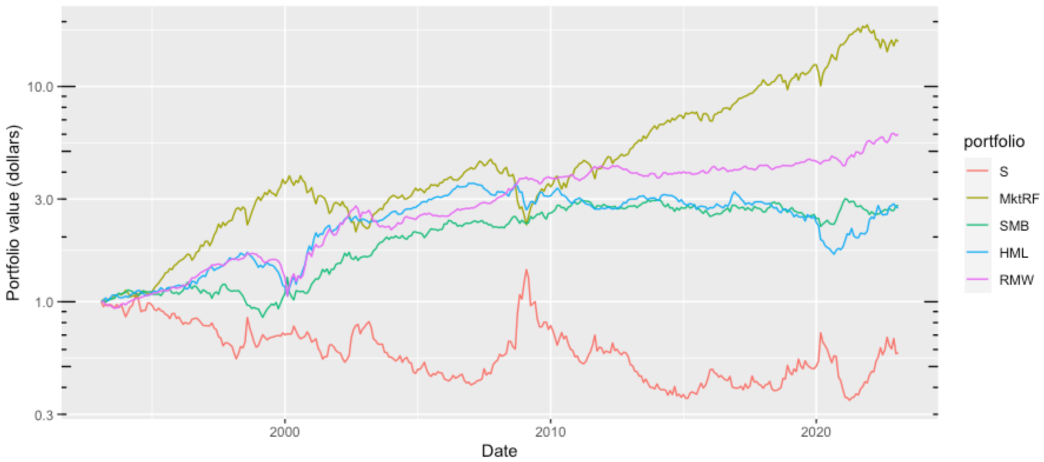

Figure A.7: Cumulative return of long-short 1,25th – 98,75th percentile portfolio with Fama-French five-factors

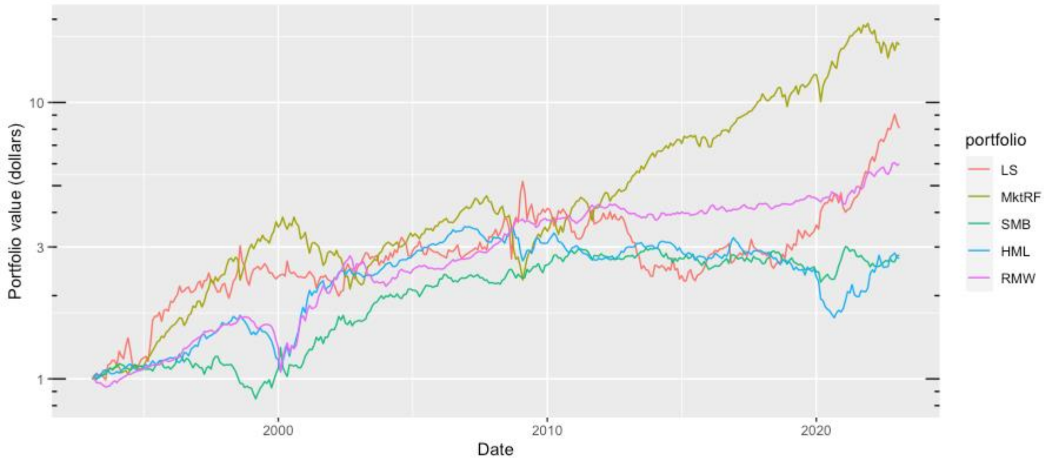

Figure A.8: Cumulative return of long 1,25th percentile portfolio with Fama-French five-factors

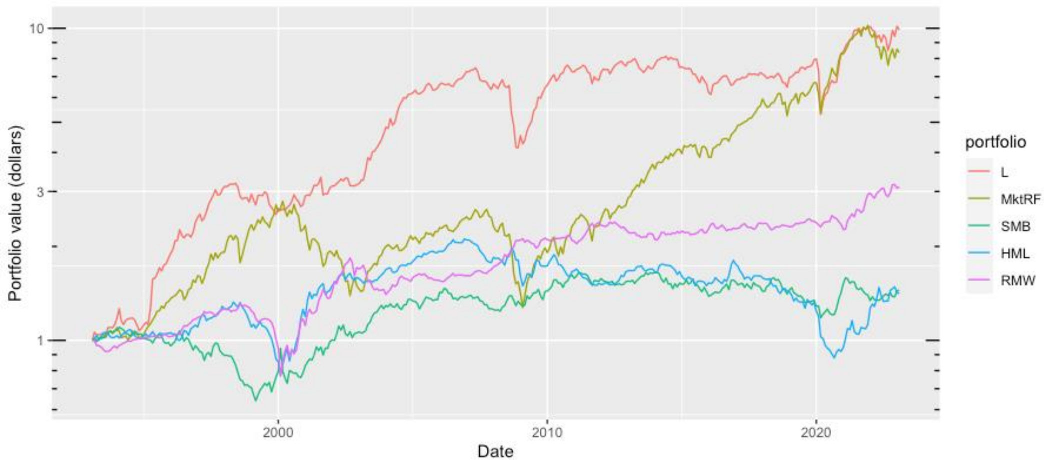

Figure A.9: Cumulative return of Short 98,75th percentile portfolio with Fama-French five-factors

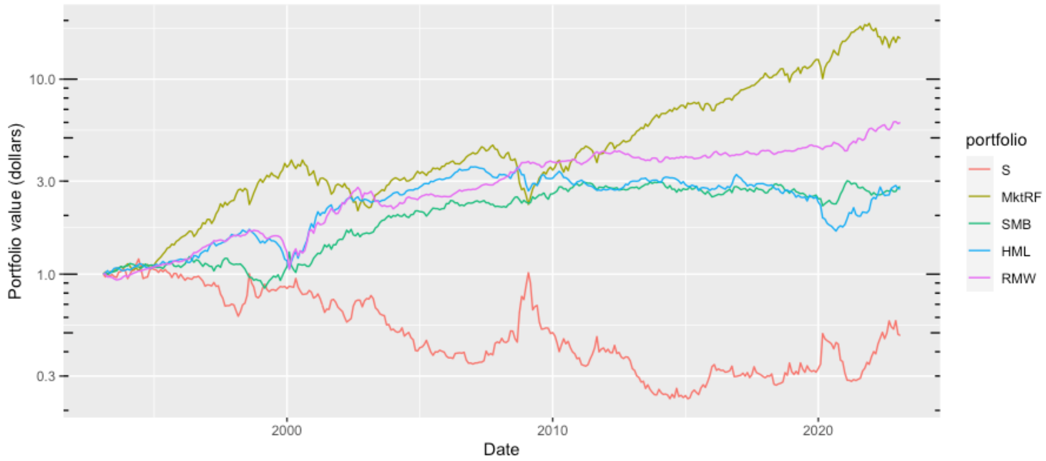

Table A.1: T-test for intercept comparison

Portfolios	t-statistic	p-value
Short 90th	0,6890	0,4913
Long 10th	-0,1958	0,8449
Long - Short 10th – 90th	0,3139	0,7538

Table A.1 presents the t-statistics and corresponding p-values from a t-test that examines whether the intercept in the Fama-French five-factor regression significantly differs from the intercept in the Fama-French three-factor regression. The portfolio labeled as "Short 90th" comprises stocks from the portfolio with the highest short interest ratio (SIR) in month t−1. Similarly, the portfolios labeled as "Long 10th" consist of stocks from the portfolio with the lowest SIR in month t−1.