LUND UNIVERSITY
School of Economics and Management

Returning to Quality

An Empirical Investigation of Short Sale Constraints Effect on a Quality Investment Strategy

Master Thesis

Carl Johan Ingvarsson

supervised by
Prof. Hossein Asgharian

June 3, 2021

Acknowledgements

I would like to extend my sincerest gratitude towards my supervisor Hossein Ashgarian for providing insightful and enthusiastic advice. I am grateful to have had the opportunity to worked with Hossein for a second time, his help greatly improved the quality of this paper.

Abstract

This paper examines an investment strategy relying on underlying characteristics of stocks and how market efficiency drives its returns. By using the recently published quality-minus-junk factor this paper attempts to explain the abnormal performance of portfolios sorted on their quality score by using a natural experiment which interferes with market efficiency. Using data from the U.S. and the SEC SHO pilot program it is shown that the returns associated with quality investing is significantly affected by varying degrees of short sale constraint. This effect is negative in size indicating that a quality investment strategy fares better in times of low market efficiency.

Keywords: Quality, Portfolio performance, Difference-in-Difference, Short-sale constraints,
Limits to arbitrage.

Contents

1 Introduction

The purpose of this paper is to investigate how market efficiency affects the returns associated with portfolios sorted on quality, as defined in Asness, Frazzini & Pedersen's 2017 paper. In their paper they find that quality only explains asset prices to a limited extent. This limitation is traced to 3 possible candidates: a) the market uses quality measures superior to the ones proposed in the paper, b) quality is linked to some unidentified risk factor or c) limited market efficiency. Their paper does not provide a definitive answer to this question, although it makes a convincing argument against explanation a) and b). This paper aims to investigate the remaining explanation, c), by exploiting a natural experiment. In an effort to investigate the effectiveness of short sale constraint on the stock market, the US securities and exchange commission (SEC) selected every third stock from the Russell 3000 index after being sorted for trading volume to be part of a pilot program (SEC, 2005). This creates a periodic gap in market efficiency (Boehmer & Wu, 2013) depending on whether the security in question is included in the pilot program or not. As such the pilot program enables a Difference-in-Difference analysis and provides a window into investigating causality in stock returns (Angrist & Pischke, 2008).

Stock quality are characteristics that investors should be willing to pay a higher price for, everything else being equal (Asness, Frazzini & Pedersen, 2017). This very broad definition creates the foundation for how quality will eventually be quantified. The quantification of quality draws upon an extensive amount of literature in its construction. Broadly, quality consists of three pillars whose connection to stock returns has an extensive literary backing. They are as follows:

- Profitability (Campbell, Hilscher & Szilagyi, 2008; Chen, Novy-Marx & Zhang, 2011; Novy-Marx, 2012; Haugen & Baker, 1996; Sloan, 1996; Piotroski, 2000)
- Growth (Ahmad & Safdar, 2018; Liu, Warner & Zhang, 2003)
- Safety (Altman, 1968; Ohlson ,1980; Campbell, Hilscher, & Szilagyi 2008; Frazzini & Pedersen, 2013; Novy-Marx, 2013; Black, Jensen & Scholes 1972; George & Hwang 2010; Penman, Richardson & Tuna, 2007)

Together they form an amalgamation of different proven positive characteristics. While quality as a concept has a long and rich history, it has come under many different definitions. As Pedersen indicates in his book (2019) Quality is in a sense linked to value investing. Both strategies aim to looking at some underlying characteristic which determines the value of the company and then going long the company if it is undervalued and shorting it if it is overvalued. The way a value investor does this is by looking at the book-to-market ratio. The way a quality investor does this is by looking at their aggregate quality score. A quality investor is a bargain hunter, believing that the price of a stock fluctuates around its fundamental value. As such assets should be purchased when the quality score is high, i.e. the firm is profitable, grows and is safe given its fundamentals, while shorting firms with low quality scores, firms that are relatively unprofitable, have poor growth and are unsafe. This is also where market efficiency enters the equation. If these fundamentals determine the price, an efficient market will not let the price fluctuate from its fundamental value enough for a bargain hunter to collect a premium. When news does hit the market so that prices fluctuate, an efficient market will move to the new equilibrium quickly and thus not leave enough time for a bargain hunter to make the analysis and collect the premium.

The natural experiment utilized in the paper, the SEC SHO pilot program, has been gainfully used to investigate other return anomalies by Chu, Hirshleifer and Ma (2020). They investigate several different return anomalies using a similar methodology as in this paper. They also examine the effect on an aggregate of all these return anomalies, documenting a significant negative relationship associated with the participation in the pilot program. Their paper serves as a methodological guide throughout this paper and a lot of inspiration is drawn from their examples. Our findings can contribute a lot to the public utilization of quality strategies. If the hypothesized relationship between return and market efficiency holds empirically it can provide a road map for investors as to when to implement the strategy and when not to. It also builds upon the literature using market inefficiencies, and limits to arbitrage specifically, to explain return anomalies and performance. Understanding the return mechanism of the strategy itself also provides an understanding that make markets themselves more efficient. Times of high market inefficiency are corrected by investors exploiting arbitrage opportunity which yield returns when the market corrects itself. As such an investor utilizing a quality strategy may help the market achieve higher efficiency by building upward (downward) pressure on assets that are undervalued (overvalued), thereby alleviating large mispricings which can hurt investors when the market corrects.

In this paper all stocks in Russell 3000 index during the time of the pilot program have their quality score estimated. Portfolios are then formed using these scores. The analysis is then conducted on two portfolios, one consisting of stocks included in the pilot program and one consisting of stocks not included in the pilot program. The main findings of the paper indicate a robust relationship between market efficiency and the returns associated with quality sorted portfolios. When market efficiency increases, through the removal of short sell constraints, the returns associated with quality sorted portfolios drop. These results remain robust through all placebo test conducted in this paper.

The rest of the paper is laid out as following. In the next section, the data and methodology will briefly be discussed and presented as well as some background surrounding the pilot program the paper draws heavily upon. After that, the results of the empirical tests will be explored extensively before moving to discussion where these results will be put into their theoretical framework and analyzed compared with results from other papers. Finally, the paper concludes with some brief remarks regarding the relationship explored throughout the paper.

2 Literary review

In this section the connecting literature will be briefly touched upon to set the background for the eventual empirical analysis.

2.1 Short selling, limits to arbitrage and the SEC pilot program

Short sale is a transaction in which the seller does not own the stock but borrows it from somebody else and sells the asset under the obligation to buy the stock back at some point in the future. This allows a trader to profit from a drop in price, by selling when the price is high and buying when the price is low. These types of transactions are usually considered risky as there is no potential limit to the losses given that assets increase in value instead of falling during the short sale period. Short selling is also utilized as a mechanism behind constructing theoretical arbitrage portfolios, in which the trader is trying to create an arbitrage profit by constructing a portfolio in which the long part systematically outperforms the short part. The perception of riskiness has brought about regulation when it comes to how and when short sales are allowed to be made (SEC, 2015).

When news, i.e. the successful acquisition of another company or other pieces of information, such as quarterly earnings report, become public, they are not always reflected immediately in prices. Even if prices move in response they do not always perfectly aggregate all information immediately. Events such as mergers or quarterly earnings report do typically move the prices up and down but the price movement is typically to small. Therefore, price has a tendency to drift following these announcements (Ball & Brown, 1968). One potential reason behind these drifts are so called limits to arbitrage. Limits to arbitrage are road blocks which basically prevent sophisticated investors from entering a position in the spirit of arbitrage pricing theory (Roll & Ross, 1980). These investors do enter positions which attempt to extract arbitrage profits from these scenarios according to Shleifer & Vishny, but due to limits to arbitrage they are not enough to adjust prices fully (Shleifer & Vishny, 1997; Shleifer 2000). Real world limits to arbitrage can take many different forms. As an example idiosyncratic movements in an asset, i.e. natural disasters, terrorist attacks or the heart attack of a CEO, can make a position as described in arbitrage pricing theory impossible to hold in reality as these movements in price may run whoever enters them bankrupt by freak accident. These limits have a history of making valuable contributions to the understanding of different pricing anomalies (Shleifer & Vishny, 1997; Shleifer, 2000; Gromb & Vayanos, 2010; Li, Sullivan & Garcia-Feijoo, 2014; Lewellen ,2011; Stambaugh, Jianfeng & Yuan, 2015). A constructed such limit to arbitrage is the uptick rule which limits short sales.

In the public eye, short selling is looked at with a lot of scepticism, being blamed for financial crashes in the US as early as 1929 (Trotta, 2008). This led to the creation of short sale constraints to regulate such trades in favor of a more stable market. Stocks traded in the U.S. follow a specific price test indicating when they are eligible to be shorted and when not. Specifically rules governing this issue implemented on the NYSE/AMEX, price tests, stated that short sales were only to be conducted on upticks, or zero-plus ticks. An uptick is when a security sells at a price higher than the preceding sale. A zero-plus tick is when a security sells at the same price as the preceding sale, where the preceding sale was an uptick. As such it is impossible to enter a short sale on these markets following a downtick in a certain stock, i.e. one that already has downward momentum. The SEC, wanting to evaluate the effectiveness of price-tests as short-sale constraints, adopted a pilot program which allowed short sales to be made on certain stocks in absence of the regular price test. The pilot program itself consisted of 1000 stocks from the Russell 3000 stock index and lasted from 2005-05 to 2007-08. Participants in the pilot program were chosen in a quasi-random fashion. All components of the Russell 3000 index were first sorted according to their average daily trading volume in the prior year. Then every third stock, starting with the second was chosen to participate in the pilot program. This pilot program, which is freed from the short sale restriction, lives in a market parallel to the other stocks that is potentially more efficient as it is freed from any form of price test. Following the termination of the pilot program the uptick rule was removed for all stocks. This was later reversed in February 2010 when the SEC reimposed a modified uptick rule (SEC, 2007, 2010).

2.2 Quality

Asness, Frazzini and Pedersen (2017) studies how assets are priced through with respect to quality. They define quality as a characteristic that, all else being equal, investors should be willing to pay more for. As discussed in detail in the data and methodology chapter, quality is made a combination of three underlying characteristics with a vast amount of literary support. Novy-Marx (2013) explores the profitability side of quality and finds that a profitability oriented strategy has a high return in its own right, while also improving other strategies such as value oriented strategies. The way quality does this is by helping a value investor distinguish between undervalues stocks and stocks that are cheap for a good reason. So while value is a good price signal for the investor, quality helps the investor pick the actual stocks to trade on. Similarly the findings of Liu, Warner and Zhang (2003) indicate that the growth rates in sales and profitability variables can be used in a similar manner. Liu, Warner and Zhang document that increases in growth are associated with significantly positive returns. The last characteristic, Safety, has a very diverse literary backing, stemming partly from the very measures of Safety used in its construction. Frazzini Pedersen (2013), document significant positive returns associated with betting against assets which covaries with the market. Having a low beta thus not only enhances returns but hedges against market risk. Similarly George and Hwang (2010) find that low distress risk, which can be measured through Altman's Z-score or Ohlson's O-score, and low leverage are associated with higher risk-adjusted returns. This extensive list of asset aspects, which all in their own right generate high risk-adjusted return, put together form an amalgamation of positive characteristics and should help investors identify which assets are over- and undervalued. Nevertheless Asness, Frazzini & Pedersen (2017) find limited support that high quality stocks actually garner a higher price than other stocks. When quality scores are used to sort assets into portfolio they do, however, document a pattern. The low quality portfolios are consistently and significantly outperformed by the high quality portfolios. This pattern is robust and remains consistent when adjusting the portfolios for their relative risk using the Fama-French model. This presents a puzzle: Why does a portfolio that goes long high quality and short low quality produce a positive risk-adjusted return. Asness, Frazzini and Pedersen (2017) discuss several potential sources for this return, as mentioned in the introduction. This paper focuses on market inefficiency as a source for these returns. The operating assumption being investigated in this paper is that two portfolios formed on quality, with differing degrees of market efficiency, will have different returns. A portfolio with higher (lower) market efficiency is believed to have a lower (higher) return associated with it. A portfolio with lower degrees of market efficiency will fluctuate more in value allowing a quality investor to identify undervalued assets and potentially increase the returns. If for example a quarterly report is announced in which profitability of an asset is substantially improved. In a perfectly efficient market the assets price is adjusted instantaneously to its new prices. If, however, there are obstacles limiting how efficient the market can be, an investor can buy the stock before the price has adjusted fully.

3 Data and Methodology

In the beginning of this chapter background on the pilot program utilized in the analysis will be presented, followed by a brief overview of the data. Next the methodology behind evaluating the quality of each company and the models used in the empirical analysis will be presented extensively.

3.1 Data

The data used in this paper primarily derives from two different sources. The first being Compustat who provides firm level financial data and provides the bulk of the data used in constructing most of the factors. Most of the data is on a quarterly level and as the analysis is conducted on a monthly level the missing data is interpolated using the last publicly available information. The other primary datasource, CRSP, provides information on stock prices and returns. This data is collected on both a daily level, for estimating quality factors, and monthly for the portfolio construction used in the analysis. The composition of the Russell 3000 index is gathered from Bloomberg and information about participation in the pilot program is gathered from the SEC (2005). The Fama-French factors (Fama & French 1993) are obtained from Ken French's website. The sample itself consists of 1559 stocks for which sufficient data was available through Compustat and CRSP to construct a quality score. The portfolios are constructed using returns running from 1990-2007 for the primary sample and 1990-2010 for the full sample. All dataitems and their sources are listed in the appendix.

3.2 Constructing quality

The origin of this model of Quality stems from Gordon's growth model which has been rewritten by dividing both sides by the book value. This results in the following equation:

$$\frac{\text{Price}}{\text{Book}} = \frac{\text{Profitability} \times \text{Payout-ratio}}{\text{Required return} - \text{Growth}} \tag{1}$$

The reason behind adjusting Gordon's growth model is to make it more stationary over time and in the crossection (Pedersen, 2019). The model when it comes to calculating the quality score follows the methodology of Asness, Frazzini & Pedersen (2017). As such quality is the average of three characteristics, which themselves are averages of normalized metrics. This makes the amount of inputs into constructing the score extensive and alleviates concerns that any one metric may have undue influence over the composite score. The first characteristic, Profitability, is constructed as the average of six different metrics all having shown power in determining future stock returns, the construction of the metrics themselves follow the methodology of Asness, Frazzini & Pedersen (2017).

- Gross profits over assets (GPOA) (Novy-Marx, 2012, 2013).

$$\text{GPOA} = \frac{\text{Revenue} - \text{Cost of Goods Sold}}{\text{Total Assets}}$$

- Return on equity (ROE) (Novy-Marx, 2012, 2013; Haugen & Baker, 1996).

$$\text{ROE} = \frac{\text{Net income}}{\text{Book Equity}}$$

- Return on assets (ROA) (Novy-Marx, 2013).

$$\text{ROA} = \frac{\text{Net income}}{\text{Total Assets}}$$

- Cash flow over assets (CFOA) (Novy-Marx, 2012, 2013; Sloan, 1996).

$$\text{CFOA} = \frac{\text{Net income} + \text{Depreciation} - \Delta\text{Working Capital} - \text{Capital Expenditures}}{\text{Total Assets}}$$

- Gross margin (GMAR) (Novy-Marx, 2012, 2013; Piotroski, 2000).

$$\text{GMAR} = \frac{\text{Revenue} - \text{Cost of Goods Sold}}{\text{Total Sales}}$$

- Low accruals (ACC) (Novy-Marx, 2012, 2013; Sloan, 1996).

$$\text{ACC} = -\frac{\Delta\text{Working Capital} - \text{Depreciations}}{\text{Total Assets}}$$

Once each metric has been calculated they are cross-sectionally ranked, the mean is determined and the standard deviation is calculated. These are then normalized to a z-score by applying the following formula:

$$z_i = \frac{rank_i - \mu}{\sigma} \tag{2}$$

By using ranks instead of the raw score undue influence of extreme values are once again addressed in the metric construction. Finally this results in the following equation for the Profitability score.

$$\text{Profitability} = Mean(z_{GPOA} + z_{ROE} + z_{ROA} + z_{CFOA} + z_{GMAR} + z_{ACC}) \tag{3}$$

The next characteristic, Growth, is defined as the five-year growth in all profitability metrics, except for accruals. It builds upon the findings of Mohanram (2005) and Novy-Marx (2013) that firms with high growth in profitability outperform those with poor.

- Five-year growth in GPOA

$$\Delta\text{GPOA} = \frac{\text{Gross Profits}_t - \text{Gross Profits}_{t-5}}{\text{Total Assets}_{t-5}}$$

- Five-year growth in ROE

$$\Delta\text{ROE} = \frac{\text{Net income}_t - \text{Net income}_{t-5}}{\text{Book Equity}_{t-5}}$$

- Five-year growth in ROA

$$\Delta\text{ROA} = \frac{\text{Net income}_t - \text{Net income}_{t-5}}{\text{Total Assets}_{t-5}}$$

- Five-year growth in CFOA

$$\Delta\text{CFOA} = \frac{\text{Cash Flow}_t - \text{Cash Flow}_{t-5}}{\text{Total Assets}_{t-5}}$$

- Five-year growth in GMAR

$$\Delta\text{GMAR} = \frac{\text{Gross Profits}_t - \text{Gross Profits}_{t-5}}{\text{Total Sales}_{t-5}}$$

Due to the limited pool of stocks included in the SEC pilot program the five year time frame has been relaxed to a two-year time frame where necessary to achieve adequate data. Once each metric has been calculated the same methodology as in (2) is applid to aggregate the scores into the final

Growth score.

The final of the three characteristics is Safety. Safety, as with the others, is measured through the use of several metrics. They are also aimed at different aspects of safety. Some, Ohlson's O-score and Altman's z-score, are aimed at estimating bankruptcy risk. Others are aimed at measuring stability, low return on equity volatility, or market exposure, low beta. Finally leverage is also considered. resulting in the following list:

- Ohlson's O-score (O) (Ohlson, 1980; Campbell, Hilscher & Szilagyi 2008).

  $$\begin{aligned} O = &- (-1.32 - 0.407 \times log(\text{Adjusted Assets/CPI}) + 6.03 \times \text{ Book value of debt /Adjusted} \\ &\text{Assets} - 1.43 \times (\text{Current Assets} - \text{Current Liabilities})/\text{Adjusted Assets} + 0.076 \times \\ &\text{Current Liabilities/Current Assets} - 1.72 \times \text{Dummy}_1 - 2.37 \times \text{Net Income/Total Assets} \\ &- 1.83 \times \text{Pre-tax Income/Total Liabilities} + 0.285 \times \text{Dummy}_2 - 0.521 \times \Delta\text{Net Income}) \end{aligned}$$

  Where the first dummy is equal to 1 if total liabilities exceeds total assets. The second is equal to 1 if Net Income is negative for current and prior fiscal year.

- Altman's Z-score (Z) (Altman, 1968, 2000; Campbell, Hilscher & Szilagyi 2008).

  $$\begin{aligned} Z = &(1.2 \times \text{Working Capital} + 1.4 \times \text{Retained Earnings} + 3.3 \times \text{Earnings before interest} \\ &\text{and taxes} + 0.6 \times \text{Market Equity} + \text{Sales})/\text{Total Assets} \end{aligned}$$

- Return on equity volatility (EVOL) (Novy-Marx, 2013).
  EVOL is estimated as the standard deviation of quarterly return on equity over the past 60 quarters, as with the growth score the time frame is adjusted where necessary.

- Low beta (BAB) (Black, Jensen & Scholes 1972; Frazzini & Pedersen, 2013).
  BAB is equal to minus market beta. The beta estimation follows the methodology of Frazzzini & Pedersen (2013) and is the product of the rolling one-year daily standard deviation of the market-portfolio, as collected from Ken French's data library, and the 3-day correlation over a rolling 5 year window.

- Low leverage (LEV) (George and Hwang 2010; Penman, Richardson, & Tuna 2007).
  LEV is defined as minus total debt over total assets.

Finally these are, as with the previous metrics, aggregated into one score for safety following the methodology of equation (2).

When all three characteristics have been calculated they are turned into the finally quality score. This is done by taking the average of all three underlying characteristics. As such the final Quality score is not tilted in the favour of any one of the aspects but appears as a combination of all three where each one is given equal weight.

$$\text{Quality} = Mean(z_{Profitability} + z_{Growth} + z_{Safety}) \tag{4}$$

3.3 Asset pricing models

In order to asses the performance of portfolios constructed in the paper it will become necessary to construct a model delivering this. The model considered in this paper is the standard Fama-French 3-factor model as well as the 5-factor model for the sake of robustness. Fama and French (1993) expand upon the Sharpe(1964)-Lintner(1965) CAPM model by adding new factors to the equation. The first factor in the model is the standard CAPM excess market return factor. It is constructed as the market rate of return, measured as the value weighted return of all CRSP firms listed on the New York stock exchange, AMEX or NASDAQ, minus the risk free rate, measured as the one month Treasury bill rate:

$$Mkt = Rm - Rf \tag{5}$$

The coefficient obtained by running a regression on a portfolio return using the CAPM factor can be interpreted as the market exposure of the portfolio. That is to say how much it co-varies with the market. This is followed by the first two Fama and French factors: small-minus-big (*SMB*) and high-minus-low (*HML*). *SMB* refers to the return spread of small minus large stocks and *HML* to the spread of cheap minus expensive stock based on their book-to-market ratio. *SMB* is constructed dividing the market into two groups based on the market capitalization of the company. These are then orthogonalized to the *HML* factor by constructing a small (big) portfolio that is consistent of 30% low 40% neutral and 30% high book-to-market. A position is then taken going long the small portfolio and shorting the big portfolio. For the exact equations behind constructing these factors, please check the appendix.

Next item is the aforementioned *HML* factor. It is constructed in a similar manner by first dividing the market into big and small stocks based on market capitalization. These two groups are then divided into two groups based on their book-to-market ratio, high and low. Just as before the *HML* factor is then orthogonalized against the *SMB* factor.

Together these three factors make up the Fama-French 3-factor model. It's interpretation is quite straightforward beyond the CAPM beta we also get 1 beta each for the other two factors. Positive coefficients in this context indicate a positive factor loading i.e. a positive (negative) *SMB* coefficient would indicate that the portfolio being evaluated covaries with small (big) stocks. Same for *HML*, positive (negative) coefficient indicates covariation with high (low) book-to-market ratio stocks.
In a later paper Fama and French expanded upon this 3-factor model by adding two additional factors (2014). These new factors relate to the return spread on the most profitable minus least profitable firms (RMW) and the return spread on those that invest conservatively minus aggressively (CMA). The construction of the *SMB* factor was also changed to orthogonalize it to the other factors. For a full list of equations detailing the construction of each of these factors check the appendix.

3.4 Portfolio evaluation

In the previous section factors were presented which will form the basis for the model that will be used to evaluate the portfolios. Before the portfolios can be evaluated however they need to be created. In order to establish the quality effect, each stock will be sorted in accordance with their quality score. In total four portfolios will be formed each consisting of 25% of value weighted stocks available at that month. Providing four portfolios with differing quality scores, two among them are of particular interest, the extreme portfolios. By choosing the extreme portfolios and going long the high quality one while shorting the low quality will result in a factor mimicking portfolio. Once these portfolios have been formed their return will be regressed on the asset pricing model. This will illustrate the factor loading of the portfolio, giving insight into what kind of risk, if any, a quality investment strategy faces. Beyond that it also determines whether there is any abnormal positive return generated by the portfolio. Abnormal positive returns are those not explained by the risk factors included in the asset pricing model and manifests itself through a positive intercept, α. Asness, Frazzini & Pedersen (2017) documented a positive and significant α associated with high quality portfolios and a portfolio going long high quality while shorting low quality. The aim of this evaluation is to provide similar results, thereby indicating the presence of the quality effect, before exploring the return mechanism further. All standard errors are Newey-West corrected with 9-lags to account for heteroskedasticity and autocorrelation in the error term (Newey & West, 1987)

3.5 Difference-in-Difference estimation

Difference-in-Difference estimation is used to explore the return mechanism. As the portfolios are time-series of returns divided into groups based on their participation in the pilot program they have the functional form of panel data. This means that in practice there are two different states in which two portfolios will be watched.

$Y_{0,0}$ = Not participating in the pilot program, before the pilot programs beginning.

$Y_{1,0}$ = Participating in the pilot program, before the pilot programs beginning.

$Y_{0,1}$ = Not participating in the pilot program, during the pilot program.

$Y_{1,1}$ = Participating in the pilot program, during the pilot program.

As such the Difference-in-Difference estimator itself will be the difference of the difference in return of the two portfolios experience before and during the pilot program:

$$\beta = [E(Y_{1,1}) - E(Y_{1,0})] - [E(Y_{0,1}) - E(Y_{0,0})]$$

3.5.1 Naive Difference-in-Difference

This will be enhanced also by adding the Fama-French risk factors introduced earlier to the regression. As such the first model that will be utilized has the following functional form:

$$Y_{i,j} = \alpha + \gamma \times During + \beta_1 Pilot_i \times During + \beta_2 Pilot_i + \sum \beta_i \times covariates + \varepsilon \tag{6}$$

In this model, Y is the return of the portfolio i at time j where j indicates whether or not the return occurred during the pilot program or not. α refers to the common intercept, γ denotes the effect of being in the period where the pilot program is active. *During* is therefore a dummy equal to 1 if the month was during the pilot program and 0 otherwise. β_1 will be the main variable of interest namely the effect that the pilot program had on the participants. $Pilot \times During$ also being a dummy equal to 1 if the return in question is from a pilot program portfolio during the pilot program and 0 otherwise. β_2 will capture the effect of simply being in the pilot program, as *pilot* will be a dummy equal to 1 if the portfolio is formed on pilot stocks and 0 otherwise. The covariates in the regression are comprised of the Fama-French 5-factor model, which will capture any residual variance explained by their risk model. A significant β_1 here would indicate that there is indeed a significant effect on the returns of stocks associated with the pilot program. In particular a negative coefficient would indicate that market inefficiency plays a role in determining the returns of a quality investment strategy.

3.5.2 Difference-in-Difference, time fixed effects & entity effects

Next the more robust setup of the Difference-in-Difference regression to be adopted is a fully fledged fixed effects regression model. In this model the time invariant effects will not simply be a dummy indicating whether the return was before or during the program. Instead there will be time invariant effects for each month resulting in the following regression:

$$Y_{i,t} = \gamma_t + \beta_1 Pilot_i \times During + \beta_2 Pilot_i + \varepsilon \tag{7}$$

Where $Y_{i,t}$ is the return of portfolio i at time t, γ_t denotes time fixed effects, *Pilot* is a dummy equal to 1 if the portfolio was formed on pilot program stocks and 0 otherwise. *During* is a dummy equal to 1 if during the time the pilot program is active. γ_t, the time fixed effects, will captures macroeconomic variables and other fixed effects, such as the returns of the different risk factors, affecting the pilot program portfolio as well as the other portfolio. This makes it pointless to add the covariates to the regression as they will completely be subsumed by the time fixed effects. The *pilot* dummy, the entity variable β_2, is used to differentiate the between the effect of the pilot program on returns from the aggregate difference in returns between the portfolios. β_1 will be the main variable of interest in this regression as it will indicate the effect the pilot program has on stock returns.

3.5.3 Difference-in-Difference, robustness tests

The robustness of the model will also be checked by elongating the sample and looking at the post-pilot program time period. The model will be specified as the following:

$$Y_{i,t} = \gamma_t + \beta_1 Pilot_i \times During + \beta_2 Pilot_i + \beta_3 Pilot_i \times Post + \varepsilon \tag{8}$$

Where $Pilot \times Post$ is a dummy taking the value of 1 if the return comes a portfolio formed on pilot program stocks after the end of the pilot program and 0 otherwise. The idea behind this regression is that once the pilot program is over and the stocks return to an equal footing in terms of regulatory restrictions they should revert to same returns. As such the added term of being in the pilot program in the post-pilot program period should be insignificant and thus not affect the outcomes of the portfolios, whereas β_2 ought to remain significant.

Furthermore robustness is further checked by implementing placebo-tests on the sample, following the methodology in Chu, Hirshleifer and Ma (2020). A potential problem with Difference-in-Differences estimation is that results may be driven by unobserved shocks that affect the portfolios differently. To test whether it is possible to falsify potential results from the Difference-in-Difference estimation placebo tests will be implemented. In this robustness checks placebo test periods are created. These placebo tests are three two-year periods prior to the real pilot program. Three subsamples are formed beginning in 1990-01 and running up until the end of each placebo test. As such the following subsamples are used:

- 1990-01 to 2003-06 with placebo program 2001-06 to 2003-06
- 1990-01 to 2001-06 with placebo program 1999-06 to 2001-06
- 1990-01 to 1999-06 with placebo program 1997-06 to 1999-06

The regression model follows that of the main Difference-in-Difference regression:

$$Y_{i,t} = \gamma_t + \beta_1 Pilot_i \times \textit{Pseudo-During} + \beta_2 Pilot_i + \varepsilon \tag{9}$$

If the coefficient for *Pseudo-During* remains insignificant for each of the placebo programs it provides some reassurance that any documented effect is not simply a coincident but only exists in the true pilot program period.

4 Results

In the following chapter the results of the model and tests presented in the methodology chapter will be presented, starting out granular and gradually working towards the main model of interest for the paper. Finally some robustness checks will be performed on the analysis as a confirmation of the results obtained throughout the paper.

4.1 Summary statistics

Table 1: Summary statistics

In the following table summary statistics are presented for both portfolios used in the the analysis later in the paper. The sample period is the full available sample for both of the portfolios, 1990-01 to 2011-01.

Portfolio	Raw return	Standard-deviation	Low Q-score	High Q-score
High - Low, pilot stocks	0.476	4.739	-0.656	0.686
High - Low, not pilot stocks	0.383	5.151	-0.641	0.669

From Table 1 it is apparent that the portfolios have slightly different raw returns. Standard deviations however are in line with each other. Overall the average quality scores fall well in line of each other indicating that neither the top nor bottom q-score differ to any larger extent. This provides some sense of comfort considering the balance of the pilot program as there does not seem to be any tendency for one portfolio to be of substantially junkier or higher quality. Concluding from the summary statistics both portfolios seem comparable in terms of performance although the pilot stocks have historically performed better both in terms of raw return and volatility.

4.2 Portfolio evaluation

Next we look at the portfolio evaluation side of the portfolios as means to determine the presence of the quality return, and comparing the factor loadings of the portfolios.

Table 2: Portfolio evaluation

In the following table the mean excess return, Fama-French 3-factor α and 5-factor α are presented for two different portfolios formed on the Russell 3000 index between 1990 up to 2011. Each portfolio is constructed as High minus Low quality portfolio where a long position is taken in the highest scoring quantile and a short in the lowest scoring. Each portfolio contains different stocks, the pilot program portfolio, as the name suggests, is constructed using all stocks in the sample included in the 2005 SEC pilot program. The second portfolio is constructed using all other stocks, those not participating in the pilot program, in the index. The Q-score columns refer to the average computed quality score of the bottom and top quantiles used to form the portfolios. All portfolios are value weighted and the standard errors are Newey-West corrected using nine lags to account for heteroskedasticity and autocorrelation. T-statistics indicating significance levels are included below each estimate.

Portfolio	Excess return	3-Factor α	5-Factor α
High - Low, pilot stocks	0.456	0.600*	0.632*
		1.882	*1.845*
High - Low, not pilot stocks	0.362	0.529**	0.523**
		2.021	*1.991*

Table 2 shows the results from the initial portfolio evaluation over a sample period from 1990-01 up to 2011-01. The raw return column indicates the mean return of the portfolio over the sample period. Both portfolios had similar results with the pilot program portfolio outperforming the other portfolio. Looking at the Fama-French 3-factor α coveys a similar story with similar coefficients and t-statistics. The results generally hold up as well when changing to the 5-factor model not impacting neither coefficient nor t-statistic to a large degree. Clearly present in the sample is however the abnormal returns associated with quality sorting portfolios documented in Asness, Frazzini & Pedersen (2017).

The next table expands upon the results of the Fama-French 3-factor and 5-factor regression in Table 2 by providing the full list of covariate factor loadings. From Table 3 the first two columns seem to indicate similar factor loading between the two portfolios. Both have a positive abnormal return, as reported in Table 2. Beyond that they all have a negative exposure to the market portfolio, which is not surprising given that negative market exposure was one of the factors employed in the portfolio construction. Furthermore all portfolios seemingly have a negative exposure to *HML* in the 3-factor model which is subsequently subsumed in significance, either partly or wholly, in the 5-factor model. The other factors, *SMB*, *RMW* and *CMA* all have factor loadings and significance levels within range of each other. Indicating that a quality investment strategy is not overly exposed to any of these risk factors, but rather risk neutral. Going into the Difference-in-Difference estimation these results do provide some reassurance in that the two groups used to produced the portfolios do not perform radically different, nor do they have any major differences in terms of factor loading coefficients or significance levels.

Table 3: Fama-French portfolio evaluation

In the following table the factor loadings of a Fama-French 3-factor and 5-factor regression are presented for two different portfolios formed on the Russell 3000 index between 1990 up to 2011. Each portfolio is constructed as High minus Low quality portfolio where a long position is taken in the highest scoring quantile and a short in the lowest scoring. Each portfolio contains different stocks, the pilot program portfolio, as the name suggests, is constructed using all stocks in the sample included in the 2005 SEC pilot program. The second portfolio is constructed using all other stocks, those not participating in the pilot program, in the index. All portfolios are value weighted and the standard errors are Newey-West corrected using nine lags to account for heteroskedasticity and autocorrelation. T-statistics indicating significance levels are included below each estimate.

	Dependent variable: Excess returns			
	(FF3-Pilot)	(FF3-Not Pilot)	(FF5 Pilot)	(FF5 Not pilot)
α	0.600*	0.529**	0.632*	0.523**
	1.851	*2.001*	*1.845*	*2.003*
Mkt-RF	-0.241^{***}	-0.202^{***}	-0.261^{***}	-0.194^{**}
	-3.289	*-2.820*	*-2.741*	*-2.239*
HML	-0.197^{**}	-0.228^{***}	-0.141	-0.254^{*}
	-2.383	*-2.658*	*-0.947*	*-1.690*
SMB	0.113	-0.022	-0.034	0.118
	1.378	*-0.282*	*1.368*	*-0.436*
RMW			-0.008	-0.023
			-0.050	*-0.166*
CMA			-0.125	0.077
			-0.581	*0.353*

Note: t-statistics in *italics* *p<0.1; **p<0.05; ***p<0.01

4.3 Difference-in-Difference estimation

Table 4 conveys the Difference-in-Difference regression results. Beginning with the Naive model, looking at the main variable of interest, *Pilot* × *During*, we see that there is a somewhat significant negative effect. Beyond the main variable of interest itself there is not much significance to be found. The only coviariate able to generate any significant results is the return of the market portfolio. The *HML* factor does have a mild negative significance. Neither the *During* variable, which can be seen as a time-fixed effect controlling for differences during and before the experiment, nor the *Pilot* variable, which is the entity fixed effect, garner any significance. This of course is not of any significance to the interpretation of the results. The insignificance of the entity effects merely indicates that there is no significant differences in the returns between the portfolios which is not attributable to the period the pilot program itself was active. *During* only indicates that there was no significant difference before and during the program which is not attributable to the pilot program. The coviariates impact both portfolios equally and as such does not affect the estimate of *Pilot* × *During*. It does however distinguish between some time effects which is somewhat necessary to account for in the naive model as the naive model does not have time-fixed effects for each individual time-period. Moving to the time-fixed effects model, where we actually do employ time-fixed effects for each time-period, we note that all variables are consumed. First of all the *Pilot* variable is employed as the entity variable and the date, year-month, is employed as time-fixed effects. This consumes all other covariates as previously mentioned as they have the same value for both entities in each time-period. As such only the variable of interest remain and with a more robust estimate of it's actual effect. The estimates are the same for both models but we note an increase in the t-statistic and the significance level of the estimate indicating *Pilot* × *During* persistence in a more robust model. The third and forth model conduct the Difference-in-Difference analysis on the short leg and the long leg separately. The long-leg regression while being insignificant is not far from being significant at the 90% level. The short leg on the other hand is significantly positive, which of course turns into a negative return once the short position is entered in the arbitrage portfolio.

Next up are the robustness checks. Starting of with the post-pilot program test in Table 5. The models presented here differ from the one in Table 4 due to the inclusion of the post-pilot-program period. Obvious from the naive regression is that the *Post* dummy seems to have had little effect on returns overall. The estimate is quite small in size and far from significance. The *Post* × *Pilot*

Table 4: Difference-in-Difference regression

In the following table the main results will be presented on Difference-in-Difference regressions. All regressions are made on monthly returns of two portfolios constructed from stocks in the Russel 3000-index. The two portfolios, one participating in the 2005 SEC pilot program and one not, are constructed based on their quality-score. Each portfolio has a net-zero investment as it goes long high quality stocks while shorting low quality stocks. Both portfolios run from 1990-01 up to the end of the SEC pilot program in 2007-08. The second model (Time-fixed effects model, Long-leg, Short-leg) is panel data model with both entity (pilot) and time fixed effects (year-month).

	Dependent variable: Excess returns			
	Naive model	Time-fixed effects model	Long-Leg	Short-Leg
α	0.471			
	1.604			
Pilot	0.481			
	0.997			
Pilot × During	-1.723*	-1.723**	-0.910	0.747**
	-1.885	*-2.022*	*-1.515*	*2.022*
During	0.044			
	0.077			
Mkt-RF	-0.246***			
	2.895			
HML	-0.258*			
	-1.727			
SMB	0.045			
	0.661			
RMW	0.004			
	0.029			
CMA	-0.001			
	-0.006			

Note: t-statistics in *italics* *p<0.1; **p<0.05; ***p<0.01

dummy is also insignificant which is promising as this points to the fact that during the post-pilot program period, pilot stocks neither under- nor outperformed the other stocks. Other estimates in the naive model are either affected negligibly or not at all. The Time-fixed effects mirrors these results and contributes by lowering the t-statistic somewhat more.

The other robustness test conducted, the placebo test, is presented in Table 6. Each regression corresponds to a different placebo test. (1) uses a placebo pilot program running from 2001-06 to 2003-06. From the regression results it is possible to see that the main variable of interest, *Pilot* $\times$ *Pseudo-During*, has no significance. Overall the results, apart from those relating to the pseudo period, seem to correlate quite nicely with those obtained in the main regression table of the paper, Table 4. Both in terms of signage and significance level. The time-fixed effects model applied on the same period yields similar results with a highly insignificant estimate. The next regression (2) delivers a largely similar result yet differing in some key aspects. First of all the market rate of return (*Mkt-RF*) has lost its significance and its effect size has shrunk to half. *HML* on the other has grown in effect size with a t-statistic slightly more significant than in (1). While these results are interesting in principle, and maybe correlated with the fact that this period incorporates some of the dot-com bubble, they do not affect results relevant to this paper. The main variable of interest, still remains highly insignificant in this model, regardless of model choice. (3) Follows the trend of (2) closely with insignificant estimates for the placebo pilot program. The only difference to model (2) lies in that *RMW*'s estimate has increased both in terms of size and significance.

Table 5: Difference-in-Difference regression, Post test inclusion

In the following table results will be presented on Difference-in-Difference regressions. Both regressions are made on monthly returns of two portfolios constructed from stocks in the Russel 3000-index. The two portfolios, one participating in the 2005 SEC pilot program and one not, are constructed based on their quality-score. Each portfolio has a net-zero investment as it goes long high quality stocks while shorting low quality stocks. Both portfolios run from 1990-01 up to 2010-02. The second model (Time-fixed effects model) is panel data model with both entity (pilot) and time fixed effects (year-month).

	Dependent variable: Excess returns	
	Naive model	Time-fixed effects model
α	0.476	
	1.594	
Pilot	0.481	
	0.985	
Pilot × During	-1.723*	-1.723**
	-1.853	*-2.018*
Post × Pilot	-0.980	-0.980
	-0.814	*-0.478*
Post	0.513	
	0.448	
During	0.036	
	0.060	
Mkt-RF	-0.246***	
	-3.435	
HML	-0.194*	
	-1.701	
SMB	0.038	
	0.596	
RMW	-0.034	
	-0.299	
CMA	-0.043	
	-0.265	
	Note: t-statistics in *italics*	*p<0.1; **p<0.05; ***p<0.01

Table 6: Difference-in-Difference regression, Placebo test

In the following table results will be presented on Difference-in-Difference regressions. All regressions are made on monthly returns of two portfolios constructed from stocks in the Russel 3000-index. The two portfolios, one participating in the 2005 SEC pilot program and one not, are constructed based on their quality-score. Each portfolio has a net-zero investment as it goes long high quality stocks while shorting low quality stocks. Both portfolios run from 1990-01 up until their respective end dates. For (1) this means 1999-06 ,(2) 2001-06, (3) 2003-06. Each regression is Newey-west corrected using nine lags to account for heteroskedasticity and autocorrelation. At the bottom the estimate using a model with time-fixed effects can be found.

	Dependent variable: Excess returns		
	(1)	(2)	(3)
α	0.609	0.667*	0.166
	1.611	*1.743*	*0.405*
Pilot	0.691	0.777	0.877*
	1.176	*1.502*	*1.734*
Pilot × Pseudo-During	-0.361	-0.324	-0.707
	-0.372	*-0.155*	*-0.480*
Pseudo-During	-0.841	-1.250	1.070
	-1.507	*-1.033*	*0.975*
Mkt-RF	-0.256***	-0.130	-0.0822
	-2.260	*-1.131*	*-0.710*
HML	-0.259*	-0.414*	-0.215
	-1.509	*-1.785*	*-1.220*
SMB	0.055	0.092	0.091
	0.727	*1.056*	*0.720*
RMW	-0.008	0.088	0.621**
	-0.057	*0.534*	*2.032*
CMA	-0.003	0.379	0.098
	0.015	*1.241*	*0.309*
Pilot × Pseudo-During	-0.361	-0.324	-0.361
with Time-fixed effects	*-0.304*	*-0.144*	*-0.303*

Note: t-statistics in *italics* *p<0.1; **p<0.05; ***p<0.01

5 Conclusion

In the following section of the paper the results obtained in the results part are summarized and the key findings put into context.

The results of the empirical tests in the paper largely indicate the following: 1) Short sale constraint, thereby market inefficiency, has an effect on the return of a quality investment strategy. 2) Market inefficiency has a negative effect on the returns of a quality investment strategy. 3) These results are robust when controlling for the post-pilot program period. 4) The results are not falsifiable using placebo tests.

Hence, the findings fall in line with those of Asness, Frazzini & Pedersen (2017). Namely that mispricing of assets, through market inefficiency, shows promise in being able to explain returns associated with their strategy. It also mirrors the findings of Chu, Hirshleifer & Ma (2020), although the effect sizes documented in this paper are far larger than those found in theirs. The primary estimate in this paper has a size of -1.723 where Chu, Hirshleifer & Ma has an estimate of -0.72 for their combination of return anomalies. While the estimates for the short leg fall in line with those in Chu, Hirshleifer & Ma, the effect size in the long-leg is far greater and more significant than all their estimates. The one caveat being Chu, Hirshleifer & Ma's findings when running the analysis using stocks sorted on Composite Equity Isseus which seems to mirror the findings in this paper but with smaller effect sizes. As such their findings are based in the short-leg almost exclusively whereas the findings here seem to implicate the long-leg as well. Once again this speaks to the findings of Asness, Frazzini & Pedersen (2017).

The result from the robustness tests follow the general findings of Chu, Hirshleifer & Ma (2020), in that they do not seem to implicate any fault in the models. In the post period test, where the two portfolios converged in terms of regulatory oversight, participation in the pilot program made no significant impact on the stock returns as theorized. Neither where the results falsifiable, with highly insignificant placebo test coefficients. All this speaks to the robustness and the certainty between the connection of market inefficiency and quality as an investment strategy. When markets are inefficient, the returns of a quality investment strategy decline.

The pilot program utilized in this paper is an example of a clear break in market efficiency between the two portfolios. While the results seem clear and robust they do not fully comply to the significance level where it is beyond doubt that the hypothesized connection between market inefficiency and quality truly is present. Limitations in significance level may stem from three sources: 1) The period is to short and the sample to small for markets to fully adapt causing the portfolio to underperform, but not to the full extent of its relationship. 2) While short selling is an important tool in maintaining market efficiency the uptick rule is not enough to change the underlying market efficiency to such a degree that the full effect becomes visible. 3) The hypothesized connection between market efficiency and quality investing is simply wrong and another unknown driver might affect the returns.

Quality investment has a curious relationship to market efficiency. While the analysis has shown results indicating that shifts in market efficiency cause large shifts in returns, the result are not beyond reproach. The models presented in the paper provide results balancing on the edge of significance. The robustness checks provide reassurance that the relationship is not merely a coincidence, through failure to falsify and robustness with different sample periods. As such, the results support the theorized relationship between quality investing and market inefficiency. But as with all models such as the one presented in this paper, it is impossible to draw conclusions about causality. The model can show us that cause happened before effect, but that is not enough to prove causality.

So while the quest to find the return driver behind quality is far from over, this paper provides a significant contribution. It provides a robust empirical backing that largely falls in line with the theories of the original creators. Quality distinguishes itself by seemingly providing an additional dynamic to it lacking in other return anomalies tested using a similar method. This indicates that relative to other return anomalies, quality has a more complex and dynamic relationship to market inefficiencies.

6 References

Ahmad, A. & Safdar,I. (2018), Dissecting stock price momentum using financial statement analysis, *Accounting & Finance*, Vol.58, no.1, pp.3-43.

Altman, E. (1968), Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *The Journal of Finance*, Vol.23, no.4, pp.589-609.

Angrist, J. & Pischke, J-S. (2008), Mostly Harmless Econometrics: An Empiricist's Companion, *Princeton University Press*.

Asness, C., Frazzini, A. & Pedersen, L. (2019), Quality Minus Junk, *Review of Accounting Studies*, Vol.24, no.1, pp.34-112.

Asness, C., Frazzini, A., Ronen, I., Moskowitz, T. & Pedersen, L. (2018), Size matters, if you control your junk, *Journal of Financial Economics*, Vol.129, no.3, pp.479-509.

Ball, R. & Brown, P. (1968), An Empirical Evaluation of Accounting Income Numbers, *Journal of Accounting Research*, Vol.6, No.2, pp.159-178.

Black, F., Jensen, M. & Scholes, M. (1972) The Capital Asset Pricing Model: Some Empirical Tests, *Studies in the Theory of Capital Markets*, Praeger Publishers Inc.

Boehmer, E. & Wu, J. (2013), Short Selling and the Price Discovery Process, *The Review of Financial Studies*, Vol.26, no.2, pp.287-322.

Campbell, J., Jilscher, J. & Szilagyi,J. (2008) In Search of Distress Risk, *The Journal of Finance*, Vol.63, no.6.

Chen, L., Novy-Marx, R. & Zhang, L. (2014) An Alternative Three-Factor Model, *Available at SSRN*

Chu, Y., Hirshleifer, D. and Ma, L. (2020), The Causal Effect of Limits to Arbitrage on Asset Pricing Anomalies, *The Journal of Finance*, Vol. 75, no.5, pp.2631-2672.

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

Fama, E., & French, K. (2015), A five-factor asset pricing model, *Journal of Financial Economics*, Vol.116, no. 1, pp. 1-22.

Frazzini, A. & Pedersen, L. (2014) Betting against Beta, *Journal of Financial Economics*, Vol. 111, no.1, pp.1-25.

George, T. & Hwang, CY. (2010) A resolution of the distress risk and leverage puzzles in the cross section of stock returns, *Journal of Financial Economics*, Vol.96, no.1, pp.56-79.

Gromb, D. & Vayanos, D. (2010) Limits of Arbitrage: The State of the Theory, *Annual Review of Financial Economics*, Vol.2, pp.251-275.

Haugen, R.,Baker N. (1996), Commonality in the determinants of expected stock returns, *Journal of Financial Economics*, Vol.41, no.3.

Lintner, J. (1965), The Valuation of Risky Assets and the Selection of Risk Investments in Stock Portfolios and Capital Budgets. *Review of Economics and Statistics*, Vol.47, pp. 13–37.

Lewellen, J. (2011), Institutional investors and the limits of arbitrage, *Journal of Financial Economics*, Vol.102, no.1, pp.62-80.

Li, X., Sullivan, R. & Garcia-Feijoo, L. (2013) The Limits to Arbitrage and the Low-Volatility Anomaly, *Financial Analysts Journal*, Vol.70, no.1, pp.52-64.

Liu, L., Warner,J. & Zhang, L. (2003), Economic Fundamentals, Risk, and Momentum Profits, *Review of Financial Studies.*

Mohanram, P. (2005), Separating Winners from Losers among Low Book-to-Market Stocks using Financial Statement Analysis, *Review of Accounting Studies*, Vol. 10, pp.133-170.

Newey, W. & West, K., (1987), A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, Vol. 55, pp. 703–708.

Novy-Marx, R. (2012) Is momentum really momentum?, *Journal of Financial Economics*, Vol.103, no.3.

Novy-Marx, R. (2013) The other side of value: The gross profitability premium, *Journal of Financial Economics*, Vol.108, pp.1-28.

Piotroski, J. (2000) Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers, *Journal of Accounting Research*, Vol.38. pp.1-41

Ohlson, J. (1980), Financial Ratios and the Probabilistic Prediction of Bankruptcy, *Journal of Accounting Research*, Vol.18, no.1, pp.109-131.

Sharpe, W. (1964), Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance*,Vol.19 , pp. 425–442.

Pedersen, L. (2019) Efficiently inefficient: How Smart Money Invests and Market Prices Are Determined, *Princeton University Press*

Penman, S., Richardson, S. & Tuna, I. (2007), The Book-to-Price Effect in Stock Returns: Accounting for Leverage, *Journal of Accounting Research*, Vol.45, No.2, pp.427-467.

Roll, R. & Ross, S. (1980), An Empirical Investigation of the Arbitrage Pricing Theory, *The Journal* *of Finance*, Vol.35, no.5, pp.1073-1103.

SEC (2005), Regulation SHO - Pilot Program, *http://www.sec.gov/spotlight/shopilot.htm*

SEC (2007), Regulation SHO and Rule 10a-1, *https://www.sec.gov/rules/final/2007/34-55970.pdf.*

SEC (2010), SEC Approves Short Selling Restrictions, *https://www.sec.gov/news/press/2010/2010-26.htm.*

SEC (2015), Key Points About Regulation SHO, *http://www.sec.gov/investor/pubs/regsho.htm.*

Shleifer, A. (2000), Inefficient Markets: An Introduction to Behavioural Finance, *Oxford University Press UK.*

Shleifer, A. & Vishny, R. (1997), The Limits of Arbitrage, *The Journal of Finance*, Vol.52, no.1, pp.35-55.

Sloan, R. (1996) Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings?, *The Accounting Review*, Vol.71, no.3.

Stambaugh, R., Jianfeng, Yu., Yuan, Y. (2015), Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle, *The Journal of Finance*, Vol.70, no.5, pp. 1903-1948.

Trotta, D. (2008), Short sellers have been the villain for 400 years, *Reuters edge*, https://www.reuters.com/article/reutersEdge/idUSTRE48P7CS20080926?PageNumber=2virtual BrandChannel=0sp=true.

7 Appendix

7.1 Fama-French equations

Fama-French 3-factor SMB:

$$SMB = \frac{(S_{low} + S_{neutral} + S_{high}) - (B_{low} + B_{neutral} + B_{high})}{3} \tag{10}$$

Fama-French 3-factor HML:

$$HML = \frac{(H_{small} + H_{big}) - (L_{small} + L_{big})}{2} \tag{11}$$

Fama-French 5-factor SMB:

$$SMB_{HML} = \frac{(S_{high} + S_{neutral} + S_{low}) - (B_{high} + B_{neutral} + B_{low})}{3} \tag{12}$$

$$SMB_{RMW} = \frac{(S_{robust} + S_{neutral} + S_{weak}) - (B_{robust} + B_{neutral} + B_{weak})}{3} \tag{13}$$

$$SMB_{CMA} = \frac{(S_{conservative} + S_{neutral} + S_{aggressive}) - (B_{conservative} + B_{neutral} + B_{aggressive})}{3} \tag{14}$$

The average return of these 3 portfolios are then used to create the ultimate factor that is later on used in the Fama-French 5-fcator model.

$$SMB = \frac{(SMB_{HML} + SMB_{RMW} + SMB_{CMA})}{3} \tag{15}$$

Fama-French 5-factor RMW:

$$RMW = \frac{(S_{robust} + B_{robust}) - (S_{weak} + B_{weak})}{2} \tag{16}$$

Fama-French 5-factor CMA:

$$CMA = \frac{(S_{conservative} + B_{conservative}) - (S_{aggressive} + B_{aggressive})}{2} \tag{17}$$

Table 1: Data-table

In this table each data item and its respective source is reported. For detailed information about how this data is used in the paper check in particular in chapter 3.2

Item	Name in Database	source
Price	PRC	CRSP
Return	RET	CRSP
Shares outstanding	SHROUT	CRSP
Consumer price index	CPI	U.S. Bureau of Labor statistics
Revenue total	REVT	Compustat
Cost of goods sold	COGS	Compustat
Total assets	AT	Compustat
Stockholde equity	SEQ	Compustat
Net income	NI	Compustat
Depreciation and amortization	DP	Compustat
Capital expenditures	CAPX	Compustat
Current liabilities	LCT	Compustat
Cash and short term investments	CHE	Compustat
Debt in current liabilities, total	DLC	Compustat
Sales/turnover, net	SALE	Compustat
Common shares outstanding	CSHO	Compustat
Price, closing	PRCC_C	Compustat
Debt long term	DLTT	Compustat
Total assets, current	ACT	Compustat
Income before extraordinary items	IB	Compustat
Pretax income	PI	Compustat
Working capital	WCAP	Compustat
Retained earnings	RE	Compustat
Retained earnings, unadjusted	REUNA	Compustat
Non-controlling interest	MIBT	Compustat
Preferred/preference capital stock, total	PSTK	Compustat
Income before extraordinary items, available for common	IBCOM	Compustat