The Power of Long-Term Profitability and Compounding Effects: Empirical Evidence from the US American Market

Felipe Gómez de Larrain and Joachim Bersztel[$]

Handelshögskolan i Stockholm

Autumn 2016, Stockholm

Master Thesis

Abstract

For the US American market, we examine whether a ten-year non-rebalancing investment strategy that picks stocks largely based on historical return on equity and the degree to which a company reinvests earnings can beat the market. Delisted companies are sold and the proceeds reinvested along dividends at the market return, which we approximate by the Russell 3000 TR index. For ten-year periods starting between 1998 and 2005, we observe outperformance over the market in all periods – all but one being statistically significant at the 95% level or higher. The strategy performs especially well when the market performs poorly. As we exclude financial stocks from our portfolios, we argue that part of the outperformance in the periods finishing during or just after the financial crisis can be attributed to this decision. Assuming the existence of a small company size effect, part of the performance of our portfolios could be owed to their predominant picking of small capitalisation companies. Yet, we also observe alpha after applying factor analysis models which show the highest loadings on the size (SMB) factor and to a lesser extent on the value (HML) factor. Loadings on the profitability (RMW), investment (CMA) and momentum (MOM) factors are low or ambiguous.

Keywords: investment strategy, long-term horizon, return predictability, compounding effects

JEL classification: G02, G11, G12

[$] Felipe Gómez de Larrain, e-mail 40748@student.hhs.se. Joachim Bersztel, e-mail 40731@student.hhs.se. We would like to thank Michael Halling, our thesis tutor, for his helpful comments and great responsiveness throughout the writing of this paper.

Contents

List of Tables

List of Figures

List of Abbreviations

B/M	Equity book-to-market valuation ratio
CAPM	Capital Asset Pricing Model
CF/P	Cashflow-to-price valuation ratio
CMA	Conservative Minus Aggressive (Investment Activity factor)
E/A	Book equity-to-asset ratio
E/P	Earnings-to-price valuation ratio
EMH	Efficient Market Hypothesis
ETF	Exchange Traded Fund
FF3F	Fama-French Three Factor Model
FF5F	Fama-French Five Factor Model
HML	High Minus Low (Valuation or Value factor)
IRR	Internal Rate of Return
MOM	Momentum factor
P/E	Price-to-earnings valuation ratio
Rm-rf	Market risk factor
RMW	Robust Minus Weak (Operating Profitability factor)
ROE	Return-on-(book)equity
ROIC	Return-on-invested-(book)capital
SMB	Small Minus Big (Company Size factor)

1. Introduction

At the latest since the 1970s there has been a discussion in the academic finance community whether markets are efficient or not. Proponents of efficient markets have argued that as all available information immediately gets reflected in the market price, it is impossible for investors to perform better than the average market. They have attributed empirical evidence that strategies based on a variety of valuation, profitability, timing and other factors beat the market to higher risk exposure. Opponents of efficient markets have instead argued that the outperformance of such strategies is due to mispricing of securities.

In this paper, we back-test a self-designed investment strategy and compare its results to the market's. Our main motivation is to see whether a simplistic, profitability based, long term strategy that does not rebalance the portfolio, as it tries to benefit from compounding effects but consequently relies on partially very outdated accounting data, can beat the market. We test our strategy against the market to see whether we generate a statistically significant outperformance and also scrutinise our performance applying the standard Fama and French three factor model (FF3F) (Fama and French, 1993) and models using additional factors such as momentum, investment and profitability to our portfolio. We skip to test for the Capital Asset Pricing Model (CAPM), as it seems to largely be invalidated by empirical studies (e.g. Fama and French, 1996a).

Due to the non-rebalancing of the portfolio and the implied use of accounting data that in its extreme is ten years outdated, an outperformance over the market that could mostly be explained by Fama and French's model should likewise be considered an interesting result. Also, as long as there is no proven intuitive interpretation for the outperformance of small over big and high equity book-to-market (B/M) over low B/M stocks it remains unclear whether a causal relationship or mere correlation is the reason for a portfolio's loading on these factors.

What we do *not* aim at with this paper is to uncover the final solution to the debate whether stock markets are efficient or not. In the literature section of this paper we are going to portray the efficiency debate and show what arguments for and against the interpretation of abnormal returns as mispricing or as risk compensation exist. While we may state our opinion on those arguments and on our results, the latter cannot ultimately prove either efficiency or inefficiency. The primary reason is that, while empirical results can be observed and individual arguments about the correct interpretation of such results can be regarded intuitive or counter-intuitive, following the argumentation of Fama and French (2006, 2008) empirical results and tests based on the valuation equations of companies cannot determine whether the cause is mispricing or risk.

1.1. Hypothesis

Specifically, the hypothesis we aim to test is that a mechanical investment strategy based primarily on past return on equity (ROE) and to a lower extent on past pay-out ratio can beat the market over a time-span of ten years. We suspect that such companies that over the previous years in terms of these metrics have outperformed other companies in the same market, should continue to on average do so for the mid-term future and that this – aided by compounding effects on the capital not paid out – should also lead to an average return higher than the broad market return over the same period. The strategy does not rebalance the portfolio over the investment period and relies on accounting metrics of the previous ten years. It is worth mentioning that we invest once for ten years, while in the finance literature a five-year non-rebalancing strategy is already considered long-term (Lakonishok et al., 1994). To be able to account for compounding effects it is a crucial element of our strategy to estimate future earnings based on aforementioned metrics, as opposed to merely ranking them on their past average ROE and pay-out ratio. The strategy differs from (price) momentum investing (e.g. Jegadeesh and Titman, 1993) in that it looks at past accounting and not stock performance, does so for the long period of ten years and invests without rebalancing for another long period of ten years. If differs from earnings momentum investing (e.g. Myers et al., 2007) in that it evaluates profitability and not trends in absolute earnings (be it on aggregate or EPS level).

Analysts and media often broach the growth of companies' sales or earnings. We thus would not be surprised to find investors overstretch the meaningfulness of current growth rates for the medium to long-term future. In fact, we find affirmation for such a presumption in the work of Lakonishok et al. (1994) and Chan et al. (2003) providing empirical support for the thesis that investors irrationally extrapolate recent growth into the future. The same applies for the findings by La Porta et al. (1997) indicating that growth investors seem to be disappointed at the earnings announcements following their investments.

ROE is chosen as the central metric of our approach as we expect it to be a relatively sustainable indicator of profitability and quality of a company over longer periods of time. Koller et al. (2015) look at empirical data to show that over the last 50 years companies have on average been fairly successful in maintaining their rates of return on invested capital (ROIC). They draw the conclusion that when a company has found a strategic formula that creates a sustainable advantage for it, the management is usually quite successful at defending and renewing it. Competition will eventually drive capital returns down but the reversion to mean on average is much slower than for growth and many companies can maintain high ROICs for decades. We prefer ROE over ROIC as the measure for our study as it represents what the common stockholder actually earns. As Ozkan (2001) and Flannery and Rangan (2006) show, most companies at least have a rough target financial leverage ratio. Therefore, the conclusions of Koller et al. (2015) about ROIC should generally hold for ROE as well. We use the average ROE of the ten previous years to reduce the weight of outliers and to ensure that we buy companies which are able to maintain a high ROE over longer periods of time.

The second important pillar of our hypothesis is that compounding effects – as they are non-linear – are extremely powerful over longer periods in time and we consider it likely that many analysts and investors do not take the power such effects can have on a company's equity over longer periods in time into account when predicting future cash-flows. This is why we complement ROE information with the average pay-out ratio of the past to estimate what part of a company's earnings is kept in the company and compounded at the company's ROE. Of course, doing so implies the assumption that the company will be able to maintain its ROE, i.e. increase its earnings at the same pace as the additional equity each year. However, as we rely on pay-out ratios and ROE from the same past years, we have the assurance that in the past the company has been able to do so. Also, by using data from ten years, our selection of companies should not be biased by exceptional short-term growth rates that cannot be maintained for the long-term.

Most successful value investors would probably not consider such a simplistic strategy on its own but look at additional aspects, both qualitative and quantitative of the firm. The strong focus on ROE and the idea of trying to figure out what earnings will be ten years into the future is said to be one element in the security analysis process of Warren Buffett (Buffett and Clark, 1999). Yet, Buffett may use a similar approach as a heuristic but he will likely consider many other aspects that will rule out buying companies which earn high ROE only due to extreme financial leverage. To ensure that our portfolio does not consist of highly levered companies we include a maximum leverage ratio threshold. However, we try to make this ratio not the predominant aspect of our strategy, as will be explained in the methodology section. Moreover, we only look at companies that over the previous three years on average have reported positive net earnings. While this is primarily due to technical reasons, it implies a qualitative judgement of companies.

It bears highlighting that our strategy does only take the initial valuation of a company into account in the form of its current stock price, which is used to determine the Internal Rate of Return (IRR) to our estimated future value. In the calculation of an IRR over ten years the potential over- or undervaluations thus carry little weight. We relinquish to account for the initial valuation in a more powerful way, e.g. by introducing a B/M or price-to-earnings (P/E) measure. This seems fairly rare in the literature for value strategies that based on certain metrics try to outperform the market. A simplistic strategy that combines our focus on capital returns with a valuation metric in the form of an earnings yield comes from Greenblatt (2005). The author blends the return-on-tangible-capital measure with the earnings yield based on earnings before interest and tax over the current stock price. He then buys the 30 stocks that show the best combined score and rebalances his portfolio yearly. Such a strategy would have earned a large outperformance in the back-testing period of 1988-2004 in the US market. While we would expect the inclusion of a valuation ratio to considerably improve the returns of our strategy, we exclude it as we do not want to subtract weight from our principal hypothesis.

The choice not to rebalance differs from most literature on trading strategies. We do so for several reasons that will be explained in detail in the methodology section. Generally, we argue that not rebalancing saves costs and for many investors taxes. Besides, as stated above, more than in developing the maximum market-outperforming strategy, we are interested in providing empirical support for the idea that over the long-term companies that are highly profitable and can reinvest these earnings at similar rates perform better than the market. With rebalancing we would not be tracking such companies over the long-term in many cases. Finally, we find it interesting whether a 'buy and go away' strategy relying over the time increasingly on outdated accounting data can perform better than the market.

Ideally, our strategy can find outperformance over the market that is not explained by the FF3F and newer models. Ex-ante we expect this at least for the FF3F model to likely be the case, as both our principal measures used – ROE and pay-out-ratio – are directly or indirectly related to the factors that Fama and French (2015a, 2015b) suggest to add to FF3F to improve the model's performance in capturing stock returns, i.e. to the profitability of a company and its level of investment. While the profitability factor is positively related, the investment factor is rather inversely related to our hypothesis. In any case, we expect both these recent additions to the FF3F model to be more closely linked – one way or another – to the performance of our portfolio. Yet, as we only evaluate ROE and pay-out ratio combined and not on a separate basis, we expect the link to both of the factors to be limited as well.

1.2. Principal findings and structure of the study

Implementing our strategy for eight different investment periods, namely all ten-year periods starting between 1998 and 2005, we exclusively find positive annualised returns of between 3.56% and 7.98%. These correspond to outperformances over the market over the respective periods of between 1.13%p.a. and 8.67%p.a. Moreover, we find that the outperformance for all implementations is driven by at least 57% of the stocks in the portfolio (and a maximum of 75%). For seven out of eight periods the outperformance is statistically significant at the 95% level. We observe a negative correlation between the annualised index performance and outperformance of the portfolios over the periods. While the companies in our portfolios that are delisted during the ten-year period on average show positive returns and outperform the index, they do so by less than the average portfolios. Our portfolios to large extent are comprised of small stocks and B/M ratios are neither extremely high nor low. Finally, we find that allowing for more leverage, choosing a smaller portfolio and excluding companies that are outliers in terms of past ROE figures does not change the overall evaluation of our strategy.

Our factor model analysis shows that our portfolios generate considerable alpha. The range of statistically significant annualised alphas across portfolios in the Fama-French five-factor model is 2.71% to 6.40%. Market beta generally is around one for all models. The most heavily loaded factor across all models is the size factor, followed by the value factor. Loadings on other factors are either much lower or ambiguous.

This paper proceeds as follows: The next section gives a review of important observations and theories in the relevant literature. Section 3 goes on to explain our methodology and the decisions and trade-offs involved in conceptualising and implementing it. Sections 4 and 5 describe and explain our empirical results and test them for statistical significance as well as robustness. Section 6 analyses our portfolios' exposure to factor loadings in the setting of different factor models. The last section summarises our work and seeks to draw conclusions from it, also highlighting what further research in the area would be desirable.

2. Literature Review

2.1. The Efficient Market Hypothesis

A major question in the financial literature is whether markets are efficient and, if so, to what extent. Building on previous work, the Efficient Market Hypothesis (EMH) as such was conceptualised by Eugene Fama in the 1960s. In his 1970 paper "Efficient Capital Markets: A Review of Theory and Empirical Work" he exhibits the idea that financial markets are efficient in the sense that stocks are always priced at their fair value and it is impossible to outperform the market by individually picking stocks (Fama, 1970). Consequentially, the only way to beat the market other than by chance would be for the investor to buy a portfolio that is riskier than the average market, for which he would be rewarded by higher returns – a point that is particularly reinforced by Malkiel (2003).

The EMH requires all actors on a market to take all available information about the future into account and information to be essentially freely available to all market participants. It does not require market participants to act rationally from an objective perspective, though – it is possible for investors to over- or underreact to news on a given stock and thus cause a temporary over- or undervaluation. The expected risk-adjusted returns, however, should be equal for all investors, i.e. it is only possible to judge whether a stock at a point in time was over- or undervalued when looking at past stock prices and any patterns that can be found in past prices are not reliable for predicting future price developments. (Fama, 1970 and Malkiel, 2003).

Fama in his article differentiated among three different degrees of efficiency. Weak form efficiency means that the market price of a stock reflects all information that can be deducted from historical prices. Under this form, technical analyses would be of no use but fundamental analysis might be successfully employed to find undervalued stocks. Semi-strong form efficiency means that all types of publicly available information are incorporated into market prices, leaving only insider information as a means to beat the market. The strong-form efficiency rules out all possibilities of systematically generating an outperformance, as even private information is reflected by the market price. (Fama, 1970). In the remainder of this section we will only be concerned with challenges to the weak and semi-strong forms of efficiency.

Initially, finance academia largely used the CAPM to test for market efficiency. Based on Markowitz' Modern Portfolio Theory, the CAPM is an asset pricing model developed by Sharpe (1964) and Lintner (1965). It assumes that the expected return for a share is composed of the return on a risk-free asset and the individual risk premium of the asset, which depends on the market risk premium and the historical volatility of the asset relative to the market's volatility. The model assumes that volatility is an appropriate proxy for risk and follows the rationale that investors in stocks with higher than average volatility must be compensated for this higher risk taking by higher returns.

2.2. Empirical cracks in the CAPM

2.2.1. The 'traditional' anomalies

Over the years many empirical apparent inconsistencies with the CAPM were discovered by finance academia and investment professionals. In 1985 DeBondt and Thaler found a reversal in long-term returns that cannot be explained by the CAPM (DeBondt and Thaler, 1985). Stocks that over the previous one to three years have underperformed the market tend to outperform the market during up to as much as five years after the portfolio formation period over the sample period from 1931-1982. On the contrary, stocks that have outperformed the market, tend to underperform the market over the same period, adding to the profits of a potential long-short portfolio. Fama and French (1988) conclude that up to 40% of long holding period returns can be predicted by a negative correlation with the preceding years. Poterba and Summers (1988) find evidence for the so-called mean reversion of stock prices over longer periods of time and with a sample period starting in the 19th century.

In contrast to this, Jegadeesh and Titman (1993) find that in the period 1965 to 1989 profits that cannot be explained by the CAPM could have been earned by a simple strategy that buys stocks that have performed well over the previous twelve months and sells stocks that have performed poorly over the same period. Such abnormal returns could have been earned over holding periods of three to twelve months. Their finding was confirmed over the following years in similar studies (e.g. Lo and MacKinlay, 1999).

Another prominently investigated return anomaly evolves around the size – usually measured as market capitalisation – of a company. Banz (1981) displays evidence that for New York Stock Exchange (NYSE) listed stocks abnormal returns could have been earned over at least forty years by investing in small rather than big firms. The effect seems to be strongest around very small companies and limited when comparing average-sized to large companies. Keim (1983) shows that independent of market risk since 1926 small companies in the U.S. produced annual return rates of over one percentage point higher than large companies. Fama and French (1993) confirm this finding for the period 1963-1990 by sorting the market into decile portfolios based on size. In a 2008 paper (Fama and French, 2008), re-analysing the size effect, they reinforce Banz' finding that the effect centres on the smallest stocks. As micro caps by their definition only account for around 3% of the total NYSE market cap but are a much larger percentage of the companies listed, they argue that it makes sense not simply to sort the market by deciles but to analyse whether the size effect persists once micro caps are excluded from the analysis. In doing so, they find that the small company effect is essentially attributable to micro caps alone.

Finally, a considerable number of studies and research has been focused on the abnormal returns of investment strategies in so-called "value" stocks – as opposed to "glamour" or "growth" stocks. The predominantly used metric to classify stocks as value or glamour is book-to-market equity value. Following research by Stattman (1980) and Rosenberg et al. (1985), Fama and French (1992) examine the performance of portfolios built on such B/M ranking of the U.S. American market between 1963 and 1990. For each year all stocks are ranked according to their B/M ratio and the market is split into deciles. The highest B/M portfolio on average shows a monthly return of 1.83% percent, whereas the bottom portfolio on average only gains 0.3% per month. The differential is not only strongly statistically significant but also twice as large as the one the authors find for the size effect in the same study. Lakonishok et al. (1994) confirm the abnormal outperformance of the strategy over holding periods of five years after portfolio formation. Their findings show that the value stock decile of the market outperforms the glamour one by 10.5 percentage points a year over such periods when looking at data from 1968-1989. The authors highlight the importance for long-term investors of the detected sustained outperformance using a buy-and-hold strategy. Chan and Lakonishok (2004) update the evidence of high B/M outperformance and confirm persistent validity of the strategy's abnormal returns for more recent periods, except for some holding periods during the new economy bubble which they attribute to irrational pricing caused by exaggerated investor optimism for some sectors. In Fama and French (1998) the value effect is verified for validity outside of the U.S. market, both for high B/M and profitability-to-price metrics, and is found applicable in almost every country examined.

2.2.2. Extensions of the value strategy and purely accounting based anomalies

Once the considerable outperformance of high B/M strategies had been amply documented, the focus of some in the field shifted towards the possibility to further add to the performance of such strategies by including accounting data. In this spirit Piotroski (2000) combines the high B/M strategy with accounting based fundamental data. Setting out from the observation that the average outperformance of previous high B/M portfolios is only attributable to 44% of the stocks in these portfolios, he applies nine measures of profitability (e.g. ROA), financial leverage, liquidity and operating efficiency to companies in the top quintile of the B/M ranking over the period 1976 to 1996. On each of the nine variables he awards companies a "good" or "bad". Piotroski then labels companies with eight or nine times "good" as strong value and finds that a long only strategy in these firms earns 7.5 percentage points more annually than the simple B/M strategy. When additionally short-selling companies that earn eight or nine times "bad" the strategy's results are even more pointed. The results are robust to tests for previously documented anomalies such as momentum, mean reversion and the size effect. However, Piotroski finds that the mean returns of his portfolio are concentrated on small- to medium-sized companies, stocks with low trading volume and stocks with little analyst coverage. He also notes that financial distress or profitability changes alone are not sufficient to consistently achieve the high documented performance. Multiple items of accounting data together seem to be a clearer indicator of historical performance.

Another example of expanding the B/M strategy comes from Bartov and Kim (2004). They argue that, on average, low accounting accruals indicate conservative accounting and high accruals stand for aggressive accounting. Yet, they expect the earnings on average to revert to mean, thereby leading to positive surprises for value stocks (those that have low accruals) and the opposite for glamour stocks. Therefore, stocks should be ranked jointly by B/M and their level of accounting accruals. The study concludes that a joint classification on B/M and the level of accounting accruals performs better than a single classification on either of the two factors.

Besides the 'value' anomalies based on the B/M ratio and accounting based amplifications of this concept, there is a considerable amount of research finding and analysing effects that are solely rooted in accounting data. Sloan (1996) first analysed the effect of the level of accounting accruals on stock returns that was later used in combination with B/M by Bartov and Kim (2004). Haugen and Baker (1996), as well as Cohen et al. (2002) find that more profitable firms have higher average stock returns. In contrast, Fairfield et al. (2003) and Titman et al. (2004) find empirical confirmation for the hypothesis that firms that invest more exhibit lower stock returns. Fama and French (2008) curtail this anomaly to primarily micro caps and to a lesser extent small caps, whereas it appears to be non-existent for big companies. Other studies (e.g. Daniel and Titman, 2006 and Pontiff and Woodgate, 2008) show that there is a negative correlation between net stock issues and average returns. Companies that are net issuers of stock will on average underperform the market. Importantly, this effect seems to be non-existent from 1926 to 1963 and only exist thereafter.

More anomalies are documented, which we do not elaborate on in this section. They are excluded for different reasons. Many anomalies have been found in relation to investing in certain months, weekdays or seasons. While the stability of these anomalies over time is questionable, we exclude them primarily as they are barely linked to the value topic of this paper and in general there seems to be little connection to other of the anomalies found. Also for the limited relatedness in the context of this paper, we do not enter into details about studies that have focused on analysing the timeliness of the stock market in reacting to new information by analysing returns around earnings announcement dates.

The reason not to look in detail at profitability-to-price value strategies, however, is different. Fama and French (1996) argue that many of the CAPM return anomalies like B/M, earnings-to-price (E/P), cash-flow-to-price (CF/P) or past sales growth are related. They find evidence that B/M has the highest explanatory power and that the other factors in combination with B/M add little value. Ever since, research has not focused on these other metrics, as finance academia in general seems to accept Fama and French's findings.

2.3. The transition from CAPM to multi-factor models

Already in 1980, Grossman and Stiglitz (1980) had highlighted an inner contradiction of the EMH. The hypothesis requires investors to seek to beat the market and in doing so to instantly incorporate all news into the market price. Yet, according to the hypothesis there is no point in trying to beat the market as all information is always immediately priced in. The investor should thus just choose stocks based on his risk preferences. The question that arises then is, if no one tries to beat the market assuming that all information is accounted for in the price already, how does the information get incorporated into the market price? A possible answer could be that markets are efficient most of the time and as soon as someone realises an inefficiency it will disappear, returning the prices to their correct levels. Yet, as famous investor Warren Buffett notes, there is another reason to question the joint theoretical framework that the EMH and the CAPM provide. Buffett states that market beta is not intuitive and provides a practical example: In 1973 the Washington Post Co. traded at \$80m. Based on its assets the company could have easily been sold to different potential buyers for \$400m, though. If the stock had declined even further to represent a valuation of \$40m the beta would have increased. The reduction in price would have made the stock look riskier in terms of beta, whereas the risk of overpaying would have further declined and the probability of earning larger returns would have risen. (Buffett, 1984).

Finally, with empirical evidence for abnormal returns piling up, the CAPM seemed to exhibit ever growing cracks when used to test market efficiency and in the early 1990s was increasingly exposed to attacks. The apparent failure of the CAPM led to the development of multi-factor models, which extend the CAPM to account for more factors that are considered crucial in the pricing of assets. By assuming that these factors actually represent different risks that investors need to be compensated for when buying a stock, such multi-factor models built on empirical findings can be brought in line with the assumption of efficient markets. The arguably most famous such model was developed by Fama and French (1992, 1993). The authors argue that a large share of the abnormal returns under the CAPM can be explained by a model of three factors, which complements the CAPM's factor of relative volatility compared to the market by a factor for the size and one for the B/M ratio ("value"-factor) of a company – the FF3F model. Conducting further research, Fama and French (1996b) reinforce that the vast majority of abnormal return factors under the CAPM in the past could have been explained by their three factor model. As the model also captures mean reversion, the only major anomaly not covered (and then known) would be momentum.

Despite the apparent empirical invalidation of the CAPM, Fama and French (1996a, 2004) state their appraisal for the theoretical foundation the model provides in linking higher returns to more risk exposure. In contemplation of the size effect largely being linked to micro caps and the many accounting based anomalies that have been documented since the turn of the millennium, Fama and French (2006, 2008) look at the explanation of abnormal returns from a different perspective. Setting out from the valuation based relation between B/M and expected returns, they argue that when controlling for B/M, higher expected returns largely depend on higher expected net cash-flows. While they attribute unique explanatory power to momentum and several of the recently found accounting-based anomalies, they argue that those are essentially proxies for expected cash-flows. For others, as e.g. the accrual accounting effect, they argue that they mostly disappear when controlling for net expected cash flow and B/M. Thus, in these publications Fama and French attribute returns largely to the expected cash-flows and the B/M ratio of a stock. Building on this work Fama and French (2015a, 2015b) finally propose a model that extends their FF3F model by a factor for investment and one for profitability, converting it into a five-factor model and narrowing down the gap between real returns and those predicted by the model. Upon publishing their new five-factor model, some were surprised that Fama and French had not included momentum as one of the new factors. Attributed to Mark Carhart (1997), there had long been an extension to the FF3F model which incorporated momentum, effectively making it a four-factor model.

2.4. Interpretations of the empirical anomalies

Generally, since the 1990s there seems to be little disagreement left on the existence of possibilities to beat the market on a risk-adjusted basis as defined by the CAPM in the past. Even Fama and French, the long-time proponents and defendants of the EMH, acknowledge this by developing the FF3F model. Chan and Lakonishok (2004) affirm that especially for the outperformance of value strategies there generally seems to be agreement among the academic community. Yet, they underline that there is much less unanimity when it comes to the interpretations of the empirical findings. Whereas one sort of literature argues for a risk-based explanation that would align the apparent abnormal returns with market efficiency, another set of papers focuses on the dependability on the empirical findings and their applicability in investing. A third branch prefers an explanation that attributes abnormal returns to mispricing and is thus inconsistent with market efficiency.

The first group argues that markets are efficient and that the reason the CAPM does not work is that it does not capture all relevant risk factors of a stock. The predictability of future returns that a model like the FF3F model or other multi-factor models allow would thus be of little value to the shareholder, as the reason for him to earn higher returns is the higher risks he incurs. Fama and French (1992, 1993) already when developing their model propose this interpretation. Smaller companies and companies with high B/M ratios would thus be riskier investments than large and low B/M ratio companies. While for a small company the reason to be riskier could for instance be an average lower liquidity, regarding the value factor Fama and French (1996b) argue that a value stock's higher returns are related to a higher exposure to a financial distress factor. The rationale here would be that stocks that seem cheap on a B/M basis are so for good reason. Any abnormal returns that could be earned under a multi-factor model would be considered to represent some unknown additional priced element of systematic risk. Although still sympathising with a risk-based explanation, in their more recent work Fama and French (2006, 2008) do not take a clear position anymore on whether deviations of actual returns from the expected returns that can be deducted from a company's B/M and expected net cash flows (or proxies for such) are attributable to differences in risk or to mispricing. Instead, they allude to any tests based on the valuation equations not being helpful in determining whether the cause is mispricing or risk, thereby making it very difficult to rule out a risk-based explanation.

The second wide-spread interpretation of the anomalies detected is that they are not dependable or might not be useful in real investment practice. Many researchers – whether inclined towards market efficiency or inefficiency – contemplate the following issues as possible, at least partial, explanations. Schwert (2001) suggests that the apparent disappearance or weakening of some effects after they are published might be caused by practitioners quickly adapting to exploit any such pattern, thereby correcting any potential mispricing. He also contemplates that as researchers are constantly skimming through data, they will eventually find patterns – whether or not those have any meaning or are mere coincidence. Malkiel (2003) supports this view and adds that some strategies may on average have worked in the past but not in all years. He doubts that any of the effects found was ever strong and dependable enough to reliably earn money on it. Regarding the size effect, he notes that it did not prevail between the mid-1980s and the end of the 1990s and similarly he considers that an outperformance based on the B/M ratio might be unique to the historical period Fama and French (1993) looked at – a point Fama and French (1998) themselves had previously contemplated. As for mean reversion Malkiel argues that such would be perfectly consistent with an efficient market if it happened as a reaction to fluctuations in interest rates. Lastly, for momentum he refers to Odean (1999) who analysed the returns of investors implementing momentum strategies in the 1990s and finds that they underperformed the market due to the transaction costs involved. Chan et al. (1996) establish a possible link between momentum and findings that firms reporting higher than expected earnings outperform firms that do the opposite and that this outperformance persists for up to six months (Latane and Jones, 1979; Bernard and Thomas, 1989; Bernard et al., 1995).

Another possible explanation for abnormal returns is that markets are not efficient and that there is mispricing in markets. Literature arguing this way often attributes this mispricing to behavioural factors influencing market participants' decisions. DeBondt and Thaler (1985) were early in attributing the outperformance of the mean reversion they discovered in stock prices to behavioural factors, referring to "*waves of optimism and pessimism*" that cause prices to "*systematically differ from their fundamental value*". Jegadeesh and Titman (1993) equally hypothesise about over- and underreacting investors. Lakonishok et al. (1994) – already after the publication of the FF3F model – argue that the outperformance of high B/M stocks is likely owed to temporarily depressed prices on these stocks as a consequence of investors overreacting to prior poor performance and maintaining too negative expectations about the future performance. Piotroski (2000), in his aforementioned study merging B/M and accounting data to identify stocks, reasons that the decreasing outperformance of his strategy for larger companies, those followed by more analysts and those most traded indicates that such environments might be better for spreading information fast. As among the high B/M companies the healthiest appear to generate the most return, he also posits that conceptually a risk-based explanation is not attractive as the "*firms with the strongest subsequent return performance appear to have the smallest amount of ex ante financial and operating risk.*" Although Piotroski does not entirely rule out a risk-based explanation, he is rather appealed by an explanation based on Lee and Swaminathan (2000a, 2000b). They argue that early stage momentum losers that continue to generate poor earnings can become subject to strong pessimism and be shunned by investors. Eventually, the average late-stage momentum loser does recover and turn into an early stage momentum winner. Piotroski's strong value firms, he argues, have the same characteristics as late stage momentum losers. Bartov and Kim (2004) make another interesting contribution, finding that in their sample the outperformance of the value portfolio is only significant for stocks with a price under 10$. They conclude that this must be due to mispricing associated with "*unsophisticated*" ownership for such stocks as many investment companies were not able to hold them.

2.5. Testing and comparing 'value' centred arguments for and against market efficiency

Chan and Lakonishok (2004) make the point that it is unlikely that the value premium was attributable to data snooping, as there is a logical explanation for it. Assuming that the value strategy's outperformance is in fact not linked to a statistical flaw as e.g. data snooping, risk and mispricing based explanations remain.

Analysing the risk-based motivation, Chan and Lakonishok (2004) bring forward that previous studies have usually not found market beta to be related to returns and that if risk differences were to explain abnormal returns, outperforming strategies should do worse than the market in extremely adverse economic scenarios. Looking at bad months for the stock market and for the overall economy they find that in both cases a value portfolio significantly outperforms the market and that the outperformance is even stronger for the worst months. On the opposite, in positive environments value stocks at least match glamour stocks' performance. Siegel (2006) shares this argument, stating that in the decade up to 2006 there had been the "*tech bubble, 9/11, a recession, major corporate scandals and wars in Afghanistan and Iraq*" and during all of them value and small stocks outperformed glamour and large cap stocks by wider margins than in the past. From a conceptual standpoint Chan and Lakonishok (2004) criticise that by ascribing high B/M stocks' higher returns to risk, Fama and French were essentially saying that new economy companies with very little book value were less risky than, for instance, utilities. Moreover, they criticise that the risk-based explanation for value stocks' outperformance was only brought up once the effect had been documented.

When it comes to behavioural arguments, assuming a non-crisis economic environment, one of the most popular cases for an outperformance of value stocks is that investors might extrapolate past growth figures too far into the future. For instance, Lakonishok et al. (1994) showed that high B/M stocks on average had poor recent growth in earnings, cash-flow and sales. Rationally, though, it should be assumed that the B/M ratio of a stock represents expectations about the *future* growth of a company and is independent of past growth figures (as long as they do not represent a good proxy for future growth). Consequently, investors should value lowly companies that they expect to grow little, resulting in a high B/M, and highly those that they expect to outgrow the market, resulting in a low B/M. Chan et al. (2003) test this presumption. From 1951-1998 stocks are ranked into deciles by the growth of income over a five-year period and B/M at inception and end of the period is measured. If initial B/M was derived from rational expectations about future fundamentals growth, a low B/M should stand for high future growth and vice versa. Yet, the authors find that companies with the highest growth at the inception of the five-year period have a slightly over average B/M ratio. Interestingly, the B/M at the end of each five-year period has a much closer relationship with growth. Those companies that had high growth over the previous five years show a low B/M ratio at the end of the period and vice versa. So, in line with Lakonishok et al. (1994), Chan et al. (2003) provide support for the behavioural argument that investors are irrationally influenced by past growth which they extrapolate into the future. From this finding Chan and Lakonishok (2004) deduct the expectation that when actual growth rates materialise, prices should adjust to a justified level. They find confirmation for this in La Porta et al. (1997). For portfolios constructed by B/M, they uncover that at earnings announcement one year after portfolio formation the growth investors seem to be disappointed, leading to a cumulative event return of -0.5%. At the same time value investors seem positively surprised, as suggested by an event return of 3.5%. Also at the following two years' earnings announcements the differences are found to be large and statistically significant.

A further behavioural argument looks not at the investors but at their agents. Analysts have their own interests in recommending past winners, as they are easier to sell – generating commissions and investment banking business. Often growth stocks are also in more exiting industries making them more apt for marketing in reports and media. (Chan and Lakonishok, 2004).

2.6. Evaluating empirical evidence from investors' performance

Finally, it is worth to look at some empirical argumentation that states that active investing cannot beat the market. Malkiel (2003) introduces the section about mutual fund returns of his paper stating that real investors' results for him are the most convincing test of market efficiency. Looking at previous studies about the returns of mutual funds, he argues that on average after fees they have underperformed the market. Also, winners over one period usually are losers over the next. The first argument is not problematic for the investor choosing the right fund. The second argument, however, seems troublesome for all investors. Yet, there are two problems with Malkiel's argument. He looks only at those funds that performed best, to see whether their performance is reversed in the following period. Such focus on potential outliers is always dangerous. Also, the period he specifically focuses on is 1998-1999 to screen for winners and 2000-2001 to screen for reversal – i.e. he focuses on the period before and after the internet bubble burst. Another interesting thing to mention is that he only examines mutual funds, leaving many other types of investors out. One such investor would be Warren Buffett.

Warren Buffett agrees with Malkiel's argument that real investors' returns are a convincing test of market efficiency. In a famous speech (Buffett, 1984) he makes an empirical argument about the efficiency of markets. Buffett presents a group of investors he has known for many years and who he claims to have selected not based on their performance with hindsight. They all largely and consistently beat the market and their returns are audited. Their portfolios are of different sizes, have very different approaches towards diversification and contain very different titles. All they have in common is Benjamin Graham's value investing mind-set to look for cheap businesses and buy parts of them. Buffett argues that they have achieved their outperformance assuming much less risk than average, having much better records especially for years when the market was weak. "*If you buy a dollar bill for 60 cents it's riskier than if you buy a dollar bill for 40 cents, but the expectation of reward is greater in the latter case. The greater the potential for reward in the value portfolio, the less risk there is.*" Clearly, Buffett's argument requires the reader to trust him, when it comes to the selection of the nine investors he presents.

2.7. Perceptions of the efficient market debate

Although less extreme in recent years, the idea of efficient markets has long been the predominant view in finance academia. While we do not want to judge here whether the EMH is correct or not, we think it bears mentioning how the surrounding debate has long been perceived by some in the academic community.

Guay (2000) opines that the "*the interpretation of the anomaly and trading strategy literature is heavily influenced by the strong and diffuse priors held by the academic community.*" Ball (1992) at an early stage underlined that the "*conclusion that markets are inefficient emerges from failing to reject a specific inefficiency hypothesis, not by a process of eliminating all other known explanations for the evidence.*" Siegel (2006) states his view that prices can be influenced by speculators and others often buying stocks for reasons not related to their fundamental value, e.g. diversification, liquidity and taxes. With respect to the efficient market debate at that time, he notes that it reminds him of the "*astronomers in the 16th century who attempted to save the earth-centred Ptolemaic view of the universe*" and who were forced to add more and more twists to the model to uphold it. Practitioner Buffett expressed his criticism on the matter in 1984: "*Ships will sail around the world but the Flat Earth Society will flourish.*" (Buffett, 1984).

3. Methodology

In this section of the paper, we first motivate our decision not to rebalance the portfolio over the investment period of ten years. We then go on to explain the choice of the US market and of the Russell 3000 TR index as the proxy for it. Next, we precisely explain the specific steps involved in the picking of stocks according to our mechanical investment strategy and in the following evaluation of their real performance. Along the way we motivate decisions we had to make in the implementation of the strategy. Subsequently, we go on to discuss how we analyse and test the empirical results we find for our back-tests. At last, the section describes and evaluates the importance of potential biases or problems that we do not account for.

Our strategy employs accounting data of 11 years (t=-11 to t=-1) preceding a base year to estimate the parameters that are then used to build a portfolio in which we invest in the beginning of the base-year (at T=0) and to which we hold on for a period of 10 years (t=1 to t=10) without rebalancing the portfolio. We use the general methodology described in this section to test our strategy for eight different periods in time, namely all ten-year investment time spans starting between 1998 and 2005 and ending between 2007 and 2014.

As mentioned in the introduction, at first sight there may appear to be similarities between our approach and momentum investing in the sense that our hypothesis partially builds on the expectation that companies which in the previous years have outperformed other companies in the same market will generally continue to do so for the mid-term future. It is important to notice, though, that, as opposed to most momentum literature, in our hypothesis we do not refer to past outperformance in terms of stock returns but in terms of *accounting* performance. Only as the consequence of such accounting outperformance, we expect to see stocks outperform the market in the future in terms of returns for the shareholder. Furthermore, in contrast to momentum trading, we do not look at all at the past price development of a stock and buy stocks for a long period of ten years, whereas most momentum literature, especially when it comes to momentum in factor models, is about short-term momentum trading (e.g. Carhart, 1997).

Yet, there are studies looking at accounting performance in terms of “earnings momentum”, instead of price momentum (e.g. Myers et al., 2007 and Hou et al., 2009). These studies indicate that buying companies that previously have shown a trend of rising earnings is likely to lead to a return outperformance versus the market. Also, these studies often look at much longer time horizons than price momentum strategies do. The difference between such studies and our strategy is that we do not barely look for a trend of rising earnings but for a sustained high profitability. A company could increase earnings marginally each year while still earning low capital returns. We, however, look for companies that earn a high ROE and can reinvest it at a similarly high profitability level. Thus, our approach should much closer be related to a potential profitability momentum strategy, picking stocks that have in the past shown high ROE. The profitability factor of Fama and French is to some extent related to this idea. However, it only assesses profitability at one point in time and it does not evaluate the profitability conjointly with the reinvestment level of earnings.

3.1. The non-rebalancing of our portfolio

In most literature regarding investment strategies portfolios are regularly rebalanced. This is to ensure that at each time the stocks that best fulfil the strategy's criteria are considered in the portfolio. We do not rebalance our portfolio for several reasons.

While if a strategy works it may seem appealing to refocus the portfolio regularly, the returns that can be obtained by the rebalancing have to be traded off against the transaction costs an investor incurs and against the taxation that in most countries the average investor faces upon realisation of a gain.

The tax benefits of investing long-term can be immense and are two-fold. Taxation on all gains is postponed into the future, which due to the time value of money is beneficial. Also, the interruption of compounding effects on the capital invested is impeded. As compounding effects are non-linear, having the same pre-tax returns on a portfolio, it is much more valuable to pay taxes of x% on returns after 10 years than to pay x% on returns every year and then reinvest the capital. The benefits of rebalancing would have to be very large to offset the tax disadvantages.[1]

Another argument not to rebalance the portfolio is that we are interested in providing empirical support for our hypothesis that over the long-term companies that are highly profitable and can reinvest these earnings at similar rates perform better than the market. If we rebalanced our portfolio, part of the portfolio would regularly change, thus making it difficult to analyse over the entire investment period of ten years the evolution of the stock-prices of companies that over a long time manage to compound their capital at attractive rates.

Finally, not rebalancing has the appeal to be able to observe how the returns of a 'buy and go away' strategy – relying over the time increasingly on outdated accounting data – would compare to those of the average market. It would be an interesting finding to see a strategy using highly outdated input data to perform better than the market.

3.2. Choice of the market and the index used as proxy for it

For the purpose of this paper, we choose the US American market. Being in Sweden, we initially considered implementing and testing our strategy for the Swedish market or, alternatively, for the Nordic market. However, we rely on data availability for each company for at least 11 years (the first year only being used in the calculation of ROE) and the number of companies listed in

[1] The longer the investment period without rebalancing, the stronger the non-linear compounding advantages of deferring the taxation into the future become.

Sweden is much lower than that of companies listed in the USA. Filtering out all companies that do not fully meet our data requirements plus at best applying some additional filters would have left us with a small sample of companies over which it would have been very difficult to find any meaningful results unless our strategy was to beat the market by vast margins.

An option to increase the size of the sample would have been to allow companies with data available for less than the entire period previous to investment (the 'data-gathering' period) to enter our observations. This would have made it difficult, though, to treat companies 'fairly' and would have resulted in less clear results. The same, although to a lower degree, applies to the Nordic market. This option additionally carried the downside that we would have faced different accounting standards and several stock markets that underlie different legislations which might have introduced differences in measurement.

Eventually, we decided to use the US market. Being it the largest and most sophisticated stock market with good availability of historical data and one single underlying legislation, it seems the most suitable choice.

For all comparisons with the market in this study we use the Russell 3000 Total Return Index as a proxy for the market. Specifically, we use it at three times when implementing our strategy: First, we use it in the forecasting process of the IRR of each stock over the investment period to assess the future value of dividends. In this case we use the historical index IRR over the ten years prior to the base year. Next, we use the index returns during the investment period to carry forward all actual dividends and the exit returns of companies that are delisted to the end of the investment period. Lastly, in the performance evaluation, we use the index return as a benchmark to which we compare the returns of our portfolio to determine the relative success of our strategy.

The choice of the Russell 3000 Total Return index as our market benchmark is attributable to its broad composition and the inclusion of dividend returns. Whereas indices like the Dow Jones 30, the Standard & Poor's 500 or the NASDAQ are either much smaller in number of companies included or have a focus on certain industries and are thus less suitable to compare our strategy focusing on the broad market to, the Russell 3000 includes 3000 companies, has no sector preferences and is a better proxy of the overall US American stock market's performance. As the dividends shall not be neglected in our strategy, the choice of a total return index is the logical consequence.

As the Russell 3000 TR index excludes from the set of stocks out of which its components are selected companies that are non-domestic in the USA as well as companies that are not listed on either the New York Stock exchange, the NASDAQ or the American Stock exchange, we also exclude such stocks to allow for a better comparability.

3.3. Preparation of the portfolio selection

In a first step, we compile past accounting and trading data for each company in the US market over the eleven periods prior to investing. The raw data-points we use in our strategy are, on the accounting side, book equity, total assets, net income, dividends paid out and shares outstanding. Additionally, we deduct the return on book equity, the pay-out ratio and as a leverage indicator the book equity to total assets ratio. This data is compiled from the CapitalIQ/Compustat North America database.

For each company in the US American market, we check the availability of data for the calculation of ROE and pay-out ratio for the years t=-10 to t=-1. In this process we calculate the ROE as a year's net income over the book equity of the company at inception of that year (thus the ROE of t=-10 equals the net income of that year over the book equity per year end t=-11). The pay-out ratio we calculate as the dividend paid in a year over the same year's net income. For any year in which a company pays a dividend despite reporting a net loss the pay-out ratio is assumed to be 100%. We then compute the average ROE and pay-out ratio over the ten-year period for each company.

Companies that during any of the data-gathering years or at inception of the investment period have a negative equity book value pose a problem to our strategy. A return on a negative equity book value does not yield a meaningful result. Likewise, projecting a future equity book value by means of compounding the current book value at the historical ROE as our strategy requires is only applicable if the starting book value is positive. Therefore, we exclude all companies for which the data to obtain both ROE and the pay-out ratio is either not available or not meaningful for a full ten years.

The motivation to exclude companies with negative equity book value is purely technical and we do not expect it to significantly add to the performance of our portfolio in a systematic way. The reasons for a company to report a negative equity book value can be manifold. Besides indicating that a company is in financial distress, a negative equity book value can also point towards a company that engages heavily in share buy-backs, which, depending on the price shares are bought for, can be a highly value-enhancing practice for shareholders (Babenko et al., 2012). Research shows that another practice that can lead to negative equity book values are cumulative R&D expenses over longer periods of time (Jan and Ou, 2012).

The usage of data over ten years and the requirement for companies to have data available over the full period to be eligible to enter our portfolio, makes it impossible for companies that have not already been listed for at least ten years to enter the portfolio. While this might lead to the systematic exclusion of companies in some newer industries, we do not consider it to be problematic, since our strategy in general disregards the industries of the companies it picks. It might appear to be a problem in the setting of the new economy, though, as many of the stocks that lost most during the bubble could not have entered our portfolios. However, given that these companies barely had equity and in many cases were accumulating losses, they would not have entered our portfolios in any case – be it due to excessive leverage or too low profitability.

In terms of trading data, we limit ourselves to the split adjusted share prices and – again – shares outstanding. Trading data originates from the CRSP North America data base. Throughout the entire paper we use as the year-end price for each company not the actual year-end price but the average monthly closing price of a stock over three months. We do so to not exclusively rely on one data point and thus mitigate the effects of outlier prices. Also, by using three monthly closing prices we reduce the impacts which year-end effects as for example window-dressing could have on the stock price (Lakonishok et al., 1991). Our aim is not to test whether our strategy works specifically at year-end, but rather whether it generally can earn an outperformance over the market.

3.3.1. Exclusions of stocks for reasons not evaluating the companies' quality

From all domestic US American companies listed on the above mentioned stock exchanges and having full data availability, we further exclude those that do not have a financial year equalling the calendar year. This is purely for the reason that the inclusion in the analysis of companies that follow a different financial year than the calendar year would introduce several difficulties. Past accounting data would not be from the exact same time-span for companies in that case. In case of the outbreak of an economic crisis in the very beginning or end of the time-span analysed companies would be affected differently by this event. Also, the investment in companies would not be done at the same point in their financial year and would in some cases be based either on accounting data lying further back or being partially forecasted.

For our strategy that chooses a large number of stocks and does so based on few data inputs and taking into account data points from ten years the problems seem manageable. We opt for excluding these stocks anyway, as the sample of US listed companies is large enough and we do not expect to introduce a systematic bias by doing so. No literature seems to support the argument that stocks that do not follow the calendar year with their financial year should on average over- or underperform the market.

We also exclude stocks that are classified as financial by the CompuStat North America database. The exclusion of financial stocks follows a trade-off between the comparability of ROE for such stocks to non-financial companies, on the one side, and the bias that is introduced by the exclusion of the sector when comparing with the Russell 3000, which includes them, on the other side.

The business model of a bank or an insurance company is entirely different from that of an industrial or services company and is on average much more directly and stronger affected by a change in the interest rate environment. The rationale behind our strategy is not linked to central bank policy, though, but to the reasoning that companies with a sustained high ROE are intrinsically better than other companies and that this is not always sufficiently appreciated by markets. Performing cross-checks among the years we look at in our back-tests and looking at historical sector ROE data from the website of Damodaran (2016), supports this point. Financial companies generally have average ROEs. Yet, their low equity as a share of total assets makes their ROE very prone to outliers in years extremely favourable or negative for the financial sector.

Although there is a good reason to exclude the sector from a conceptual standpoint, the minimum equity-to-asset ratio of 30% we apply would exclude almost all financial companies in any case.

The remaining set of stocks after applying these additional exclusion criteria is the total investment universe of stocks we choose from and of which we buy the ten percent that best fulfil the criteria of our strategy.

3.3.2. Exclusion criteria comprising a judgement of companies

Whereas all criteria so far have been applied to exclude stocks from our investment universe for technical or comparability reasons, the following criteria comprise a qualitative judgement of companies and do not reduce the number of investable companies any further but rather merely prevent certain companies from entering our portfolio.

To avoid loading our portfolio heavily on companies with extreme financial leverage we incorporate a cut-off threshold based on the common financial analysis measure of equity-to-assets (E/A) into our strategy. The level we choose is a minimum requirement of 30% equity financing which we set with opposing goals in mind.

On the one hand we do not want to have a portfolio of highly levered companies as this would likely increase the odds of defaults. Also, by applying the leverage filter we exclude some companies for which ROE is a measure of little meaningfulness as they have low equity requirements and thus show highly volatile ROEs. A good example for such a company would be Rightmove plc, a company running a classifieds webpage. Their balance sheet contains very small equity and consists mainly of working capital. While this is no financial debt, it increases total assets and thus leads to a low E/A ratio and a highly volatile ROE.

On the other hand, we want to avoid the leverage ratio to be the dominant factor in our strategy. It is rather supposed to correct the pitfalls that our strategy contains when based exclusively on ROE and pay-out ratio without a qualitative evaluation of the companies, i.e. when implemented as an automated strategy. Though we cannot support the choice of 30% by literature dealing with a similar application framework for this leverage metric, from our previous work and academic experience 30% seems an adequate level to us – not too extreme in either direction.

Moreover, we exclude companies from our portfolio that during the three years preceding the base year have on average reported negative earnings. While the exclusion of such companies certainly is a criterion that excludes companies of lower quality, for us it is also one of practical nature: Since we use the net income average of these three years for the calculation of a price earnings trading multiple at inception of our investment period (which we later use as the exit trading multiple at the end of our investment period) we need to obtain a positive multiple.[2] At the same time we expect the constraint to have limited implications for the performance of our portfolio, as in most cases a company that for at least three out of ten years evaluated has a negative ROE is unlikely to make it into the portfolio.

3.4. Evaluating stocks and choosing a portfolio

For those companies that fulfil all criteria this far, we employ their equity book value at inception of the investment period (at T=0) as the starting value in our estimations of their future value. On the equity book value, we apply our historical ROE average for the respective company to find an estimate of net income for the following year (i.e. t=1). To this net income estimate we apply the historical average pay-out ratio to determine what part of earnings is paid out and what part is kept in the company. The part we assume is kept in the company is then added to the historical equity book value of the company to give an estimate of the equity book value at T=1. For the following year (i.e. t=2) this estimated book value is used as the foundation on which once again the empirical historical ROE and pay-out ratio averages are applied to come to an estimate of

[2] A negative multiple could be used but would lead to an estimated negative future value of the stock which is not sensible and would likewise exclude the stock from the portfolio.

retained earnings which is added to the previous – now estimated – equity book value at T=1. This estimation procedure is applied nine times to yield an equity book value estimate after nine years of investment (at T=9). We then apply the historical average ROE one more time to find an earnings-estimate also for the tenth year. Instead of also applying the pay-out ratio and finding a book value estimate at the end of year ten, this time we apply a P/E ratio on our earnings estimate to find a hypothetical market value of the company at the end of our investment period at T=10.

The price-to-earnings ratio we apply differs from company to company and is based on the assumption that there is no P/E multiple expansion or contraction over the ten-year investment span – i.e. that the trading multiple of each stock at inception of the investment period will be unchanged at the end of the investment period.

There are several options to estimate an exit trading multiple. One would be to use average industry trading multiples from the past. The problem with doing so is that companies that are higher quality than the average of their industry suffer a disadvantage here, while companies that are below average would likely be aided to make it into the portfolio.

Another option would be to use a past average trading multiple specific to each company in the market. By using an average past multiple we expect that in some cases we might account for a possible overpricing of a stock at inception of our investment period by assuming a multiple contraction to a historically lower multiple and likewise would account for a potentially undervalued stock by assuming a multiple expansion to return to a historically higher average trading multiple.

While we would expect this to positively affect the selection and subsequent performance of our portfolio, it would make it more difficult to assess our primary hypothesis that companies that manage to permanently compound their capital at high ROEs, over the long-term and on average, will outperform the market. A substantial part might be attributable to assumed changes in the multiples of companies when selecting the portfolio. Hence, while we are aware of the importance of the assumption we make, we consider it the most suitable in the setting of this study.

Combining accounting and trading data, we calculate the P/E trading multiple at T=0 based on the average earnings of the previous three years. The choice of using the earnings of three years instead of the current one only is made to reduce the exposure to earnings figures that are outliers for the respective company. We choose to use not more than three years, though, to keep the earnings figure recent.

In each of the ten years of the investment period dividends might be paid out. As described above, we already account for this, for each year only adding the part of net income that according to the historically observed habits of each company should be kept in the company to the book equity (estimates) in the beginning of a year.

However, thereby we neglect the value the paid out money has for the security holder. Thus, for each year we assume that the part of estimated earnings that is paid out as a dividend is carried forward at the market's internal rate of return during the ten years previous to investing. For simplification and as the effects on the overall idea of the strategy should be negligible, we only account for full years, meaning e.g. that a dividend paid during investment year nine will only be compounded during investment year ten (i.e. for one full year), as if it had been paid on the last day of year nine. This assumption corresponds to a joint annual investment per year-end of all dividends received over the year in an Exchange Traded Fund (ETF) tracking the Russell 3000 TR performance. To plough back earnings into the company paying them out is not a viable option in the forecasting process as this would require the knowledge of future stock prices at the time of investment.

Finally, both the estimated future market capitalisation at T=10 and the sum of the future values of all dividends estimated to be paid over the investment period are added together. Comparing this estimated total future value of each security with the actual market capitalisation at inception of the investment period, the expected IRR over the ten-year investment period can be calculated.

Consequentially, we rank all stocks that we have not previously excluded as investment opportunities – either due to lack of data or by applying a filter – by their IRR and buy an equally-weighted portfolio consisting of the top 10% of all stocks that enter our investment universe, as defined earlier in this section.

We could have also chosen 5% or 20% for example. The reasoning we follow here is not to dilute the potential effect of our strategy too much and converge to the average market but also not to buy too small of a portfolio. As our strategy is simplistic and based on few accounting numbers, we expect a too narrowly defined portfolio to be too susceptible to outliers and chance. We need a certain share of the stocks available in the market to test whether our strategy on average works.

3.5. Measuring the portfolio performance

Having finished the portfolio selection based on our strategy, the next step is to back-test how such a portfolio would have performed in the market and whether it is possible to observe a significant out- or underperformance versus the general market. We do the entire performance evaluation of our portfolio on a per share basis and using stock prices adjusted for share splits. Only by evaluating performance on a per share basis we can account for splits while at the same time accounting for share repurchases or other measures during the investment period that affect capital (e.g. a capital increase) in a sensible way.

Same as in the portfolio building process, dividends paid are compounded and carried forward for full years only. Whilst previously we had to use a historical market return rate for this, we can now use the actual market return rate for each specific year in the investment period. We do also account for the effects of stock splits on dividends. This means that if a stock split occurs of say 2:1 then we will not receive one but two per stock dividends from that point in time going forward.

Companies in the portfolio that are delisted over the holding period pose a major challenge. There are three main reasons why a company would be delisted. It could either be insolvent, be bought out by another company or an investor or taken private by its shareholders. If the reason for delisting is that a company is bought out, we would not want to stick in a company as minority shareholder or, in the case this happens in the setting of a merger, potentially with the shares of a different company than we initially invested in. In the case of a going-private transaction it would also make sense to sell our shareholdings as we would not want to be stuck in a company that is not traded any longer. Lastly, in the case of a company being insolvent our shareholdings would be wiped out. In all cases, it seems reasonable to assume an exit from the respective stock.

Thus, we do not distinguish among the above scenarios and make the decision when to exit a company in any of the stated situations an automated one: We exclude a stock from our portfolio when it is delisted and we assume that what we get in return for our shares is the average price of the last 20 trading days. This exit return is reinvested per respective year-end at the market rate of return for the time remaining until the end of year ten – same as we do with dividends. Stock prices and last trading dates are sourced from the CRSP North America data base.

Choosing to use the average price of the last 20 trading days of a stock could be interpreted as a divestment in steps over this time frame. In doing so, we evaluate a stock not only at one single date (i.e. the last day of trading). Thereby, we make the performance contribution to our portfolio less prone to extreme outlier prices just before the delisting of the stock. If the reason for delisting is a take-over, such shifts would in many cases be beneficial to the performance of our portfolio as prices commonly face a run-up to the take-over price before the delisting or, based on speculation, even above it (Bagnoli and Lipman, 1996). For reasons of conservatism we stick to the 20-day average, though, being aware that also this average will in many cases be already significantly higher than the stock prices a few months earlier.

Why do we not use the prices over a longer period of time then? – First, this would require that the delisting date always be known months in advance. But even if, for the sake of performance conservatism, we ignored this logical constraint in our back-test, we think that there are good arguments for including a large part of possible take-over announcement driven performance: For one thing, such performance is real performance our portfolio generates. For another, if our portfolio happened to load more than average on companies that get bought out this might well be a structural quality of our strategy. One common reason for an acquirer to go ahead with an acquisition is that he sees an opportunity to buy a good company at an interesting price.

As stated earlier, the carrying forward of dividends and delisting returns in our strategy corresponds to the assumption that all cash inflows during a year – from the perspective of the investor – are reinvested at year end in an ETF which tracks the performance of the Russell 3000 TR index. As opposed to the forecasting process, for the performance evaluation we know the actual stock prices and could assume that dividends are reinvested by the investor into the company paying them out. Delisting returns meanwhile could be invested in the remaining stocks or in stocks that a new application of the strategy would identify.

The reason for us to not do so is twofold: On the one hand, investing at each year-end into different securities would significantly increase transaction costs versus the option of investing all proceeds paid out into a single ETF. On the other hand, the aim of our strategy is to keep things as simple as possible for the potential investor. By assuming an investment in an ETF the investor essentially only has to spend time on the portfolio composition at inception of the investment period and afterwards issue one buy order for one and the same ETF at the end of each year during the investment period comprising the sum of all payments received during that year.

Finally, for all stocks that entered our portfolio, we compare the per share price at inception of the investment period to the sum of the price at the end of the period and the future values of all dividend payments made during the ten-year time span. As stated, in the case of a delisted company, we compare the initial price to the delisting price carried forward at market return to year ten plus the future values of all dividends that were paid previous to the delisting. We determine the IRR over the ten-year period for each stock, calculate the excess returns over the market during this period and compute the average for each of them to evaluate the portfolio's overall performance. Based on the aggregate numbers, we evaluate how the portfolio compares to the market.

3.6. Analysing and testing the portfolio performance

In addition to the above portfolio statistics, we calculate the standard deviation across IRRs for the portfolio and use it along with the portfolio excess IRR over the market and the number of observations in the portfolio to test for the statistical significance of the results. We do so assuming as null hypothesis that our portfolio will not beat the market and test whether we can reject this hypothesis.

To see what type of companies it is that our strategy selects and to be able to draw more specific conclusions, also with regard to the subsequent factor analyses, we look in detail at the overlap across the portfolios for the different implementations we do, the evolution of the initially equal weights of the stocks in the portfolio and the typical size and valuation (in terms of B/M) of companies both at T=0 and T=10. We also look at the industries companies belong to and whether some are more frequent than others.

Furthermore, we test the robustness of our results. We test for the effects on our results of allowing for more leverage, specifically lowering the minimum E/A threshold from 30% to 20%. Also, we consider whether buying a more concentrated portfolio of only 5% of the investable stocks would increase our returns and how the stocks that are outliers in terms of ROE, with average ROE of greater than 50%, compare to the average portfolios' returns.

In order to paint a more comprehensive picture, we also analyse our portfolios based on factor models. We do so using the monthly returns of the stocks in our portfolio, based solely on monthly price data, and aggregating the data to get a monthly price return for each portfolio. These returns are then regressed with regard to various factor model portfolios: the three- and five-factor models of Fama-French, Carhart's four-factor extension, as well as our own modified three-factor model.

The resulting coefficients from the regressions show how each of our portfolios correlates to the factor model portfolios and whether our strategy generates alpha. These factors include market beta ("Rm-rf": the correlation to overall market movements), company size in terms of market capitalisation ("SMB": Small Minus Big), valuation in the form of a B/M ratio ("HML": High Minus Low, high indicating "value" and low indicating "growth"), price momentum ("MOM": a stock that is increasing in price is expected to keep increasing), profitability ("RMW": Robust Minus Weak operating profitability), and investment pattern ("CMA": Conservative Minus

Aggressive expected growth in book value). The excess return that cannot be explained by any of the factors is referred to as alpha and is presented in an annualised form.

The coefficients that are generated for each factor as well as the alpha are then evaluated using a p-value from the regression in order to determine if those coefficients are statistically significant. A p-value of between 0 and 0.1 would be considered statistically significant, whereas higher values indicate that the results are less reliable and should not be used to base conclusions upon.

Finally, a summary statistic for each factor model analysis is produced by running a regression across all portfolios over their respective holding periods in one consolidated regression. This results in factor coefficients that are almost exclusively statistically significant and should be a good general indication on what types of companies our investment strategy selects.

The issue of stocks in our portfolios that get delisted during the holding periods leads to some discrepancies in the factor model analyses. Since the Fama-French factor portfolios are rebalanced quarterly and we do not rebalance our portfolios at all during the entire ten-year holding periods, we expect a full ten-year factor analysis of any of our portfolios to be skewed with a bias towards those companies that were not delisted during the holding period. In order to account for this, we perform factor analysis on each portfolio broken down into two five-year periods in addition to the actual ten-year period. Each portfolio therefore has two methods of evaluation: an overall analysis for the entire ten-year period and a comparative analysis for the two five-year periods. The results of the five-year analyses can be found in Appendix 3.

Worth noting is that the portfolios that we use for factor model analyses are not identical to our actual implementation portfolios. Our implementation reinvests proceeds from dividends and delistings into a Russell 3000 TR tracking ETF in order to account for the time-value of the dividends paid out and the proceeds from delistings. These Russell ETF investments have the effect of bringing the behaviours of our returns to closer replicate those of the general stock market, which would obscure the factor loadings of the stocks that are actually in our portfolios. Therefore, in the case of delistings in the factor model version of our portfolios we have chosen to simply remove those stocks from the portfolio without replacing them, effectively increasing the portfolio weightings of the remaining stocks evenly. (The factor analyses for our portfolios including index reinvestments can be found in Appendix 4.)

3.7. Potential biases or problems we do not account for

3.7.1. Share repurchases and capital increases during the ROE measurement period

Share buy-backs conducted over the ten years preceding the investment period, i.e. the years for which we gather the ROE figures and then average them, will lead to higher returns on equity for companies undertaking them. As they decrease their equity by repurchasing their own shares, they increase their likelihood of entering our portfolio via increased ROE figures. Whereas for the investment period we can account for share repurchases by adjusting dividends and evaluating the final performance on a per share basis instead of market capitalisation, there is no similarly clear way to properly account for the effects of repurchases on historical ROE rates.

One possibility would be to assume share repurchases had not taken place. This, however, would render several further assumptions necessary. A decision would have to be taken regarding the alternative usage of the funds that were actually paid out in the form of share repurchases. An option would be to assume that they were kept in the company instead. This would increase the following years equity and – all else equal – reduce that years ROE. Thus, an assumption would also have to be made about the returns that could be earned on the capital not paid out to increase the numerator of the ROE equation as well.

Another treatment of the funds would be to assume that they were paid out as dividends instead. The problem here is that this would have a major impact and lead to distortions on the pay-out ratio – which we also use as input for our strategy. This is problematic as a dividend and a share repurchase in terms of our strategy have one major difference: From the perspective of an investor who does not sell his shares to the company when the latter repurchases shares, a share repurchase is a reinvestment of the company’s earnings in the own business model and capital return generation capabilities. A dividend from the same shareholder’s perspective, though, is just capital that leaves the company.[3]

What appears to be an entirely different method is to accept that share repurchases have happened and to account for this by calculating ROE on a per share base. If computing ROE on a net earnings per share over equity per share basis, one would initially expect the profitability measure not to be affected. Notwithstanding, this does not work in practice for a simple reason: The way stock repurchases are accounted for, reducing equity by their repurchase price and not

[3] This is essentially the same reasoning that stands behind the use of the pay-out ratio in our strategy. Since we do not sell to the company when share repurchases occur during the investment period, in such cases we implicitly also treat share repurchases as reinvestments in the company’s business model and return generation qualities.

by the book value that corresponds to each share. It also seems difficult to imagine a different way of accounting for them, as the price paid is the sum that actually leaves the company.

One clear option to deal with the problem that comes to our mind is to exclude all companies that over the years preceding the investment period do any share repurchases. This seems appealing as it solves the problem at once. We do not opt for this solution as stock repurchases are common in the USA and such a decision would have significantly reduced the number of companies available for investment. Julio and Ikenberry (2004) find that "*in 1997, the total dollar amount of cash distributed as stock buybacks exceeded total dividend payments for the first time in U.S. history*". Weigand and Baker (2009) add that "*this gap continues to widen each year*" – a finding that is updated and supported by the most recent FactSet "Buyback-Quarterly" report evidencing that since 2005 in almost all quarters more than 300 out of 500 companies in the S&P500 have repurchased own shares (FactSet, 2016).

Thus we chose not to take any action against the problem, being aware that it is one. While during the periods we gather data for stock repurchases were still not as common as they are nowadays, they were most likely strong enough to lead to the inclusion of some companies in our portfolios that would have otherwise not entered them. This is probably one of the weakest spots of our investment strategy. However, in case our strategy does work and outperform the market, any bias created by this decision should work against our strategy. This is because the inclusion of companies in our portfolios that only enter due to engaging in share buy-backs would mean the inclusion of stocks that have a much less extraordinary profitability than what they appear.

As opposed to share buy-backs, capital increases should not pose a problem to our investment strategy's functioning. If a company during the period over which we estimate ROE and pay-out ratios increases its capital, this will lead to a lower ROE in the following year unless earnings increase proportionally. This effect is sensible and justified though, as for all subsequent years earnings will need to be divided among more shareholders.

3.7.2. The omission of transaction costs

Throughout our paper we omit transaction costs. While at first sight this might appear to be a drag on the performance of our portfolio that we omit, there are good reasons for omitting transaction costs. First of all, we only invest once in our portfolio and after that would only incur transactions costs when companies get delisted, triggering us to sell them, and when we reinvest delisting returns and dividends at the end of each year into an ETF tracking the Russell 3000 TR index. Therefore, transaction costs when spread over the total ten years of investment seem relatively small. Besides, it must be taken into account that any investor who wanted to replicate the Russell 3000 TR returns without being exposed to an institution's default risk (as might be the case with buying a certificate), would do so cheapest by investing into an ETF on this index. While ETFs charge relatively low fees, there is a management fee involved in such an investment. iShares, for instance, charges 0.2% per annum on their ETF tracking the Russell 3000 index.[4] An ETF needs to cover the transaction costs of rebalancing the entire portfolio each year. On the other hand, it has a size advantage and will deal at better conditions than a smaller investor implementing our strategy individually.

In line with these considerations, we expect the transaction costs incurred when implementing our strategy for anything but a small private investor's portfolio to be of manageable dimensions and do not think that they will be decisive in determining the chances of success of our strategy. Compared with an ETF instead of the actual market, we expect the cost differences for all but small portfolios to be limited. For comparability, besides transaction costs we also omit the annual management fee of the ETF. We do so both when looking at market returns and when looking at the reinvestments of dividends and delisting returns in our portfolios.

4. Empirical results

4.1. Observations and their interpretation

We implement our strategy for eight different periods. The first implementation assumes investment from 1998-2007 and the last one from 2005-2014. As the number of investable companies according to our earlier definition of such varies from period to period, so does the number of companies in our portfolio of the top ranked ten percent of stocks at inception. For the first six time frames our portfolios contain between 102 and 111 stocks; for the last two with 133 and 138, respectively, the number is larger (see Table 1, Panel A).

Looking at the IRRs estimated in the selection process of the companies that enter our portfolios, we find that the average estimated IRR across portfolios lies between a minimum of 33% (for the fourth portfolio) and a maximum of 42% (for the second portfolio). As there are some outliers that skew the picture, the median expected IRR for each portfolio is below the average one. It moves between a low of 28% for the second and third portfolio and a high of 34% for the fifth portfolio. There seems to be no relationship between these figures and the number of companies present in a portfolio. The highest expected IRR across portfolios moves in the range

[4] The iShares Russell 3000 ETF annual fee is taken from the iShares product information webpage as of per end of October 2016. It can be verified under: https://www.ishares.com/ us/products/239714/ishares-russell-3000-etf.

of 83% and 726% and the lowest – which is synonymous with the threshold of expected IRR required to enter a portfolio – moves in the much smaller range of 22% to 25%. The appearance of unrealistically high expected IRR rates can be attributed to the fact that our strategy leaves the initial valuation of companies out of account by assuming a constant trading multiple over the ten-year investment time-span and by the way it is constructed benefits companies that most recently have performed worse than over a longer historical horizon. We will discuss this aspect in detail later in this section.

Considering the two primary variables of our strategy that lead to the above stated anticipated IRRs, for the historical ten-year mean ROE of the companies in our portfolios we find that on average it ranges between 25% and 33%. The portfolio medians for historical ROE are between a low of 22% and a high of 27%. As expected, there seems to be a clear relationship to the expected IRR of portfolios, with the portfolio showing the lowest average estimated IRR also evidencing the lowest average ROE and the portfolio exhibiting the highest average expected IRR also representing the highest average historical ROE. The medians show a similar pattern.

The highest historical ROE of portfolios ranges between 93% and 452% and the lowest between 2% and 6%. As we will show in section 5, the number of companies in the portfolios that exhibit historical ROE of greater than 50% is very small. Still, for companies reporting a ROE of several hundred percent it might be questionable how suitable of a measure ROE is. However, the high ROE of these companies in many cases is owed to high past stock repurchase activity and a hence low equity book-value. Companies with exceptionally low ROE seem to be picked for the above mentioned and later explained tendency of our strategy to pick up stocks with worse three-year than ten-year historical performance in combination with the assumption of a constant trading multiple from inception to end of the investment period.

For the historical pay-out ratio over ten years, we find that across portfolios the portfolio average moves in the range of 14% to 27%. The median is much lower with a pay-out ratio of 0% to 11%. To benefit a stock in our selection process, the historical pay-out ratio should be low. Looking at the indicated numbers, this is exactly what we find in our portfolios, evidencing the importance of retaining earnings in our strategy and showing that the selection is not merely driven by the ROE criterion. The extreme portfolio values for the pay-out ratio variable lie between 107% and 320% on the upper end and are 0% for all portfolios on the lower end. This reinforces the importance of both variables – ROE and pay-out ratio – for the strategy. We observe that a high pay-out ratio does not disqualify a stock. While a low pay-out ratio alone will not lead to a stock entering our portfolios, a stock with an attractive ROE will benefit strongly from a low pay-out ratio in the selection process.

Moving on to the evaluation of the performance of our strategy, the return figures for our eight portfolios are given in Panel A of Table 1. All portfolios exhibit positive IRRs over the ten-year investment period. Our portfolios achieve annualised performances of between 3.56% and 7.98%. In seven of eight cases the median IRR of the stocks in the portfolio is lower than the average IRR, hinting at exceptional performance of a sub-set of the portfolio. In combination with this, for each time-span we also measure the single highest and lowest IRR in the respective portfolio. With the highest IRRs ranging from 27.41% to 43.77% and the lowest from -26.21% to -14.77% we observe one of the favourable characteristics of stocks: While the upside is unlimited, only 100% loss are possible (an IRR of -26.21% over ten years equates to a loss of over 95%).

More decisive for the evaluation of our strategy is the outperformance we generate over the index. Panel B of Table 1 exhibits that all our portfolios outperform the index. While the first implementation outperforms the market by an annual 3.35%, the next three do so by between 6.22% and 8.67%. Over time-spans five to seven our strategy annually beats the market by 3.8%, 2.11% and 2.29%, respectively, and for the final one we achieve 1.13% p.a. more than the market. As these numbers evidence, there clearly is a trend to increasing outperformance over the first periods which then reverses. As two consecutive portfolios are picked with only one year difference in time and nine of the investment years are identical, this seems expectable.

As with the absolute IRRs of our portfolios, the median outperformance for the first seven periods is lower than the portfolio one. Though, as Panel B shows, for all time-spans the majority of the stocks in the portfolio contribute to its outperformance over the market. While for the first implementations, which achieve higher average outperformances, more than 60% or even 70% of the observations in the respective portfolio perform better than the market, for the last three portfolios it is still close to 60%. This is interesting as it shows that the high performance of our portfolio does not rely on picking a few stocks with extremely positive performance to save a portfolio of mostly losers. It is also interesting when contrasting it to Piotroski's (2000) observation that for high B/M portfolios the outperformance over the market on average relies on only 44% of the stocks in the portfolio.

Testing our results for statistical significance as specified in the methodology section, we observe that seven out of eight implementations are statistically significant at the 95% level, six of them even at the 99% level (see Table 1, Panel C). The only one that fails to generate a statistically significant result is the last portfolio which has a t-stat of 1.52. With standard deviations ranging from 8.58 pps to 12.5 pps and the number of observations moving between 101 and 138, the statistical significance depends largely on the outperformance achieved by each portfolio. As the last portfolio is the one to least clearly beat the market, it comes as no surprise that the corresponding t-stat is the lowest.

Table 1: Empirical Back-Test Results

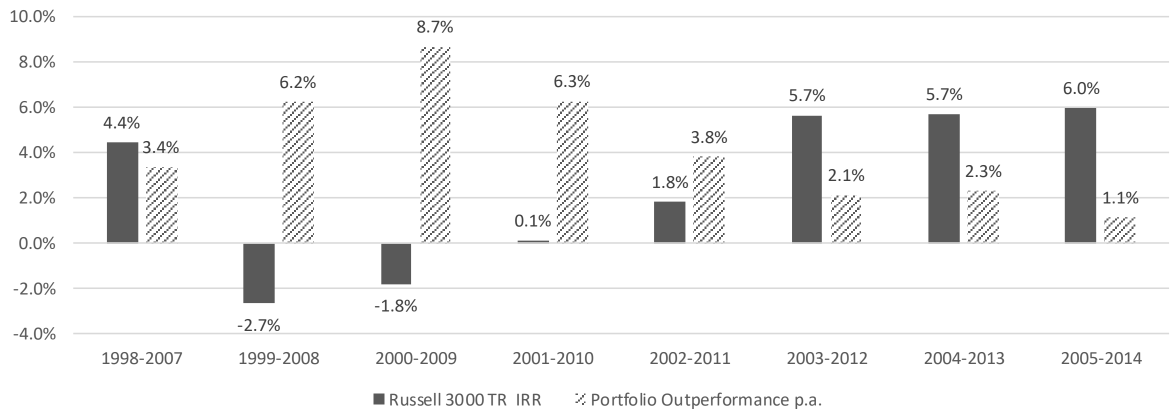
This table reports the results of the back-tests of our strategy for all eight implemented ten-year investment periods, starting between 1998 and 2005 and ending between 2007 and 2014. Panel A shows the number of companies out of which we choose our portfolio. Such are all companies listed on the New York Stock Exchange, the American Stock exchange or the NASDAQ that are domestic in the USA, are classified as non-financial by the CompuStat North America database, have a financial year equalling the calendar year and for which there is full and meaningful accounting data available for the entire 'Accounting Data Period'. Panel A also shows the number of companies comprising our portfolio and the portfolio p.a. return statistics. The median, highest and lowest IRR measures refer to individual observations in each portfolio, an observation being either a stock or a delisted stock carried forward to the end of the investment period at the index return. Panel B shows the same measures for the outperformance versus the Russell 3000 TR index and indicates what percentage of the observations in each portfolio out- and underperformed the index. Panel C shows the portfolio standard deviations, t-stats of statistical significance tests and the corresponding levels of significance. Tests for statistical significance are performed as the squareroot of the number of observations times the annualised portfolio outperformance over the market and divided by the portfolio standard deviation.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
PANEL A: Portfolio Size and Returns									
Investable Co.	1092	1067	1022	1009	1037	1106	1333	1384	1131
Invested Co. (10%)	109	107	102	101	104	111	133	138	113
Portfolio IRR	7.8%	3.6%	6.9%	6.4%	5.6%	7.8%	8.0%	7.1%	6.6%
Median IRR	5.6%	1.6%	6.5%	5.2%	4.9%	7.0%	7.4%	7.7%	5.7%
Highest IRR	39.8%	35.4%	32.8%	43.4%	27.4%	43.8%	30.5%	29.2%	35.3%
Lowest IRR	-23.1%	-19.2%	-20.8%	-26.2%	-23.0%	-18.8%	-14.8%	-17.9%	-20.5%
PANEL B: Outperformance over Russell 3000 TR									
Russell 3000 TR IRR	4.4%	-2.7%	-1.8%	0.1%	1.8%	5.7%	5.7%	6.0%	2.4%
Portfolio Outperformance p.a.	3.4%	6.2%	8.7%	6.3%	3.8%	2.1%	2.3%	1.1%	4.2%
Median Outperformance p.a.	1.2%	4.2%	8.3%	5.1%	3.0%	1.3%	1.7%	1.7%	3.3%
Highest Outperformance p.a.	35.4%	38.0%	34.6%	43.3%	25.6%	38.1%	24.9%	23.3%	32.9%
Lowest Outperformance p.a.	-27.5%	-16.5%	-19.0%	-26.3%	-24.8%	-24.4%	-20.5%	-23.9%	-22.9%
% of Co. Better than Market	61%	73%	75%	73%	66%	57%	59%	58%	65%
% of Co. Worse than Market	39%	27%	25%	27%	34%	43%	41%	42%	35%
PANEL C: Statistical Significance									
Standard Deviation (pps)	10.7%	11.1%	11.6%	12.5%	8.8%	10.0%	8.6%	8.8%	10.2%
t-stat	3.26	5.81	7.58	5.04	4.41	2.23	3.08	1.52	4.12
Significance	at 99%	at 99%	at 99%	at 99%	at 99%	at 95%	at 99%	insig.	

The comparison of our portfolios' outperformance with the corresponding annualised market return rate yields another interesting observation. The Russell 3000 TR index over two of the eight ten-year periods yields a negative return and over the two following periods, the fourth and fifth one, yields a small positive return. For the first implementation it is above 4%p.a. and for the last three it is above 5%p.a. (see Table 1, Panel B). This pattern is exactly the opposite of what we described for the outperformance of our portfolios. Figure 1 shows that for the periods the index performed best our strategy achieved the lowest outperformance and vice versa.

Figure 1: Comparison of Annualised Portfolio Outperformance and Russell 3000 TR Returns

For each ten-year investment period, this figure shows the annualised portfolio outperformance over the Russell 3000 TR index and contrasts it with the annualised Russell 3000 TR performance over the same period. It shows that there is an empirical relationship between the index performance across investment periods and the outperformance over the index that the respective portfolio yields. The input data is also exhibited in Table 1, Panel B.



If our portfolio consisted of value stocks this observation could be seen as supporting the evidence of Chan and Lakonishok (2004) that in adverse economic environments or when the overall stock market has not performed well value stocks tend to show the strongest outperformance over the market and that in a positive market they match the market return. Consequently, from the outset one would expect to see lower outperformance in the later tested periods, as in those the strong market following the global financial crisis, supported by an economic upswing but also by large quantities of cheap money, carries more weight. However, Chan and Lakonishok (2004) use high B/M ratio as the definition for value stocks in their study and thus a different metric than we do.

The primary explanation for the high outperformance of especially portfolios two to four with investment periods ending in 2008, 2009 and 2010 could instead be the exclusion of the financial sector from our investment universe. As we mentioned earlier, if not excluded per se, most financial stocks would have fallen prey to the financial leverage cut-off ratio we use – and in many cases with good reason as the severe undercapitalisation of the industry that became evident in the course of the crisis shows. In any case, a comparison of the Russell 3000 Index with the Dow Jones U.S. Financial Services Index shows that the financial sector fell by c. 65% from its pre-crisis high, whereas the overall market fell by c. 45%. While these numbers are only indicative as they depend on the specific index used for comparison, they evidence the much more severe hit of financial stocks as compared to the overall market.

Contemplating the heavy weight of the financial sector of up to 20% in indices just before the crisis, it seems not unlikely that the higher outperformance for the portfolios expiring during the three years after the crisis is entirely or largely attributable to this effect. This applies even when taking into account that a substantial share of all portfolios at the end of their investment period (on average 38%) is comprised of the Russell 3000 TR index and thus accounts for the financial sector (see section 4.3. for a detailed analysis).

A remarkable aspect of our strategy is that it beats the market, even though in the selection of stocks it ignores the valuation of a company at inception of the investment.[5] As we use the entry P/E multiple of a company to our portfolio as the assumed exit multiple for the respective company after ten years, the valuation of the company is neutralised for the strategy. While the valuation has little effect on a company's chances to be selected, disproportionately many companies with extremely high (and hardly meaningful) P/E ratios make it to the upper parts of our expected IRR based stock ranking for all time periods.

There appears to be a rational explanation for this. We do not look at companies that constantly lose money, as for the P/E multiple we use the average earnings of the previous three years and exclude all companies for which this number is negative. If a company usually is a better performer and overall has not lost money over the last three years but has had one or more bad years, the company valued on such a P/E ratio will appear very expensive. Investors valuing the company on normalised instead of current earnings would lead to a high current P/E. Such a P/E should return to a normal level in the future if the company recovers its past profitability or, if poor performance persists, as investors lose confidence in the company.

Maintaining the P/E constant in our calculation that is based on better profitability figures from the past ten years, the company will benefit more than can reasonably be expected from an assumed earnings recovery as the off-setting effect of a contracting multiple is not given. On the other hand, companies that have been poor performers for a longer period and thus are in the lower part of our stock ranking, will likely be awarded lower P/E rankings, as investors do not expect them to improve their performance in the future.

---

[5] The valuation is taken into account but only in the sense that the share price at inception of the investment period affects the IRR estimate we calculate. Over ten years, the effects on the IRR are limited, though.

Companies that have only slightly positive earnings will have a high (and volatile) P/E for any positive valuation. Such a company, depending on the performance over the previous ten years, might be highly or lowly ranked.

4.2. The impact of delisted companies on our results

Delistings of companies in our portfolio play an important role for the evaluation of our results. As Table 2 shows, from close to 30% to just over 35% of the companies in our portfolios get delisted over the ten-year holding period. This seems a high number. Yet, Diodge et al. (2015) find that in 1997 there were 7,313 U.S. companies listed in the domestic market and that between 1997 and 2012 8,327 companies were delisted, 4,957 of them due to merger activity.

This data is remarkable in two ways. Firstly, because it shows that without new listings taking place over the same period in time there would not be any publicly traded companies left in the U.S. market today. Secondly, as it shows that more than 50% of the delistings were merger related and not driven by bankruptcy (or other reasons as going private). Although the time period cited does not exactly match the years we perform our study for, in light of its results it seems that the apparently high share of companies in our portfolios that are delisted over the ten-year period should not be considered exceptionally high.

Table 2 also evidences that delisted companies in all eight of our portfolios on average leave the portfolio at a higher price than they entered it. The share of delisted companies leaving their respective portfolio with a positive absolute return is above two thirds for two portfolios and is superior to 75% for the other six portfolios.

Carrying the delisting returns forward until the portfolio liquidation as described in the methodology section, we find that for seven of eight portfolios the delisted companies on average exhibit an outperformance over the market. However, it is less pointed than the overall portfolio outperformance in all but one case. While generally the delisted companies in our portfolios add to the portfolios' positive performance and outperformance, they are a drag on the overall outperformance of the portfolios. Considering that these stocks subsequent to delisting are carried forward at market return, it does not seem surprising that the return figures of these observations after ten years are closer to the index return than those of the overall portfolio.

As many of the delisted companies get bought out, they possibly tend to be cheaper relative to other stocks and thus represent attractive acquisition targets.[6] If so, then a value investor might argue that their stock performance might have lagged behind the development of

[6] Other reasons for the acquisitions could e.g. be synergies or overpayment by empire building managers.

their intrinsic value when they were bought out and that in the future this gap would likely eventually have closed resulting in attractive returns for the shareholder. As a result of the merger, the benefit of the closing of such valuation gap would be exclusively for the acquirer instead of benefitting our portfolio.

Table 2: Delisted Companies

This table shows the initial number of companies in each portfolio, the number of companies that get delisted during the ten-year investment period (in absolute terms and as % of the initial number) and what part of the delistings in each portfolio is at a higher or a lower price than at portfolio inception. The table also shows the average ten-year IRR of all observations that start out as stock investments but are delisted and then reinvested in a Russell 3000 TR tracking index fund. For each time-frame this return rate is compared to the annualised index return over the respective period and the out- or underperformance versus the index is measured.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
Invested Co.	109	107	102	101	104	111	133	138	113
Delisted Portfolio Co.	39	39	31	29	36	36	46	43	37
Delisted Portfolio Co. (%)	36%	36%	30%	29%	35%	32%	35%	31%	33%
% of Co. Delisted at Pos. Return	72%	77%	77%	79%	78%	83%	76%	65%	76%
% of Co. Delisted at Neg. Return	28%	23%	23%	21%	22%	17%	24%	35%	24%
Average 10 year IRR delisted Co.	4.3%	1.0%	5.0%	8.2%	5.5%	5.9%	7.0%	6.3%	5.4%
Russell 3000 TR IRR	4.4%	-2.7%	-1.8%	0.1%	1.8%	5.7%	5.7%	6.0%	2.4%
Average 10 year excess IRR	-0.1%	3.7%	6.8%	8.0%	3.7%	0.3%	1.3%	0.4%	3.0%

4.3. Portfolio characteristics over time and across portfolios

Our portfolios set out from a situation of equally-weighted stock investments. As the number of stocks varies across portfolios, the initial weight per stock varies between 0.72% and 0.99% across our eight portfolios, the average being 0.9%.

Looking at the evolution of the weights of stocks and the ETF tracking the Russell 3000 TR index, we find that in the 2001-2010 portfolio the weight of the ETF in year ten is highest with a share of 44% and in the 2003-2012 portfolio it is lowest with a 29% share in the portfolio. With an average portfolio share of 38%, the ETF is not dominant in any portfolio compared to the overall share of stocks but is the largest single investment (see Table 3).

Considering that in all periods we beat the market, the fact that on average more than one third of the portfolio ends up being invested in the index represents a drag on our performance towards market returns. We have a clear motivation for not rebalancing the portfolio, but would expect that in the case that a stock gets delisted and we thus have to sell it, the performance of our strategy could be improved by 'refilling' the empty spot in our portfolio. For instance, for each year, cash-flows to the investor from all stocks delisted over the year could be reinvested equally-weighted in the same number of new stocks from the top of a then current ranking according to our strategy.[7]

With regard to the weights of individual stocks at the end of investment year ten, we see that on average 18 companies, or 23.9% of the companies left in each portfolio, account for more than 1% of the respective portfolio's overall value. The most extreme portfolio in this regard is the first one, where 22 or 31.7% of companies account for more than 1% each and the one with the lowest respective numbers is the last one with 14 or 13.7% of companies. Applying the same procedure to find companies that account for more than 2% of their portfolio at the end of year ten, we find the number of companies to be between four and seven and in every case to be below 10% of the remaining stocks in the portfolio. We find between one and four stocks per portfolio that account for more than 3% and there are between zero and two stocks that account for more than 5%. In two cases we observe that a stock ends the investment period with a weight of slightly over 10% in its portfolio. While, considering that all portfolios set out from more than 100 stocks, this is a strong weight, stocks with such weightings are extreme outliers and overall most portfolios are, except for the ETF, in large parts diversified. The clearest exception here is the portfolio ending in 2010 which at that point in time is invested to 44% in the ETF and to 19.1% in two stocks.

In any case, the degree of diversification of our portfolios is not a central issue for us. In fact, we would prefer to invest in less companies to save transaction costs but opt for the presented portfolio sizes as our strategy is simplistic and thus prone to outliers and we only expect it to work on average – not for individual companies. Therefore, from the outset it can be expected that some stocks over time will become more heavy weight in the portfolios. Also, we do not see it as a problem if part of our portfolio is delisted and has to be reinvested (other than that by reinvesting in an ETF the portfolio performance is dragged towards the index performance).

The detailed data in Table 3 underlines our point about the degree of concentration of portfolios and in addition provides the precise portfolio weights of stocks accounting for more than 3% of their portfolio at the end of year ten, also showing to how much weight measured not relative to the entire portfolio but only to the part held in stocks this corresponds.

With regard to these figures it is important to observe that depending on the portfolio stocks have a different initial weight, which makes it more or less likely for them to end up having a high share in the portfolio after ten years. Yet, while the percentage of remaining companies

[7] Disregarding for such investments stocks that are already represented in the portfolio.

that achieve high weights goes down slightly for the last portfolios, which initially are comprised of more companies, there is no clear overall trend in the data explaining closely the variations in portfolio concentration at the end of the investment period.

Table 3: Evolution of Portfolio Weights

This table shows the number of stocks in each portfolio at inception and end of the investment period and shows how much weight initially corresponds to each stock in the equally weighted portfolios. It also shows what part of the portfolio at T=10 is invested in a Russell 3000 TR ETF and what part remains invested in stocks. Furthermore, for each investment period it details how many stocks at T=10 have a weight in the portfolio of more than 1%, 2% and 3% and what percentage of the remaining stocks it is that has such weights. For each time-span there are 1-4 stocks with weights in the portfolio at T=10 of >3%. The rounded individual weights of such stocks are provided in descending order. The second figure provided for each such stock is the corresponding weight when calculated only over the part of the portfolio that remains in stocks at T=10.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
# stocks at T=0	109	107	102	101	104	111	133	138	113
# stocks at T=10	70	68	71	72	68	75	87	95	76
Initial weight in portfolio per stock	0.92%	0.93%	0.98%	0.99%	0.96%	0.90%	0.75%	0.72%	0.90%
% of portfolio in ETF at T=10	34%	36%	38%	44%	40%	29%	39%	41%	38%
% of portfolio in stocks at T=10	66%	64%	62%	56%	60%	71%	61%	59%	62%
# stocks accounting for 1%+	22	21	21	12	21	18	15	14	18
% stocks accounting for 1%+	31.7%	30.8%	29.4%	16.1%	30.7%	22.9%	16.1%	13.7%	23.9%
# stocks accounting for 2%+	7	5	7	4	7	6	7	4	6
% stocks accounting for 2%+	9.3%	7.4%	8.5%	5.4%	9.0%	7.5%	7.6%	4.2%	7.4%
# stocks accounting for 3%+	3	4	4	3	3	4	2	1	3
% stocks accounting for 3%+	4.0%	5.8%	5.2%	3.3%	4.0%	4.6%	2.1%	0.6%	3.7%
Weight of over 3% - stock #1	7%/10%	7%/11%	5%/8%	10%/18%	5%/8%	11%/15%	3%/5%	4%/6%	
Weight of over 3% - stock #2	5%/ 8%	5%/8%	4%/7%	9%/16%	4%/6%	9%/12%	3%/5%	-	
Weight of over 3% - stock #3	5%/7%	5%/8%	4%/7%	3%/6%	4%/6%	4%/5%	-	-	
Weight of over 3% - stock #4	-	5%/7%	4%/6%	-	-	3%/5%	-	-	

To analyse how much of our performance across portfolios is caused by the same stocks entering them and how much is unique to each portfolio, we look at the frequency of stocks repeating in two, three and four consecutive portfolios. We also count in how many of our portfolios each stock is represented overall.

As Table 4 shows, from the second to the eighth portfolio, we find that between a low of 46% and a high of 72% of stocks were also represented in the previous year's portfolio, the average across portfolios being 58%. For the third to the last portfolio we observe that stocks repeat portfolios three times in a row with frequencies ranging from 33% to 51% and with an average over portfolios of 40%. The respective numbers for stocks that enter four consecutive portfolios (starting with the fourth implementation) are 23% to 37% and an average of 28%.

Table 4: Stock Presence in Consecutive Portfolios

This table for each investment period specifies what percentage of the stocks in the current portfolio are also present in the previous investment period's portfolio. Similar percentage figure is specified for the share of a portfolio's companies that are present in the two and three previous portfolios.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
Stock in the previous PF		68%	72%	56%	46%	56%	49%	56%	58%
Stock in the 2 previous PFs			51%	46%	33%	38%	34%	36%	40%
Stock in the 3 previous PFs				37%	29%	25%	23%	27%	28%

Table 5 shows that out of a total of 398 different stocks that enter our portfolios, 199 or 50% have unique appearances, 78 or 20% appear twice, 37 or 9% appear three times and equally many appear four times. The remaining 12% are in five or more portfolios and the average number of portfolios that companies get into is 2.3. A total of twelve companies make it into all eight portfolios. Interestingly, these companies come from very different industries, including industrials, software development, pharma, semiconductors and consumer goods. To give the reader an idea, the most famous names are The Coca-Cola Co., Intel Corp., Schering Plough Corp., Franklin Electric Co. and Fastenal Co.

As our hypothesis sets out from the empirical evidence that companies can often maintain their capital returns for much longer than their growth figures, we are not surprised to see many companies appear repeatedly in our portfolios over several iterations. While it is remarkable if a company manages to enter eight consecutive portfolios, looking at fewer periods it clearly plays a role that the accounting data to large extent is the same, as we use data averaged over ten years as the input. Thus, for instance, for two consecutive periods the accounting data used to evaluate companies is to 90% identical. In light of this, we would rather expect more than on average just 58% of companies repeating in two consecutive portfolios.

For our strategy it is of little importance whether companies repeat in portfolios or not. However, stocks repeating in several portfolios reduce potential additional returns that could be achieved by rebalancing each portfolio regularly, as it shows that many of the companies continue to fulfil the criteria to make it into the portfolio in subsequent periods. Stocks repeating in many portfolios may also be seen as a hint that superior profitability in fact can be sustained over longer periods by some companies.

The downside of stocks repeating across portfolios is that it makes the overall evaluation of whether our strategy reliably works more difficult. As portfolios are exposed to the same companies and their individual corporate events (e.g. bankruptcies or takeovers at a specific price), the need to test the strategy over a longer period of time to gain a more reliable impression of its performance is increased.

Table 5: Overall Stock Presence in Portfolios

This table shows the frequencies at which stocks repeatedly enter the eight portfolios. For each possible frequency of appearance across portfolios (i.e. 1-8) the table returns the corresponding number of unique stocks. It is also specified what share of the overall 398 unique stocks corresponds to each frequency. The average appearance per unique stock is of 2.3 times.

Appearances across portfolios	1	2	3	4	5	6	7	8
# stocks	199	78	37	37	19	12	4	12
% of unique stocks	50%	20%	9%	9%	5%	3%	1%	3%

Looking at our portfolio compositions using the North American Industry Classification System, we observe that our strategy overweighs companies within the manufacturing categories (see Table 6). For reference, as of 30th of September 2016, a Russell 3000 ETF had a 20% weight in financials, making it the largest category (Vanguard Group, 2016). Technology was the second largest at 18%, followed by consumer discretionary, health care and producer durables. We did not expect the industry weights in our portfolios to mimic the weights in the Russell 3000 closely, especially since we excluded financial stocks from our portfolios. Yet, we did not expect an outsized share of stocks from the manufacturing sector, which across portfolios on average is higher than 50%.

We suggest that an explanation for this is that we partially select companies based on consistently high performance of ROE, which is perhaps more likely to occur among manufacturing companies that may have competitive advantages that are harder to erode than those of companies in other industries such as construction and wholesale trade. Furthermore, manufacturing companies may be less likely to suffer from price changes in ways that mining, agriculture and other companies providing commoditised goods are.

Though not expected, the fact to on average be invested more than 50% in manufacturing companies is not of concern to us. As our portfolios are comprised of a large number of companies and the manufacturing sector comprises many different product areas, we iterate our belief that the diversification of our portfolios is sufficient.

Table 6: Industries Selected

This table shows the initial weights of different industries in our portfolios, as selected by our strategy. The North American Industry Classification System and the classifications as provided by the CompuStat North America database are used to group companies into sectors.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
Agriculture, Mining, Energy	13%	8%	11%	4%	3%	8%	8%	5%	7%
Construction	4%	2%	3%	1%	1%	1%	2%	3%	2%
Manufacturing	56%	57%	53%	62%	50%	50%	46%	55%	54%
Wholesale Trade, Transport	7%	8%	7%	6%	12%	13%	13%	8%	9%
Information	4%	3%	5%	4%	6%	4%	7%	7%	5%
Real Estate, Leasing and Other	9%	13%	12%	13%	12%	9%	13%	10%	11%
Professional Services	5%	7%	6%	5%	7%	6%	6%	5%	6%
Education & Health Care	1%	2%	2%	2%	5%	4%	2%	4%	3%
Entertainment, Accommodation, Food Services	2%	1%	2%	3%	4%	5%	3%	3%	3%

Analysing the size of the companies in our portfolios, we find that many of them are small in terms of market capitalisation. Table 7 shows that at initial investment the 50th percentile of the stocks in our portfolios on average corresponds to a market capitalisation of c. \$0.5bn and the 75th percentile to a market capitalisation of about \$2.2bn. When looking only at the initial size of those stocks that are not later delisted over the holding period, the 50th percentile remains at c. \$0.5bn market capitalisation, whilst the 75th percentile increases to an average \$2.7bn.

At the end of year ten the stocks remaining in the portfolio have much higher market capitalisations than at inception, both comparing with the initial size of all companies invested in and of only those that are also present in the portfolio at the end of year ten. The average 50th percentile after ten years comes in at \$1.6bn and the 75th percentile at \$6.5bn of market value. This finding does not only apply overall, but is the same for each single portfolio. Further, in seven of eight cases the picture is the same when instead looking at the change in the mean market capitalisation of each portfolio.

While we did not plan to specifically target small companies, we see a probable explanation for the observation that many of them enter our portfolio. Our strategy focuses on companies that earn high ROE and can reinvest those returns at similarly high capital return rates. Big companies that generate exceptionally high ROE may generally be less likely than small companies to have the possibility to reinvest a high share of their earnings at sustained high capital return rates.[8] In such case, a big company would either have to increase pay-outs or invest at lower ROE rates.

The fact that companies at T=10 are bigger than at inception comes as no surprise, considering that our portfolios all earn positive returns and the companies in our portfolios generally have low pay-out ratios. It is thus to be expected that companies after ten years are clearly larger in terms of book value and therefore in most cases also in terms of market value.

The heavy weight of small capitalisation stocks in our portfolios was not planned. Yet, we see no problem in the fact that our strategy seems on aggregate to choose small capitalisation stocks. If such are the ones that best can meet the criteria we impose, they should be in the portfolios. We have to make one reservation, though. For a large investor, e.g. a pension fund seeking to employ a similar investment strategy as we do, the selection of many of the stocks in our portfolios would be impracticable due to their size. Besides regulations that may prohibit such investors to hold stocks of small size or the following limited liquidity, an investor seeking to buy into small stocks would inevitably bid up the market price.

[8] Be it due to the lower absolute sums to be invested or due to a smaller company for instance operating in a niche sector where entry for new players is difficult or unattractive but expansion is still possible.

Table 7: Size of Stocks

This Table shows the market capitalisation of the stocks in our portfolios at T=0 and T=10 in million USD. Panel A deals with market capitalisations at T=0 and Panel B does so for T=10. Panel A shows two alternative ways of looking at the portfolios. The first one includes all companies in the respective portfolio at T=0. The second one looks only at those companies that are not delisted over the holding period to provide for a better comparison with the data for T=10 presented in Panel B . For both options in Panel A and for Panel B, for each portfolio the table specifies the average, highest and lowest market capitalisation and provides the market capitalisations corresponding to the 75th, 50th and 25th percentile.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
PANEL A: Size at T=0									
Average - All	4,353	6,939	6,315	9,744	5,847	4,587	7,552	6,722	6,507
Highest - All	79,858	83,520	143,327	235,515	221,105	114,295	268,363	196,341	167,790
75th percentile - All	1,398	2,070	1,409	1,707	1,908	2,748	2,501	3,681	2,178
50th percentile - All	451	376	372	287	417	502	853	788	506
25th percentile - All	90	98	84	51	112	124	229	229	127
Lowest - All	2	2	4	3	6	5	4	13	5
Average - Remaining	5,576	9,252	7,766	12,103	7,360	5,515	10,180	8,694	8,306
Highest - Remaining	79,858	83,520	143,327	235,515	221,105	114,295	268,363	196,341	167,790
75th percentile - Remaining	2,050	2,834	1,461	2,852	1,958	3,215	3,046	4,144	2,695
50th percentile - Remaining	324	459	405	260	419	442	893	1,080	536
25th percentile - Remaining	80	95	85	74	114	126	243	281	137
Lowest - Remaining	2	2	4	3	6	5	4	13	5
PANEL B: Size at T=10									
Average - Remaining	12,473	10,410	9,957	8,029	8,050	11,535	14,413	16,024	11,361
Highest - Remaining	143,200	162,551	174,541	116,920	125,967	165,964	197,857	196,174	160,397
75th percentile - Remaining	9,955	5,092	4,544	3,810	4,048	6,273	9,909	8,217	6,481
50th percentile - Remaining	2,168	1,247	1,327	1,118	1,211	1,083	2,051	2,585	1,599
25th percentile - Remaining	398	309	330	305	260	327	763	595	411
Lowest - Remaining	10	5	7	12	9	7	17	12	10

Applying a similar analysis to our portfolios as for the market capitalisation, we find that the initial valuation of stocks in terms of B/M is on average 1.12 for the 75th, 0.63 for the 50th and 0.36 for the 25th percentile (see Table 8). The average of the portfolios' mean B/M is of 1.02. Yet, we observe that the mean B/M ratio differs strongly across portfolios, ranging from 0.64 to 1.38. The same volatility of B/M across portfolios also applies to all percentile values. Attributing more value to the median (50th percentile) than to the mean, our portfolios are rather low than high B/M, although they comprise stocks of the entire range from very low to very high B/M. The initial B/M ratios excluding stocks that are delisted over the holding period are similar.

Looking at the valuation levels at the end of year ten for the then remaining stocks, we observe for all metrics we look at, both averaged across portfolios and separately for each portfolio, lower B/M levels than at inception, with an average 75^{th} percentile of 0.82, 50^{th} percentile of 0.47, 25^{th} percentile of 0.32 and mean of 0.63. While there clearly appears to be a systematic element that causes B/M levels in our portfolios to be lower after ten years than at inception, the magnitude of the effect seems to be limited.

The observation that the stocks in our portfolios are rather low than high B/M comes as no surprise as companies that have high profitability and reinvest earnings at similar rates should be considered growth companies and, if this performance is expected to be sustainable in the future and pricing is rational, should show low B/M ratios. We would expect to see more pointed low portfolio B/M ratios if our portfolios did not comprise some recent underperformers as a consequence of our constant entry and exit multiple assumption and its picking of stocks in some cases that have had a poorer performance over the past three years than over the past ten years (see section 4.1.). Such companies should mostly have high B/M ratios and thus shift the statistics for the overall portfolio, especially the average, up.

Furthermore, we suggest that the somewhat lower B/M at T=10 is also related to the inclusion in the portfolios of most recent underperformers with high B/M ratios. Some of these companies will over time recover their historical performance and a corresponding higher valuation, leading to a reduction in their B/M ratio and to lower overall portfolio B/M figures.

Table 8: Valuation of Stocks

This Table shows the B/M ratio of the stocks in our portfolios at T=0 and T=10. Panel A deals with B/M at T=0 and Panel B does so for T=10. Panel A shows two alternative ways of looking at the portfolios. The first one includes all companies in the respective portfolio at T=0. The second one looks only at those companies that are not delisted over the holding period to provide for a better comparison with the data for T=10 presented in Panel B. For both options in Panel A and for Panel B, for each portfolio the table specifies the average, highest and lowest B/M and provides the B/M corresponding to the 75th, 50th and 25th percentile.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
PANEL A: Valuation at T=0									
Average - All	1.19	1.22	1.35	1.38	0.80	0.88	0.70	0.64	1.02
Highest - All	12.39	13.19	10.40	17.43	3.62	4.91	5.28	3.43	8.83
75th percentile - All	1.28	1.24	1.65	1.32	0.95	1.04	0.77	0.73	1.12
50th percentile - All	0.60	0.63	0.81	0.67	0.63	0.68	0.52	0.49	0.63
25th percentile - All	0.34	0.32	0.36	0.37	0.36	0.43	0.34	0.34	0.36
Lowest - All	0.03	0.04	0.08	0.05	0.08	0.08	0.06	0.09	0.06
Average - Remaining	1.52	1.53	1.57	1.49	0.81	0.94	0.74	0.64	1.15
Highest - Remaining	12.39	13.19	10.40	17.43	3.62	4.91	5.28	3.43	8.83
75th percentile - Remaining	2.01	1.96	2.16	1.40	0.91	1.07	0.77	0.73	1.38
50th percentile - Remaining	0.70	0.77	0.95	0.74	0.63	0.70	0.58	0.48	0.69
25th percentile - Remaining	0.35	0.33	0.36	0.36	0.35	0.50	0.35	0.29	0.36
Lowest - Remaining	0.03	0.05	0.08	0.05	0.08	0.19	0.06	0.09	0.08
PANEL B: Valuation at T=10									
Average - Remaining	0.53	0.80	0.76	0.70	0.69	0.62	0.46	0.47	0.63
Highest - Remaining	2.93	3.18	6.63	4.68	3.67	2.69	3.68	1.87	3.67
75th percentile - Remaining	0.77	1.10	0.93	0.89	0.87	0.81	0.60	0.61	0.82
50th percentile - Remaining	0.39	0.56	0.56	0.51	0.51	0.54	0.38	0.34	0.47
25th percentile - Remaining	0.30	0.39	0.38	0.34	0.37	0.30	0.25	0.22	0.32
Lowest - Remaining	-1.02	-1.30	-0.58	-0.13	0.02	-0.66	-0.10	-0.09	-0.48

5. Verifying the robustness of our results

As it is likely that companies with extreme high historical ROE will enter our portfolio, we test for the effect of companies that are outliers in terms of ROE on the overall performance. As explained in the methodology section, high ROE can be owed mostly to financial leverage. In such a case we would expect the highest ROE companies to perform worse than the average portfolio, as ROE for such companies has little meaningfulness. Using a maximum leverage ratio, we expect that there should not prevail such an effect and that while very high ROE companies will likely be exposed to stronger mean reversion on their ROE than other companies, on average they should yield better returns than the overall portfolio.

Testing the outperformance versus the market of stocks that show a historical ROE of over 50%, for each of our portfolios, and comparing this subsets' performance to the respective overall portfolio outperformance, we find that in seven of eight cases the highest ROE companies perform better than average (see Table 9). Although the results confirm our reasoning, the group of portfolio companies that showed ROE greater than 50% in many cases is smaller than ten. Consequently, the results are not statistically significant and could be unique to the periods we look at. Also, some of the companies in the group are the same across portfolios.[9]

Although, considering the simplicity of our strategy, we deem ten percent of the market as the cut-off level in the building of our portfolios an appropriate choice of size, we analyse whether more concentrated five percent portfolios can enhance returns. Over our eight sample periods we do not find evidence for systematic out- or underperformance of such portfolios compared to the ten percent portfolios. In five periods the five percent portfolios perform better (the first three and the last two samples) and in the remaining three the ten percent portfolios perform better (see Table 9). With shrinking portfolio size, on the whole the statistical significance of the results obtained decreases. The full results for the five percent portfolios can be found in Appendix 1.

Finally, we modify our strategy to use a minimum E/A financing of 20% instead of the previously used 30%. We want to determine how allowing for increased leverage would have affected our portfolios in the past. For all eight periods we find that the modified strategy, using lower minimum equity financing, yields better returns than the regular version. The outperformance over the market of the modified strategy is displayed in Table 9. It is between 0.14pps and 1.65pps higher than under the regular strategy and on average beats the latter with 0.75pps. The full results for the modified leverage portfolios can be found in Appendix 2.

Apparently, when allowing for liability financing of 80% instead of 70% the benefits of increased leverage still overcompensate the disadvantages. Despite most of our portfolios being exposed to the financial crisis of 2008, the outperformance of the modified version is slightly higher. While the modification increases the returns, we consider a leverage of up to 80% highly risky as there is no guarantee that what has worked in the past continues to do so in the future.

There are other factors we want to mention that affect the performance of the strategy but that we do not explicitly test for. Using the last 20 trading days as the assumed exit price of

[9] Following the methodology of this paper, this test assumes that dividends for all stocks are reinvested at market return. – We do not account for differences in the average pay-out ratio of extreme ROE companies versus the overall portfolio here.

stocks that are delisted affects portfolio returns positively compared to averaging prices over a longer period and negatively compared to a shorter period. For the reasons explained earlier, we consider the choice of 20 days appropriate. Another factor, regarding both the treatment of delisted companies and dividend payments, is that we only carry the capital flows to the investor forward as of the beginning of the following year. On average this is a conservative treatment and to the disadvantage of our portfolio. Only in cases in which the index in the meantime falls this is to the benefit of our strategy.

Table 9: Results of Robustness Tests

The table shows the results of robustness tests we perform in the form of variations to our strategy for each period. It first shows the IRR over ten years only of the subset of stocks in each portfolio that on average earned >50% ROE over the 'Accounting Data Period' and provides the IRR of the overall portfolio for comparison. Next, it gives the outperformance over the Russell 3000 TR index of a portfolio picked according to our strategy but only investing in the top 5% of investable companies and provides the base case (10% portfolio) results for comparison. Similarly, the outperformance over the index that a modified strategy allowing for a minimum E/A ratio of only 20% could earn is compared to the base case (minimum of 30% E/A). Appendices 1 and 2 present more detailed results for portfolios of decreased size or of lower minimum E/A.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
IRR stocks >50% ROE	11.2%	5.0%	9.6%	7.8%	8.0%	10.8%	10.6%	5.7%	8.6%
IRR overall portfolio	7.8%	3.6%	6.9%	6.4%	5.6%	7.8%	8.0%	7.1%	6.6%
5% PF - Outperformance p.a.	3.1%	6.2%	7.7%	7.2%	4.7%	2.1%	1.8%	0.7%	4.2%
10% PF - Outperformance p.a.	3.4%	6.2%	8.7%	6.3%	3.8%	2.1%	2.3%	1.1%	4.2%
Deviation from base case (pps p.a.)	-0.2%	0.0%	-1.0%	0.9%	0.9%	0.0%	-0.5%	-0.5%	0.0%
Min. E/A 20% - Outperf. p.a.	3.9%	7.3%	8.9%	7.9%	4.6%	3.6%	2.5%	1.3%	5.0%
Min. E/A 30% - Outperf. p.a.	3.4%	6.2%	8.7%	6.3%	3.8%	2.1%	2.3%	1.1%	4.2%
Deviation from base case (pps p.a.)	0.6%	1.0%	0.2%	1.6%	0.8%	1.5%	0.2%	0.1%	0.8%

6. Testing our results for factor loadings

As discussed in the methodology chapter, for the factor analyses we evaluate all eight portfolios both based on a full ten-year holding period as well as by splitting the holding period into two five-year periods. We do so to illustrate variations in portfolio compositions that result from delistings and to observe in which periods we potentially generate alpha and what the corresponding factor loadings are. From the outset we would expect that any factor loadings that could be expected beforehand from a strategy like ours, e.g. a loading on the profitability factor, should be stronger in the early years, as companies' characteristics may change over time and the companies we pick at T=0, after 10 years may not be related to the criteria they were initially picked for any longer. To benefit the clarity of the paper the results for the sub-periods are postponed to Appendix 3.

Another thing worth noting is that the factor model analysis presented below is not performed on the implemented versions of our portfolios. Since parts of our portfolios end up being reinvested in the Russell 3000, we aim to prevent this fact from distorting the factor loadings of our core portfolios, meaning the stocks we actually picked. The portfolios we use for factor model analysis therefore only consist of the stocks in each portfolio and have no holdings of the Russell 3000 ETF. A version of the factor model analyses with portfolios that do reinvest in the Russell 3000 index can be found in Appendix 4. Overall, they are very similar in characteristics, but show slightly toned down alphas.

Beginning with the most interesting part, our portfolios generate alpha across all variations of factor model analysis. With the exception of a few statistically insignificant results, no negative annualised alphas are observed in any of the portfolios, regardless of which model we use to evaluate them. The lowest observed annualised alpha for an entire ten-year period is the statistically insignificant -0.66%, which is found in the Fama-French five factor (FF5F) analysis of the 1998-2007 portfolio. On the other end of the range, the 2000-2009 portfolio generated an alpha of 7.65%p.a. when evaluated with the FF3F model.

For each of the different factor models, the measured market betas (Rm-rf) remain in the range of 0.84 and 1.12, which is fairly close to 1.0 and is to be expected from portfolios of the sizes that we use. Since our portfolio sizes range between 101 and 138 stocks and the portfolios comprise investments across multiple industries, our portfolio selection method leads to a fairly high degree of diversification and thus a high deviation from the overall market (in terms of volatility in the form of beta) would have been surprising.

As Appendix 3 shows, not much can be concluded from separating the portfolios into two five-year periods, except that it becomes clearer in which time periods the annualized alphas are generated. Looking across all portfolios, the second five-year period performs much better for the FF5F model. The opposite is true for all the other models, although the differences between the time periods are smaller in those cases. Looking at Appendix 4 suggests that our overall portfolios including reinvestments face a more difficult task in generating alpha in the second period, as delistings and reinvestments in the Russell 3000 ETF make the share of stocks in the portfolios smaller and drag the overall performance towards the market. The second period here – to a higher extent than is the case without reinvestments – performs worse than the first in most cases.

Looking at the FF3F model in Table 10, we find that our portfolios load up fairly heavily on both the Small Minus Big (SMB) and the High Minus Low (HML) factors. In other words, the model indicates that our investment strategy favours smaller companies (SMB) and companies with a high B/M ratio (HML). Overall, HML is the dominant factor in the first two portfolios, while SMB becomes highly dominant in the subsequent six portfolios. Both coefficients remain statistically significant in all eight portfolios, based on the fact that all p-values are below 0.1.

Across all portfolios, an annualized alpha of above four percent is generated, with a market beta close to one and fairly high but even loadings on SMB and HML. The occurrence of high SMB coefficients persists throughout all different tested variations of the factor models, while HML coefficients do so to a lesser extent.

Especially the high SMB loadings were somewhat unexpected, as our strategy is not designed to specifically select smaller companies. However, the loadings are consistent with that in the empirical analysis of our portfolios we find that many of the companies entering them are indeed very small or small companies. While there is significant growth among them over the ten-year investment period, most of them remain small capitalisation stocks even at T=10. As we explained in section 4.3., the likely reason for our portfolios' loading on small companies is that such are more likely to be able to maintain high ROE while investing and not paying out their earnings.

We did not expect the loading on the HML factor. Considering that we look for companies that earn high profits and at the same time reinvest them, our portfolio should rather consist of stocks with low B/M ratios - assuming that pricing on average is correct. The inclusion in our portfolios of some companies that have on average underperformed over the past three years compared to the past ten years, caused by our constant entry and exit P/E multiple assumption as explained in sections 4.1. and 4.3., probably leads to investment in some rather high B/M companies. Yet, the HML factor loadings are much stronger than what we would expect both theoretically and from looking at our empirical results, which do not show particularly high initial B/M ratios.

Looking at the evolution of B/M ratios in our portfolios from inception to divestment, it seems likely that on average recent poor performers in terms of profitability manage to return to their more long-term historical performance. If this was the case, we would essentially be observing what is known as mean reversion. We select stocks based on accounting data and not on prices, but prices will on average and overall follow these figures. This argumentation of linking the B/M factor to mean reversion would be in line with Fama and French (1996b) who conclude that the FF3F model covers mean reversion.

Table 10: Fama French Three-Factor Model Analysis

This table shows the Fama French three-factor model analysis on all of our portfolios when no reinvestment of dividends or delisted stocks is performed. The factors have been calculated by running regressions on the monthly price performance of our portfolios with respect to each of the factor model portfolios. Coefficients for each factor have been calculated based on the performance of each portfolio for the respective ten-year holding period. P-values were calculated for each coefficient in order to determine statistical significance (p-values below 0.1 represent a statistically significant coefficient at the 90% level). Finally, an overall column has been calculated by running factor model regressions on all of our portfolios together to show a summary of how the portfolios perform and what overall characteristics they show. Appendix 3 complements the data presented here by an analysis of the first and the second five years of the holding period separated from each other.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Overall
Alpha (annual)	1.31%	5.29%	7.65%	6.25%	4.71%	1.81%	3.02%	2.87%	4.33%
p-value	0.59	0.01	0.00	0.01	0.00	0.25	0.03	0.07	0.00
Rm-rf	0.94	0.91	0.98	1.12	1.05	1.05	1.05	1.03	1.01
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SMB	0.41	0.33	0.38	0.64	0.63	0.73	0.68	0.68	0.46
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HML	0.60	0.45	0.37	0.31	0.17	0.13	0.23	0.22	0.43
p-value	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00

The Carhart four-factor model in Table 11 adds an interesting bit of data to the FF3F model. With statistical significance in each case, the model shows that all eight portfolios load up negative coefficients for the momentum factor. Since MOM is described as the tendency for a stock to continue rising if it is going up, and vice-versa if it is going down, it appears as though our strategy in some cases leads us to invest in stocks that have recently been falling and thus have a negative price momentum.

However, the loadings are small and it is questionable how much importance can be attributed to them as momentum, as defined by the momentum factor, is a rather short term characteristic of a stock. As we do not rebalance our portfolios for ten years, it is unlikely that the stocks in our portfolios have much in common with the ones in the factor portfolios already after a minor part of our holding period. Our results for the two five-year periods support this thought. For the portfolios for which both the first and the second period are statistically significant, the first period is in all cases the one with the higher negative loadings.

One explanation for the small negative momentum loadings could be derived in connection with the previous discussion about loadings on the HML factor and mean reversion. As stated, in some cases we pick up companies that have recently been weaker performers than they used to (profitability wise). As we buy these stocks in a 'bad' moment, the price at that point in time will have fallen from earlier levels. It may well be possible that for some time these stocks on aggregate continue to fall, giving rise to negative momentum for the portfolio. Some of them – the majority as our empirical findings indicate – will eventually recover whereas others will not.

While the effects of and loadings on the MOM factor are limited and questionable, we would expect our strategy to exhibit more long-term momentum effects, meaning that stocks that over a longer period have gone up will on average and over a longer time-span continue to do so in the future. The reason such effects would not surprise us but rather be expected is that our strategy looks to identify firms of high quality that have managed to earn extraordinary equity returns and been able to sensibly reinvest them over a period of ten years. As shown earlier, high capital returns are in some cases very durable – sometimes for much longer than ten years (e.g. The Coca-Cola Co.). As the stock price should follow the performance of a company over time, we would expect such companies to have a good past long-term performance record and expect them to continue to have it for the future.

The existence of such effects, though, would generally be in contrast to the literature on mean reversion over the long-term. However, Jegadeesh and Titman (2001) and Cooper et al. (2004) show that significant return reversals exist in the medium- to long-term especially following large short-term momentum effects. As our portfolios do not seem to exhibit strong but rather small and negative momentum effects, potentially overall[10] our portfolios are less likely to show mean reversion.

Concerning SMB and HML, we do not observe great changes in any of the portfolios when adding MOM. The market beta, however, decreases for all portfolios and drops from 1.01 to 0.94 across all portfolios. Worth noting, however, is the shift in weightings from the first time period to the second, as the SMB correlation coefficient increases considerably. We cannot make sense of this finding – which we also observed for SMB in the FF3F model – as it should be expected to see the opposite. Loadings should be lower in the second period as companies in our portfolios grow over time and some stocks that were small in the first period are much less so in the second period.

[10] As mentioned, some stocks in our portfolios are short-term profitability – and thus likely also stock return – underperformers compared to their more long-term historical performance. For this part of each portfolio it is likely to observe some mean reversion.

Table 11: Carhart Four-Factor Model Analysis

This table shows the Carhart four-factor model analysis on all of our portfolios when no reinvestment of dividends or delisted stocks is performed. The factors have been calculated by running regressions on the monthly price performance of our portfolios with respect to each of the factor model portfolios. Coefficients for each factor have been calculated based on the performance of each portfolio for the respective ten-year holding period. P-values were calculated for each coefficient in order to determine statistical significance (p-values below 0.1 represent a statistically significant coefficient at the 90% level). Finally, an overall column has been calculated by running factor model regressions on all of our portfolios together to show a summary of how the portfolios perform and what overall characteristics they show. Appendix 3 complements the data presented here by an analysis of the first and the second five years of the holding period separated from each other.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Overall
Alpha (annual)	3.40%	6.30%	7.56%	6.27%	4.86%	1.92%	3.44%	3.26%	4.84%
p-value	0.14	0.00	0.00	0.00	0.00	0.19	0.01	0.03	0.00
Rm-rf	0.85	0.84	0.90	1.04	1.00	1.00	1.01	0.99	0.94
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SMB	0.46	0.37	0.42	0.63	0.66	0.75	0.71	0.69	0.49
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HML	0.56	0.42	0.35	0.34	0.16	0.08	0.16	0.15	0.40
p-value	0.00	0.00	0.00	0.00	0.00	0.15	0.00	0.01	0.00
MOM	-0.17	-0.12	-0.11	-0.13	-0.08	-0.11	-0.14	-0.13	-0.13
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

The FF5F model in Table 12 adds two factors that we would expect to have effects on our portfolios: an operating profitability factor (Robust Minus Weak, RMW) and an investment factor (Conservative Minus Aggressive, CMA). Of the two, it is the RMW factor that has the more substantial loading overall and also in each of the portfolios, as we would have expected. Yet, the loadings are relatively low and much more so than we would have expected. Noteworthy is also that in many of the cases the CMA coefficient comes out negative but statistically insignificant.

The five-factor model reiterates a propensity for our investment strategy to select small companies while maintaining a loading on companies with a high B/M ratio, indicated by the SMB and HML loadings. Analysing the HML factor loading, it shows that some part of what we ascribed to it in the FF3F model has shifted to the profitability and to some extent also the investment factor when looking at the overall factors across portfolios.

This shift overall seems reasonable for profitability, as it is one of the key variables in our selection process. The overall loading on the investment factor is more difficult to evaluate. Few of the measured portfolios have statistically significant coefficients for CMA, although those that do, show a limited but existent loading on CMA. The overall factor across portfolios is similarly low but existent.

The investment factor focuses on companies with conservative expected growth in book value, which is to some extent related to the pay-out ratio criterion in our stock selection process. Albeit the pay-out ratio being secondary to the ROE criterion, our idea of selecting companies that with low pay-out ratios reinvest their net income in the company, thus growing book value, should rather put us in the aggressively investing companies and not the conservative ones. The last five portfolios do show a negative correlation, but not in a statistically significant way. While somewhat unexpected, it could simply be the case that since we employ pay-out ratio only in combination with ROE, we do not pick the average heavily investing company and thus the behaviour of the stocks in our portfolios is not related to that of the typical aggressively investing company.

Following the same line of argument, it is possible that as we only evaluate ROE in combination with the pay-out ratio of a company, we do not pick the average company earning very high ROE. Therefore, while many of the companies in our portfolios are highly profitable, it is possible that our portfolio does not mimic the returns of the profitability factor portfolio very well.

Table 12: Fama French Five-Factor Model Analysis

This table shows the Fama French five-factor model analysis on all of our portfolios when no reinvestment of dividends or delisted stocks is performed. The factors have been calculated by running regressions on the monthly price performance of our portfolios with respect to each of the factor model portfolios. Coefficients for each factor have been calculated based on the performance of each portfolio for the respective ten-year holding period. P-values were calculated for each coefficient in order to determine statistical significance (p-values below 0.1 represent a statistically significant coefficient at the 90% level). Finally, an overall column has been calculated by running factor model regressions on all of our portfolios together to show a summary of how the portfolios perform and what overall characteristics they show. Appendix 3 complements the data presented here by an analysis of the first and the second five years of the holding period separated from each other.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Overall
Alpha (annual)	-0.66%	3.38%	5.90%	6.38%	3.87%	2.07%	2.72%	3.10%	2.56%
p-value	0.77	0.08	0.01	0.01	0.02	0.21	0.06	0.06	0.00
Rm-rf	1.04	1.01	1.06	1.12	1.09	1.04	1.06	1.03	1.08
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SMB	0.53	0.39	0.45	0.64	0.67	0.75	0.70	0.68	0.54
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HML	0.24	0.26	0.23	0.32	0.15	0.17	0.25	0.24	0.22
p-value	0.01	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00
RMW	0.19	0.28	0.22	-0.01	0.16	0.00	0.08	-0.03	0.27
p-value	0.08	0.00	0.01	0.93	0.06	0.98	0.42	0.80	0.00
CMA	0.19	0.21	0.15	-0.03	-0.04	-0.18	-0.11	-0.08	0.12
p-value	0.08	0.02	0.12	0.77	0.69	0.10	0.23	0.52	0.00

Surprised by our portfolios' relatively low and sometimes even negative (yet statistically insignificant) loadings on the RMW factor along with the abovementioned behaviour of the CMA factor and considering that the interaction between the different factors could make it more difficult to observe the effects of a subset of them, we test our empirical results for a three-factor model that excludes the SMB and HML factors as presented in Table 13. Ideally, we would have liked to exclude beta as well, since we agree with Warren Buffett's previously mentioned view that market beta is counterintuitive as a risk measure (Buffett, 1984), especially considering our very long-term approach. This turns out to not be an option, however, since the factor model would need a beta factor to account for market fluctuations, without which the RMW and CMA factors receive coefficients that are neither meaningful nor comparable with the other models.

The results of this model are not what we anticipated. Looking at the RMW factor, we find it to be positive but to a much smaller extent overall than before. Instead, in addition to the market beta increasing, the CMA factor takes on a much heavier loading.

As mentioned, the high CMA, i.e. conservative investment patterns, seem to be somewhat contradictory to our focus on finding companies with low pay-out ratios that focus on reinvesting in their respective highly profitable businesses. Again, it could be expected that with ROE being more dominant of a factor in our strategy, companies in our portfolios do actually pay out a large part of their earnings, as these high ROE companies cannot find equally high return investment opportunities. However, looking at our portfolios we do not find evidence for this with the average pay-out ratio across our portfolios ranging between a historical 14% and 27%.

Table 13: Modified Three-Factor Model Analysis

This table shows our modified three-factor model analysis on all of our portfolios when no reinvestment of dividends or delisted stocks is performed. The model comprises market beta, the profitability factor and the investment factor. The factors have been calculated by running regressions on the monthly price performance of our portfolios with respect to each of the factor model portfolios. Coefficients for each factor have been calculated based on the performance of each portfolio for the respective ten-year holding period. P-values were calculated for each coefficient in order to determine statistical significance (p-values below 0.1 represent a statistically significant coefficient at the 90% level). Finally, an overall column has been calculated by running factor model regressions on all of our portfolios together to show a summary of how the portfolios perform and what overall characteristics they show. Appendix 3 complements the data presented here by an analysis of the first and the second five years of the holding period separated from each other.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Overall
Alpha (annual)	2.12%	6.84%	10.43%	10.76%	6.17%	4.21%	4.19%	3.59%	5.35%
p-value	0.48	0.00	0.00	0.00	0.01	0.08	0.06	0.12	0.00
Rm-rf	0.99	0.99	1.07	1.26	1.22	1.19	1.23	1.17	1.17
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RMW	0.21	0.12	0.00	-0.04	0.09	-0.22	-0.13	-0.32	0.11
p-value	0.01	0.06	0.98	0.75	0.42	0.11	0.35	0.02	0.00
CMA	0.47	0.42	0.39	0.30	0.25	0.09	0.14	0.18	0.43
p-value	0.00	0.00	0.00	0.01	0.04	0.54	0.30	0.22	0.00

Overall, the results of the factor model analyses are somewhat different from what we expected. Since our focus is on profitability and high reinvestment rates, we would have expected RMW to be the most heavily loaded factor, but cannot find support for this expectation.

Given our search for high ROE companies reinvesting their earnings at attractive capital returns, we expected to see limited loadings on the HML factor as companies picked should rather be priced at a low than a high B/M ratio given that they are growing. While some of the loading can be explained by picking companies that have most recently performed less well than they generally historically have, the loading on the factor remains surprisingly high, also when comparing it to our empirical findings.

The dominance of SMB was also an unexpected finding, as we did not factor company market capitalisation into our selection process at all. The loading is in line with our empirical findings, though, and we suggest that the reason for the heavy weights of small companies in our portfolios is that they are more likely to be able to generate high ROE which they can reinvest at similar returns.

The CMA factor shows a negative correlation in a majority of the portfolios, albeit in statistically insignificant manners. The overall factor of 0.12 for CMA in the FF5F is not in line with our expectation to see, if anything, a negative relationship between our portfolios and the factor. Yet, the loading is low and we do not evaluate the investments of a company themselves but only the pay-out ratio[11] and only do so in combination with its profitability, thus making interpretations difficult.

Alphas and market betas are generally in line with our findings in the results section, given that our portfolios generate statistically significant excess returns while consisting of a sufficient quantity of stocks such that a fairly wide diversification generates a beta of around one. Worth noting with regard to the annualized alphas overall is that we cannot be certain of what effect the delistings have on our alphas when compared in two five-year periods. On the one hand, the delistings could hurt us if it is the best companies that are being delisted, but on the other hand, it could be that alpha is improved by companies getting removed from our portfolios that were set to underperform in the remaining years to T=10.

The overall meaningfulness of market beta – on which all presented factor analyses rely – in terms of determining risk for our portfolios can be questioned, since our long-term investment horizon makes the portfolios practically insensitive to short-term risks in the form of price volatility. Over a ten-year holding period our main concerns regard operational underperformance and risk of the companies in our portfolios. These factors should on average determine the stock prices after ten years.

7. Conclusion

The aim of this paper is to test our hypothesis that on average the long-term past profitability of successful companies carries predictive value for the future earnings of those companies, inspired by a suspicion that the power of compounding at high capital returns is systematically underestimated by analysts and investors. Whereas many market participants focus on the short

[11] Not all investments are financed by earnings. A company may also finance growth by capital increases or debt.

term, we expect that value and a certain predictability of returns can be created by a long-term investment strategy.

We back-test our hypothesis by means of a simplistic, profitability based, ten-year investment strategy, that does not rebalance the portfolio and primarily relies on historical ROE and pay-out ratio. The strategy makes qualitative judgements of companies, excluding those that show a E/A ratio of more than 30% and, mostly for technical reasons, companies that over the past three years had negative average earnings. Importantly, we do not apply a metric for the current valuation of companies.

For all periods, starting between 1998 and 2005, we exclusively find positive annualised returns and over all periods outperform the market. For seven of eight periods this outperformance is statistically significant at the 95% level. Moreover, we find that in all cases the outperformance is driven by the majority of stocks in the portfolio. There is evidence for assuming that delisted companies within our portfolios do not bias our results to make them appear better than they are. Allowing for more leverage, choosing a smaller portfolio and excluding companies that are outliers in terms of past ROE figures does not change the overall interpretations of our results, although allowing for 80% leverage increases performance for all portfolios.

We observe a negative correlation between the annualised index performance and outperformance of the portfolios over the periods. The fact that we achieve small outperformances when the market performs best and large outperformances over periods during which the market performs worst could be seen as evidence for the often observed finding that the performance differential of value stocks over the market is largest in adverse market environments. However, these observations are generally based on the B/M criterion and looking at the composition of our portfolios we do not find evidence for our portfolios on average comprising stocks with high B/M ratios. We propose that the explanation lies rather in our exclusion of the financial services industry, which in view of the financial crisis is strongly to the advantage of our portfolio in the periods ending 2008-2011.

Applying factor model analyses in the form of the Fama-French three- and five-factor models and the Carhart version with four factors, we generally find market betas around one. Throughout all analyses the most dominant factor is the size factor which is in line with our empirical finding that small companies make up a large part of our portfolios. While not expected, we deem it possible that big companies have more difficulties with fulfilling the criteria of our model, i.e. earning high returns on their equity while at the same time being able to reinvest them at similar returns. Less pointed so, we also find loadings on the B/M factor. The explanation for this seems more certain. Albeit not the aim of our strategy, its mechanics lead to a not dominant but disproportionate share of our portfolios consisting of stocks that had poorer average profitability over the three years prior to investment than over the ten-year period we use for estimating their future returns. These stocks in our selection process benefit more than can reasonably be expected from an assumed earnings recovery by investors. The momentum factor has slightly negative loadings which could also be related to picking up companies that most recently have performed worse than they used to. Searching the market for companies that have been high quality for a long time and which we expect to remain so, we would assume that our strategy exhibits some *long-term* momentum – the momentum factor does not account for this, though. The results for the profitability and investment factors are ambiguous. In the first periods they have positive and significant loadings but then turn negative and insignificant. The profitability factor is stronger than the investment factor which we expected as ROE is the primary metric in our strategy. In an alternative three-factor model comprising market beta, RMW and CMA the latter turns predominant, which we cannot explain, and the profitability strongly decreases and turns mostly insignificant. In any case, the degree of relation of these factors to our strategy is unclear, as we only apply profitability and the somewhat investment factor related pay-out ratio combined and not individually.

Many empirical CAPM anomalies have been detected over time. They can be narrowed down to a couple of factors that subsume all others. One of the strongest and most accepted ones is the B/M factor, often times called the 'value factor'. We would argue that the profitability of a company, especially combined with its possibilities to extend it to incremental equity, are just as much of a value factor. While there certainly are no professional value investors that leave the valuation out of sight, the profitability is an optional criterion that distinguishes the purely bargain hunting value investor from the value investor looking for good businesses at the right price. In an automated strategy we expect the B/M factor to be more decisive, but as it has been widely analysed, are more interested in looking at the combination of profitability with reinvestment opportunities detached from it. Interestingly, the type of company we aim to select should on average, if pricing is rational, be considered a growth stock by means of the HML factor.[12] However, we would argue that there are different types of growth stocks and that depending on their industry and the type of growth they pursue they are more 'sexy' to investors or less[13] and

[12] High ROE and earnings that are not paid out should lead to growth, which – assuming that this performance can be expected to continue in the future and that pricing is rational – should lead to a low B/M ratio.
[13] For instance, we find that the manufacturing sector has a heavy weight in all our portfolios. It is doubtable whether industrial products generally are the type of product that creates enthusiasm among investors.

thus achieve lower or higher valuations. As a result, we do not expect the high B/M portfolio to include only low growth and the low B/M portfolio to include only high growth stocks.

As becomes evident from our argumentation, we are highly sceptical of market efficiency. The risk explanation seems counter-intuitive to us when based on market beta or in light of the findings that high B/M stocks often perform better in adverse markets than the average market. As shown in the literature section, at the same time there are appealing behavioural arguments for a mispricing-based explanation. Chan et al. (2003) provide empirical support for the hypothesis that investors are irrationally influenced by past growth which they extrapolate into the future. Chan and Lakonishok (2004) show that this evidence aligns with La Porta et al. (1997) who measure outperformance of value stocks (high B/M) and underperformance of growth stocks (low B/M) around subsequent earnings announcement dates, indicating that the former seem to meet or exceed expectations and the latter seem to fail to do so. Considering that most analysts write reports about the stock price within one year, it seems likely that at times factors that are crucial for the long-term prospects of a company, as its profitability and ability to sustainably reinvest at high capital returns, as well as potential long-term non-rebalancing compounding benefits for the investor get out of sight, are misinterpreted or misvalued. Another behavioural argument is based on principal-agent theory, hinting that it might be easier and more appealing to analysts to recommend recent high growth companies in attractive industries (e.g. Chan and Lakonishok, 2004).

We also consider comprehensible the impression of Guay (2000) and others that the discussion regarding the EMH appears to be influenced by strong priors held by parts of the academic community. Yet, we agree with Fama and French (2006, 2008) that the empirical findings in themselves cannot ultimately rule out a risk-based explanation as risk might exist in a form that has not yet been uncovered.

Being bound by the data availability for the Russell 3000 TR index, we could not look at earlier time-spans. However, given our supposition that part of our outperformance is attributable to the exclusion of the financial services sector and that there is a substantial overlap of stocks across our portfolios, we would welcome further tests of our strategy for the past, using a broad index with a longer history, or for future periods. Alternatively, it would be interesting to see the strategy implemented for the same periods we looked at, but using a broad index that excludes the financial services industry. Furthermore, a test of the strategy in combination with a valuation metric, as e.g. B/M, would be appealing.

References

Babenko, I., Tserlukevich, Y. and Vedrashko, A. (2012). The Credibility of Open Market Share Repurchase Signaling. *Journal of Finance and Quantitative Analysis*, 47(5), 1059-1088.

Bagnoli, M., Lipman, B.L. (1996). Stock price manipulation through takeover bids. *The RAND Journal of Economics*, 27(1), 124-147.

Ball, R. (1992). The earnings-price anomaly. *Journal of Accounting and Economics*, 15(2-3), 319-345.

Banz, R.W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1), 3-18.

Bartov, E. and Kim, M. (2004). Risk, Mispricing, and Value Investing. *Review of Quantitative Finance and Accounting*, 23(4), 353-376.

Bernard, V.L. and Thomas, J.K. (1989). Post-earnings-announcement drift: delayed price response or risk premium?. *Journal of Accounting Research*, 27, 1-36.

Bernard, V.L., Thomas, J.K. and Wahlen, J. (1995). Accounting-based stock price anomalies: Separating market inefficiencies from research design flaws. *Working paper, University of Michigan*.

Bondt, W.F. and Thaler, R. (1985). Does the stock market overreact?. *The Journal of Finance*, 40(3), 793-805.

Buffett, M. and Clark, D. (1999). Buffettology: The Previously Unexplained Techniques That Have Made Warren Buffett the World's Most Famous Investor. *Fireside*, New York.

Buffett, W.E. (1984). The Superinvestors of Graham-and Doddsville. *Hermes: The Columbia Business School Magazine*, New York.

Carhart, M.M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance,* 52(1), 57-82.

Chan, L.K., Jegadeesh, N. and Lakonishok, J. (1996). Momentum strategies. *The Journal of Finance*, 51(5), 1681-1713.

Chan, L.K., Karceski, J. and Lakonishok, J. (2003). The level and persistence of growth rates. *The Journal of Finance*, 58(2), 643-684.

Chan, L.K. and Lakonishok, J. (2004). Value and Growth Investing: Review and Update. *Financial Analysts Journal*, 60(1), 71-86.

Cohen, R.B., Gompers, P.A. and Vuolteenaho, T. (2002). Who underreacts to cash-flow news? Evidence from trading between individuals and institutions. *Journal of Financial Economics*, 66(2), 409-462.

Cooper, M.J., Gutierrez Jr., R.C. and Hameed, A. (2004). Market States and Momentum. *The Journal of Finance*, 59(3), 1345-1365.

Damodaran, A. (2016). Price and Value to Book Ratios and ROE by Industry Sector. *http://pages.stern.nyu.edu/~adamodar/New_Home_Page/dataarchived.html*.

Daniel, K. and Titman, S. (2006). Market reactions to tangible and intangible information. *The Journal of Finance*, 61(4), 1605-1643.

Doidge, C., Karolyi, G.A. and Stulz, R.M. (2015). The U.S. listing gap. *NBER Working Paper 21181*.

FactSet (2016). FactSet Buyback Quarterly: September 20, 2016. *http://www.factset.com/websitefiles/PDFs/buyback*.

Fairfield, P.M., Whisenant, J.S. and Yohn, T.L. (2003). Accrued earnings and growth: Implications for future profitability and market mispricing. *The Accounting Review*, 78(1), 353-371.

Fama, E.F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25(2), 383-417.

Fama, E.F. and French, K.R. (1988). Permanent and temporary components of stock prices. *The Journal of Political Economy*, 96(2), 246-273.

Fama, E.F. and French, K.R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2), 427-465.

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

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

Fama, E.F. and French, K.R. (1996b). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1), 55-84.

Fama, E.F. and French, K.R. (1998). Value versus growth: The international evidence. *The Journal of Finance*, 53(6), 1975-1999.

Fama, E.F. and French, K.R. (2004). The Capital Asset Pricing Model: Theory and Evidence. *The Journal of Economic Perspectives*, 18(3), 25-46.

Fama, E.F. and French, K.R. (2006). Profitability, investment and average returns. *Journal of Financial Economics*, 82(3), 491-518.

Fama, E.F. and French, K.R. (2008). Dissecting anomalies. *The Journal of Finance*, 63(4), 1653-1678.

Fama, E.F. and French, K.R. (2015a). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22.

Fama, E.F. and French, K.R. (2015b). International tests of a five-factor asset pricing model. *Fama-Miller Working Paper*.

Flannery, M.J. and Rangan, K.P. (2006). Partial adjustment toward target capital structures. *Journal of Financial Economics*, 79(3), 469-506.

Greenblatt, J. (2005). The Little Book that Beats the Market. *Wiley*, Hoboken.

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

Guay, W. (2000). Discussion of Value Investing: The use of historical financial statement information to separate winners from losers. *Journal of Accounting Research*, 38, 43-51.

Haugen, R.A. and Baker, N.L. (1996). Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, 41(3), 401-439.

Hou, K., Xiong, W. and Peng, L. (2009). A tale of two anomalies: The implications of investor attention for price and earnings momentum. *SSRN Paper 976394*.

Jan, C.-L. and Ou, J.A. (2012). Negative-Book-Value Firms and Their Valuation. *Accounting Horizons*, 26(1), 91-110.

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

Jegadeesh, N. and Titman, S. (2001). Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. *The Journal of Finance*, 56(2), 699-720.

Julio, B. and Ikenberry, D. (2004). Reappearing Dividends. *Journal of Applied Corporate Finance*, 16(4), 89-100.

Keim, D.B. (1983). Size-related anomalies and stock return seasonality: Further empirical evidence. *Journal of Financial Economics*, 12(1), 13-32.

Koller, T., Goedhart, M. and Wessels, D. (2015). Valuation: Measuring and Managing the Value of Companies, 6th Edition. *Wiley*, Hoboken.

La Porta, R., Lakonishok, J., Shleifer, A. and Vishny, R. (1997). Good news for value stocks: Further evidence on market efficiency. *The Journal of Finance*, 52(2), 859-874.

Lakonishok, J., Shleifer, A., Thaler, R. and Vishny, R. (1991). Window Dressing by Pension Fund Managers. *NBER Working Paper 3617*.

Lakonishok, J., Shleifer, A. and Vishny, R.W. (1994). Contrarian investment, extrapolation, and risk. *The Journal of Finance*, 49(5), 1541-1578.

Latane, H.A. and Jones, C.P. (1979). Standardized unexpected earnings 1971-1977. *The Journal of Finance*, 34(3), 717-724.

Lee, C. and Swaminathan, B. (2000a). Do stock prices overreact to earnings news?. *Working paper, Cornell University*.

Lee, C. and Swaminathan, B. (2000b). Price momentum and trading volume. *The Journal of Finance*, 55(5), 2017-2069.

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

Lo, A.W. and MacKinlay, A.C. (1999). A Non-Random Walk Down Wall Street. *Princeton University Press*, Princeton.

Malkiel, B.G. (2003). The Efficient Market Hypothesis and Its Critics. *Journal of Economic Perspectives*, 17(1), 59-82.

Myers, J.N., Myers, L.A. and Skinner, D.J. (2007). Earnings momentum and earnings management. *Journal of Accounting, Auditing & Finance*, 22(2), 249-284.

Odean, T. (1999). Do investors trade too much?. *American Economic Review*, 89(5), 1279-298.

Ozkan, A. (2001). Determinants of Capital Structure and Adjustment to Long Run Target: Evidence from UK Company Panel Data. *Journal of Business Finance & Accounting*, 28(1-2), 175-198.

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

Pontiff, J. and Woodgate, A. (2008). Share issuance and cross-sectional returns. *The Journal of Finance*, 63(2), 921-945.

Poterba, J.M. and Summers, L.H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics*, 22(1), 27-59.

Rosenberg, B., Reid, K. and Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, 11(3), 9-16.

Schwert, G.W. (2001). Anomalies and Market Efficiency, Handbook of the Economics of Finance. *North Holland*, Amsterdam.

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

Siegel, J.J. (2006). The 'Noisy Market' Hypothesis. *Wall Street Journal*, June 14.

Sloan, R. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings?. *The Accounting Review*, 71(3), 289-315.

Stattman, D. (1980). Book values and stock returns. *The Chicago MBA: A journal of selected papers*, 4(1), 25-45.

Titman, S., Wei, K.J. and Xie, F. (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis*, 39(4), 677-700.

Vanguard Group (2016). Vanguard Russell 3000 ETF Product Overview. *https://institutional.vanguard.com/VGApp/iip/site/institutional/investments/productoverview?fundId=3354.*

Weigand, R.A. and Baker, H.K. (2009). Changing Perspectives on distribution policy: The evolution from dividends to share repurchase. *Managerial Finance*, 35(6), 479-492.

Appendices

Appendix 1

Appendix 1 provides the detailed results for modifying our strategy to buy the upper five instead of ten percent of companies in our ranking of investable companies. Otherwise all aspects of our strategy remain unchanged.

Appendix 1 - Table 1: Empirical Back-Test Results (5% portfolio)

This table reports the results of the back-tests of our strategy for all eight implemented ten-year investment periods, starting between 1998 and 2005 and ending between 2007 and 2014. Panel A shows the number of companies out of which we choose our portfolio. Such are all companies listed on the New York Stock Exchange, the American Stock exchange or the NASDAQ that are domestic in the USA, are classified as non-financial by the CompuStat North America database, have a financial year equalling the calendar year and for which there is full and meaningful accounting data available for the entire 'Accounting Data Period'. Panel A also shows the number of companies comprising our portfolio and the portfolio p.a. return statistics. The median, highest and lowest IRR measures refer to individual observations in each portfolio, an observation being either a stock or a delisted stock carried forward to the end of the investment period at the index return. Panel B shows the same measures for the outperformance versus the Russell 3000 TR index and indicates what percentage of the observations in each portfolio out- and underperformed the index. Panel C shows the portfolio standard deviations, t-stats of statistical significance tests and the corresponding levels of significance. Tests for statistical significance are performed as the squareroot of the number of observations times the annualised portfolio outperformance over the market and divided by the portfolio standard deviation. Due to the lower number of observations in each portfolio, for several of the investment time-spans the statistical significance of the results decreases compared to the base case of our strategy investing in ten percent of the available companies.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
PANEL A: Portfolio Size and Returns									
Investable Co.	1092	1067	1022	1009	1037	1106	1333	1384	1131
Invested Co. (5%)	55	54	51	51	52	56	67	69	57
Portfolio IRR	7.6%	3.5%	5.9%	7.3%	6.5%	7.8%	7.5%	6.6%	6.6%
Median IRR	5.5%	1.1%	5.8%	4.3%	6.5%	7.3%	7.2%	7.7%	5.7%
Highest IRR	34.0%	29.5%	32.8%	43.4%	23.8%	40.8%	26.7%	29.2%	32.5%
Lowest IRR	-23.1%	-18.7%	-20.8%	-20.0%	-12.3%	-15.2%	-9.6%	-17.9%	-17.2%
PANEL B: Outperformance over Russell 3000 TR									
Russell 3000 TR IRR	4.4%	-2.7%	-1.8%	0.1%	1.8%	5.7%	5.7%	6.0%	2.4%
Portfolio Outperformance p.a.	3.1%	6.2%	7.7%	7.2%	4.7%	2.1%	1.8%	0.7%	4.2%
Median Outperformance p.a.	1.1%	3.7%	7.6%	4.1%	4.7%	1.6%	1.5%	1.8%	3.3%
Highest Outperformance p.a.	29.6%	32.2%	34.6%	43.3%	22.0%	35.2%	21.0%	23.3%	30.1%
Lowest Outperformance p.a.	-27.5%	-16.1%	-19.0%	-20.2%	-14.1%	-20.9%	-15.3%	-23.9%	-19.6%
% of Co. Better than Market	56%	72%	67%	67%	67%	59%	57%	58%	63%
% of Co. Worse than Market	44%	28%	33%	33%	33%	41%	43%	42%	37%
PANEL C: Statistical Significance									
Standard Deviation (pps)	10.7%	10.9%	13.2%	15.0%	8.6%	10.3%	7.8%	9.0%	10.7%
t-stat	2.17	4.19	4.16	3.42	3.92	1.55	1.91	0.63	2.74
Significance	at 95%	at 99%	at 99%	at 99%	at 99%	insig.	at 90%	insig.	

Appendix 1 - Table 2: Delisted Companies (5% portfolio)

This table shows the initial number of companies in each portfolio, the number of companies that get delisted during the ten-year investment period (in absolute terms and as % of the initial number) and what part of the delistings in each portfolio is at a higher or a lower price than at portfolio inception. The table also shows the average ten-year IRR of all observations that start out as stock investments but are delisted and then reinvested in a Russell 3000 TR tracking index fund. For each time-frame this return rate is compared to the annualised index return over the respective period and the out- or underperformance versus the index is measured.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
Invested Co.	55	54	51	51	52	56	67	69	57
Delisted Portfolio Co.	21	22	15	15	19	20	25	22	20
Delisted Portfolio Co. (%)	38%	41%	29%	29%	37%	36%	37%	32%	35%
% of Co. Delisted at Pos. Return	81%	82%	73%	87%	79%	85%	76%	68%	79%
% of Co. Delisted at Neg. Return	19%	18%	27%	13%	21%	15%	24%	32%	21%
Average 10 year IRR delisted Co.	3.6%	1.4%	4.9%	11.4%	7.3%	6.7%	7.2%	7.2%	6.2%
Russell 3000 TR IRR	4.4%	-2.7%	-1.8%	0.1%	1.8%	5.7%	5.7%	6.0%	2.4%
Average 10 year excess IRR	-0.9%	4.1%	6.7%	11.3%	5.5%	1.1%	1.5%	1.2%	3.8%

Appendix 1 - Table 3: Evolution of portfolio weights (5% portfolio)

This table shows the number of stocks in each portfolio at inception and end of the investment period and shows how much weight initially corresponds to each stock in the equally weighted portfolios. It also shows what part of the portfolio at T=10 is invested in a Russell 3000 TR ETF and what part remains invested in stocks. Furthermore, for each investment period it details how many stocks at T=10 have a weight in the portfolio of more than 1%, 2% and 3% and what percentage of the remaining stocks it is that has such weights. Compared to the 'normal' version of our strategy (buying ten percent of the available companies in the market) in the five percent case each of the companies in a portfolio has a higher initial weight and is more likely to have a high weight at T=10.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
# stocks at T=0	55	54	51	51	52	56	67	69	57
# stocks at T=10	34	32	36	36	33	36	42	47	37
Initial weight in portfolio per stock	1.82%	1.85%	1.96%	1.96%	1.92%	1.79%	1.49%	1.45%	1.78%
% of portfolio in ETF at T=10	30%	39%	36%	45%	46%	33%	43%	42%	39%
% of portfolio in stocks at T=10	70%	61%	64%	55%	54%	67%	57%	58%	61%
# stocks accounting for 1%+	19	14	17	12	20	16	22	22	18
% stocks accounting for 1%+	57.1%	44.5%	45.4%	32.5%	59.6%	44.2%	52.3%	45.2%	47.6%
# stocks accounting for 2%+	13	11	11	4	12	10	8	7	10
% stocks accounting for 2%+	37.9%	33.4%	27.8%	10.9%	34.7%	25.1%	17.8%	13.4%	25.1%
# stocks accounting for 3%+	6	6	6	4	3	6	6	3	5
% stocks accounting for 3%+	18.8%	20.1%	13.9%	8.4%	8.6%	16.0%	12.5%	4.8%	12.9%

Appendix 1 - Table 4: Stock presence in consecutive portfolios (5% portfolio)

This table for each investment period specifies what percentage of the stocks in the current portfolio are also present in the previous investment period's portfolio. Similar percentage figure is specified for the share of a portfolio's companies that are present in the two and three previous portfolios.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
Stock in the previous PF		67%	61%	57%	37%	46%	34%	41%	49%
Stock in the 2 previous PFs			41%	39%	25%	27%	24%	14%	28%
Stock in the 3 previous PFs				25%	19%	16%	15%	10%	17%

Appendix 1 - Table 5: Overall stock presence in portfolios (5% portfolio)

This table shows the frequencies at which stocks repeatedly enter the eight portfolios. For each possible frequency of appearance across portfolios (i.e. 1-8) the table returns the corresponding number of unique stocks. It is also specified what share of the overall 246 unique stocks corresponds to each frequency. The average appearance per unique stock is of 1.8 times.

Appearances across portfolios	1	2	3	4	5	6	7	8
# stocks	151	44	18	17	8	3	4	1
% of unique stocks	61%	18%	7%	7%	3%	1%	2%	0%

Appendix 1 - Table 6: Size of stocks (5% portfolio)

This Table shows the market capitalisation of the stocks in our portfolios at T=0 and T=10 in million USD. Panel A deals with market capitalisations at T=0 and Panel B does so for T=10. Panel A shows two alternative ways of looking at the portfolios. The first one includes all companies in the respective portfolio at T=0. The second one looks only at those companies that are not delisted over the holding period to provide for a better comparison with the data for T=10 presented in Panel B . For both options in Panel A and for Panel B, for each portfolio the table specifies the average, highest and lowest market capitalisation and provides the market capitalisations corresponding to the 75th, 50th and 25th percentile.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
PANEL A: Size at T=0									
Average - All	4,594	6,574	10,137	13,585	8,839	5,362	11,210	4,421	8,090
Highest - All	79,858	83,160	143,327	235,515	221,105	114,295	268,363	78,301	152,990
75th percentile - All	1,619	2,044	1,580	3,414	2,338	1,197	2,378	3,232	2,225
50th percentile - All	639	390	166	211	442	395	878	771	486
25th percentile - All	68	54	44	47	111	69	301	267	120
Lowest - All	5	4	4	3	11	5	5	13	6
Average - Remaining	5,277	7,940	12,251	16,686	11,047	6,702	16,443	5,034	10,173
Highest - Remaining	79,858	83,160	143,327	235,515	221,105	114,295	268,363	78,301	152,990
75th percentile - Remaining	3,046	1,333	2,470	4,707	2,034	975	2,991	3,662	2,652
50th percentile - Remaining	219	252	185	221	435	329	912	1,101	457
25th percentile - Remaining	62	54	43	47	118	69	348	369	139
Lowest - Remaining	5	4	4	3	11	5	7	13	7
PANEL B: Size at T=10									
Average - Remaining	12,855	8,383	11,097	9,756	10,816	10,244	15,615	9,596	11,045
Highest - Remaining	143,200	76,681	108,619	116,920	125,967	100,899	197,857	120,868	123,876
75th percentile - Remaining	8,286	4,904	5,413	4,279	4,702	7,321	8,270	5,906	6,135
50th percentile - Remaining	2,374	1,247	1,283	984	1,246	1,025	1,618	1,880	1,457
25th percentile - Remaining	398	168	286	148	256	281	792	620	369
Lowest - Remaining	10	5	7	12	9	7	22	17	11

Appendix 1 - Table 7: Valuation of stocks (5% portfolio)

This Table shows the B/M ratio of the stocks in our portfolios at T=0 and T=10. Panel A deals with B/M at T=0 and Panel B does so for T=10. Panel A shows two alternative ways of looking at the portfolios. The first one includes all companies in the respective portfolio at T=0. The second one looks only at those companies that are not delisted over the holding period to provide for a better comparison with the data for T=10 presented in Panel B. For both options in Panel A and for Panel B, for each portfolio the table specifies the average, highest and lowest B/M and provides the B/M corresponding to the 75th, 50th and 25th percentile.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
PANEL A: Valuation at T=0									
Average - All	1.14	1.17	1.39	1.64	0.84	0.97	0.65	0.58	1.05
Highest - All	6.28	7.90	10.40	17.43	3.62	4.91	5.28	1.84	7.21
75th percentile - All	1.26	1.34	1.77	1.36	1.01	1.19	0.70	0.72	1.17
50th percentile - All	0.54	0.66	0.70	0.67	0.67	0.72	0.47	0.49	0.61
25th percentile - All	0.32	0.34	0.24	0.30	0.34	0.40	0.32	0.31	0.32
Lowest - All	0.03	0.04	0.08	0.05	0.08	0.08	0.12	0.09	0.07
Average - Remaining	1.52	1.59	1.61	1.69	0.83	1.02	0.64	0.52	1.18
Highest - Remaining	6.28	7.90	10.40	17.43	3.62	4.91	5.28	1.84	7.21
75th percentile - Remaining	2.29	2.03	2.10	1.19	0.90	1.15	0.62	0.67	1.37
50th percentile - Remaining	0.79	1.08	0.61	0.64	0.62	0.72	0.45	0.46	0.67
25th percentile - Remaining	0.35	0.48	0.29	0.18	0.34	0.46	0.30	0.27	0.33
Lowest - Remaining	0.03	0.05	0.08	0.05	0.08	0.19	0.12	0.09	0.09
PANEL B: Valuation at T=10									
Average - Remaining	0.59	0.88	0.68	0.56	0.64	0.65	0.50	0.47	0.62
Highest - Remaining	2.93	3.18	1.75	1.58	2.16	2.69	3.68	1.25	2.40
75th percentile - Remaining	0.79	1.22	0.94	0.80	0.74	0.97	0.71	0.74	0.86
50th percentile - Remaining	0.46	0.63	0.57	0.43	0.48	0.51	0.37	0.35	0.47
25th percentile - Remaining	0.31	0.41	0.40	0.33	0.36	0.28	0.25	0.24	0.32
Lowest - Remaining	-0.04	0.10	0.07	-0.13	0.11	-0.66	0.02	0.03	-0.06

Appendix 2

Appendix 2 provides the detailed results for modifying our strategy to exclude companies in the process that do not fulfil the requirement of a E/A ratio of at least 20% instead of the 30% used in the base scenario. All other aspects of our strategy remain unchanged.

Appendix 2 - Table 1: Empirical Back-Test Results (min. E/A of 20%)

This table reports the results of the back-tests of our strategy for all eight implemented ten-year investment periods, starting between 1998 and 2005 and ending between 2007 and 2014. Panel A shows the number of companies out of which we choose our portfolio. Such are all companies listed on the New York Stock Exchange, the American Stock exchange or the NASDAQ that are domestic in the USA, are classified as non-financial by the CompuStat North America database, have a financial year equalling the calendar year and for which there is full and meaningful accounting data available for the entire 'Accounting Data Period'. Panel A also shows the number of companies comprising our portfolio and the portfolio p.a. return statistics. The median, highest and lowest IRR measures refer to individual observations in each portfolio, an observation being either a stock or a delisted stock carried forward to the end of the investment period at the index return. Panel B shows the same measures for the outperformance versus the Russell 3000 TR index and indicates what percentage of the observations in each portfolio out- and underperformed the index. Panel C shows the portfolio standard deviations, t-stats of statistical significance tests and the corresponding levels of significance. Tests for statistical significance are performed as the squareroot of the number of observations times the annualised portfolio outperformance over the market and divided by the portfolio standard deviation. It can be observed that overall results improve as compared to the base case of our strategy using a higher minimum E/A ratio of 30%.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
PANEL A: Portfolio Size and Returns									
Investable Co.	1092	1067	1022	1009	1037	1106	1333	1384	1131
Invested Co. (10%)	109	107	102	101	104	111	133	138	113
Portfolio IRR	8.4%	4.6%	7.1%	8.0%	6.4%	9.2%	8.1%	7.2%	7.4%
Median IRR	5.8%	3.2%	6.3%	6.7%	5.1%	7.5%	7.7%	7.7%	6.3%
Highest IRR	39.8%	42.6%	32.8%	43.4%	45.4%	40.8%	30.5%	29.2%	38.1%
Lowest IRR	-23.1%	-19.2%	-20.8%	-20.0%	-23.0%	-15.2%	-14.8%	-28.6%	-20.6%
PANEL B: Outperformance over Russell 3000 TR									
Russell 3000 TR IRR	4.4%	-2.7%	-1.8%	0.1%	1.8%	5.7%	5.7%	6.0%	2.4%
Portfolio Outperformance p.a.	3.9%	7.3%	8.9%	7.9%	4.6%	3.6%	2.5%	1.3%	5.0%
Median Outperformance p.a.	1.4%	5.8%	8.1%	6.6%	3.3%	1.8%	2.0%	1.8%	3.9%
Highest Outperformance p.a.	35.4%	45.3%	34.6%	43.3%	43.5%	35.2%	24.9%	23.3%	35.7%
Lowest Outperformance p.a.	-27.5%	-16.5%	-19.0%	-20.2%	-24.8%	-20.9%	-20.5%	-34.6%	-23.0%
% of Co. Better than Market	61%	77%	74%	75%	68%	60%	62%	61%	67%
% of Co. Worse than Market	39%	23%	26%	25%	32%	40%	38%	39%	33%
PANEL C: Statistical Significance									
Standard Deviation (pps)	11.5%	11.0%	11.9%	12.7%	9.9%	10.3%	8.1%	9.2%	10.6%
t-stat	3.55	6.81	7.57	6.29	4.75	3.64	3.49	1.63	4.72
Significance	at 99%	at 99%	at 99%	at 99%	at 99%	at 99%	at 99%	insig.	

Appendix 2 - Table 2: Delisted Companies (min. E/A of 20%)

This table shows the initial number of companies in each portfolio, the number of companies that get delisted during the ten-year investment period (in absolute terms and as % of the initial number) and what part of the delistings in each portfolio is at a higher or a lower price than at portfolio inception. The table also shows the average ten-year IRR of all observations that start out as stock investments but are delisted and then reinvested in a Russell 3000 TR tracking index fund. For each time-frame this return rate is compared to the annualised index return over the respective period and the out- or underperformance versus the index is measured.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
Invested Co.	109	107	102	101	104	111	133	138	113
Delisted Portfolio Co.	38	39	32	28	34	35	44	43	37
Delisted Portfolio Co. (%)	35%	36%	31%	28%	33%	32%	33%	31%	32%
% of Co. Delisted at Pos. Return	71%	77%	72%	82%	79%	80%	75%	70%	76%
% of Co. Delisted at Neg. Return	29%	23%	28%	18%	21%	20%	25%	30%	24%
Average 10 year IRR delisted Co.	4.2%	0.9%	4.8%	9.1%	5.9%	6.2%	7.1%	6.8%	5.6%
Russell 3000 TR IRR	4.4%	-2.7%	-1.8%	0.1%	1.8%	5.7%	5.7%	6.0%	2.4%
Average 10 year excess IRR	-0.3%	3.6%	6.6%	9.0%	4.1%	0.5%	1.4%	0.9%	3.2%

Appendix 2 - Table 3: Evolution of portfolio weights (min. E/A of 20%)

This table shows the number of stocks in each portfolio at inception and end of the investment period and shows how much weight initially corresponds to each stock in the equally weighted portfolios. It also shows what part of the portfolio at T=10 is invested in a Russell 3000 TR ETF and what part remains invested in stocks. Furthermore, for each investment period it details how many stocks at T=10 have a weight in the portfolio of more than 1%, 2% and 3% and what percentage of the remaining stocks it is that has such weights.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
# stocks at T=0	109	107	102	101	104	111	133	138	113
# stocks at T=10	71	68	70	73	70	76	89	95	77
Initial weight in portfolio per stock	0.92%	0.93%	0.98%	0.99%	0.96%	0.90%	0.75%	0.72%	0.90%
% of portfolio in ETF at T=10	36%	35%	42%	42%	35%	29%	40%	43%	38%
% of portfolio in stocks at T=10	64%	65%	58%	58%	65%	71%	60%	57%	62%
# stocks accounting for 1%+	22	21	19	15	17	19	14	13	18
% stocks accounting for 1%+	30.7%	30.2%	26.4%	19.8%	24.0%	23.7%	14.9%	13.4%	22.9%
# stocks accounting for 2%+	6	4	5	5	5	8	6	4	5
% stocks accounting for 2%+	8.4%	5.5%	7.1%	5.6%	6.9%	9.4%	6.3%	4.1%	6.7%
# stocks accounting for 3%+	3	3	3	3	3	5	1	1	3
% stocks accounting for 3%+	3.8%	4.0%	3.7%	2.9%	3.2%	5.4%	0.7%	0.5%	3.0%

Appendix 2 - Table 4: Stock presence in consecutive portfolios (min. E/A of 20%)

This table for each investment period specifies what percentage of the stocks in the current portfolio are also present in the previous investment period's portfolio. Similar percentage figure is specified for the share of a portfolio's companies that are present in the two and three previous portfolios.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
Stock in the previous PF		68%	69%	59%	45%	58%	49%	53%	57%
Stock in the 2 previous PFs			50%	48%	33%	39%	35%	31%	39%
Stock in the 3 previous PFs				35%	29%	29%	23%	25%	28%

Appendix 2 - Table 5: Overall stock presence in portfolios (min. E/A of 20%)

This table shows the frequencies at which stocks repeatedly enter the eight portfolios. For each possible frequency of appearance across portfolios (i.e. 1-8) the table returns the corresponding number of unique stocks. It is also specified what share of the overall 407 unique stocks corresponds to each frequency. The average appearance per unique stock is of 2.2 times.

Appearances across portfolios	1	2	3	4	5	6	7	8
# stocks	212	81	37	31	16	8	8	14
% of unique stocks	52%	20%	9%	8%	4%	2%	2%	3%

Appendix 2 - Table 6: Size of stocks (min. E/A of 20%)

This Table shows the market capitalisation of the stocks in our portfolios at T=0 and T=10 in million USD. Panel A deals with market capitalisations at T=0 and Panel B does so for T=10. Panel A shows two alternative ways of looking at the portfolios. The first one includes all companies in the respective portfolio at T=0. The second one looks only at those companies that are not delisted over the holding period to provide for a better comparison with the data for T=10 presented in Panel B . For both options in Panel A and for Panel B, for each portfolio the table specifies the average, highest and lowest market capitalisation and provides the market capitalisations corresponding to the 75th, 50th and 25th percentile.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
PANEL A: Size at T=0									
Average - All	4,542	6,165	6,323	9,821	6,509	4,686	8,236	8,098	6,797
Highest - All	79,858	83,520	143,327	235,515	221,105	114,295	268,363	196,341	167,790
75th percentile - All	1,398	1,748	1,594	1,454	2,173	2,401	3,664	3,545	2,247
50th percentile - All	398	376	376	244	442	500	970	777	510
25th percentile - All	93	108	89	48	122	128	378	217	148
Lowest - All	2	4	4	3	11	5	4	13	6
Average - Remaining	5,776	8,066	7,273	11,784	8,078	5,711	10,986	10,298	8,497
Highest - Remaining	79,858	83,520	143,327	235,515	221,105	114,295	268,363	196,341	167,790
75th percentile - Remaining	1,674	2,139	1,549	1,438	2,253	2,816	4,367	3,735	2,496
50th percentile - Remaining	230	459	453	274	442	472	1,169	1,003	563
25th percentile - Remaining	74	105	93	71	135	137	427	261	163
Lowest - Remaining	2	4	4	3	11	5	4	13	6
PANEL B: Size at T=10									
Average - Remaining	12,814	8,367	8,533	8,458	9,459	12,972	16,748	16,156	11,688
Highest - Remaining	143,200	76,681	108,619	116,920	125,967	165,964	197,857	196,174	141,423
75th percentile - Remaining	9,538	6,783	6,339	4,314	5,057	10,038	10,445	8,199	7,589
50th percentile - Remaining	2,137	1,633	1,592	1,153	1,657	2,365	2,913	2,511	1,995
25th percentile - Remaining	406	420	327	322	337	473	1,041	595	490
Lowest - Remaining	10	5	7	12	9	7	22	12	11

Appendix 2 - Table 7: Valuation of stocks (min. E/A of 20%)

This Table shows the B/M ratio of the stocks in our portfolios at T=0 and T=10. Panel A deals with B/M at T=0 and Panel B does so for T=10. Panel A shows two alternative ways of looking at the portfolios. The first one includes all companies in the respective portfolio at T=0. The second one looks only at those companies that are not delisted over the holding period to provide for a better comparison with the data for T=10 presented in Panel B. For both options in Panel A and for Panel B, for each portfolio the table specifies the average, highest and lowest B/M and provides the B/M corresponding to the 75th, 50th and 25th percentile.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Average
PANEL A: Valuation at T=0									
Average - All	1.26	1.20	1.41	1.55	0.86	1.08	0.70	0.65	1.09
Highest - All	12.39	8.13	10.40	17.43	4.58	6.38	5.33	4.81	8.68
75th percentile - All	1.39	1.32	1.87	1.65	0.95	1.17	0.77	0.72	1.23
50th percentile - All	0.62	0.64	0.81	0.75	0.62	0.74	0.49	0.46	0.64
25th percentile - All	0.35	0.35	0.36	0.37	0.33	0.44	0.32	0.31	0.35
Lowest - All	0.03	0.04	0.07	0.05	0.08	0.08	0.06	0.01	0.05
Average - Remaining	1.60	1.54	1.64	1.69	0.91	1.17	0.74	0.66	1.24
Highest - Remaining	12.39	8.13	10.40	17.43	4.58	6.38	5.33	4.81	8.68
75th percentile - Remaining	2.11	2.41	2.54	1.65	0.94	1.18	0.78	0.70	1.54
50th percentile - Remaining	0.78	0.92	0.93	0.77	0.61	0.75	0.49	0.45	0.71
25th percentile - Remaining	0.38	0.40	0.36	0.39	0.34	0.50	0.32	0.27	0.37
Lowest - Remaining	0.03	0.05	0.08	0.05	0.08	0.19	0.06	0.01	0.07
PANEL B: Valuation at T=10									
Average - Remaining	0.54	0.80	0.78	0.67	0.64	0.61	0.45	0.48	0.62
Highest - Remaining	2.93	3.18	6.63	4.68	3.67	2.69	3.68	1.99	3.68
75th percentile - Remaining	0.77	1.15	0.99	0.82	0.80	0.78	0.58	0.60	0.81
50th percentile - Remaining	0.42	0.61	0.58	0.47	0.47	0.53	0.38	0.33	0.47
25th percentile - Remaining	0.31	0.35	0.38	0.35	0.35	0.33	0.25	0.21	0.32
Lowest - Remaining	-1.02	-1.30	-0.58	-0.13	-0.80	-0.66	-0.10	-0.09	-0.58

Appendix 3

This appendix provides the results of all factor analyses we run, assuming no reinvestments into the index (i.e. the Russell 3000 TR) of dividends and delisting returns. Besides the results for the analyses over the entire ten-year investment time-spans, it comprises the results for all variations of factor analysis we use for the first five and for the second five years independent of each other. This is to show how due to the non-rebalancing of our portfolios the characteristics of the stocks in them over time may drift away from what they were at initial investment.

Appendix 3 - Table 1: Fama French Three-Factor Model Analysis

This table shows the Fama French three-factor model analysis on all of our portfolios when no reinvestment of dividends or delisted stocks is performed. The factors have been calculated by running regressions on the monthly price performance of our portfolios with respect to each of the factor model portfolios. Coefficients for each factor have been calculated both based on the performance of each portfolio for the entire respective ten-year holding period (Panel A), as well as split into two five-year holding periods (Panels B and C) in order to further analyse possible changes in each portfolio's performance and behavior over the holding period. P-values were calculated for each coefficient in order to determine statistical significance (p-values below 0.1 represent a statistically significant coefficient at the 90% level). Finally, an overall column has been calculated by running factor model regressions on all of our portfolios together to show a summary of how the portfolios perform and what overall characteristics they show.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Overall
PANEL A: Full 10-Year Holding Period (t=1 to t=10)									
Alpha (annual)	1.31%	5.29%	7.65%	6.25%	4.71%	1.81%	3.02%	2.87%	4.33%
p-value	*0.59*	*0.01*	*0.00*	*0.01*	*0.00*	*0.25*	*0.03*	*0.07*	*0.00*
Rm-rf	0.94	0.91	0.98	1.12	1.05	1.05	1.05	1.03	1.01
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.41	0.33	0.38	0.64	0.63	0.73	0.68	0.68	0.46
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.60	0.45	0.37	0.31	0.17	0.13	0.23	0.22	0.43
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.03*	*0.00*	*0.00*	*0.00*
PANEL B: First 5 Years of Holding Period (t=1 to t=5)									
Alpha (annual)	0.39%	9.44%	14.74%	8.31%	4.41%	1.98%	0.30%	2.19%	4.02%
p-value	*0.93*	*0.01*	*0.00*	*0.01*	*0.08*	*0.31*	*0.87*	*0.38*	*0.00*
Rm-rf	0.97	0.88	0.96	1.05	0.97	0.96	1.06	1.08	1.00
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.42	0.29	0.31	0.58	0.56	0.67	0.55	0.70	0.41
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.66	0.45	0.34	0.20	0.11	0.08	0.26	0.20	0.48
p-value	*0.00*	*0.00*	*0.00*	*0.03*	*0.26*	*0.40*	*0.00*	*0.01*	*0.00*
PANEL C: Last 5 Years of Holding Period (t=6 to t=10)									
Alpha (annual)	3.53%	2.58%	1.76%	5.52%	5.91%	2.52%	5.70%	4.74%	3.90%
p-value	*0.14*	*0.15*	*0.35*	*0.09*	*0.01*	*0.32*	*0.01*	*0.03*	*0.00*
Rm-rf	0.90	0.93	1.00	1.14	1.06	1.06	1.02	0.96	1.04
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.49	0.50	0.65	0.71	0.76	0.82	0.84	0.69	0.67
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.24	0.16	0.14	0.33	0.14	0.11	0.23	0.21	0.22
p-value	*0.04*	*0.04*	*0.03*	*0.00*	*0.06*	*0.16*	*0.00*	*0.03*	*0.00*

Appendix 3 - Table 2: Carhart Four-Factor Model Analysis

This table shows the Carhart four-factor model analysis on all of our portfolios when no reinvestment of dividends or delisted stocks is performed. The factors have been calculated by running regressions on the monthly price performance of our portfolios with respect to each of the factor model portfolios. Coefficients for each factor have been calculated both based on the performance of each portfolio for the entire respective ten-year holding period (Panel A), as well as split into two five-year holding periods (Panels B and C) in order to further analyse possible changes in each portfolio's performance and behavior over the holding period. P-values were calculated for each coefficient in order to determine statistical significance (p-values below 0.1 represent a statistically significant coefficient at the 90% level). Finally, an overall column has been calculated by running factor model regressions on all of our portfolios together to show a summary of how the portfolios perform and what overall characteristics they show.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Overall
PANEL A: Full 10-Year Holding Period (t=1 to t=10)									
Alpha (annual)	3.40%	6.30%	7.56%	6.27%	4.86%	1.92%	3.44%	3.26%	4.84%
p-value	*0.14*	*0.00*	*0.00*	*0.00*	*0.00*	*0.19*	*0.01*	*0.03*	*0.00*
Rm-rf	0.85	0.84	0.90	1.04	1.00	1.00	1.01	0.99	0.94
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.46	0.37	0.42	0.63	0.66	0.75	0.71	0.69	0.49
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.56	0.42	0.35	0.34	0.16	0.08	0.16	0.15	0.40
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.15*	*0.00*	*0.01*	*0.00*
MOM	-0.17	-0.12	-0.11	-0.13	-0.08	-0.11	-0.14	-0.13	-0.13
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
PANEL B: First 5 Years of Holding Period (t=1 to t=5)									
Alpha (annual)	3.75%	10.83%	13.98%	7.92%	4.41%	2.13%	1.78%	1.83%	4.96%
p-value	*0.33*	*0.00*	*0.00*	*0.01*	*0.08*	*0.28*	*0.33*	*0.38*	*0.00*
Rm-rf	0.86	0.77	0.86	0.94	0.97	0.95	1.02	1.00	0.89
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.47	0.34	0.40	0.57	0.55	0.68	0.58	0.68	0.47
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.60	0.40	0.36	0.33	0.11	0.08	0.19	0.10	0.46
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.29*	*0.40*	*0.02*	*0.16*	*0.00*
MOM	-0.19	-0.16	-0.15	-0.17	0.00	-0.03	-0.14	-0.17	-0.17
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.97*	*0.56*	*0.00*	*0.00*	*0.00*
PANEL C: Last 5 Years of Holding Period (t=6 to t=10)									
Alpha (annual)	3.73%	2.93%	1.62%	4.91%	5.29%	1.82%	5.16%	4.72%	3.98%
p-value	*0.12*	*0.11*	*0.37*	*0.10*	*0.01*	*0.42*	*0.00*	*0.03*	*0.00*
Rm-rf	0.89	0.92	0.97	1.09	1.03	1.01	0.98	0.96	1.00
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.51	0.51	0.64	0.72	0.77	0.82	0.85	0.69	0.68
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.24	0.14	0.10	0.24	0.05	0.04	0.16	0.21	0.16
p-value	*0.04*	*0.08*	*0.12*	*0.02*	*0.50*	*0.63*	*0.01*	*0.03*	*0.00*
MOM	-0.04	-0.03	-0.07	-0.14	-0.13	-0.13	-0.12	0.00	-0.11
p-value	*0.53*	*0.46*	*0.03*	*0.00*	*0.00*	*0.00*	*0.00*	*0.96*	*0.00*

Appendix 3 - Table 3: Fama French Five-Factor Model Analysis

This table shows the Fama French five-factor model analysis on all of our portfolios when no reinvestment of dividends or delisted stocks is performed. The factors have been calculated by running regressions on the monthly price performance of our portfolios with respect to each of the factor model portfolios. Coefficients for each factor have been calculated both based on the performance of each portfolio for the entire respective ten-year holding period (Panel A), as well as split into two five-year holding periods (Panels B and C) in order to further analyse possible changes in each portfolio's performance and behavior over the holding period. P-values were calculated for each coefficient in order to determine statistical significance (p-values below 0.1 represent a statistically significant coefficient at the 90% level). Finally, an overall column has been calculated by running factor model regressions on all of our portfolios together to show a summary of how the portfolios perform and what overall characteristics they show.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Overall
PANEL A: Full 10-Year Holding Period (t=1 to t=10)									
Alpha (annual)	-0.66%	3.38%	5.90%	6.38%	3.87%	2.07%	2.72%	3.10%	2.56%
p-value	*0.77*	*0.08*	*0.01*	*0.01*	*0.02*	*0.21*	*0.06*	*0.06*	*0.00*
Rm-rf	1.04	1.01	1.06	1.12	1.09	1.04	1.06	1.03	1.08
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.53	0.39	0.45	0.64	0.67	0.75	0.70	0.68	0.54
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.24	0.26	0.23	0.32	0.15	0.17	0.25	0.24	0.22
p-value	*0.01*	*0.00*	*0.00*	*0.00*	*0.02*	*0.01*	*0.00*	*0.00*	*0.00*
RMW	0.19	0.28	0.22	-0.01	0.16	0.00	0.08	-0.03	0.27
p-value	*0.08*	*0.00*	*0.01*	*0.93*	*0.06*	*0.98*	*0.42*	*0.80*	*0.00*
CMA	0.19	0.21	0.15	-0.03	-0.04	-0.18	-0.11	-0.08	0.12
p-value	*0.08*	*0.02*	*0.12*	*0.77*	*0.69*	*0.10*	*0.23*	*0.52*	*0.00*
PANEL B: First 5 Years of Holding Period (t=1 to t=5)									
Alpha (annual)	-3.97%	6.03%	12.97%	6.69%	2.76%	1.75%	-0.77%	2.73%	1.95%
p-value	*0.36*	*0.09*	*0.00*	*0.04*	*0.26*	*0.37*	*0.70*	*0.30*	*0.08*
Rm-rf	1.06	0.98	1.02	1.15	1.09	0.98	1.09	1.06	1.11
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.53	0.34	0.34	0.62	0.59	0.69	0.59	0.68	0.50
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.22	0.24	0.13	0.05	-0.02	0.09	0.26	0.21	0.17
p-value	*0.14*	*0.06*	*0.30*	*0.69*	*0.83*	*0.40*	*0.00*	*0.03*	*0.00*
RMW	0.45	0.25	0.14	0.23	0.30	0.04	0.18	-0.11	0.31
p-value	*0.00*	*0.03*	*0.29*	*0.11*	*0.01*	*0.79*	*0.24*	*0.53*	*0.00*
CMA	0.26	0.24	0.18	0.18	0.19	-0.08	-0.11	0.00	0.24
p-value	*0.14*	*0.09*	*0.19*	*0.17*	*0.09*	*0.61*	*0.41*	*0.99*	*0.00*
PANEL C: Last 5 Years of Holding Period (t=6 to t=10)									
Alpha (annual)	2.88%	1.84%	1.47%	8.50%	7.08%	4.00%	6.57%	4.89%	4.09%
p-value	*0.21*	*0.32*	*0.45*	*0.01*	*0.00*	*0.16*	*0.00*	*0.03*	*0.00*
Rm-rf	0.97	0.92	0.97	1.04	1.02	1.02	1.00	0.97	1.02
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.60	0.54	0.66	0.69	0.78	0.82	0.85	0.72	0.68
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.13	0.17	0.11	0.40	0.23	0.16	0.29	0.28	0.19
p-value	*0.28*	*0.03*	*0.12*	*0.00*	*0.01*	*0.10*	*0.00*	*0.04*	*0.00*
RMW	0.31	0.07	0.01	-0.51	-0.11	-0.12	-0.03	0.10	0.00
p-value	*0.07*	*0.61*	*0.93*	*0.02*	*0.45*	*0.46*	*0.79*	*0.45*	*0.98*
CMA	-0.01	-0.20	-0.21	-0.24	-0.30	-0.19	-0.20	-0.12	-0.17
p-value	*0.96*	*0.11*	*0.11*	*0.21*	*0.04*	*0.27*	*0.16*	*0.54*	*0.00*

Appendix 3 - Table 4: Modified Three-Factor Model Analysis

This table shows our modified three-factor model analysis on all of our portfolios when no reinvestment of dividends or delisted stocks is performed. The model comprises market beta, the profitability factor and the investment factor. The factors have been calculated by running regressions on the monthly price performance of our portfolios with respect to each of the factor model portfolios. Coefficients for each factor have been calculated both based on the performance of each portfolio for the entire respective ten-year holding period (Panel A), as well as split into two five-year holding periods (Panels B and C) in order to further analyse possible changes in each portfolio's performance and behavior over the holding period. P-values were calculated for each coefficient in order to determine statistical significance (p-values below 0.1 represent a statistically significant coefficient at the 90% level). Finally, an overall column has been calculated by running factor model regressions on all of our portfolios together to show a summary of how the portfolios perform and what overall characteristics they show.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Overall
PANEL A: Full 10-Year Holding Period (t=1 to t=10)									
Alpha (annual)	2.12%	6.84%	10.43%	10.76%	6.17%	4.21%	4.19%	3.59%	5.35%
p-value	*0.48*	*0.00*	*0.00*	*0.00*	*0.01*	*0.08*	*0.06*	*0.12*	*0.00*
Rm-rf	0.99	0.99	1.07	1.26	1.22	1.19	1.23	1.17	1.17
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
RMW	0.21	0.12	0.00	-0.04	0.09	-0.22	-0.13	-0.32	0.11
p-value	*0.01*	*0.06*	*0.98*	*0.75*	*0.42*	*0.11*	*0.35*	*0.02*	*0.00*
CMA	0.47	0.42	0.39	0.30	0.25	0.09	0.14	0.18	0.43
p-value	*0.00*	*0.00*	*0.00*	*0.01*	*0.04*	*0.54*	*0.30*	*0.22*	*0.00*
PANEL B: First 5 Years of Holding Period (t=1 to t=5)									
Alpha (annual)	-1.26%	10.44%	19.18%	14.54%	5.71%	2.20%	2.41%	5.18%	6.38%
p-value	*0.81*	*0.01*	*0.00*	*0.00*	*0.09*	*0.43*	*0.40*	*0.16*	*0.00*
Rm-rf	0.96	0.88	0.92	1.12	1.10	1.14	1.20	1.22	1.09
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
RMW	0.21	0.08	-0.06	-0.04	0.09	-0.33	-0.02	-0.33	0.10
p-value	*0.05*	*0.27*	*0.39*	*0.80*	*0.51*	*0.07*	*0.90*	*0.18*	*0.01*
CMA	0.48	0.40	0.31	0.23	0.35	-0.08	0.05	0.32	0.46
p-value	*0.00*	*0.00*	*0.01*	*0.09*	*0.02*	*0.70*	*0.81*	*0.15*	*0.00*
PANEL C: Last 5 Years of Holding Period (t=6 to t=10)									
Alpha (annual)	3.65%	4.56%	3.82%	10.28%	8.82%	6.67%	5.21%	3.06%	5.48%
p-value	*0.21*	*0.08*	*0.23*	*0.02*	*0.02*	*0.11*	*0.14*	*0.31*	*0.00*
Rm-rf	1.14	1.02	1.14	1.24	1.21	1.20	1.23	1.12	1.19
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
RMW	0.05	-0.13	-0.18	-0.69	-0.24	-0.26	-0.29	-0.30	-0.18
p-value	*0.80*	*0.47*	*0.39*	*0.02*	*0.30*	*0.28*	*0.17*	*0.08*	*0.01*
CMA	0.01	-0.08	0.05	0.26	0.16	0.14	0.15	0.11	0.11
p-value	*0.96*	*0.65*	*0.81*	*0.27*	*0.41*	*0.50*	*0.46*	*0.58*	*0.10*

Appendix 4

This appendix provides the results of all factor analyses we run, assuming that all dividends and delisting returns are reinvested into the index (i.e. the Russell 3000 TR). Albeit the factor analysis results with reinvestments into the index being less relevant for showing the factor loadings achieved by the stocks that we actually choose for our portfolios, we include this version of factor analyses here as it shows the actual factor loadings our strategy generates considering that our portfolios do reinvest. For all variations of factor analysis that we use, the appendix comprises analyses over the entire ten-year investment time-spans and for the first five and the second five years independent of each other. This is to show how due to the non-rebalancing of our portfolios the characteristics of the stocks in them over time may drift away from what they were at initial investment.

Appendix 4 - Table 1: Fama French Three-Factor Model Analysis (Reinvesting)

This table shows the Fama French three-factor model analysis on all of our portfolios when proceeds from dividends and delisted stocks are reinvested in the Russell 3000 TR index. The factors have been calculated by running regressions on the monthly price performance of our portfolios with respect to each of the factor model portfolios. Coefficients for each factor have been calculated both based on the performance of each portfolio for the entire respective ten-year holding period (Panel A), as well as split into two five-year holding periods (Panels B and C) in order to further analyse possible changes in each portfolio's performance and behavior over the holding period. P-values were calculated for each coefficient in order to determine statistical significance (p-values below 0.1 represent a statistically significant coefficient at the 90% level). Finally, an overall column has been calculated by running factor model regressions on all of our portfolios together to show a summary of how the portfolios perform and what overall characteristics they show.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Overall
PANEL A: Full 10-Year Holding Period (t=1 to t=10)									
Alpha (annual)	-0.12%	3.99%	7.16%	5.51%	3.82%	1.25%	1.80%	3.01%	3.52%
p-value	*0.96*	*0.02*	*0.00*	*0.00*	*0.01*	*0.30*	*0.09*	*0.08*	*0.00*
Rm-rf	0.95	0.91	0.96	1.09	1.02	1.04	1.04	0.90	0.98
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.32	0.28	0.33	0.57	0.53	0.58	0.53	0.50	0.37
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.50	0.41	0.32	0.24	0.15	0.09	0.17	0.17	0.35
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.04*	*0.00*	*0.01*	*0.00*
PANEL B: First 5 Years of Holding Period (t=1 to t=5)									
Alpha (annual)	-1.33%	7.80%	14.07%	8.08%	4.10%	1.80%	0.32%	1.78%	3.81%
p-value	*0.73*	*0.02*	*0.00*	*0.01*	*0.09*	*0.29*	*0.84*	*0.45*	*0.00*
Rm-rf	0.97	0.89	0.96	1.05	0.97	0.98	1.05	1.01	0.98
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.34	0.27	0.29	0.57	0.52	0.60	0.46	0.55	0.36
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.56	0.42	0.33	0.19	0.12	0.07	0.20	0.12	0.41
p-value	*0.00*	*0.00*	*0.00*	*0.03*	*0.19*	*0.41*	*0.01*	*0.14*	*0.00*
PANEL C: Last 5 Years of Holding Period (t=6 to t=10)									
Alpha (annual)	2.19%	1.42%	0.88%	3.57%	3.89%	1.31%	3.40%	6.85%	2.49%
p-value	*0.20*	*0.27*	*0.54*	*0.15*	*0.02*	*0.48*	*0.03*	*0.00*	*0.00*
Rm-rf	0.93	0.96	1.01	1.10	1.05	1.05	1.01	0.74	1.01
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.35	0.35	0.48	0.55	0.55	0.59	0.61	0.49	0.48
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.17	0.11	0.10	0.25	0.12	0.09	0.18	0.16	0.17
p-value	*0.04*	*0.06*	*0.04*	*0.00*	*0.04*	*0.12*	*0.00*	*0.10*	*0.00*

Appendix 4 - Table 2: Carhart Four-Factor Model Analysis (Reinvesting)

This table shows the Carhart four-factor model analysis on all of our portfolios when proceeds from dividends and delisted stocks are reinvested in the Russell 3000 TR index. The factors have been calculated by running regressions on the monthly price performance of our portfolios with respect to each of the factor model portfolios. Coefficients for each factor have been calculated both based on the performance of each portfolio for the entire respective ten-year holding period (Panel A), as well as split into two five-year holding periods (Panels B and C) in order to further analyse possible changes in each portfolio's performance and behavior over the holding period. P-values were calculated for each coefficient in order to determine statistical significance (p-values below 0.1 represent a statistically significant coefficient at the 90% level). Finally, an overall column has been calculated by running factor model regressions on all of our portfolios together to show a summary of how the portfolios perform and what overall characteristics they show.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Overall
PANEL A: Full 10-Year Holding Period (t=1 to t=10)									
Alpha (annual)	1.74%	4.94%	7.09%	5.52%	3.92%	1.32%	2.10%	3.39%	3.96%
p-value	*0.37*	*0.00*	*0.00*	*0.00*	*0.00*	*0.25*	*0.03*	*0.03*	*0.00*
Rm-rf	0.87	0.85	0.89	1.01	0.99	1.00	1.01	0.86	0.92
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.37	0.32	0.37	0.56	0.54	0.60	0.54	0.51	0.40
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.46	0.38	0.31	0.27	0.14	0.06	0.13	0.09	0.33
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.17*	*0.00*	*0.12*	*0.00*
MOM	-0.15	-0.12	-0.09	-0.11	-0.06	-0.08	-0.10	-0.13	-0.11
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.01*	*0.00*	*0.00*	*0.00*	*0.00*
PANEL B: First 5 Years of Holding Period (t=1 to t=5)									
Alpha (annual)	1.66%	9.15%	13.32%	7.67%	4.11%	1.94%	1.37%	1.50%	4.65%
p-value	*0.63*	*0.00*	*0.00*	*0.01*	*0.09*	*0.26*	*0.39*	*0.48*	*0.00*
Rm-rf	0.88	0.78	0.85	0.93	0.96	0.96	1.02	0.95	0.88
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.39	0.31	0.38	0.55	0.52	0.61	0.49	0.54	0.41
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.51	0.37	0.35	0.33	0.13	0.07	0.15	0.04	0.40
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.20*	*0.41*	*0.04*	*0.63*	*0.00*
MOM	-0.17	-0.16	-0.15	-0.18	-0.01	-0.03	-0.10	-0.13	-0.15
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.92*	*0.56*	*0.01*	*0.00*	*0.00*
PANEL C: Last 5 Years of Holding Period (t=6 to t=10)									
Alpha (annual)	2.33%	1.63%	0.79%	3.11%	3.43%	0.80%	3.00%	6.93%	2.56%
p-value	*0.18*	*0.23*	*0.58*	*0.18*	*0.02*	*0.63*	*0.02*	*0.00*	*0.00*
Rm-rf	0.92	0.96	0.99	1.07	1.02	1.01	0.99	0.74	0.98
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.36	0.36	0.48	0.56	0.56	0.59	0.61	0.50	0.49
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.17	0.10	0.07	0.17	0.05	0.03	0.13	0.16	0.12
p-value	*0.04*	*0.10*	*0.13*	*0.03*	*0.37*	*0.53*	*0.01*	*0.11*	*0.00*
MOM	-0.03	-0.02	-0.04	-0.10	-0.09	-0.10	-0.09	-0.02	-0.09
p-value	*0.56*	*0.56*	*0.08*	*0.00*	*0.00*	*0.00*	*0.00*	*0.82*	*0.00*

Appendix 4 - Table 3: Fama French Five-Factor Model Analysis (Reinvesting)

This table shows the Fama French five-factor model analysis on all of our portfolios when proceeds from dividends and delisted stocks are reinvested in the Russell 3000 TR index. The factors have been calculated by running regressions on the monthly price performance of our portfolios with respect to each of the factor model portfolios. Coefficients for each factor have been calculated both based on the performance of each portfolio for the entire respective ten-year holding period (Panel A), as well as split into two five-year holding periods (Panels B and C) in order to further analyse possible changes in each portfolio's performance and behavior over the holding period. P-values were calculated for each coefficient in order to determine statistical significance (p-values below 0.1 represent a statistically significant coefficient at the 90% level). Finally, an overall column has been calculated by running factor model regressions on all of our portfolios together to show a summary of how the portfolios perform and what overall characteristics they show.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Overall
PANEL A: Full 10-Year Holding Period (t=1 to t=10)									
Alpha (annual)	-1.82%	2.17%	5.49%	5.45%	2.99%	1.55%	1.54%	3.74%	2.02%
p-value	*0.36*	*0.21*	*0.01*	*0.01*	*0.04*	*0.22*	*0.17*	*0.04*	*0.00*
Rm-rf	1.03	1.01	1.04	1.09	1.06	1.02	1.04	0.88	1.04
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.43	0.34	0.39	0.57	0.55	0.59	0.55	0.48	0.44
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.19	0.22	0.18	0.24	0.11	0.12	0.20	0.19	0.17
p-value	*0.02*	*0.00*	*0.01*	*0.00*	*0.03*	*0.01*	*0.00*	*0.01*	*0.00*
RMW	0.36	0.26	0.20	0.01	0.14	-0.03	0.07	-0.15	0.21
p-value	*0.00*	*0.00*	*0.02*	*0.91*	*0.05*	*0.68*	*0.36*	*0.19*	*0.00*
CMA	0.18	0.22	0.18	0.00	0.02	-0.12	-0.09	-0.09	0.12
p-value	*0.06*	*0.01*	*0.04*	*0.99*	*0.77*	*0.15*	*0.20*	*0.48*	*0.00*
PANEL B: First 5 Years of Holding Period (t=1 to t=5)									
Alpha (annual)	-5.17%	4.72%	12.42%	6.48%	2.48%	1.67%	-0.58%	2.48%	1.88%
p-value	*0.18*	*0.17*	*0.00*	*0.05*	*0.28*	*0.34*	*0.73*	*0.33*	*0.07*
Rm-rf	1.05	0.98	1.00	1.14	1.07	0.99	1.07	0.99	1.08
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.44	0.31	0.31	0.60	0.55	0.61	0.50	0.53	0.44
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.18	0.23	0.19	0.04	-0.01	0.07	0.20	0.12	0.13
p-value	*0.17*	*0.05*	*0.12*	*0.74*	*0.93*	*0.41*	*0.01*	*0.18*	*0.00*
RMW	0.39	0.22	0.11	0.23	0.28	0.02	0.15	-0.15	0.28
p-value	*0.00*	*0.04*	*0.37*	*0.11*	*0.01*	*0.88*	*0.24*	*0.40*	*0.00*
CMA	0.23	0.22	0.18	0.17	0.22	-0.05	-0.09	0.01	0.24
p-value	*0.13*	*0.10*	*0.19*	*0.17*	*0.04*	*0.70*	*0.41*	*0.94*	*0.00*
PANEL C: Last 5 Years of Holding Period (t=6 to t=10)									
Alpha (annual)	1.73%	0.82%	0.61%	5.86%	4.70%	2.30%	3.88%	6.78%	2.78%
p-value	*0.30*	*0.54*	*0.68*	*0.02*	*0.01*	*0.27*	*0.02*	*0.01*	*0.00*
Rm-rf	0.98	0.96	0.99	1.03	1.02	1.02	1.00	0.74	0.99
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
SMB	0.42	0.38	0.49	0.53	0.57	0.59	0.62	0.47	0.49
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
HML	0.09	0.12	0.14	0.30	0.18	0.13	0.22	0.12	0.16
p-value	*0.28*	*0.04*	*0.01*	*0.00*	*0.01*	*0.08*	*0.00*	*0.41*	*0.00*
RMW	0.22	0.06	0.02	-0.38	-0.07	-0.08	0.00	-0.07	-0.03
p-value	*0.08*	*0.53*	*0.84*	*0.02*	*0.50*	*0.53*	*0.98*	*0.61*	*0.53*
CMA	-0.01	-0.15	-0.20	-0.21	-0.22	-0.14	-0.13	0.07	-0.14
p-value	*0.94*	*0.11*	*0.06*	*0.16*	*0.04*	*0.27*	*0.23*	*0.72*	*0.00*

Appendix 4 - Table 4: Modified Three-Factor Model Analysis (Reinvesting)

This table shows our modified three-factor model analysis on all of our portfolios when proceeds from dividends and delisted stocks are reinvested in the Russell 3000 TR index. The model comprises market beta, the profitability factor and the investment factor. The factors have been calculated by running regressions on the monthly price performance of our portfolios with respect to each of the factor model portfolios. Coefficients for each factor have been calculated both based on the performance of each portfolio for the entire respective ten-year holding period (Panel A), as well as split into two five-year holding periods (Panels B and C) in order to further analyse possible changes in each portfolio's performance and behavior over the holding period. P-values were calculated for each coefficient in order to determine statistical significance (p-values below 0.1 represent a statistically significant coefficient at the 90% level). Finally, an overall column has been calculated by running factor model regressions on all of our portfolios together to show a summary of how the portfolios perform and what overall characteristics they show.

Accounting Data Period	1987-1997	1988-1998	1989-1999	1990-2000	1991-2001	1992-2002	1993-2003	1994-2004	
Investment Period	1998-2007	1999-2008	2000-2009	2001-2010	2002-2011	2003-2012	2004-2013	2005-2014	Overall
PANEL A: Full 10-Year Holding Period (t=1 to t=10)									
Alpha (annual)	0.40%	5.14%	9.41%	9.46%	4.90%	3.25%	2.67%	4.05%	4.30%
p-value	*0.87*	*0.01*	*0.00*	*0.00*	*0.02*	*0.08*	*0.12*	*0.06*	*0.00*
Rm-rf	1.00	0.99	1.05	1.20	1.16	1.15	1.17	0.98	1.11
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
RMW	0.19	0.12	0.00	-0.04	0.08	-0.21	-0.09	-0.36	0.08
p-value	*0.00*	*0.04*	*0.94*	*0.69*	*0.41*	*0.06*	*0.38*	*0.01*	*0.00*
CMA	0.40	0.39	0.38	0.26	0.26	0.08	0.10	0.11	0.37
p-value	*0.00*	*0.00*	*0.00*	*0.01*	*0.01*	*0.49*	*0.32*	*0.41*	*0.00*
PANEL B: First 5 Years of Holding Period (t=1 to t=5)									
Alpha (annual)	-2.93%	8.63%	18.20%	14.09%	5.29%	2.06%	2.04%	4.36%	5.79%
p-value	*0.53*	*0.02*	*0.00*	*0.00*	*0.09*	*0.40*	*0.40*	*0.17*	*0.00*
Rm-rf	0.97	0.89	0.91	1.11	1.08	1.13	1.16	1.10	1.06
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
RMW	0.19	0.08	-0.06	-0.04	0.09	-0.31	-0.03	-0.33	0.08
p-value	*0.04*	*0.25*	*0.38*	*0.80*	*0.49*	*0.06*	*0.87*	*0.13*	*0.02*
CMA	0.42	0.37	0.30	0.22	0.37	-0.05	0.03	0.22	0.42
p-value	*0.00*	*0.00*	*0.01*	*0.09*	*0.01*	*0.77*	*0.84*	*0.25*	*0.00*
PANEL C: Last 5 Years of Holding Period (t=6 to t=10)									
Alpha (annual)	2.28%	2.74%	2.38%	7.27%	5.90%	4.19%	2.83%	5.82%	3.73%
p-value	*0.27*	*0.14*	*0.32*	*0.03*	*0.04*	*0.16*	*0.27*	*0.03*	*0.00*
Rm-rf	1.10	1.02	1.10	1.18	1.15	1.15	1.17	0.84	1.12
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
RMW	0.03	-0.08	-0.13	-0.52	-0.17	-0.18	-0.19	-0.32	-0.16
p-value	*0.80*	*0.53*	*0.40*	*0.02*	*0.32*	*0.31*	*0.23*	*0.03*	*0.01*
CMA	0.01	-0.06	0.03	0.17	0.13	0.11	0.14	0.15	0.07
p-value	*0.97*	*0.61*	*0.86*	*0.34*	*0.37*	*0.47*	*0.36*	*0.37*	*0.17*