Norwegian School of Economics
Bergen, Spring 2020

Investing in Equity Mutual Funds

A study of the Norwegian Fund Market

Andreas Dobloug and Per Haakestad

Supervisor: Tommy Stamland

Master thesis, Economics and Business Administration

Major: Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This thesis is written as a part of our Master of Science in Economics and Business Administration at the Norwegian School of Economics (NHH).

The choice of research topic is based on both authors interest in asset allocation. We wanted learn more about the investment decision of investors and the practicability of the models we have learned in school. In the writing process we have experienced how challenging it can be to produce a valid empirical study and the efforts it takes to make the results presentable. The work has been demanding and given us valuable insights in data handling and finance.

First, we would like to thank our supervisor, Tommy Stamland, for his feedback and guidance, both in selecting our topic and during the writing process. We would also like to thank Timothy B. Riley for clarifications regarding his methodology which we based part of our paper on. Lastly, we would like to thank Børsprosjektet at NHH and VFF for providing data on Norwegian mutual funds.

Norwegian School of Economics

Bergen, June 2020

Per Haakestad | Andreas Steen Dobloug

Per Haakestad | Andreas Steen Dobloug

Abstract

In this thesis we have analyzed Norwegian equity funds over the last eleven year period. We investigate if the performance of individual funds can be attributed to the skillset of managers, if investors can achieve abnormal returns by betting on funds with historical good performance, and if applying an optimization framework within previous winners provide additional benefits to the average of these funds. We use a data set free of survivorship-bias with monthly and daily net returns for 55 actively managed Norwegian mutual funds in the period 2009-2019.

We find that Norwegian equity mutual funds, on aggregate, are able to cover their costs, but do not deliver any abnormal performance over their benchmark. To test the skillset of managers in individual funds we apply a bootstrap procedure from Kosowski et al. (2006). We are unable to find sufficient evidence to claim any presence of skill, or lack of skill, among fund managers in the best and worst performing funds. Inspired by Riley (2019), we then turn to a portfolio approach based largely on persistence in performance among previous winners. With monthly rebalancing we find that optimal portfolios from the Treynor and Black (1973) model achieve positive alphas before transaction costs across several formation parameters, but do not deliver any added performance over the average fund in the same portfolio. Despite the alphas being positive, we do not find enough evidence to claim the strategy deliver a performance better than the passive benchmark for an investor. We also test the long-run persistence in performance for the portfolios and find that monthly rebalancing is necessary in order to maintain a positive alpha.

All taken together, our results indicate that actively managed Norwegian equity mutual funds do not add value for investors compared to an equivalent passive investment. This holds both when funds are evaluated individually and as portfolios consisting of past winners.

Keywords –

Contents

List of Figures

List of Tables

1 Introduction

Two important choices to make for investors is to decide what type of asset category is best to achieve their investment goal, and subsequently, to choose which assets within this category to invest in. For this thesis, we will focus on Norwegian equity mutual funds and try to answer questions through the eyes of the investor. One important question relates to the everlasting debate on the value of active management, and another is how to, ex ante, pick out the funds that will perform well in the future. In a similar manner to why investors should hold a portfolio of stocks, we also argue investors should take a portfolio approach when investing in actively managed funds.[1]

The topic of mutual fund performance has been of long-standing interest in the academic literature – and the conflicting findings amongst researchers have led to debate on whether actively managed mutual funds add value for their investors. For example, studies on the US market such as Carhart (1997), Fama and French (2010) and Davis (2001) concludes there seem to be little or no evidence of skill in active management, while studies by Kosowski et al. (2006), Barras et al. (2010) and Kacperczyk et al. (2014) claim there is meaningful evidence of both fund manager skill, and perhaps most importantly lack of skill, in the extreme left and right tails of the performance distribution, respectively.

In Norway, Sørensen (2009) examined all Norwegian equity mutual funds from 1982 to 2008 and found, on aggregate, no abnormal performance. At the individual fund level he finds no clear evidence of superior performing funds but provide evidence of inferior performance for the worst performers. Additionally, he finds no evidence of performance persistence for either winner or loser funds. Contrary to Sørensen (2009), and perhaps the paper closest to ours, Gallefoss et al. (2015) finds evidence of fund manager skill among both the best and worst performing funds using a daily data set of Norwegian equity mutual funds in the period 2000-2010. Furthermore, he finds short term persistence in the performance of the worst performing funds and persistence in relative performance for the best performing funds. The short term persistence in the worst performers could

[1]Even though mutual funds already are diversified across a number of stocks, they could still have differences in strategies and manager ability. In our sample we find the aggregated fund's return series to be explained reasonably well by the benchmark. However, factor exposures and performance vary significantly from the mean for some funds. Following Riley (2019), we find signs that might suggest a better risk/reward from holding a portfolio of actively managed mutual funds compared to any individual fund, despite all our funds sharing the same asset class and investment universe.

indicate managers learn from their mistakes or perhaps more plausibly, that funds induce costs when they shift their portfolio, or for other reasons have temporary high transaction levels.

Inspired by such varying conclusions about mutual fund performance and thus also the value of active management, we have devoted part of our thesis to contribute to the literature on mutual fund performance. We believe this is relevant for a couple of reasons. First and foremost because there is no clear consensus on the value of active management, and especially so for Norwegian funds. Secondly, by performing a study with more recent data than previous studies it will be interesting to see if we will come to the same conclusions as studies in the US have indicated the selected time period might influence the found performance of mutual funds. To shed light on these issues we will attempt to answer the following question:

i) Is there superior or inferior performance among Norwegian equity mutual funds that can be attributed to skill, or lack of skill, among fund managers?

We find that actively managed Norwegian equity mutual funds, on aggregate, are not able to generate risk-adjusted returns sufficient to justify the fees they put on investors. When looking at individual funds we find fat tails in both ends of the performance distribution but are unable to attribute any of these observations to either skill, or lack of skill, amongst fund managers. From an investors point of view our results suggest investors should not expect an investment in any actively managed fund to perform better than a passive investment in the benchmark.

The previous literature on mutual fund performance tends to focus on the performance of individual funds. An investor is, however, able to invest in multiple funds at the same time, and from the perspective of an investor who considers investing in actively managed mutual funds we thus believe the analysis should focus on portfolios of these funds. Because of variation in strategies, ideas and trading behavior, a portfolio of actively managed funds should provide the investor with a better risk-reward trade-off than any individual actively managed fund, similar to how a portfolio stocks offers a better trade-off compared to an individual stock.

Using the Treynor and Black (1976) model on a sample of actively managed US mutual funds, Riley (2019) finds that the resulting optimal portfolio is superior to both an equal weight portfolio of the same funds and to a passive investment in the benchmark. The model seeks to maximize alpha relative to idiosyncratic volatility and is thus heavily reliant on the ability to accurately measure alpha. As illustrated by the disagreement about the value of active management, the task of measuring alpha is not straight forward. However, his findings suggest the Treynor and Black (1976) model could still provide value to investors given that we are able to forecast alpha "good enough". Inspired by Riley (2019), we consider this portfolio approach to our sample of actively managed Norwegian equity mutual funds. The questions we will attempt to answer is the following:

ii) Can you achieve abnormal returns in the Norwegian market by betting on actively managed Norwegian equity funds with recent outperformance?

iii) Can a portfolio of these actively managed funds, constructed such that its appraisal ratio is maximized, provide the investor with additional abnormal returns compared to the average fund in this portfolio?

Replicating Riley (2019) we find that our portfolios delivered positive but not significant alphas before accounting for transaction costs. However, we are not able to distinguish the optimal and equal weight portfolios, which speaks against any added value from the Treynor and Black (1976) model and support previous critiques about the model being too sensitive to alpha forecasts. These results hold only when the portfolios are rebalanced monthly. In general we find the performance persistence to be short lived, making a strategy based on historical performance likely to be unprofitable because of the high transaction costs from frequent rebalancing. This suggest past performance is a bad portfolio formation metric for mutual funds implying other ways of forming portfolios is needed in order to find feasible strategies for investors.

The questions we seek to answer should be of interest to a large number of investors, ranging from amateurs with little or no knowledge of financial markets to large institutional investors. In general, we are under the impression that the research on Norwegian mutual funds is quite scarce compared to research on funds operating in bigger markets like the United States, Germany, and France. That being said, the invested NOK amount in Norwegian funds is not negligible, which in our opinion makes the need for specific studies on the Norwegian market important. According to the Norwegian Fund and Asset

Management Association (VFF, 2019), the allocation to mutual funds registered with the VFF surpassed 1200 billion NOK in 2019, whereof about half of this is placed in equity funds. Furthermore, they report that private consumers allocate 90 percent of their funds to actively managed funds, and only 10 percent to passive funds. For their sake, let us hope active management keep their word and in fact are able to deliver abnormal returns for its investors, or at least cover their costs. In total we believe shedding light on the value of active management is still a highly relevant and important topic in finance.

The remainder of this paper is organized as follows. Section 2 explains the most important theoretical concepts and present existing literature on mutual fund portfolios and performance. Section 3 describes our data set and collection process. Section 4 introduces the empirical methodology, and section 5 presents the results of our analysis. Section 6 discusses limitations to our paper and areas for further research. Section 7 concludes.

2 Background and Literature review

The purpose of this section is to provide the relevant background information and theory for understanding the applied methodology and our results. In section 2.1 we will give a brief introduction to mutual funds and the Norwegian mutual industry. Section 2.2 will present and discuss existing literature on mutual fund performance and formation of fund portfolios, while section 2.3 explains the key theoretical framework we use in our analysis.

2.1 Structure of Mutual Funds

We will now explain the structure of Norwegian mutual funds, which for the most part is equal to the structure in other countries, but with some specific features. At a glance, mutual funds are investment units that brings money from a number of investors together and invests these money in stocks, bonds, money-market instruments, other type of securities, or some combination of these investments. Before investing any money, the fund needs to decide what type of assets it will invest in, and make this decision clear to investors. The exact composition of the portfolio is decided by the investment goals set by the fund management. Furthermore, every mutual fund is required to have a separate legal entity, which is managed by an investment company with a concession in the country where the fund is registered. In Norway, all funds report the price of their shares to Oslo Stock Exchange every business day, typically after all major exchanges on which the fund owns securities are closed. The price, the per-share value of a mutual fund's assets minus its liabilities, is called the Net Asset Value (NAV). The NAV is what an investor needs to pay in order to get one share in the mutual fund. For open-ended mutual funds, which is the focus of our thesis, there is no restriction on the number of shares the fund can issue, meaning any investor can buy as many shares as they want. When buying a share, the money is added to the same pot, shared by all investors.

To make it easier to compare mutual funds, the Norwegian Fund and Asset Management Association (VFF), divide funds into different groups and sub-groups. The four main groups are equity funds, debt funds, money market funds and hybrid funds. Equity funds are further divided into sub groups, depending on their investment mandate and actual asset allocation. For example, equity funds investing primarily in Norway are classified as

Norwegian equity funds, funds investing primarily in the Nordics are classified as Nordic equity funds etc. Also, funds investing primarily in a particular industry, may be classified within an industry mark. Besides equity funds, there are also bond and money market funds which both are investing in debt instruments, and are recognized for having both lower volatility and expected returns than equity and hybrid funds.

Norwegian equity mutual funds, as classified by VFF, are required to have an exposure of 80 percent or more in domestic equities. It follows from "Verdipapirloven" (VPL), that a mutual fund only can allocate a maximum of 5 percent to a single security. However, 10 percent is allowed if the total sum of the allocations does not exceed 40 percent. As a consequence, the number of securities in mutual funds' portfolios will always be 16 or more. Additionally, mutual funds are not allowed to short stocks or take part in futures or option markets. However, under certain regulations given by the Ministry of Finance, they may use derivatives. The above restrictions impose limitations on the mutual fund managers, and as such, making it harder to for managers of actively managed funds to beat a passively managed index fund. The limitations also raise questions about what benchmark that should be used to evaluate fund manager performance. We will return to this issue later.

As mentioned, Norwegian equity mutual funds can be either passively or actively managed. The aim of an actively managed fund is to beat the benchmark index, meaning that the fund manager must use his ability and time to produce analysis and strategies in order to deliver excess returns. The costs imposed for this effort are often quite sizeable, and investors thus need the excess return to also cover these fees. Contrary, passive management aims to track a given benchmark index, resulting in lower costs compared to actively managed funds. The choice between active or passive management is important for investors but unfortunately there is not yet a clear consensus on which alternative is the best.

2.2 Norwegian Mutual Fund Industry

Mutual funds is not something new and have been around for a long time in a number of financial markets. This goes especially for the largest and most established markets such as the United States. On the contrary, the Norwegian mutual fund industry has a considerably shorter history, but has been growing fast during the last decades. For the fund category we focus on in our thesis, Norwegian equity mutual funds, the assets under management (AUM) has grown from 82 million NOK in 1982 to 153 billion NOK at the end of 2019.

Table 2.1: Aggregate Development of Norwegian Equity Mutual Funds

The table presents annual summary statistics for Norwegian Equity Mutual Funds in the period from 1994 through 2019. The input data is obtained from The Norwegian Fund and Asset Management Association (VFF). Column 1 shows the average number of customers per fund in a given year. Column 2 report the average number of assets under management. Column 3 shows the average net inflow. Column 4 and 5 presents aggregated assets under management for all funds as percent of the total Norwegian equity fund market and of the total fund market, respectively. Average AUMs and Average net inflows are reported in million NOK.

Year	Average customers	Average AUM	Average net inflow	% of total equity fund market	% of total fund market
2019	2,878	1594	17	22.6	11.5
2018	3,373	1529	16	23.5	11.4
2017	3,823	1749	59	22.9	11.7
2016	4,162	1451	93	22.9	11.1
2015	3,877	1112	-50	19.9	9.6
2014	4,138	1090	-25	20.9	10.2
2013	4,634	1087	-13	22.4	12.3
2012	5,745	945	-10	24.5	12.2
2011	6,017	832	-18	24.6	12.5
2010	6,281	1062	60	26.6	15.6
2009	6,874	821	-4	24.8	13.9
2008	6,571	358	-1	19.7	8.7
2007	6,726	746	-44	23.1	12.9
2006	6.175	635	16	24.5	14.8
2005	6,854	504	-61	26.2	14.0
2004	8,342	421	-52	31.8	16.8
2003	9,281	351	-1	35.9	17.3
2002	9,024	215	-11	37.1	15.8
2001	11,302	374	-11	37.0	20.7
2000	11,537	459	-23	38.3	24.6
1999	14,255	573	7	46.1	30.8
1998	15,878	403	4	67.3	38.4
1997	14,858	604	140	80.1	47.8
1996	13,354	422	99	86.1	41.4
1995	9,689	227	7	91.9	34.1
1994	10,987	235	8	92.0	37.0

Table 2.1 shows some of the developments in the time period from 1994 to 2019. Looking at column 1, we observe a steady decrease in the average number of customers in Norwegian

equity mutual funds. Column 5 and 6 displays the percentage share of Norwegian equity mutual funds of all equity funds and of the total fund market, respectively. As for the average number of customers, these percentage shares are decreasing quite substantially. Domestic equity funds share of the total fund market in terms of AUM, decreased from 37 percent in 1994 to 11.4 percent in 2019, while within the equity fund category, domestic fund only constituted 22.4 percent in 2019 compared 92 percent in 1994. These patterns point towards a larger preference for international equity and other types of asset classes among investors. The reasons for this preference change could be many. For example, different fund types have been becoming increasingly more available along with the technological development, while at the same time, investors might find investing abroad to give diversification benefits to their portfolio. Part of the reason could also relate to the available assets among investors becoming to large for all to be placed in the Norwegian market, and as such, forcing investors to look for other alternatives. Column 2 reveal, however, that the preference for other fund types only is in relative terms, given that the average assets under management for Norwegian equity mutual funds is nearly 7 times as high in 2019 as in 1994, while at the the same time the number of domestic funds has increased.

Figure 2.1: Asset Allocation of Norwegian Mutual Funds

The figure shows the asset allocation (in percent of the total fund market) of Norwegian Mutual Funds in the time period 1998-2019. Data on assets under management for each fund category is obtained from The Norwegian Fund and Asset Management Association (VFF).

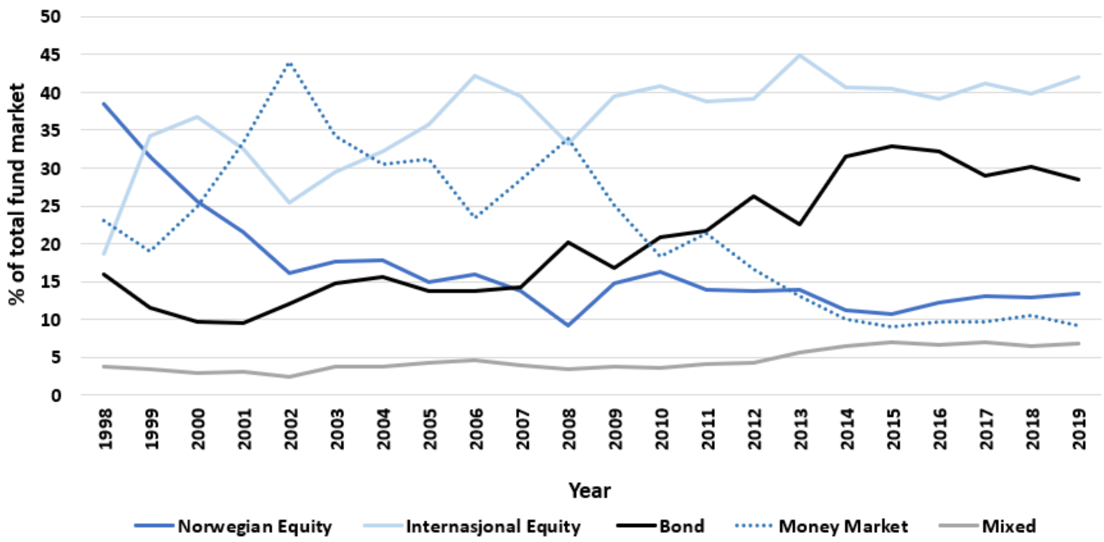


Figure 1.1 displays the asset allocation between the major fund types and groups discussed above; Norwegian equity, International equity, Bond, Money Market and Hybrid funds. Looking at the development from 1998 to 2019, it becomes clear that the percentage allocation of capital in Norwegian equity mutual funds has shrinked considerably, as also was evident in 2.1. Much of this capital seems to have gone into international equity mutual funds instead, which constituted less than 20 percent of the total fund market in 1998 but more than 40 percent in 2019. The black line shows increasing popularity of bonds during the same period, and maybe especially so in the years after the financial crisis in 2008, while the money market fund have gone in the opposite direction. In general, the preference among investors seems to have shifted towards more international equity and bonds during the past 20 years, but growth in total assets available for investment still have constituted for strong growth in the NOK amount placed in domestic equity funds.

2.3 Existing Literature

Research on mutual funds has been devoted large attention in academic literature. Initial studies focused on explaining and improving measures to evaluate performance, and over the past decades a debate has evolved about whether fund manager skill leads to persistent out-performance by funds or not. There has also been some research on methods to construct portfolios of funds instead of evaluating them as individual securities. This subsection will outline the evolution of previous research in mutual funds.

The majority of prominent studies on mutual funds have been performed in United States. Starting in the 1960s, and building on the introduction of the Capital Asset Prising Model (CAPM), Treynor (1965) , Sharpe (1966) and Jensen (1968) all created models to evaluate fund performance, whereof Jensen (1968) is the one who has gained the most attention in later literature with his famous Jensen's Alpha. By regressing a funds excess return on the excess return of the market, he introduced the alpha as a measure of fund manager ability, represented by the intercept from the mentioned regression. Using a sample of 115 US mutual funds, Jensen (1968) found that managers under-perform a passive investment in the market portfolio after accounting for management fees, with only one of the 115 funds having significant abnormal performance. The model of Jensen (1968) would later lay the foundation for the development of the multifactor performance models we know today.

A couple of decades later, contradictory studies emerged. In an updated study of Jensen (1968), Ippolito (1989) presented evidence of fund manager skill using a sample of 143 mutual funds with 20 years of data. More specifically, he found that 12 of the funds in his sample had positive alphas after fees, and that actively managed funds, on aggregate, outperformed the S&P500 index by 0.88 percent. However, Elton et al. (1993) later discarded these results, arguing that Ippolito (1989) used a faulty benchmark. Grinblatt and Titman (1989) found significant positive alphas in aggressive growth funds and funds with limited assets, implying mutual funds could possess qualities to offset expenses.

Recent studies also show conflicting evidence of skill. Studies done by Carhart (1997), Davis (2001), and Fama and French (2010) claim little evidence of skill. Carhart (1997) creates a 4-factor model expanded from the 3-factor model of Fama & French (1993) adding momentum as the 4th factor. He found that common factors explained almost all persistence in the performance of mutual funds, and that the only persistence not explained was concentrated in strong under-performance by the worst performing funds. Davis(2001) also found little evidence of skill, where the only evidence that was found was within short-run performance of the best performing growth funds and worst performing small-cap funds. Fama and French (2010) looked at the aggregate performance of the US mutual fund market, and found that if there are managers with skill to produce higher return, they are hidden in aggregate results by the performance of managers without the same skill.

On the other hand several studies claim there is meaningful evidence of skill. Kosowski et al. (2006) found that a sizable minority of managers pick stocks well enough to more than cover their costs using a sample of domestic mutual funds from 1995 to 2002. Barras et al. (2010) found a significant amount of skilled managers prior to 1996, but almost none by 2006, while Kacperczyk et al. (2014) found evidence of both stock picking skill during expansions and timing during recessions by the same managers.

While all these studies has been done on the US market, there are also some studies that have been done on the European markets, and also in Norway. Otten and Bams (2002) investigated the performance of mutual funds in the five biggest European markets. Using the Carhart 4-factor model, they found that European funds, specially small cap funds, were able to add value. They also investigated the persistence in performance, and found persistence in the UK, but not in the other four countries.

Gallefoss et al. (2015) studied the Norwegian market using daily data. They found that top funds outperform bottom funds both in stock picking and timing using the Carhart 4-factor model. They also found persistence in the bottom performers within most measures, and for the top performers in relative performance, but that actively managed funds as a whole underperforms the benchmark by approximately the management fees. They also found that the factor loadings vary significantly with time, showing the benefit of daily data. Other studies, like the unpublished works of Sørensen (2009) and Sandvik and Heitmann(2010), did not find any persistence in Norway using monthly data.

There have also been others who has illustrated the benefits of active management from a portfolio perspective before Riley (2019). Baks et al. (2001), Ľuboš Pástor and Stambaugh (2002), and Avramov and Wermers (2006) all illustrate the benefits of investing in actively managed funds by creating portfolios. The methodology to construct the portfolios vary significantly, suggesting that there isn't a clear consensus on how it should be done. Even with extremely skeptical beliefs in manager skill, allocation to actively managed funds can still be optimal. We have not found anyone that has applied similar strategies on any European markets.

2.4 Theoretical framework

This subsection will briefly present the most important theoretical frameworks underlying our analysis. We will begin by examining the Efficient Market Hypothesis (EMH) and its applicability to mutual funds. We will then explain the Treynor-Black Model used to find optimized portfolios under the assumption that markets are not perfectly efficient.

2.4.1 Efficient Market Hypothesis

A prerequisite for active management to provide value is the existence of mispriced securities. To better understand the dynamics of stock price formation in the market, we will present and discuss the EMH as laid out by Fama (1970). He defines an efficient market as a market in which prices always fully reflect available information. He further divides the market into three different forms of efficiency: weak-form, semi-strong and strong-form.

In *weak-form* efficiency, prices reflect the information contained in historic prices. If the market is weak-form efficient, it is impossible to achieve persistent abnormal returns by studying past returns. This information will already be embedded in the prices and the prices will follow a random walk.

At the *semi-strong* level of efficiency, prices reflect not only historic prices, but also all publicly available information. This means that prices will adjust immediately to new information released to the public. In this form, technical and fundamental analysis will provide no advantage and only private information is a source for abnormal returns. As such, if the markets are semi-form efficient, actively managed mutual funds will not be able to deliver any added value compared to an passively managed alternative.

In *strong-form* efficiency, prices reflects all public and private information available. It is no longer possible to have any information not already embedded in prices, making it impossible to find mispriced securities. This version is quite extreme and implies, for example, that company announcements have no impact on the stock price. Anyway, under this form it would be no way to gain any advantage even for company insiders and the market would be based on luck.

Grossman and Stiglitz (1980) introduced a paradox regarding the efficiency in markets supporting the advocates of active management. If gathering information is costly, prices cannot reflect available information. If no one is gathering this information because there is no value in it, then the information wouldn't be reflected in the prices. If someone then start analyzing the information and gains a profit, others would do it as well until the profit disappears. This represents an equilibrium, where you can analyze stocks and make money doing so, but on average not more than to cover the costs of your effort. This implies that the best managers in the market will be able to generate a significant profit through active management, while poor managers would lose money for their investors.

Berk and Green (2004) derived a model that supports the existence of skilled managers, despite the lack of evidence of persistence in fund returns. The lack of persistence in returns does not imply that no managers have skill, but investors rationally respond to past performance until diseconomies of scale offset's the managers ability to achieve abnormal returns. As the assets controlled by the manager increases, it becomes more difficult to find enough mispriced assets, making it harder to earn abnormal returns. This dynamic continues until an equilibrium between assets under management and the manager skill is reached, where the fund won't be able to deliver any abnormal return.

If the Grossman-Stiglitz Paradox and the model of Berk and Green (2004) holds true, it should be possible to identify fund manager skill based on past performance. However, we should not expect this performance to hold long into the future, since fund flows will soon make the assets under management too large for the fund manager to find enough mispriced securities.

2.4.2 The Treynor-Black Model

In Modern Portfolio Theory (MPT), investors maximize the Sharpe ratio by mixing a risk-free asset with a risky portfolio, where the risky portfolio is the market portfolio, which in most cases would be a passive investment in some index. The Treynor and Black (1973) model attempts to construct a portfolio under conditions such as the information-inefficient market equilibrium proposed by Grossman and Stiglitz (1980). Treynor and Black (1973) argue that the risky portfolio should be comprised of an investment in a passive market index and an active portfolio of mispriced securities. The model provides a framework to identify the portfolio of mispriced securities, that can be combined with the index portfolio to obtain the optimal risky portfolio. As such, the model assumes the same mean-variance criterion as in MPT, but differs in that the optimal risky portfolio now allows fund managers to take a larger position in securities they believe is not efficiently priced.

We will now provide a short review of how the Treynor-Black optimal risky portfolio is obtained. The set up is partly inspired by White (2003), but considerably less exhaustive in order to focus on the most basic insight. The Sharpe ratio of the risky portfolio, p, is given by:

$$S_p^2 = \frac{[w_A(\alpha_A + \beta_A R_M) + (1 - w_A)E[R_M]]^2}{w_A^2(\beta_A^2\sigma_M^2 + \sigma_A^2) + (1 - w_A)^2\sigma_M^2 + 2w_A(1 - w_A)\beta_A\sigma_M^2} = S_M^2 + \frac{\alpha_A^2}{\sigma_A^2} \quad (1)$$

where w_A, β_A and σ_A^2 is the weight, beta and residual variance of the active portfolio, respectively. α_A is the risk adjusted return of the active portfolio expected by the fund manager, while $E[R_M]$ and σ_M^2 is the expected return and variance of the passive market portfolio. From the right hand side of the equation, we observe that the appraisal ratio (α_A/σ_A) of the active portfolio determines its marginal contribution to the Sharpe ratio of the passive strategy. Since investors seek the highest possible Sharpe ratio, the weights to the optimal active portfolio is thus calculated such that its appraisal ratio is maximized. This is done by choosing the weight, w_i, for the i^{th} security out of n mispriced securities, to be:

$$w_i = \frac{\frac{\alpha_i}{\sigma_i^2}}{\sum_{i=1}^{n} \frac{\alpha_i}{\sigma(e_i)^2}} \quad (2)$$

Looking at (2), we first estimate the nominator for all funds and then scale the weights by dividing by the sum of all nominators for the n funds. Applying this solution to equation (1), we get

$$S_p^2 = S_M^2 + \frac{\alpha_A^2}{\sigma_A^2} = S_M^2 + \sum_{i=1}^{n} \frac{\alpha_i^2}{\sigma(e_i)^2} \quad (3)$$

which shows that the squared appraisal ratio of security i equals its marginal contribution to the risky portfolio's squared Sharpe ratio. The equation demonstrates that if there exists mispriced securities, and the forecast quality of fund managers exceed some threshold, the risky active portfolio should indeed yield superior performance compared to an single investment in a passive index.

Theoretically, the Treynor-Black model is superior to standard MPT under the assumption that fund managers are able to identify inefficiencies in security pricing. However, the model depends critically on the ability to predict abnormal returns, which has been shown to be difficult. Additionally, fund managers might have constraints in their trading mandate making the model hard to implement. For example, long-only funds are not allowed to short sell securities, which makes it necessary to impose changes to the original model. The model is luckily very flexible in regards of inferring constraints, since you only need to correctly identify one mispriced security in order for the model to yield abnormal returns compared to a passive investment.

Despite most of the discussion around the model relates to stocks, there is no reason the model cannot be applied to mutual funds as well. Similar to stocks, they are just another security paper. One difference, however, relates to the measurement of mispricing. For example, performing fundamental analysis on mutual funds would not make much sense, while it for stocks is considered "the way to go" by a large proportion of practitioners. In general, we expect the number of methods to quantify mispricing is lower for mutual funds, and that the deviation from the true price is less.

When we later apply the Treynor-Black model to our sample of mutual funds following Riley (2019), we will from the framework presented above only consider the active portfolio comprised of mispriced securities. Within the model framework this portfolio is referred to as the optimal active portfolio.

2.4.3 Appraisal ratio

Which we briefly discussed above, we construct optimal portfolios by maximizing the Appraisal Ratio (AR). The ratio compares the fund's alpha to its idiosyncratic risk. The alpha is the return achieved over the benchmark while the idiosyncratic risk is the extra risk that has been taken by diverting from the benchmark. By diverging from the market portfolio and therefore taking on idiosyncratic risk, investors should expect to receive a benefit in abnormal return. AR measures how much abnormal return they achieve compared to the additional idiosyncratic risk they take, and can thus be used as a measure of fund manager skill. AR is given by:

$$AR_p = \frac{\alpha_p}{\sigma(e_p)}$$

Where α_p is the portfolio's alpha and $\sigma(e_p)$ is the idiosyncratic risk.

.

3 Data

This section presents our main fund sample, our data sources, and the adjustments we have made.

3.1 Fund Sample

To build our sample of mutual funds we use information mainly from two sources; VFF and Børsprosjektet at NHH. We restrict our sample to funds that are registered with the Oslo Stock Exchange and classified as a Norwegian fund by the VFF, meaning at least 80 percent of a funds' assets are invested in Norwegian equities. Only funds with an active investment strategy is included and we exclude any fund that may use derivatives, leverage or any other type of instrument conflicting with a traditional long-only strategy. To avoid having duplicates of funds with different share classes we only include the primary fund of each fund family.

The final data set comprises all surviving and non-surviving Norwegian equity mutual funds registered at Oslo Stock Exchange in the period 2008 - 2019, resulting in a sample of 55 funds. Note that non-surviving funds either dies because it is liquidated, or because it is merged into another fund. The first usually happens to bad performing funds, while the latter usually is because of either bad performance or due to acquisitions in the mutual fund industry. Similar to Elton et al. (1996b), we assume that for any fund merged into another fund, the money is invested in the acquiring fund according to the merger terms, and thereafter treated as the same fund.

To compute the funds' returns, we have obtained historical daily and monthly Net Asset Value (NAV) for each fund from Oslo Stock Exchange Information Services, which is available through Børsprosjektet at Norwegian School of Economics (NHH). The NAV is calculated by taking the total value of all stocks in a funds' portfolio, deducting management fees and other ongoing expenses, and dividing this amount on the total number of shares outstanding. The NAV is thus net of management fees and costs but disregard any front- or back-load charges associated with purchase or sale of a share.

According to Oslo Børs Information Services the frequency of the reported daily NAV values corresponds to days where there is trading at the Oslo Stock Exchange. However, we find occasional days during our time period where there are missing observations for most of the funds in our sample. To deal with this problem we remove the observed price for every fund on such dates in order to get a consistent data set across all funds. Furthermore, the daily NAV values in Børsprosjektet are not adjusted for dividends. To capture the total return of our funds we thus had to collect information about dividend payments and adjust the NAV values for funds that had distributed one or more dividends during our sample period. Using these adjusted NAV values the total return between t and $t-1$ for fund i is computed as follows:

$$r_{i,t} = \frac{(NAV_t \; - NAV_{t-1})}{NAV_{t-1}}$$

For the daily frequency, this yields a total of 129 678 observations of total net returns for the 55 funds in the period 2008 - 2019, with an average of 250 days with price observations each year. We would prefer a longer time series for our analysis but because of missing values in the historical daily price series for some funds prior to 2008 we are not able to construct a complete data set until the end of 2007. All returns are calculated on the basis of the Norwegian Krone (NOK). Table 3.1 shows the investment styles of our funds as

Table 3.1: Fund Investment Styles

The table the investment styles of all 55 funds in our sample. The matrix corresponds to the Morningstar Style Matrix and the counted investment styles is based on the latest available categorization on Morningstar as of May 2019.

Styles	Growth	Blend	Value
Small-Cap	3	9	4
Mid-Cap	10	23	6
Large-Cap	0	0	0

categorized by Morningstar. As expected there are zero funds in the Large-Cap category. The reason is simply that the number of large-cap companies at Oslo Stock Exchange is very low and with the diversification requirements put on Norwegian mutual funds, it is impossible to obtain a Large-Cap style. Most funds are investing in mid-cap companies, and having either a growth strategy or a blend strategy. Table 3.2 provides a full list of all funds included in our sample.

Table 3.2: List of funds

The table shows a full list of the funds in our sample.

Fund Name	Inception Date	End Date	Status End of 2019	Asset Management Company	ISIN
Alfred Berg Gambak	11/06/1990	-	Active	Alfred Berg Kapitalforvaltning	NO0010105489
Alfred Berg Norge Classic	10/24/1990	-	Active	Alfred Berg Kapitalforv. AS	NO0010089402
Alfred Berg Aktiv	12/29/1995	-	Active	Alfred Berg Kapitalforv. AS	NO0010089444
Alfred Berg Aktiv II	9/15/1997	10/2/2012	Liquidated	Alfred Berg Kapitalforv. AS	NO0010105497
Alfred Berg Norge Classic	12/4/1997	4/23/2014	Liquidated	Alfred Berg Kapitalforv. AS	NO0010089519
Alfred Berg Humanfond	12/23/1999	-	Active	Alfred Berg Kapitalforv. AS	NO0010032055
Alfred Berg Norge Etisk	3/14/2002	4/23/2014	Liquidated	Alfred Berg Kapitalforv. AS	NO0010138373
Arctic Norwegian Value Creation	8/22/2014	-	Active	Arctic Fund Management AS	IE00BNGMYG44
Arctic Norwegian Equities	12/29/2010	-	Active	Arctic Fund Management AS	IE00B449S282
Sbanken Framgang Sammen	2/1/2016	-	Active	Alfred Berg Kapitalforv. AS	NO0010754146
C WorldWide Norge	7/12/1995	-	Active	C WorldWide Asset Management AS	NO0008001476
Danske Invest Norge II	12/30/1993	-	Active	Danske Invest Asset Management AS	NO0008000460
Danske Invest Norge I	12/30/1993	-	Active	Danske Invest Asset Management AS	NO0008000577
Danske Invest Norge Vekst	12/30/1993	-	Active	Danske Invest Asset Management AS	NO0008000486
Danske Invest Norge Aksj. Inst 1	4/13/2000	-	Active	Danske Invest Asset Management AS	NO0010047228
Danske Invest Norge Aksj. Inst 2	11/28/2006	-	Active	Danske Invest Asset Management AS	NO0010340748
DNB Norge (Avanse I)	1/13/1982	3/21/2014	Liquidated	DNB Asset Management AS	NO0003603607
DNB Norge (I)	1/11/1984	3/21/2014	Liquidated	DNB Asset Management AS	NO0005259705
DNB Norge (Avanse II)	12/7/1990	-	Active	DNB Asset Management AS	NO0008000627
DNB Norge Selektiv E	6/13/1994	-	Active	DNB Asset Management AS	NO0008000007
DNB Norge	8/4/1995	8/23/2019	Liquidated	DNB Asset Management AS	NO0010338064
DNB Norge (III)	2/6/1996	8/23/2019	Liquidated	DNB Asset Management AS	NO0010336944
DNB SMB A	3/16/2001	-	Active	DNB Asset Management AS	NO0010337819
NB-Aksjefond	9/1/1996	10/21/2013	Liquidated	Eika Kapitalforvaltning AS	NO0008001302
Eika SMB	3/31/1998	10/21/2013	Liquidated	Eika Kapitalforvaltning AS	NO0008001369
Terra Norge	4/1/1998	10/21/2013	Liquidated	Eika Kapitalforvaltning AS	NO0008001849
Eika Norge	9/4/2003	-	Active	Eika Kapitalforvaltning AS	NO0010199086
FIRST Generator S	10/15/2010	-	Active	FIRST Fondene AS	NO0010584105
Fondsfinans Norge	11/1/2002	-	Active	Fondsfinans Kapitalforvaltning AS	NO0010165764
PLUSS Markedsverdi (Fondsforvaltning)	1/16/1995	-	Active	Fondsforvaltning	NO0010606080
PLUSS Aksje (Fondsforvaltning)	12/27/1996	-	Active	Fondsforvaltning	NO0010606072
FORTE Norge	3/3/2011	-	Active	Forte Fondsforvaltning AS	NO0010601271
FORTE Trønder	4/9/2013	-	Active	Forte Fondsforvaltning AS	NO0010665441
Handelsbanken Norge	3/3/1995	-	Active	Handelsbanken NUF	SE0009696750
Holberg Norge A	12/28/2000	-	Active	Holberg	NO0010073224
KLP AksjeNorge	3/10/1999	-	Active	KLP Kapitalforvaltning AS	NO0010272388
Landkreditt Norge	6/20/2006	6/24/2016	Liquidated	Landkreditt Forvaltning AS	NO0010279011
Landkreditt Utbytte A	2/28/2013	-	Active	Landkreditt Forvaltning AS	NO0010662836
Nordea Vekst	7/13/1983	11/11/2016	Liquidated	Nordea Fondene	NO0010325707
Nordea Avkastning	9/7/1983	-	Active	Nordea Fondene	NO0010325699
Nordea Kapital	3/1/1995	-	Active	Nordea Fondene	NO0010325715
Nordea Norge Verdi	2/6/1996	-	Active	Nordea Fondene	NO0010325731
Nordea SMB	5/21/1997	11/11/2016	Liquidated	Nordea Fondene	NO0010325749
Nordea Norge Pluss	4/27/2011	-	Active	Nordea Fondene	NO0010605637
ODIN Norge C	6/24/1992	-	Active	ODIN Forvaltning	NO0008000379
Pareto Investment Fund A	1/7/1985	-	Active	Pareto Asset Management AS	NO0010040496
Pareto Aksje Norge A	9/10/2002	-	Active	Pareto Asset Management AS	NO0010160575
Storebrand Norge	9/21/1983	-	Active	Storebrand Asset Management	NO0008000783
Storebrand Vekst	9/9/1992	-	Active	Storebrand Asset Management	NO0008000841
Delphi Norge	5/26/1994	-	Active	Storebrand Asset Management	NO0010039688
Storebrand Aksje Innland	7/2/1996	-	Active	Storebrand Asset Management	NO0008000940
Delphi Vekst	10/20/1997	10/28/2013	Liquidated	Storebrand Asset Management	NO0010039704
Storebrand Verdi A	12/22/1997	-	Active	Storebrand Asset Management	NO0008000999
Storebrand Norge I	4/3/2000	-	Active	Storebrand Asset Management	NO0010044621
Storebrand Optima Norge	12/28/2000	4/15/2019	Liquidated	Storebrand Asset Management	NO0010080815

3.2 Market Proxies

Because the true market portfolio is unobservable we need a proxy for the market return in order to estimate excess market returns. Throughout our analysis we estimate alphas and residuals two times, first during portfolio formation and then when we evaluate the performance of the optimal portfolio. For the portfolio formation our benchmark should be suitable to measure fund performance, as we want to identify the top performing funds during the past 12 months. When evaluating the optimal and equal weight portfolios from an investor's point of view we believe it is most relevant to use an investable benchmark, to account for the costs of investing in the market index. The investable benchmark therefore serves as an equivalent passive strategy.

3.2.1 Optimal Portfolio formation

In Norway, the Oslo Stock Exchange Mutual Fund Index (OSEFX) serve as the benchmark for most Norwegian mutual funds registered at OSE. In addition to capture the market performance, the index is designed to meet specific regulation and diversification requirements put on the funds in compliance with the directives of UCITS. For example, Norwegian mutual funds are required to invest in at least 16 different securities and no individual security can have a weight of more than 10 percent. Therefore, it seems like an appropriate benchmark to use when forming the optimal portfolio given our goal of identifying skilled fund managers.

3.2.2 Optimal Portfolio Evaluation

As mentioned previously, our ultimate goal is to compare the optimal portfolio with an equivalent passive investment. Hence, the benchmark for portfolio evaluation must be investable and open to every investor. For this purpose a passively managed index fund is a natural choice. There is no such fund that tracks the full Norwegian equity market as defined by the securities comprised in the OSEAX. For our optimal portfolio evaluation we apply KLP AksjeNorge Indeks as an alternative investable market proxy. The fund seeks to track the Oslo Stock Exchange Benchmark Index (OSEBX), which is constructed by Oslo Stock Exchange to be representative of the Norwegian equity market. The index is revised twice a year, and from its introduction in 2001 the number of companies included has varied between 52 and 81.

As for our sample of actively managed mutual funds we obtain historical daily and monthly NAV values for all market proxies from Oslo Børs Information Services through Børsprosjektet at NHH. This includes the OSEFX and KLP AksjeNorge Indeks.

3.3 Factors and Risk-free rate

To estimate the Fama-French 3-Factor (FF3) and Carhart 4-Factor (FFC4) models employed in this study we need return series for the Small-Minus-Big (SMB), High-Minus-Low (HML) and Momentum (PR1YR) risk-factors of Fama and French (1993) and Jegadeesh and Titman (1993). Professor Bernt Arne Ødegaard at University of Stavanger has constructed similar factors by applying the same methodology for the Norwegian equity market using companies listed at Oslo Stock Exchange. His factors has been commonly used in studies of the Norwegian market and seems to be well accepted among academic researchers. We have obtained return series from his website for the SMB, HML and MOM factors for our entire sample period. The exact construction of the factors is described in his papers Ødegaard (2020b) and Ødegaard (2020a).

For the risk-free rate we follow the recommendations of Norges Bank (2013) and Ødegaard (2013). They suggest that the Norwegian Inter Bank Offered Rate (NIBOR) is the best proxy for the risk-free rate in the Norwegian market. This differs from common practice in international markets, where the use of T-bills is dominant. The reasoning behind using the NIBOR relates to the low liquidity of Norwegian T-bills compared to T-bills in bigger markets. Hence, we use the three-month NIBOR in our analysis as an estimate of the risk-free rate. Until 2013 the calculation and distribution of NIBOR was carried out by Norges Bank, and from then by Oslo Stock Exchange. The rate from 2008 to 2013 is thus collected from Norges Bank, while the rate from 2013 to 2019 is collected from the Oslo Stock Exchange database. The rate is quoted as a simple annualized rate assuming 360 (12) interest bearing days (months) in a year. To compute the daily and monthly simple rate we divide the quoted rate by 360 (12).

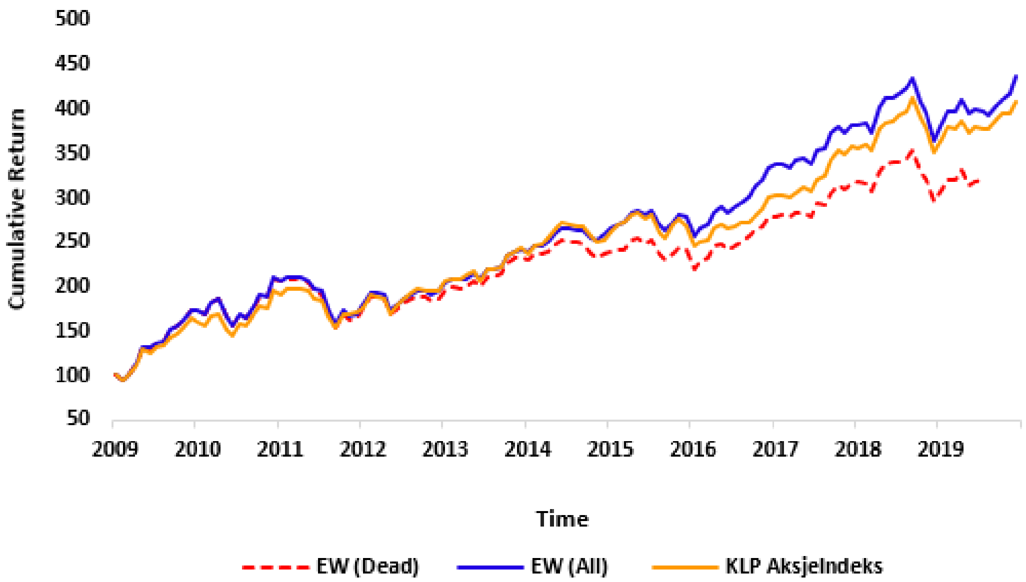
3.4 Potential biases in Mutual Fund returns

When working with mutual fund data it is important to be aware of potential biases that might arise, so they can be properly dealt with. One common bias highlighted in several previous studies is survivorship bias (see.e.g. Brown et al., 1992; Makiel, 1995; Elton et al., 1996b). Survivorship bias is the tendency of bad performing funds to be liquidated or merged by the mutual fund companies. A sample with only surviving funds will thus overestimate the returns and impose an upward bias to the average fund return in the sample. In turn, this makes the aggregated estimate of mutual fund performance higher than it should be. In order to avoid this bias it is critical to include both surviving and non-surviving funds in the data set. Figure 3.1 displays the cumulative returns of an equally weighted portfolio of all funds and of funds that died during our sample period, as well as the KLP AksjeNorge Indeks. While it is hard to distinguish the equally weighted portfolio of all funds and the KLP AksjeNorge Indeks, it is easy to see that the portfolio comprising only dead funds yields significant lower returns compared to both the KLP AksjeNorge Indeks and the portfolio of all funds. This clearly illustrates that survivorship bias would be a problem in our data set if we failed to include both surviving and non-surviving funds.

Another bias in mutual fund data is the incubation bias documented by Evans (2010) in the CRSP Survivorship-free database for US domestic equity mutual funds. The bias arises from a strategy used by some fund companies to develop new fund offerings, known as mutual fund incubation. During incubation, the fund company opens multiple funds with limited capital. At the end of the incubation period, there is a tendency that only the best performing fund(s) are opened to the public. If the return history from before the fund became open to the public are included in the database, there are return observations from a period where the fund could not be bought by the public. Additionally, Evans (2010) found that during the incubation period, the incubated funds outperformed other funds by approximately 3.5 percent per year. To our knowledge, there is no study that has addressed the presence of incubation bias in the Oslo Stock Exchange Information Services database. To account for the possible presence of a incubation bias in our data set, we apply a method proposed by Evans (2010) of removing all return observations of a fund until the fund is 3 years old, which he found to remove 95 percent of the bias. The downside of this method is that valid return observations of non-incubated funds are also removed. However, as we are not able to point out which funds are incubated and not, we don't find another way to make sure our sample is more or less free from incubation bias.

Figure 3.1: Illustration of Survivorship Bias

The figure presents the cumulative performance (in NOK) of a 100 NOK investment in different portfolios and KLP AksjeNorge Indeks from Jan. 2009 through Dec. 2019. EW(Dead) and EW(All) is based on our sample of 55 actively managed Norwegian mutual funds, where EW (Dead) is an equally weighted portfolio of funds that died during our sample period and EW (All) is an equally weighted portfolio of all funds.



4 Methodology

The methodology section consists of three subsections. The first subsection (4.1) explains the method used for performance evaluation, including model selection and a framework for bootstrap analysis. The second subsection (4.2) presents the optimization framework and process used for calculating the optimal portfolio weights. The third subsection (4.3) presents the methodology we use to assess the long-run performance of the optimal portfolio.

4.1 Performance Evaluation

An important decision when studying mutual funds is to choose an appropriate model for performance measurement. For our analysis, we first use performance measures when ranking funds and calculating weights for the optimal and equal weight portfolios, and then when evaluating ex post performance. A common way of doing this is to apply single and multi-factor models to fund returns, as done in the papers of Jensen (1968), Fama and French (1993) and Carhart (1997). The use of factor models to measure alpha certainly has it drawbacks. For example, there is a question about which factors that matters to investors, and Elton (2019) argue the lack of investable alternatives to capture the return series of some factors can lead to imprecise conclusions about performance and the value of active management. However, factor models remains popular among researchers, and using factor models make our results more comparable to previous findings, among others Riley (2019) which part of our paper is largely based upon. The next two subsections will briefly discuss our choice and implementation of multi-factor performance models.

4.1.1 Single and Multi-Factor Models

The Capital Asset Pricing Model (CAPM) as first introduced by Sharpe (1964), Lintner (1965) and Mossin (1966), seeks to explain the relationship between systematic risk and expected return of an asset. With basis in the CAPM, Jensen (1968) came up with the single factor model, which serve as the foundation for all multi-factor performance models. Jensen's alpha of fund i, α_i, is the intercept of the model as presented below in equation (1), and measure the performance relative to the market benchmark at time t:

$$\boldsymbol{r_{i,t} - r_{f,t} = \alpha_t + \beta_t \times MKT_t + e_{i,t}} \quad \textbf{(1)}$$

where $r_{i,t}$ is the return of a portfolio i in period t, $r_{f,t}$ is the risk-free rate at time t, MKT_t is the market risk premium, β_i is a fund's exposure to the market factor (non-diversifiable risk). The error term, $e_{i,t}$, has an expectation of zero and represents idiosyncratic volatility of portfolio i. Assuming the CAPM holds, a positive and significant alpha means the portfolio generates returns that are higher than expected given the portfolio's level of risk. On the contrary, a negative alpha reflects poor portfolio performance since the investors could have earned a higher risk-adjusted return by holding the market portfolio. Thus, within the framework of the single factor model, alpha can be interpreted as a measure of fund manager skill measured in terms of the positive or negative risk-adjusted return generated by the manager's portfolio.

By using the single-factor model to evaluate mutual fund performance, one implicitly assume that the market factor is sufficient to capture the investment behaviour of fund managers. In other words, if a portfolio is exposed to other risk factors than the market factor, Jensen's Alpha won't represent the true risk-adjusted return of the portfolio. Several studies have questioned the adequacy of the single-factor model in performance evaluation. Two well known extensions of the single-factor model is the three-factor model (FF3) of Fama and French (1993) and four-factor model (FFC4) of Carhart (1997). They show that the market factor is not the only relevant factor to explain the behaviour of expected stock returns. The regression specifications for the FF3 and FFC4 are shown below in equation (2) and (3), respectively:

$$\boldsymbol{r_{i,t} - r_{f,t} = \alpha_t + \beta_{1,i} MKT_t + \beta_{2,i} SMB_t + \beta_{3,i} HML_t + e_{i,t}} \quad \textbf{(2)}$$

$$\boldsymbol{r_{i,t} - r_{f,t} = \alpha_t + \beta_{1,i} MKT_t + \beta_{2,i} SMB_t + \beta_{3,i} HML_t + \beta_{4,i} PR1YR_i + e_{i,t}} \quad \textbf{(3)}$$

where SMB, HML and PR1YR are returns on value-weighted, zero investment, factor-mimicking portfolios for size, book-to-market equity and one-year momentum in stock returns. As in the single-factor model, the intercept is a measure of abnormal returns but now after controlling for a portfolio's exposure to additional risk factors. Carhart (1997) find that the four-factor model substantially improves the average pricing errors relative to the single-factor and three-factor model. Additionally, in order to correctly evaluate the performance of a fund manager, the benchmark should include risk factors that reflects all possible investment strategies of the fund manager. Table 4.1, presents summary statistics of the individual factor loading's of all 55 funds in our sample. The average loading on MKT is as expected large and strongly significant for the whole sample. The loadings on SML, HML and PR1YR factors are smaller but still significant for a large amount of the funds. It should also be mentioned that using even more factors could lead to more precise alpha estimates. For example, Næs (2009) argued that a liquidity factor was relevant for explaining the returns in the Norwegian Stock market. However, we do not find our funds to have any significant loadings in the liquidity factor, and choose therefore to leave it out of our analysis. Furthermore, including additional factors not relevant for explaining returns in the Norwegian Stock market would also introduce the risk of overfitting to become a significant problem. Based on these considerations, we believe the well documented factors of MKT, SML, HML and PR1YR is the most obvious set of factors to include, and we will take into consideration all of the three models discussed above when evaluating performance.

Table 4.1: Factor Exposure of Individual Funds

This table presents summary statistics of the individual factor exposures of all 55 funds in our sample. The factor exposures are estimated using the FFC4-model. Column 1, 2 and 3 shows the average, maximum, and minimum exposure, respectively. Column 4 and 5 shows the percentage of funds with significant coefficient estimates at the 5 % and 10 % level. The sample period is from Jan. 2009 through Dec. 2019.

	$\beta_{average}$	β_{max}	β_{min}	Significant$_{0.05}$	Significant$_{0.10}$
MKT	0.921	1.158	0.559	100.00	100.00
SML	0.076	0.382	-0.058	81.13	84.91
HML	-0.017	0.123	-0.125	73.58	75.47
PR1YR	0.021	0.094	-0.060	66.03	71.70

In the above discussion we present well known risk premiums and conclude to use a set of three factor models to evaluate performance. Since we know the SML, HML and PR1YR risk premiums exists, this is somewhat contradicting to standard financial theory, because some of the return from exposure to the omitted risk factors are likely to introduce an upward bias to alpha. In other words, evaluated against the CAPM, a fund manager can expect to increase his alpha simply by buying SML, HML and PR1YR factor exposure. The investor would then be looking at a biased alpha estimate when considering to buy this fund. However, including all three performance evaluation models in our analysis, allows us to look into the size of these potential biases that relates to model choice.

Final framework for Performance Evaluation

With our choice of factor models for performance evaluation, equation (1), (2) and (3) can be summarized to equation (4), where $f_{i,t}$ is a vector containing the different risk factors of the CAPM, FF3 and FFC4 model. Thus, risk-adjusted returns are estimated from the following equation using different sets of pricing factors:

$$\boldsymbol{r_{i,t} - r_{f,t} = \alpha_i + \sum_{i=1}^{N} \beta_i \times f_{i,t} + \epsilon_{i,t}} \qquad \textbf{(4)}$$

where $r_{i,t}$ is the return for fund i in period t, $r_{f,t}$ is the risk-free rate in period t, α_i is the alpha of fund i, and $\epsilon_{i,t}$ is the residual (idiosyncratic volatility) of fund i in period t. When estimating equation (4) we use net returns to capture the return net of fees. Additionally, the MKT is the KLP AksjeIndeks Norge and we use SMB, HML and PR1YR factor returns calculated on the Norwegian market[2].

4.1.2 Bootstrap Evaluation of Portfolio Alphas

To make more precise inferences about the significance levels of the observed performance, we apply a bootstrap procedure similar to that of Kosowski et al. (2006) and Fama and French (2010). The rationale for using bootstrapping is that financial time-series data, and in particular the cross-section return series of mutual funds, are not likely to satisfy the OLS assumptions needed for valid inference. There are several reasons for this, including non-normalities in individual fund returns and in the cross-section of mutual fund alphas. Table A1.1 in the Appendix indeed confirms the presence of non-normal monthly return series for about half of the funds in our sample, further supporting the argument for using a bootstrap to evaluate significance levels.

Implementation

The bootstrap procedure proposed by Kosowski et al. (2006) involves residual-only

[2]The returns are obtained from the website of the Norwegian professor Bent Arne Ødegaard as described in Section 3.

resampling under the null hypothesis of zero alpha. We apply the procedure to daily fund returns using the CAPM, FF3 and FFC4 models in equation (1), (2) and (3), respectively. As an example, we will explain the procedure for the FFC4 model. However, there is no difference if applied to other unconditional factor-models, except the set of included risk factors.

First, we estimate equation (3) for fund $i = \{1, 2, 3, ..., N\}$ and save the coefficient estimates $\{\hat{\alpha},\ \hat{\alpha}_{i,MKT},\ \hat{\alpha}_{i,SML},\ \hat{\alpha}_{i,HML},\ \hat{\alpha}_{i,PR1YR}\}$, the t-statistic of alpha, $\hat{t}_{\hat{\alpha}}$, and the time-series of estimated residuals $\{\hat{\epsilon}_{i,t},\ t = T_{i0}, ..., T_{i1}\}$ where T_{i0} and T_{i1} are the dates of the first and last daily return observations available for fund i, respectively. Then, for each fund, i, we draw a random sample with replacement from the residuals we saved in the step above, creating a time-series of re-sampled residuals, $\hat{\epsilon}^{\,b}_{i,t_\epsilon}$, with the same length as the initial residual vector and where b indicates the bootstrap number. The bootstrap number is simply an index of the series of bootstraps, in other words, $b = 1$ is the first bootstrap resampling, $b = 2$ is the second bootstrap and so on. Next, using the time-series of resampled residuals and the estimated factor loadings, we construct a time-series of pseudo daily excess returns for fund, i, imposing the null hypothesis of zero true performance by construction ($\alpha_i = 0$):

$$\boldsymbol{r^b_{i,t} = \hat{\beta}_{i,MKT}MKT_t + \hat{\beta}_{i,SML}SML_t + \hat{\beta}_{i,HML}HML_t + \hat{\beta}_{i,PR1YR}PR1YR_t + \hat{\epsilon}^b_{i,t_\epsilon}} \quad \textbf{(5)}$$

We then estimate the FFC4 model again, using the vector of artificial returns, $r^b_{i,t}$. If a bootstrap has drawn an abnormally high number of positive residuals, a positive alpha may result. Conversely, if the number of drawn negative residuals is abnormally high, a negative alpha may result. For a given bootstrap iteration, b, the above steps are repeated across all funds, $i = 1, ..., N$, resulting in a draw from the cross-section of bootstrapped alphas. We then order the bootstrapped alphas in a particular draw, $\{\tilde{\alpha}^b_i, i = 1, 2, ..., N\}$, from the highest to the lowest. Repeating this process for bootstrap numbers $b = 1, 2, ..., 1000$, we build a a distribution of these cross-sectional draws of alphas that results only from sampling variation, while imposing a null hypothesis of a true alpha of zero. Percentiles of this distribution are used for inference about the significance level of the ex-post realized alphas of the individual funds used for generating the distribution. For example, the distribution of alphas for the ex-post top ranked fund, is constructed using the maximum alphas from all bootstrap simulations. The ex-post realized alpha of this fund is then compared to its belonging bootstrapped distribution, to decide if its performance is solely due to sampling variation or not. More generally, if the bootstrap iterations generate far fewer extreme positive (negative) values compared to those in the actual data, then the conclusion is that sampling variation is not the only source of high (low) alphas. Similarly, if one uses the t-statistic as a performance measure, the bootstrap procedure is implemented following the same steps as above, but for each bootstrap iteration, the sorting is done on the t-statistic rather than alpha.

To make inferences about our portfolios of mutual funds, using the bootstrapped distributions described above, we include the return series of the optimal and equal weight portfolios when making the bootstrap distribution. In other words, we consider the portfolios as possible investment choices, along with the actual mutual funds we have in our initial sample. It should be stated that the feasibility of our strategies can be questioned in a real world scenario because of transaction costs, but for now we will leave this concern out of consideration.

4.2 Forming Optimal and Equal Weight Portfolios

We use the Treynor and Black (1973) model to calculate the weights of our optimal portfolio. At the start of each month, from January 2009 to December 2019, we estimate the alpha and residual variance for every fund using daily returns from the previous 12 months. As stated previously, a similar study was performed on the US Market for the period 2000-2016 by Riley (2019). From his sample of 2234 funds, he only included funds in the top 5 percent of alpha into the optimization, estimated over the past 12 months of daily returns. Considering that our sample consists of 55 funds, where the number of alive funds ranges from 39 to 48, a threshold of 5 percent would result in a portfolio consisting of only 2-3 funds. Therefore, we believe it is favourable to increase the inclusion threshold to include a higher percentage of funds. Relying on modern portfolio theory, if a fund is not perfectly correlated with another fund, there should be a diversification benefit of adding it to the portfolio, which speaks towards no threshold at all. More specifically, we will report results from portfolios formed with funds in the top 10, top 25 and top 50 percent of alpha during the previous year, as well as all funds.

The baseline Treynor-Black Model allocates a negative weight to assets with a negative alpha. Since it is not possible to sell short a mutual fund, allowing for negative weights would create a portfolio that investors cannot hold. Accordingly, our main emphasis is on portfolios without negative weights, but we will also report results of a hypothetical unconstrained portfolio since it may reveal insight about the behaviour about the best and worst performing funds. For example, Gallefoss et al. (2015) found performance persistence in alpha among the bottom ten percent of Norwegian equity funds, meaning a negative weight in these funds would yield positive abnormal returns.

We construct our optimal portfolios as explained in section 2.4.2 using equation 2. Simultaneously, we form equal-weight portfolios of funds with a non-zero weight in the optimal portfolio. The use of equal weight portfolios is common practice in mutual fund research to compare a fund's performance with the average performance of a group of funds. As such, comparing our portfolio to the equal-weight portfolios allows us analyze the value of optimizing the composition of funds using the Treynor and Black (1973) model compared to the average fund in the same percentile.

For each portfolio, this yields 132 sets of weights, one set for each month in the period from January 2009 through December 2019.

4.3 Long Run Performance

For investors, short-term trading fees can obstruct them from rebalancing the portfolio every month as done in our baseline model. Such a strategy is not likely to be feasible in real markets, and especially not so for mutual funds where transaction costs often are significant. Thus, the persistence in the portfolio performance is important for investors considering forming portfolios based on historical performance measures. Putting the investor perspective aside, persistence is also an important issue from an academic standpoint, as observed persistence would contradict the semi-strong market efficiency hypothesis. To assess long run performance we adapt the method of Riley (2019) to see how our portfolios, on average, perform in each of the subsequent 12 months after portfolio formation. We look at the performance of portfolios formed using lagged weights of up to 12 months. Previously we calculated optimal weights using the past twelve months of daily returns, and then calculated the return with these weights for the next one month, *t+1*. Now, we calculate the return series for the next twelve months as well, *t+2* through *t+13*, for each set of weights. This yields thirteen time-series of returns, where each time series is calculated using weights of a particular lag. For example, the time series, *t+2*, are made up of returns calculated using 2 month lagged weights, *t+3* is calculated using 3 months lagged weights, and so on. Having 120 set of weights from Jan. 2010 through Dec. 2019, the resulting time series exhibit 120 months of returns. To keep the time period consistent across different lags when measuring the cross-section performance, we remove the first 12 months (the full year of 2009) of observations. The resulting alphas of the portfolios can be interpreted as how they perform, on average, in each of the subsequent twelve months after the initial portfolio formation.

5 Analysis

Our analysis is divided into two main parts in order to answer our research questions. The first part (Section 5.1) is dedicated to the performance of individual Norwegian equity mutual funds. In the second part (Section 5.2), we turn our attention to portfolios of mutual funds following the methodology of Riley (2019).

5.1 Fund Sample Performance

We will begin the analysis by examining the aggregate mutual fund performance of all the funds in our sample. We then turn to bootstrap simulations to decide whether the performance of top and bottom performing funds should be attributed to pure chance, or if managerial skills actually exists. As mentioned previously, a similar analysis was performed by Gallefoss et al. (2015) for the time period from year 2000 through 2010. They found superior and inferior performance among the top and bottom funds, respectively, that could not be attributed to luck. Our analysis differs on a couple of points. Firstly, we use monthly instead of daily return data. Secondly, we take an investor's perspective by evaluating fund performance using an investable benchmark tracking the Norwegian equity market as our market factor, rather than a pure index that is considered "fair" to the fund managers. Despite pure indexes are commonly used there are some recent studies suggesting this might be a bad idea, and especially so if the aim is to advice investors on the choice between actively and passively managed funds. For example, Elton (2019) questions the relevance of using pure indexes that are not easily available for investment to the investor, as they argue using indexes not considering transaction costs when evaluating mutual fund performance might lead to wrong conclusions about the value of active management. As such, we would also prefer using SML, HML and PR1YR factor returns net of transaction fees, but we are not able to find any substitutes for these that considers the cost of implementing the strategy.

5.1.1 Aggregate Fund Performance

Table 5.1 shows aggregate regression estimates for our full sample of funds measured using three different performance models. We find that the average mutual fund performance after costs, as measured by alpha, is 0.74 percent per year using the CAPM model. For the FF3 and FFC4 model the performance is closer to zero, at -0.35 and 0.11 percent, respectively. Relying on the FFC4 alpha, the aggregated Norwegian equity mutual fund is thus able to cover its costs, but do not deliver any abnormal performance to investors. The variation in alpha estimates across models comes as a result of exposure to the included risk factors, and is thus reliant upon the choice of performance benchmark. Multi-factor models provide valuable information to the investor. For example, the coefficient estimates for the FFC4 model can provide insight about the investment strategy, which in our case show that the aggregated portfolio have a loading on the market factor at about 1, and a positive loading of 0.14 in SMB, indicating that the funds slightly favour small companies over large companies. The loadings on the HML and PR1YR factors are smaller and indistinguishable from zero, meaning fund managers do not seem to prefer stocks with high book to market values over stocks with low book to market values, or stocks that have recently outperformed the market. In fact, the slightly negative coefficients suggests the opposite.

Overall, all of the models seem to explain the aggregate returns fairly well with adjusted R-squared of about 95 percent. The high value of R-squared indicates that the average fund in our sample follows our passive index relatively close, and that the amount of idiosyncratic risk taken by our sample of funds therefore is low. This could later be a problem when implementing the strategy of Riley (2019), where idiosyncratic risk make up the denominator in the optimal weight calculation. The estimation of idiosyncratic risk might contain some noise and just a small amount of noise could have a big impact on the optimal portfolio composition if the value is small to begin with. Weights would thus be highly sensitive to very small values for noise in our estimations. We will return in more depth to this issue in the next subsection.

Table 5.1: Aggregate Fund Regressions

The table shows output from regressions estimated with the CAPM, FF3 and FFC4 performance evaluation models, using monthly data, on an equally weighted portfolio comprised of our full sample of mutual funds. The t-statistics of the intercept and coefficient estimates is reported in the parentheses, and *, ** and *** indicates significance levels of 10 %, 5 % and 1 %, respectively. Alphas are reported in percent per year.

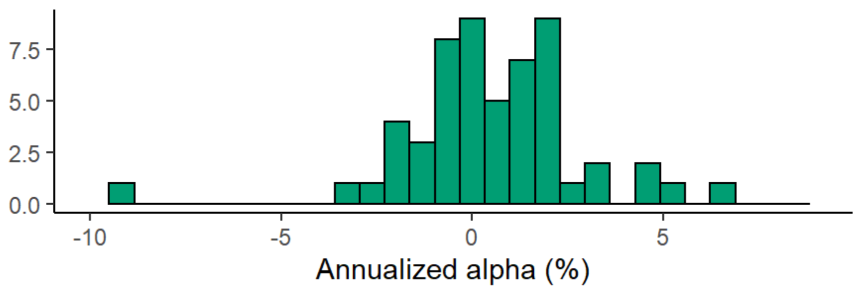
	α	β_{MKT}	β_{SMB}	β_{HML}	β_{PR1YR}	$R^2_{Adj.}$	DW
CAPM	0.74	0.992***				0.942	1.83
	(0.69)	(45.96)					
FF3	-0.355	1.062***	0.142***	-0.012		0.950	1.84
	(-0.34)	(43.24)	(4.83)	(-0.47)			
FFC4	0.114	1.053***	0.139***	-0.012	-0.026	0.950	1.83
	(0.10)	(41.11)	(4.74)	(-0.49)	(-1.2)		

Turning to figure 5.1 showing the cross-section alpha distribution for our sample, we observe the same pattern across the different models as in table 5.1. Alphas estimated using the CAPM model (Panel A) are tilted somewhat to the right, while alphas from the FFC4 model (Panel B) shifts the distribution slightly to the left. For all models there seem to be inferior and superior performers, represented by the observations in the tails of the distributions. Given the small size of our sample, we cannot conclude about the exact distribution of the alphas, which is one of the reasons why we believe a bootstrap analysis is necessary to make valid inferences. Also, from the two histograms the distributions seem to differ somewhat from the normal distribution and they exhibit fat tails, which further support the decision to use a bootstrap.

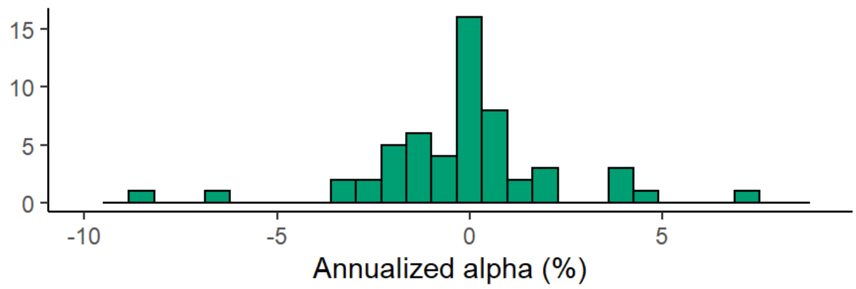
Figure 5.1: Alpha Distribution of Individual Funds

The figure displays alpha distributions of the 55 funds in our sample for the CAPM (Panel A) and FFC4 (Panel B) models. Each of the fund alphas is estimated using monthly data for as long as the fund was active during our sample period (2009-2019). The x-axis shows the annualized alpha estimates. The y-axis shows the number of funds that falls within a given alpha range.

(a) Estimation model: CAPM



(b) Estimation model: FFC4



5.1.2 Separating Luck from Skill

From the the fat tails in the cross-section alpha distributions in the previous section, we observed that some funds performed notably better than the average fund and some funds performed a lot worse. There can be two possible explanations for this, either it is a result of superior and inferior fund manager skill, or it could come from good and bad luck. To distinguish between luck and skill, we will use a bootstrap procedure similar to Kosowski et al. (2006) and Fama and French (2010) as described in section 4.2.

The results of the bootstrap simulations are presented in table 5.2. We report results using both alpha and the alpha t-statistic as ranking measures. For each panel, the first row in the top (bottom) funds table shows the actual alpha estimate of funds in the right (left) tail of the performance distribution, as well as a portfolio of the top (bottom) three funds. The second line shows the observed t-statistic that belong to each of these alpha estimates, while the third line reports the p-value of the ranking measure, which is derived from the bootstrapped distribution of this same measure. For example, when funds are ranked on the t-statistic, the p-value is the probability to observe a higher (lower) t-statistic than the observed t-statistic in the right (left) tail of the belonging bootstrapped distribution of t-statistics. When funds are ranked on alpha the interpretation is the same, expect that the p-value describe probabilities about alpha instead of the t-statistic.

Table 5.2: Performance of Top and Bottom Funds

The table present the results from a cross sectional bootstrap analysis of mutual fund performance among Norwegian equity mutual funds in the period from Jan. 2009 through Dec. 2019. To be included in the analysis we require the fund to have at least 36 months of observations. The methodology for the calculations follows Kosowski et al. (2006) and is explained in section 4.2. Panel A shows results estimated with the CAPM model. Panel B shows results estimated with FFC4 model. Actual alphas (annualized) and actual t-statistics are reported for the the three best (worst) performing funds, as well as an equally weighted portfolio of the top (bottom) three funds. Column 1-4 contain the results when funds are ranked according to their t-statistic, while column 5-8 shows results when funds are ranked according to their alpha. The p-values of the ranking measure is derived from the bootstrapped distributions of the ranking measure and are based on 10,000 resamples.

Panel A: CAPM model

Top Funds	Ranked on t-value				Ranked on alpha			
	Best	2nd	3rd	Top 3	Best	2nd	3rd	Top 3
Alpha	4.87	6.79	3.53	5.07	6.79	5.45	4.88	5.71
t-stat	2.23	1.95	1.90	2.03	1.95	1.53	1.48	1.65
p-value	0.54	0.43	0.24	0.42	0.44	0.33	0.21	0.36
Bottom Funds	**Ranked on t-value**				**Ranked on alpha**			
	Worst	2nd	3rd	Bottom 3	Worst	2nd	3rd	Bottom 3
Alpha	-9.42	-1.48	-1.70	-4.21	-9.42	-3.24	-2.70	-5.12
t-stat	-1.86	-1.45	-1.14	-1.49	-1.86	-0.68	-0.73	-1.09
p-value	0.84	0.93	0.98	0.94	0.14	0.90	0.91	0.52

Panel B: Carhart 4-factor model

Top Funds	Ranked on t-value				Ranked on alpha			
	Best	2nd	3rd	Top 3	Best	2nd	3rd	Top 3
Alpha	4.05	7.12	3.62	4.93	7.11	4.28	4.17	5.19
t-stat	2.16	1.92	1.85	1.98	1.92	0.99	1.21	1.37
p-value	0.61	0.49	0.27	0.48	0.43	0.69	0.43	0.53
Bottom Funds	**Ranked on t-value**				**Ranked on alpha**			
	Worst	2nd	3rd	Bottom 3	Worst	2nd	3rd	Bottom 3
Alpha	-8.45	-1.94	-6.52	-5.64	-8.45	-6.52	-3.33	-6.10
t-stat	-2.01	-1.35	-1.18	-1.51	-2.01	-1.18	-1.12	-1.44
p-value	0.72	0.96	0.97	0.93	0.23	0.19	0.77	0.31

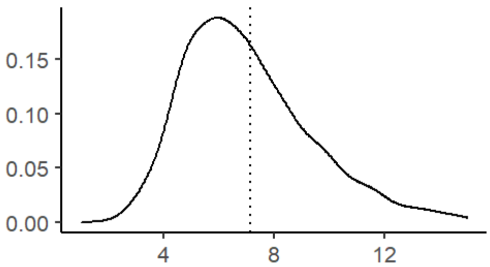
For both the CAPM (Panel A) and FFC4 model (Panel B), we can't find statistical evidence of superior (inferior) performance of the best (worst) performing funds that can be attributed to the skillset of fund managers using a 5 percent significance level. In fact, none of the results are significant even using a 10 percent significance level, with the lowest p-value across both evaluation models and ranking measures being 0.14. Thus, the result holds both when funds are ranked on their alpha t-statistic and when they are ranked on alpha itself. This result differs from the findings of some previous studies on mutual fund performance. For example, both Kosowski et al. (2006) and Fama and French (2010) find that the performance of top and bottom performing US mutual funds cannot be explained by luck, and thus concludes that there exist superior and inferior fund managers. In Norway, as we have mentioned before, Gallefoss et al. (2015) find superior and inferior actively managed Norwegian equity funds in the time period from Jan. 2000 through Dec. 2010 using daily data. There can be several explanations for why we do not come to the same conclusions. For example, in the time period 2000-2010, there was two major market crashes: the dot-com bubble and the 2008 financial crisis. While in our time period, the market has generally been trending upward. Could it be that the skillset of managers become more obvious during distressed market conditions? Some studies suggest so, see for example Kosowski et al. (2006) who finds the value of active management to be greater during recession periods.

Figure 5.2 display bootstrapped distributions of alphas for the best and worst performing funds in the cross section estimated using the FFC4 model, where the actual alpha value for each fund is plotted with a vertical dotted line. The plots provides a visual interpretation of the p-values in table 5.2. For example, (a) provides the bootstrapped distribution of alphas for the fund with the highest realized alpha across all bootstrap simulations. That is, it is the collection of the highest alphas observed in each of the 10,000 bootstrap simulations. The distribution is centered around 6.1, while the observed alpha is 7.11. From the plot it becomes clear that the number of bootstrapped alphas above 7.11 is too high (>500) for concluding that the alpha comes as a result of fund manager skill. Figure A2 in the appendix show the same plot when the CAPM model is used for performance evaluation.

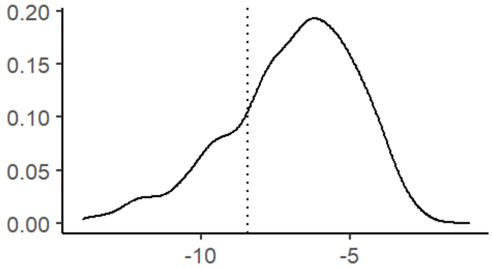
Figure 5.2: Estimated alphas vs. bootstrapped distributions of alpha for individual funds

The figure displays kernel density estimates of the bootstrapped distributions of alpha (solid line). To produce the distributions we follow the methodology of Kosowski et al. (2006) as described in section 4.2, using alpha as ranking measure and the FFC4 model for performance evaluation. The x-axis shows annualized alphas in percent and the y-axis shows the kernel density estimate. The vertical dotted line represents the actual (estimated) fund alpha. The top (bottom) fund refers to the fund with highest (lowest) alpha during our sample period (2009-2019). Similarly, the top (bottom) 3 funds refers to an equally weighted portfolio of the 3 funds with highest (lowest) alpha during the same period.

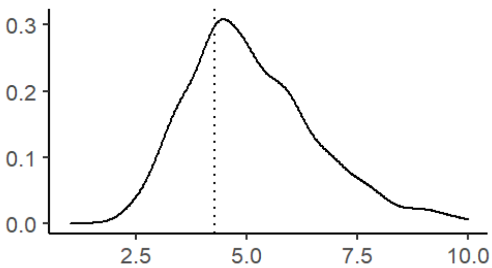
(a) Top Fund



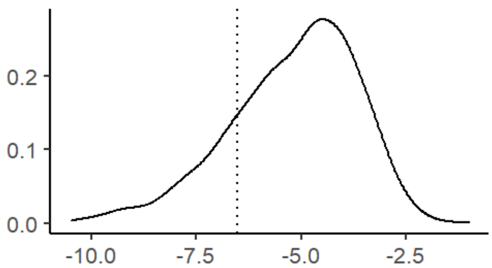
(b) Bottom Fund



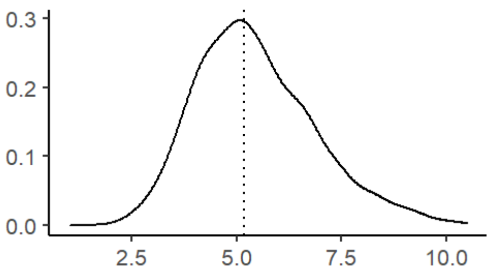
(c) Second Best



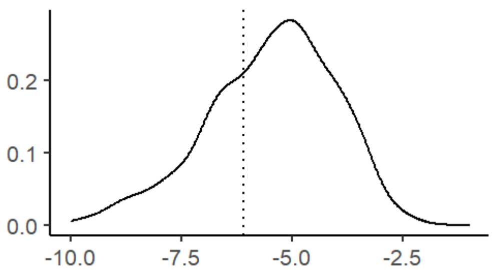
(d) Second Worst



(e) Top 3 Funds



(f) Bottom 3 Funds



For investors, our findings are bad news because it seems very hard to earn long term abnormal returns by investing in a single fund, at least over our time period of 2009-2019. On the other hand, the good news is investors should expect no loss in doing the simplest thing available, namely investing in a passive index fund. This conclusion holds both when adjusting for the risk factors included in the FFC4 model, and when we only take into consideration the market factor represented by a passive index fund. The positive and negative alphas (t-statistics) of the best and worst performers presented in table 5.2 shows that there are funds who have delivered positive and negative alphas (t-statistics), but we are not able to attribute these results to superior or inferior skill using the bootstrap methodology of Kosowski et al. (2006). This means investors should, ex-ante, not be able to identify which funds will deliver long-term abnormal returns in the future. Thus, investors might be able to earn long term abnormal returns simply by luck, but they will not be able to do so by identifying fund manager skill. Given these results, we believe investors are better off investing in a passive index fund than putting their money in a single actively managed fund.

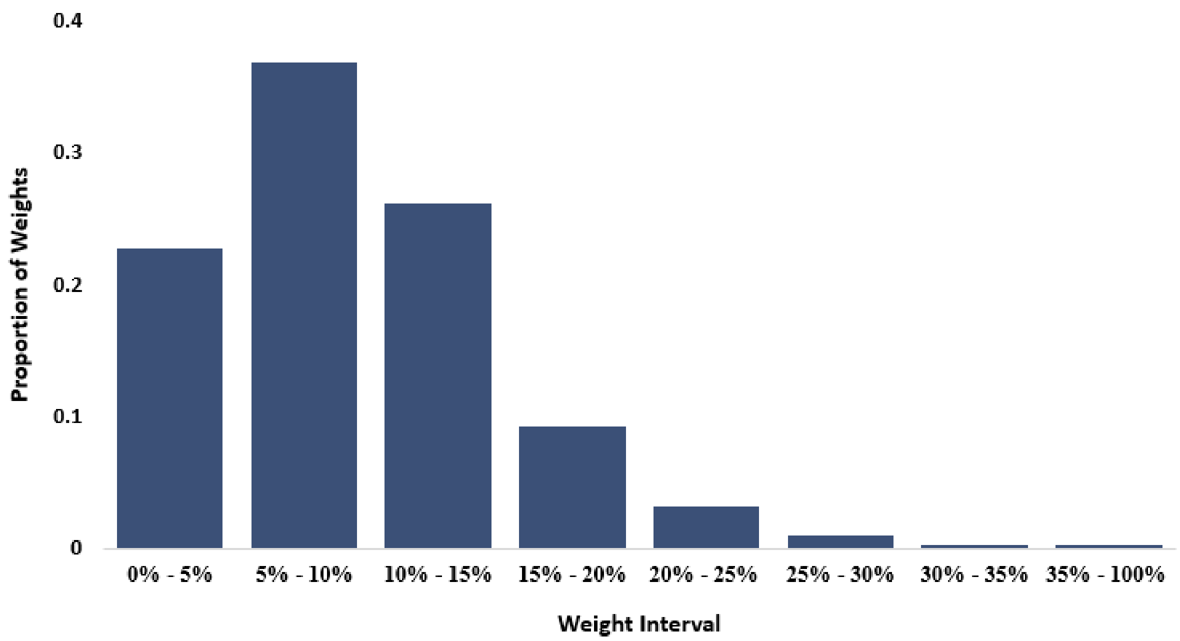
5.2 Portfolios of Actively Managed Mutual Funds

Despite the lack of performance among individual funds over the sample period, it might be that a portfolio of funds is able to deliver a positive alpha that is not a result of chance. Contrary to the extensive literature on mutual fund performance and persistence, the literature looking at the investment in mutual funds as a portfolio decision is rather limited. In this section we will look at whether portfolios of mutual funds can be superior to an investment in a passively managed index fund. Our methodology and analysis in this section is largely based upon Riley (2019), who finds that portfolios of US actively managed mutual funds formed with the Treynor-Black model are able to deliver superior performance compared to an investable market factor and to an equal weight portfolio of the funds in the optimized portfolio. The methodology for portfolio formation is described in section 4.2.

5.2.1 Performance of the Optimal Portfolio

Figure 5.3: Distribution of the 75th Percentile FF3 Optimal Weights

The figure shows the distribution of the portfolio weights for the optimal portfolio formed using the FF3 model. To calculate the weights for a given month, we estimate alpha and residual variance for each fund using daily returns for the past year as input to the Treynor-Black optimization framework. The time period is Jan. 2009 through Dec. 2019 and only funds in the top 25 percent of alpha during the previous year is included.



To begin with, we will focus our attention to the portfolio formed with the FF3 model, and with weights calculated using funds in the top 25 percent of alpha during the previous year. When evaluating the performance of our portfolios we will use the the CAPM, FF3 and FFC4 models as presented in section 4.1. In general, we find that portfolios formed with the FF3 model consistently yields the best results. This is true for both optimally and equally weighted portfolios. To ease interpretation we have annualized all measures presented throughout our analysis.

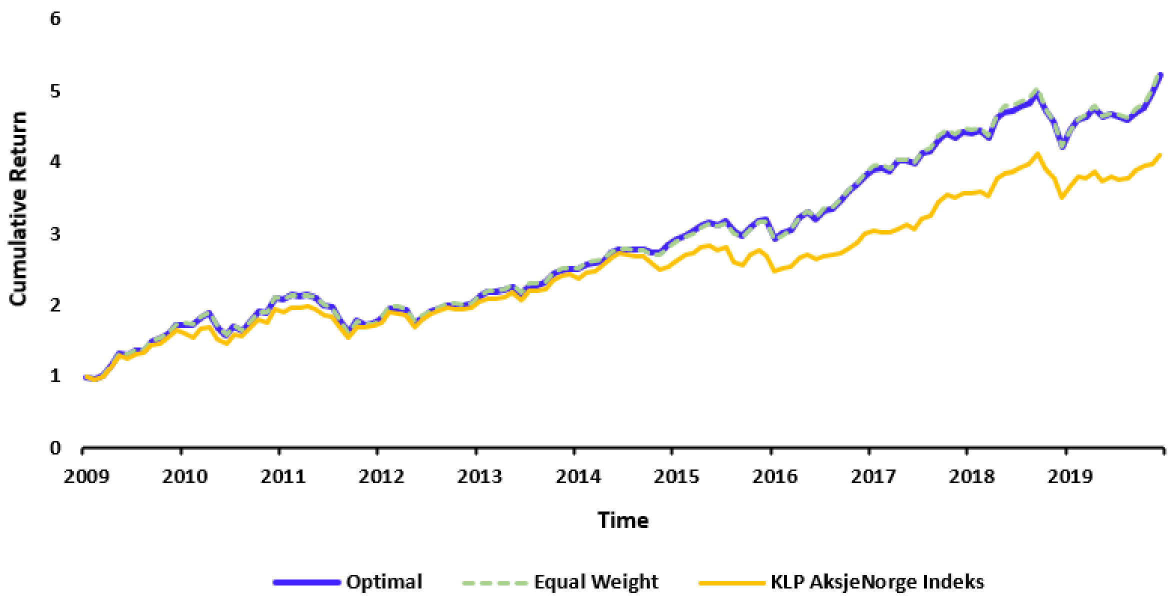
Unlike Riley (2019), who put an upper limit of 10 percent allocated to any fund, we do not put an upper limit on the weights to force the portfolio to be diversified across a minimum number of funds. However, this does not produce a portfolio with low diversification. Figure 5.3 shows the distribution of the optimal portfolio weights calculated at the beginning of each month from January 2009 to December 2019, and only weights in the top 25 percent of alpha are included. As we can see, the number of weights above 20 percent is very low, and about 85 percent of the weights are below 15 percent. Thus, the optimal portfolio seems to be well diversified across the funds available each month. The number of active funds in our investment universe is each month between 37 and 47, meaning the number of funds in the top 25 percent of alpha in each month is between 9 and 12. However, the actual number of funds with a positive weight in the optimal portfolio is between 7 and 12, since the portfolio only has positive weights in funds with a positive alpha. Compared to Riley (2019), who had between 51 and 70 funds in his portfolios each month, we have considerably fewer due to the size of our sample.

By using historical return data we seek to create portfolios with a better risk-reward profile than both our benchmark and the average fund in our portfolio. Comparing our result to the benchmark is done to test if our portfolio delivers superior performance to an equivalent passive investment, while the comparison with the equal weight portfolio is performed to test if there is value in optimally allocating between the funds with a positive alpha during the past 12 months compared to the performance of the average fund in the sample. Since the strategy is based on historical returns, it is largely dependent on performance persistence among previous winning funds. In addition to looking into the value of having a portfolio of funds, our analysis can therefore also help to answer questions about the persistence among the top performers.

The intention of using the Treynor-Black model is to maximize the Appraisal Ratio (AR) of a portfolio, not its total return. Despite this, we find that the optimal portfolio delivers better return than our benchmark in the period of our analysis. However, we are unable to distinguish the performance of the optimal portfolio from the equal weight portfolio, which suggests that optimally allocating based on past performance does not provide any significant benefits to picking the average previous winner. Figure 5.4 shows the growth of a 1 NOK investment in January 2009 held until the end of December 2019. The optimal portfolio delivers a return of 4.22 NOK, which is substantially better than the 3.09 NOK return of the benchmark, but marginally lower than the 4.29 NOK return of the equal weight portfolio. That being said, the optimal portfolio could still provide benefits in risk-adjusted measures compared to its equal weight counterpart, which we will look at next.

Figure 5.4: Performance of the Optimal and Equal Weight Portfolios

The figure shows the cumulative performance of a 1 NOK investment in the optimal portfolio and the equal weight portfolio, as well as the investable market factor (KLP AksjeNorge Indeks), from Jan. 2009 through Dec. 2019. The optimal portfolio is formed using the FF3 model and comprises only funds in the top 25 percent of alpha. The two portfolios are rebalanced monthly and the return calculation assumes no transaction costs.



Continuing the evaluation of our results, Table 5.3 shows several performance measures for the portfolios constructed from the 75th and 90th top alpha percentile. Similar to returns, we find that the optimal portfolio performs well compared to an passive investment in the benchmark (KLP AksjeNorge Indeks), but is more or less indistinguishable from the equal weight portfolio, performing marginally worse in terms of absolute return and marginally better in terms of risk. Optimally investing within the funds reduces the return series standard deviation from 14.63 % to 14.45 %, in the 75h percentile, but it also yields a slightly lower return of 16.22% compared to 16.35% for the equal weight portfolio. For the 90th percentile the numbers are slightly better, but similar in a relative comparison to the equally weighted portfolio. The Sharpe ratios for the two portfolios are virtually the same, but considerably higher than the benchmark. Since the standard deviation of the benchmark is more or less identical to those of the optimal and equal weight portfolio, this difference comes largely from lower arithmetic return of 14.82 percent. Another interesting observation is that when the alpha percentile goes form the 75th to the 90th, the standard deviation of two portfolios return series remains at more or less the same level, while the return goes up. This could imply that there is benefits to achieve with this strategy of picking previous winners, even in a small market such as the Norwegian. For a bigger market, we might even expect these differences to be larger since there would likely be more extreme observations in the tails of the performance distribution.

Table 5.3: Standard Performance Evaluation Measures

This table presents performance measures for the optimal and equal weight portfolios formed using FF3 model. Column 1-2 and column 3-4 shows metrics for portfolios in the 75th percentile and 90th percentile of alpha during the previous year, respectively. Column 5 shows metrics for our market factor benchmark (KLP AksjeNorge Indeks). All measures are calculated over the period from Jan. 2009 through Dec. 2019.

	75th Alpha Percentile		90th Alpha Percentile		
	Optimal	Equal Weight	Optimal	Equal Weight	Benchmark
Arithmetic Average (%)	17.42	17.58	18.13	18.68	14.82
Geometric Average (%)	16.22	16.35	16.93	17.43	13.66
Standard Deviation (%)	14.45	14.63	14.44	14.69	14.43
CAPM Idiosyncratic Risk (%)	4.07	4.28	5.32	5.46	0
CAPM Beta	0.962	0.971	0.931	0.946	1
Sharpe Ratio	1.007	1.005	1.051	1.065	0.853
Treynor Ratio	0.181	0.182	0.195	0.197	0.148

To draw conclusions about the differences and patterns we observed in 5.3, we need more formal statistical tests. Table 5.4 shows the alphas for the two portfolios evaluated using the CAPM, FF3 and FFC4 models, with the alpha t-statistic in the brackets below. The alphas are significantly positive at the 5 % level for both the optimal portfolio and the equal weight evaluated with the CAPM, but not significant when evaluated with FF3 and FFC4. In addition to determining statistical significance, the t-statistics have an additional interpretation here, as they have the same interpretation as the appraisal ratio in relative terms, since the formula for the t-statistic equals the appraisal ratio times the square root of our time period. The actual appraisal ratio will differ in value but the relative interpretation remains the same, meaning the optimal portfolio has a superior appraisal ratio compared to the equal weight portfolios across all models, although the improvement is small. This is due to the lower idiosyncratic risk achieved optimizing the portfolios, since the improvement in t-value is larger than the difference in return for all models. The average alpha across the different models is 1.92 % for the optimal portfolio. while the equal weight achieves 1.85 %. When evaluated with the FF3 and FFC4, the optimal portfolio is slightly better in both abnormal return and risk, while for CAPM the optimal portfolio is only superior in risk. The AR is still higher represented by the t-value due to lower idiosyncratic risk as shown in the previous paragraph.

Table 5.4: Alphas for Portfolios Formed with the Fama-French 3-factor Model
The table presents annualized percentage alphas for the optimal and equal weight portfolios estimated using three different models for performance evaluation. The optimal and equal weight portfolio is formed using the FF3 model and includes only funds in the top 25 percent alpha during the previous year. The set of evaluation models includes the CAPM, the FF3 model and the FFC4 model. All values are estimated using monthly returns in the period from Jan. 2009 through Dec. 2019. Brackets below each measure of alpha show the alpha t-statistic.

	CAPM	Fama-French 3-Factor	Carhart 4-Factor
Optimal	2.72	1.79	1.26
	[2.14]	[1.42]	[0.94]
Equal	2.77	1.66	1.12
Weight	[2.07]	[1.27]	[0.81]

The small difference between the optimal and equal weight portfolio indicates there is little value added by optimizing the portfolio, while the low t-statistics suggest investing in previous winners add little or no value compared to an investment in a passive index fund. This observation become even more obvious considering the transaction costs that would apply from monthly rebalancing. Using a shorter time period, ending in 2017, we found significantly better results, indicating that actively managed funds have struggled more over the last two years. And perhaps more importantly, that the strategy seems dependent on the time period in which it is used, implying that strong results might only happen because of luck and not from a consistently superior strategy.

5.2.2 Portfolios with alternative formation models

In this section we will consider the performance of optimal portfolios when constructed using CAPM and the FFC4 model as well, in addition to the FF3 model we reported earlier. Table 5.5 shows the alphas for the optimal portfolios formed using different formation models, with t-values in the brackets below. The methodology fails to provide significant results for the optimal portfolios formed using the CAPM and the FFC4. As mentioned previously we find the FF3 model to be superior of the three formation models in both alpha and t-statistics across all evaluation models.

Table 5.5: Performance of Optimal Portfolios from Different Formation Models
The table presents annualized percentage alphas for optimal portfolios calculated using funds in the top 25 percent of alpha during the previous year. The 'Formation Model' is the model used to calculate the prior year alphas and residuals, which are used to identify the 75th percentile and to calculate the optimal weights. The 'Evaluation Model' is the model used to measure the alphas of the resulting optimal portfolios. The set of performance models used for formation and evaluation is the CAPM, the FF3 model and the FFC4 model. The time period is Jan. 2009 through Dec. 2019. Alpha t-statistics are shown in the brackets below each measure of alpha.

		Formation Model		
		CAPM	FF 3-Factor	Carhart 4-Factor
Evaluation Model	CAPM	2.28	2.72	2.41
		[1.78]	[2.15]	[1.94]
	FF 3-Factor	1.26	1.79	1.47
		[1.00]	[1.43]	[1.20]
	Carhart 4-Factor	0.62	1.26	1.11
		[0.46]	[0.94]	[0.84]

Table 5.6 shows the alphas for the equivalent equal weight portfolios. We find that all of the equal weight portfolios have positive alphas for all evaluation models. However, they only have a significant alpha at the 5 percent level when evaluated with the CAPM. Comparing the alphas and t-statistics to that of the optimal portfolio in table 5.5, we find that the optimal weights add very little value compared to the equal weight for all methods of formation and evaluation. This is both in alpha and t-statistic. It seems clear at this point that we cannot conclude that the optimal portfolio delivers statistically significant benefits compared to an equal weight portfolio in the Norwegian market over our sample period. .

A drawback of testing the strategy in a small market like the Norwegian, is the relatively

Table 5.6: Performance of Equal Weight Portfolios from Different Formation Models

The table presents annualized percentage alphas for equal weight portfolios calculated using funds in the top 25 percent of alpha during the previous year. The 'Formation Model' is the model used to calculate weights for the optimal portfolios in 5.5. Only funds with a positive weight in the optimal portfolio is included in the equal weight portfolio. The 'Evaluation Model' is the model used to measure the alphas of the resulting equal weight portfolios. The set of performance models used for formation and evaluation is the CAPM, the FF3 model and the FFC4 model. The time period is Jan. 2009 through Dec. 2019. Alpha t-statistics are shown in the brackets below each measure of alpha.

		Formation Model		
		CAPM	FF 3-Factor	Carhart 4-Factor
Evaluation Model	CAPM	2.24	2.77	2.56
		[1.64]	[2.07]	[1.95]
	FF 3-Factor	1.02	1.66	1.38
		[0.77]	[1.27]	[1.09]
	Carhart 4-Factor	0.41	1.12	0.97
		[0.29]	[0.81]	[0.72]

few funds available to investors. All else equal, and if there actually exists any differences, a sample of more than 55 funds would increase the chance of detecting any differences in performance between the optimal and equal weight portfolios. To illustrate this point we can look at the portfolio composition of Riley (2019) and ours. While he had a portfolio comprising between 51 and 70 funds in each month using only funds in top 5 percent of alpha, our portfolios have between 7 and 12 funds using the top 25 percent of alpha. In a larger market we would expect observations that were more extreme in both ends, which could increase the effect from calculating optimal weights, similar to what was found by Riley (2019) in the US market.

In summary our portfolios provide better realized results compared to our passive benchmark before transaction costs, but the low significance levels indicate the results might only have happened because of luck. Further, we are unable to separate the results of our optimized portfolios from the performance of the average fund in the portfolios, indicating there is not much value in the optimal weights. We will now turn our attention to the long run performance of our portfolios.

5.2.3 Long-run performance

Up to this point we have used optimal weights recalculated at the beginning of each month. In the real world, constraints like short term trading fees can be put on investors which prohibit them from reforming the portfolio every month. Also, funds may have load fees when entering or exiting the fund forcing investors to hold the portfolio for longer periods to achieve economic results. For example, the portfolio in the top 25 percent of alpha formed with the FF3 model that we described in the previous subsection, have an average turnover ratio of 25 percent. If investors were to change one-fourth of their portfolio every month, this would involve considerable transaction costs pulling alpha downwards.

We will now consider how the optimal and equal weight portfolios performs when held for a longer time period. Ideally we would also have preferred to analyze the performance of the portfolios when excluding funds that have load fees, but limitations in our data is stopping us from identifying these funds. However, Riley(2019) found that excluding funds with load-fees only reduced the sample by about 15 % and that the performance still holds in a no-load sample. Furthermore, we have found no academic evidence that load fees have a positive correlation with either fund manager skill or alpha performance, which support the hypothesis that our formation strategy would yield similar results in a no-load sample.

Table 5.7 shows the alpha of the optimal portfolio formed using the FF3 model in the first month after formation, as well as the subsequent 12 months. The time period is now 2010 to 2019, and since t+1 equals our original strategy, we have found that altering the time period has given large changes in the alpha estimate when evaluated with the FFC4. In our time period the model yields abnormal return for the first three months when evaluated using CAPM and FF3, but performs worse than the benchmark with the FFC4. Using the CAPM, the alpha remains positive for five months, while the same is true for the first three months using FF3. The model yields negative alphas for all holding periods evaluated with the FFC4. However, it is worth noting that for the FF3 and the FFC4 evaluation models, all alphas are negative after four months, indicating that previous top performing funds may underperform some time in the future. This implies that the top funds in Norway struggle to consistently recreate their success, and could in the long run perform worse than a passive investment.

Table 5.7: Long Run Performance of Optimal and Equal Weight Portfolios
The table displays annualized alphas (in percent) for the optimal and equal weight portfolios formed using lagged weights. The weights are calculated using the FF3 model and only funds in the 75th percentile of top alpha during the prior year is included. Alphas are estimated using monthly returns and three different evaluation models, which include the CAPM, the FF3 model and the FFC4 model. For each set of weights, alpha is calculated in the first month after portfolio formation (t+1) and in each of the next 12 months (t+2 through t+13). The time period is Jan. 2010 through Dec. 2019. Alpha t-statistics are shown in the brackets below each measure of alpha.

	Optimal			Equal Weight		
	CAPM	FF3	FF4	CAPM	FF3	FF4
t+1	2.46	1.48	-0.10	2.45	1.35	0.00
	(1.83)	(1.12)	(-0.07)	(1.74)	(0.98)	(0.01)
t+2	1.45	0.52	-1.13	1.31	0.12	-1.31
	(1.12)	(0.41)	(-0.82)	(0.93)	(0.09)	(-0.88)
t+3	0.92	0.04	-1.73	0.96	-0.15	-1.80
	(0.71)	(0.03)	(-1.27)	(0.69)	(-0.11)	(-1.24)
t+4	0.16	-0.69	-2.44	0.18	-0.96	-2.51
	(0.12)	(-0.55)	(-1.80)	(0.13)	(-0.71)	(-1.72)
t+5	0.02	-0.93	-2.34	0.03	-1.22	-2.59
	(0.02)	(-0.75)	(-1.75)	(0.03)	(-0.90)	(-1.75)
t+6	-0.26	-1.06	-2.28	-0.52	-1.61	-2.66
	(-0.20)	(-0.83)	(-1.65)	(-0.38)	(-1.20)	(-1.81)
t+7	-0.07	-0.94	-1.79	-0.38	-1.60	-2.19
	(-0.06)	(-0.72)	(-1.25)	(-0.27)	(-1.15)	(-1.42)
t+8	-0.01	-0.72	-1.61	-0.50	-1.66	-2.12
	(-0.01)	(-0.55)	(-1.12)	(-0.36)	(-1.24)	(-1.43)
t+9	0.15	-0.71	-1.54	-0.09	-0.95	-1.00
	(0.12)	(-0.55)	(-1.12)	(-0.06)	(-0.92)	(-0.97)
t+10	-0.20	-0.91	-1.46	-0.47	-1.51	-1.55
	(-0.15)	(-0.71)	(-1.03)	(0.34)	(-1.12)	(-1.04)
t+11	0.00	-0.70	-1.29	0.09	-0.95	-1.01
	(0.00)	(-0.55)	(-0.92)	(0.07)	(-0.73)	(-0.69)
t+12	-0.16	-0.80	-1.32	-0.46	-1.47	-1.58
	(-0.13)	(-0.62)	(-0.93)	(-0.34)	(-1.11)	(-1.08)
t+13	0.06	-0.58	-1.09	-0.07	-1.12	-1.08
	(0.05)	(-0.46)	(0.77)	(-0.05)	(-0.85)	(-0.74)

This is in line with what Gallefoss et al. (2015) found in the Norwegian markets when studying the persistence in performance of the top 10 % and bottom 10 % funds. While there was significant evidence of persistence for the worst performers, the top performers showed no persistence that lasted longer than one month in terms of alpha when evaluated with the FFC4 model. We should, however, notice none of the estimates are statistically significant, which means we cannot say any values truly are different from zero.

For stocks, most of the literature on momentum strategies find that momentum is short lived, and thus requires frequent rebalancing to be profitable. Therefore, one possible explanation of the short lived persistence could be that much of the alpha generated by our strategy comes from a momentum effect, indicated by the lower significance levels using the FF4 model versus the CAPM for evaluation. Another possible explanation is the we deal with incubation bias. If the persistence of mutual funds are short lived and their best performance is before their assets under management reach the level to offset their skill in line with the Berk and Green (2004) model, we may have excluded viable funds that were expected to perform well in the three year period before they got included in the data set.

In summary, the persistence in Norwegian fund performance seems to be short lived. The results of our models vary significantly on what factors the portfolio is exposed to, and seems very reliant on the momentum factor as well as what time period the strategy is used. Regardless of model the investor must rebalance quite often, and if short term trading fees are bigger than the gain of monthly rebalancing, which is highly likely given our results, a passive investment is superior to a persistence based portfolio approach in our sample using the strategy applied by Riley (2019). Our results have, however, the same tendency as was found in the US market, but we are unable to recreate results of the same strength. In other words, we do not find statistical evidence of either short or long-term persistence. In the next sections we will alter our parameters to investigate how sensitive our results are to our formation choices.

5.2.4 Varying the top alpha percentiles

We will now consider different top alpha percentiles when constructing portfolios. Specifically, we will use funds in the top 10, 50 and 100 percent of alpha, in addition to the previous top 25 percent, and construct optimal and equal weight portfolios as before. The results are shown in table 5.8. We observe that alphas and t-statistics increase steadily when the alpha percentile is getting narrower. This pattern support our previous statement, that it might be the Norwegian market is to small for the strategy to yield significant results, as a bigger market would allow to use a percentile even longer to the right in the performance distribution.

Table 5.8: Performance of Portfolios from Different Alpha Percentiles.

This table shows annualized alphas for optimal (O) and equally weighted (EW) portfolios for four different percentiles of top alpha. Formation model is the model used to estimate the alpha and tracking error used in the Treynor-Black portfolio optimization. Evaluation model is the model used to estimate the alphas of the optimal and EW portfolios. Alpha is estimated using monthly returns over the period from Jan. 2009 through Dec. 2019. In Panel A all funds with a full year of daily returns are included in portfolios. In Panel B-D, funds not in the top 50, 25 and 10 percent, respectively, are given a weight of zero before calculating optimal weights. Only funds with a positive weight in the optimal portfolio is included in the EW portfolio. Alpha t-statistics are shown in the brackets below each measure of alpha.

Panel A: All funds

		Formation Model					
		CAPM O	CAPM EW	FF3 O	FF3 EW	FF4 O	FF4 EW
Evaluation Model	CAPM	1.35	1.17	1.71	1.39	1.77	1.44
		[1.24]	[1.03]	[1.58]	[1.26]	[1.64]	[1.30]
	FF 3-Factor	0.51	0.11	0.92	0.40	0.99	0.47
		[0.47]	[0.10]	[0.86]	[0.37]	[0.93]	[0.44]
	Carhart 4-Factor	0.16	-0.06	0.64	0.24	0.76	0.35
		[0.14]	[-0.05]	[0.56]	[0.21]	[0.67]	[0.30]

Panel B: Top 50 %

		Formation Model					
		CAPM O	CAPM EW	FF3 O	FF3 EW	FF4 O	FF4 EW
Evaluation Model	CAPM	1.55	1.52	1.80	1.59	1.88	1.65
		[1.36]	[1.26]	[1.61]	[1.37]	[1.70]	[1.42]
	FF 3-Factor	0.70	0.44	1.00	0.59	1.08	0.65
		[0.62]	[0.38]	[0.90]	[0.52]	[0.99]	[0.58]
	Carhart 4-Factor	0.30	0.15	0.67	0.32	0.81	0.44
		[0.25]	[0.12]	[0.56]	[0.26]	[0.69]	[0.36]

Panel C: Top 25 %

		Formation Model					
		CAPM O	CAPM EW	FF3 O	FF3 EW	FF4 O	FF4 EW
Evaluation Model	CAPM	2.28	2.24	2.72	2.77	2.41	2.56
		[1.78]	[1.64]	[2.15]	[2.07]	[1.94]	[1.95]
	FF 3-Factor	1.26	1.02	1.79	1.66	1.47	1.38
		[1.00]	[0.77]	[1.43]	[1.27]	[1.20]	[1.09]
	Carhart 4-Factor	0.62	0.41	1.26	1.12	1.11	0.97
		[0.46]	[0.29]	[0.94]	[0.81]	[0.84]	[0.72]

Panel D: Top 10 %

		Formation Model					
		CAPM O	CAPM EW	FF3 O	FF3 EW	FF4 O	FF4 EW
Evaluation Model	CAPM	3.62	3.65	3.72	4.00	3.29	3.45
		[2.03]	[2.04]	[2.24]	[2.35]	[2.04]	[2.08]
	FF 3-Factor	2.40	2.25	2.49	2.68	2.08	2.19
		[1.36]	[1.29]	[1.52]	[1.60]	[1.31]	[1.34]
	Carhart 4-Factor	1.43	1.25	1.78	1.90	1.62	1.68
		[0.77]	[0.67]	[1.02]	[1.06]	[0.96]	[0.96]

Another pattern relates to the relative performance between the equal weight and optimal portfolio. When all funds are included in the portfolio optimization, i.e that all funds with a positive alpha during the previous year have a positive weight in the portfolios, the optimal portfolio delivers overall higher alphas and t-statistics than the equal weight portfolio. Then, as the alpha inclusion criteria gets stricter, the equal weight portfolio gradually approaches the optimal portfolio in terms of performance. And when only funds in the top 10 percent of alpha are included, the equally weighted portfolio generally have higher alphas and t-statistics than the optimal portfolio. Despite none of these differences are statistically significant the pattern is quite consistent across all combinations of formation and evaluation models. This could mean that as the number of funds in the optimization go down, the value of optimal weights become smaller. If this should be true, one reason might be that measuring precise alphas become more difficult the further from the middle of the performance distribution one go.

5.2.5 Long/short portfolios

Until now we have considered long-only strategies in the previously best performing funds. If there is some performance persistence among the worst performing funds, a long-short strategy might be able to provide better results. We should mention the structure of mutual funds do not allow for a fund to be sold short. However, since we are looking at funds investing in stocks, there would theoretically be possible to go short in the assets a fund holds, even though this involves some obstacles. You need ongoing and up to date information about the funds transactions, which even the most transparent funds are unlikely to supply sufficiently within a reasonable time-frame.

To construct the long-short portfolios we rely largely on the same methodology as before. The optimal portfolio now comprises a long portfolio formed as previously, but we add a short portfolio formed using the same methodology on the worst performing funds. Thus, the weights in the long portfolio sum to 1 and the weights in the short portfolio sum to -1, where the long portfolio only has positive weights and the short portfolio only has negative weights. The performance of the long-short strategy is presented in table 5.9. Panel A shows the performance of portfolios formed using funds in the top and bottom 25 percent of alpha during the previous year, while the portfolios in Panel B is formed with funds in the top and bottom 10 percent of alpha. Over the line, the alphas and t-statistics are generally higher than for the long-only portfolios formed in the same alpha percentiles, as shown in table 5.8. Portfolios formed with the FF3 model achieve the best relative results of our three formation models once again. Alphas are now also statistically significant for the FF3 model in addition to the CAPM for both percentiles when portfolios are formed with FF3 and FFC4. When evaluated with the FFC4 model, however, none of the alphas are statistically significant for any formation model.

Table 5.9: Performance of Long/Short Portfolios

This table shows annualized alphas for self-financing optimal (O) and equally weighted (EW) long/short portfolios. The portfolios comprises one long portfolio with weights that sum to 1, and one short portfolio with weights that sum to -1. Further, they have positive (negative) weights in the best (worst) performing funds during the previous year in terms of alpha. . Panel A report performance for quartile portfolios, while Panel B report results for decile portfolios. Formation model is the model used to estimate the alpha and tracking error used in the Treynor-Black portfolio optimization. Evaluation model is the model used to estimate the alphas of the resulting portfolios. Alpha is estimated using monthly returns over the period from Jan. 2009 through Dec. 2019. Alpha t-statistics are shown in the brackets below each measure of alpha.

Panel A: Top / Bottom 25 %

		Formation Model					
		CAPM O	CAPM EW	FF3 O	FF3 EW	FF4 O	FF4 EW
Evaluation Model	CAPM	2.52	2.48	3.01	2.97	3.18	3.06
		[1.88]	[1.85]	[2.14]	[2.23]	[2.41]	[2.38]
	FF 3-Factor	2.95	2.81	3.16	3.55	3.31	3.51
		[2.11]	[2.06]	[2.13]	[2.42]	[2.27]	[2.68]
	Carhart 4-Factor	0.67	0.52	1.27	1.39	1.49	1.65
		[0.42]	[0.39]	[0.97]	[1.08]	[1.16]	[1.26]

Panel B: Top / Bottom 10 %

		Formation Model					
		CAPM O	CAPM EW	FF3 O	FF3 EW	FF4 O	FF4 EW
Evaluation Model	CAPM	2.96	3.10	4.51	4.67	4.25	4.61
		[1.42]	[1.52]	[2.21]	[2.29]	[1.99]	[2.26]
	FF 3-Factor	3.48	3.60	5.27	5.52	5.09	5.36
		[1.63]	[1.73]	[2.52]	[2.71]	[2.36]	[2.60]
	Carhart 4-Factor	0.92	1.02	2.02	2.12	2.19	2.27
		[0.51]	[0.64]	[1.01]	[1.06]	[1.03]	[1.10]

As expected the results using a long-short strategy are better than the results we got when only investing in previous winnersm although only slightly. This is in line with earlier findings by Gallefoss et al. (2015) of performance persistence among the worst performing Norwegian equity mutual funds. We do not, however, find enough evidence to conclude there is any value in the optimal weights compared to an equal allocation, but there is some indication that the weights could provide some value compared to our passive benchmark. These portfolios are still highly hypothetical and difficult to replicate in the real world, but with increasing amounts of ETFs it could be more viable in the future.

5.2.6 Different estimation windows

Before concluding this section we will look at the performance of portfolios when we estimate weights based on longer and shorter time-period intervals. There are a couple of reasons for this. Firstly, given the lack of long-term persistence we found with a 12 month estimation window, it could be that portfolios formed over a shorter time period will exhibit stronger persistence, making a strategy with less frequent rebalancing more feasible. This comes from the Berk and Green (2004) model which says that a fund achieves abnormal returns until enough money is added to the fund to offset the managers ability to find new assets to place them. Secondly, the rationale for using a longer formation period is that funds performing well over a longer period might imply that this outperformance is not a result of luck, and thus is more likely to continue into the future. Panel A of table 5.10 shows results for portfolios formed with 6 months of daily returns, while Panel B shows the results for portfolios formed using the past 24 months of daily returns.

Compared to our earlier findings for portfolios formed over a 12 month period, we observe that a 6 month estimation window yield superior performance both in terms of alpha and the t-value. Relying on the t-value, all portfolios have a significant alpha when evaluated with the CAPM, but not when controlling for the factors in the FF3 and FFC4 model. On the contrary, the portfolios formed over a 24 month window yield inferior performance in the absolute value of both alpha and the t-value, when compared to portfolios with both 6 and 12 month estimation windows. The finding that a shorter estimation window seems to give the better results supports our conclusion in the the previous section that outperformance among the funds in our sample is because of chance, and not a result of fund manager skill.

Table 5.10: Performance of Portfolios with Different Formation Windows

The table shows annualized alphas for optimal (O) and equally weighted (EW) portfolios formed with estimation windows of 6 (Panel A) and 24 (Panel B) months. The equally weighted portfolio comprises only funds with a positive weight in the optimal portfolio. 'Formation model' is the model used to estimate the alpha and tracking error used in the Treynor-Black portfolio optimization. 'Evaluation Model' is the model used to estimate the alphas of the optimal and equally weighted portfolios. Alpha is estimated using monthly returns over the period from Jan. 2009 through Dec. 2019. Alpha t-statistics are shown in the brackets below each alpha estimate.

Panel A: 6 Months

		Formation Model					
		CAPM O	CAPM EW	FF3 O	FF3 EW	FF4 O	FF4 EW
Evaluation Model	CAPM	3.20	3.67	3.17	3.39	2.94	3.29
		2.38	[2.53]	[2.46]	[2.39]	[2.43]	[2.49]
	FF 3-Factor	1.89	2.13	2.01	2.02	1.91	1.94
		[1.49]	[1.58]	[1.64]	[1.51]	[1.61]	[1.53]
	Carhart 4-Factor	1.64	1.97	1.78	1.83	1.74	1.86
		[1.21]	[1.36]	[1.35]	[1.28]	[1.37]	[1.38]

Panel B: 24 Months

		Formation Model					
		CAPM O	CAPM EW	FF3 O	FF3 EW	FF4 O	FF4 EW
Evaluation Model	CAPM	2.01	1.67	2.11	2.05	2.10	2.12
		[1.55]	[1.24]	[1.67]	[1.61]	[1.73]	[1.71]
	FF 3-Factor	1.11	0.52	1.12	0.96	1.19	0.97
		[0.85]	[0.40]	[0.96]	[0.77]	[0.99]	[0.81]
	Carhart 4-Factor	0.58	0.02	0.62	0.49	0.78	0.66
		[0.42]	[0.01]	[0.46]	[0.37]	[0.61]	[0.52]

To evaluate the persistence we can look to table A8 in the appendix. Comparing this table to Table 5.7, there seem to be slightly stronger long run performance for portfolios estimated using a 6 month window. The 6 month window portfolios exhibit positive alphas for 9 months after portfolio formation, while the 12 month portfolios only have 5 months before alphas become negative. That being said, the difference seem too narrow to draw bastant conclusions about a true difference in persistence between the two. Table A9 shows the persistence for the optimal and equal weight portfolio formed over a 24 month window. The short term results seem to be weaker, with alphas and t-statistics for month t+1 well below the same numbers for the portfolios with 6 and 12 month estimation intervals. However, in the long run we do not observe the same shifts towards strong negative values, at least when evaluated with the CAPM where all alphas are positive for a full year after portfolio formation. This result could be explained by the fact that longer estimation windows should identify funds more able to deliver consistent abnormal performance.

6 Limitations and Further Research

This section will provide a review of the most apparent limitations of our thesis and suggest areas of interest for further research.

6.1 Limitations

6.1.1 Market size

A clear limitation to our analysis and strategy based on the methodology of Riley (2019) is the size of the Norwegian market. With a sample of only 55 funds it is very hard to construct a portfolio which both achieves benefits in diversification and are able to isolate previous winners. Our analysis has also shown that Norwegian funds, on aggregate, take on little idiosyncratic risk to begin with, making a portfolio less viable as there is less diversification to be obtained. In the US market, Riley (2019) used only funds in the top 5 percent alpha and still had portfolios with 51 to 70 funds. We have seen in the Norwegian market as well, that portfolios of previous winners do better as we reduce the percentage of funds in the portfolio, and that the Treynor and Black (1973) framework works best compared to an equal weight allocation when there are more funds in the portfolio.

6.1.2 Accuracy of alpha estimates

The intercept (i.e. alpha) from a regression of a funds' return series on a set of risk factors seems to be both widely accepted and contested as a performance measure for mutual funds. The task of measuring alpha in an exact manner is, however, not straight forward and is affected by a number of choices and data issues. Firstly, the estimate is dependent on the choice of benchmark. Most studies seems to use pure indexes for this purpose, despite these indexes do not capture the cost of obtaining their return series. Secondly, mutual fund returns are not normally distributed and neither are the cross section alpha distribution of mutual fund alphas. These issues makes it hard to make proper inferences about the significance levels of alphas reported throughout our thesis. Additionally, the Treynor-Black model used for forming the optimal portfolio will only give theoretical correct weights if the alpha estimate is equal to the true alpha. To believe this is the case in our estimations would be quite naive, but it could still be reasonable to think the model could add value if the weights are "correct enough", i.e. that up to some threshold of correctness they are able to give a portfolio with added value compared to its equally weighted counterpart. However, we are not able to quantify neither the level of correctness or the threshold.

6.1.3 Transaction costs for portfolio strategies

At last we want to emphasize that most of the portfolio strategies presented in our thesis relies on frequent rebalancing. This would involve transaction costs which we have not taken into consideration. Additionally, because the Norwegian market is small in size, we have chosen to include funds with front load fees. Combined these costs would most probably make many of the reported alphas in our thesis smaller in size if the strategies were to be implemented. That being said, we still believe our reported results without considering these costs could be of relevance since more and more financial products become available at lower costs, for example, the growing amount of ETFs makes mutual funds available to investors at terms similar to individual stocks. Since our strategy is largely persistence based, our results also add to the literature on performance persistence in mutual funds, although our approach is not as targeted as other papers (see. Hendricks et al. (1993), Grinblatt and Titman (1989) and Fama and French (1996)) for investigating this particular question.

6.2 Further Research

In the process of writing this thesis, we have discovered a few areas that might be of interest regarding further research. Firstly, we believe more studies on what is an appropriate benchmark would be great contributions to the literature seeking to evaluate mutual fund performance. For example, Elton (2019) attempts to replicate factor returns using a set of ETFs, and Berk and Binsbergen (2015) uses Vanguard funds instead of theoretical indexes. The rationale behind their studies is that benchmarks should be investable alternatives taking into consideration all costs of obtaining their returns. An accepted approach to construct investable benchmarks to obtain factor returns has, however, not yet been discovered, and in Norway there has to our knowledge been no attempts of finding investable benchmarks for factors other than the market. We believe more insight into this area is important and could add to the debate on active management and its value.

Additionally, we believe further looking into the decision of investing in mutual funds as a portfolio choice has potential. In this thesis we replicate an approach suggested by Riley (2019), but the practical implementation of this strategy would most probably make the strategy inferior to a passive investment in an index fund, since transaction costs would eat positive alphas downward. A portfolio approach is, however, not limited to looking at persistence based strategies, and we believe investigating other ways of forming portfolios is an interesting topic for further research.

7 Conclusion

This thesis aims to provide insights about actively managed Norwegian equity mutual funds for investors. Our data set is free of survivorship-bias and contains monthly and daily net returns for 55 active Norwegian mutual funds in the period 2009-2019. We investigate if the performance of individual funds can be attributed to the skillset of managers, if investors can achieve abnormal returns by betting on funds with historical good performance, and if applying an optimization framework within previous winners provide additional benefits to the average of these funds.

Finding fat tails in the aggregate performance of Norwegian equity mutual funds, we applied bootstrap simulations following the methodology of Kosowski et al. (2006) to test if the performance of the superior and inferior performing funds could be attributed to skill or lack or skill among fund managers. We were unable to attribute the performance of the best and worst performing funds to the skillset of managers. This conclusion holds both when funds were ranked on the alpha t-statistic and when we ranked on their alpha point estimate. Thus, investors might be able to earn long term abnormal returns simply by luck, but we do not find evidence supporting they can do so by identifying fund manager skill.

Turning to funds as a portfolios, we followed Riley (2019) and used the Treynor and Black (1973) model to construct portfolios of individual funds based on their historical performance. Focused on the 75th percentile of funds, our optimal portfolios delivered positive but non-significant alphas before transaction costs compared to our passive benchmark, but we were unable to separate its performance from the average fund in the same sample represented by an equal weight portfolio. The short-lived performance persistence for both the optimal and equal weight portfolios means the results are dependent on frequent rebalancing and supports our findings of no fund manager skill in the sample.

Altering our parameters we found that the closer you get to the top percentiles of historical alpha the better is the performance of the optimal and equal weight portfolios. This indicates that the best fund managers are able to maintain their relative performance, at least one month into the future. We also found that a shorter estimation period improved the results marginally, while a longer estimation period made them slightly worse. In total our model likely suffers from the Norwegian market being relatively small, and that using it in a bigger market could lead to better results similiar to what Riley (2019) found in the US market.

All taken together, we do not find evidence supporting that there is value in active management for those looking to invest in Norwegian equity mutual funds. Our analysis instead suggest investors should expect no loss in doing the simplest thing available, namely investing in a low-cost passively managed index fund.

References

Avramov, D. and Wermers, R. (2006). Investing in mutual funds when returns are predictable. *Journal of Financial Economics*, 81(2):339–377.

Baks, K. P., Metrick, A., and Wachter, J. (2001). Should investors avoid all actively managed mutual funds? a study in bayesian performance evaluation. *The Journal of Finance*, 56(1):45–85.

Barras, L., Scaillet, O., and Wermers, R. (2010). False discoveries in mutual fund performance: Measuring luck in estimated alphas. *The Journal of Finance*, 65(1):179–216.

Berk and Binsbergen, ; Berk, J. B. J. (2015). Measuring skill in the mutual fund industry. *Journal*.

Berk, J. and Green, R. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, 112:1269–1295.

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

Davis, J. L. (2001). Mutual fund performance and manager style. *Financial Analysts Journal*, 57(1).

Elton, E. J., Gruber, M. J., Das, S., and Hlavka, M. (1993). Efficiency with costly information: A reinterpretation of evidence from managed portfolios. *The Review of Financial Studies*, 6(1):1–22.

Elton, Edvin; Gruber, M. S. A. (2019). Are passive funds really superior investments: An investor perspective. *Working paper*.

Evans, R. B. (2010). Mutual fund incubation. *Journal of Finance*, 65(4):1581–1611.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383–417.

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. (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1):55–84.

Fama, E. F. and French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. *The Journal of Finance*, 65(5):1915–1947.

Gallefoss, K., Hansen, H. H., Haukaas, E. S., and Molnár, P. (2015). What daily data can tell us about mutual funds: Evidence from norway. *Journal of Banking Finance*, 55:117–129.

Grinblatt, M. and Titman, S. (1989). Mutual fund performance: An analysis of quarterly portfolio holdings. *The Journal of Business*, 62(3):393–416.

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

Hendricks et al., Darryll; Jayendu Patel, R. Z. (1993). Hot hands in mutual funds: Short-term persistence of relative performance, 1974-1988. *Journal of Finance*, 48(1):93–130.

Ippolito, R. A. (1989). Efficiency with costly information: A study og mutual fund performance, 1965-1984. *The Quarterly Journal og Economics*, 104(2):1–23.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The Journal og Finance*, 23(2):389–416.

Kacperczyk, M., Nieuwerburgh, S. V., and Veldkamp, L. (2014). Time-varying fund manager skill. *The Journal of Finance*, 69(4):1455–1484.

Kosowski, R., Timmermann, A., Wermers, R., and Jr., H. L. W. (2006). Can mutual fund 'stars' really pick stocks? new evidence from a bootstrap analysis. *The Journal of Finance*, 61(6).

Kosowski et al., R. (2006). Can mutual fund 'stars' really pick stocks? new evidence from a bootstrap analysis. *Journal of Finance*, 61(6):2551–2595.

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

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometria*, 34(4):768–783.

Næs, Randi; Skjeltorp, J. B. A. (2009). What factors affect the oslo stock exchange. *Working paper of Norges Bank.*

Otten, R. and Bams, D. (2002). European mutual fund performance. *European Financial Management*, 8(1):75–101.

Riley, T. B. (2019). Portfolios of actively managed mutual funds. Unpublished working paper.

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

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, 39(1):119–138.

Sørensen, L. (2009). Mutual fund performance at the oslo stock exchange. *SSRN Electronic Journal.*

Treynor, J. L. (1965). How to rate management of investment funds. *Harvard Business Review*, 43(1):63–75.

Treynor, J. L. and Black, F. (1973). How to use security analysis to improve portfolio selection. *The Journal of Business*, 46(1):66–86.

VFF (2019). Kjøpte verdipapirfond for 33 milliarder i første halvår. URL: https://www.vff.no/news/2019/halvarsstatistikk, Last Accessed = 2019-12-04.

White, Halbert; Kim, T.-H. K. K. (2003). Active portfolio management: The power of the treynor-black model. *Working paper*.

Ødegaard, B. A. (2020a). Empirics of the oslo stock exchange. asset pricing results. 1980-2019.

Ødegaard, B. A. (2020b). Empirics of the oslo stock exchange. basic, descriptive results 1980-2020.

Ľuboš Pástor and Stambaugh, R. F. (2002). Investing in equity mutual funds. *Journal of Financial Economics*, 63(3):351–380.

Appendix

Figure A1: Development in risk-free rate

The figure plots the annualized one-month NIBOR interbank rate for the period from 2009 to 2019.

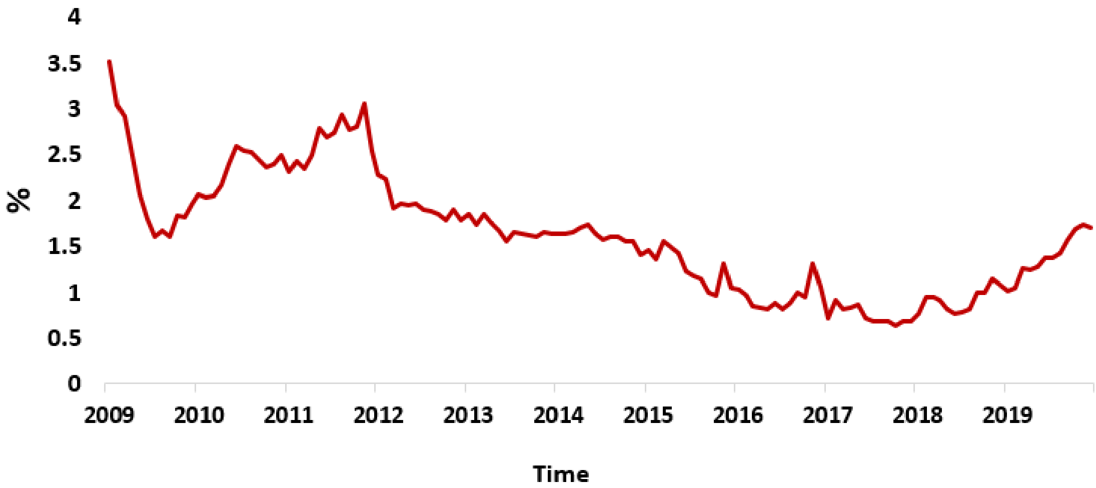


Table A1: Descriptive Statistics of Individual Fund Returns (1/2)

The table reports descriptive statistics measures for every individual mutual fund in our sample for the period 2009-2019. Column 1 shows the number of monthly observations for each fund. The observation count is excluding any observations until a fund is 3 years of age to remove possible incubation bias. Column 2 and 3 report skewness and kurtosis, while column 4 shows the p-value from a Shapiro-Wilk test for normality. If the p-value is greater than 0.05, the null hypothesis of non-normality is rejected and the return series is concluded to be normally distributed.

Fund Name	N	Skewness	Kurtosis	Shapiro-Wilk
Alfred Berg Gambak	132	-0.33	4.76	0.00
Alfred Berg Norge Classic	132	0.07	4.91	0.00
Alfred Berg Aktiv	132	-0.15	4.92	0.00
Alfred Berg Aktiv II	45	-0.13	2.98	0.76
Alfred Berg Norge $+_g ml$	63	0.06	3.57	0.14
Alfred Berg Humanfond	132	0.01	4.69	0.00
Alfred Berg Norge Etisk	63	-0.02	3.39	0.37
Arctic Norwegian Value Creation	28	-0.59	3.34	0.30
Arctic Norwegian Equities	72	-1.14	4.87	0.00
Sbanken Framgang Sammen	11	-0.02	2.30	0.97
C WorldWide Norge	132	0.01	4.28	0.00
Danske Invest Norge II	132	0.09	4.56	0.00
Danske Invest Norge I	132	0.09	4.57	0.00
Danske Invest Norge Vekst	132	0.12	3.74	0.30
Danske Invest Norge Aksj. Inst 1	132	0.05	4.67	0.00
Danske Invest Norge Aksj. Inst 2	121	-0.45	4.31	0.00
DNB Norge (Avanse I)	62	0.05	3.11	0.39
DNB Norge (I)	62	0.11	3.20	0.63
DNB Norge (Avanse II)	69	0.10	3.40	0.17
DNB Norge Selektiv E	132	0.24	3.96	0.07
DNB Norge	127	0.10	3.94	0.06
DNB Norge (III)	127	0.11	3.96	0.07
DNB SMB A	132	0.06	3.45	0.35
NB-Aksjefond	57	-0.10	3.63	0.75
Eika SMB	57	0.02	3.56	0.43
Terra Norge	57	0.06	3.43	0.75
Eika Norge	132	0.00	5.06	0.00
FIRST Generator S	75	-0.82	4.87	0.01
Fondsfinans Norge	132	0.08	3.81	0.12
PLUSS Markedsverdi	132	0.19	4.48	0.00
PLUSS Aksje	132	0.09	4.31	0.00
FORTE Norge	70	-0.19	3.37	0.51
FORTE Trønder	45	-0.31	2.89	0.92
Handelsbanken Norge	132	-0.04	4.45	0.00
Holberg Norge A	132	0.11	4.11	0.04
KLP AksjeNorge	132	0.15	4.46	0.00
Landkreditt Norge	84	-0.27	3.14	0.08
Landkreditt Utbytte A	46	-0.83	3.63	0.04

Table A2: Descriptive Statistics of Individual Fund Returns (2/2)

Fund Name	N	Skewness	Kurtosis	Shapiro-Wilk
Nordea Vekst	73	0.13	3.57	0.25
Nordea Avkastning	132	0.01	4.26	0.01
Nordea Kapital	132	0.08	4.37	0.01
Nordea Norge Verdi	132	0.20	4.57	0.00
Nordea SMB	73	0.27	3.45	0.11
Nordea Norge Pluss	69	-1.03	3.88	0.00
ODIN Norge C	132	-0.43	4.70	0.00
Pareto Investment Fund A	132	0.12	4.45	0.00
Pareto Aksje Norge A	132	0.20	4.47	0.02
Storebrand Norge	132	0.11	4.38	0.00
Storebrand Vekst	132	0.37	4.73	0.01
Delphi Norge	132	-0.26	4.08	0.01
Storebrand Aksje Innland	132	0.13	4.15	0.03
Delphi Vekst	57	-0.25	3.21	0.30
Storebrand Verdi A	132	0.00	4.04	0.06
Storebrand Norge I	132	0.13	4.18	0.02
Storebrand Optima Norge	132	0.11	3.78	0.08

Table A3: Individual Fund Regressions (1/2)

The table shows output from individual fund regressions estimated with the FFC4 performance evaluation model using monthly. Regressions are estimated using all observations for each fund available during our sample period from 2009 through 2019. Column 1 refers to the number of observations for each fund and column 2 shows the annualized alpha estimated from these observations. Columns 3-6 show the loading on the MKT, SML, HML and PR1YR factors, while column 7 report the adjusted R-squared.

	N	α	β_{MKT}	β_{SMB}	β_{HML}	β_{PR1YR}	$R^2_{Adj.}$
Alfred Berg Gambak	1.53	1.03	0.21	-0.02	0.10	0.80	
Alfred Berg Norge Classic	0.90	1.03	0.07	-0.02	0.03	0.94	
Alfred Berg Aktiv	0.25	1.08	0.14	-0.02	0.04	0.89	
Alfred Berg Aktiv II	-1.88	1.18	0.22	-0.06	-0.05	0.94	
Alfred Berg Norge $+_g ml$	0.44	1.10	0.15	-0.03	-0.05	0.96	
Alfred Berg Humanfond	-1.94	1.03	0.05	-0.05	0.03	0.92	
Alfred Berg Norge Etisk	-1.63	1.13	0.14	-0.02	-0.02	0.96	
Arctic Norwegian Value Creation	4.17	0.97	0.01	0.02	-0.17	0.85	
Arctic Norwegian Equities	-1.57	0.85	0.09	-0.04	0.12	0.81	
Sbanken Framgang Sammen	0.02	0.74	0.13	0.01	0.01	0.80	
C WorldWide Norge	-1.45	1.05	0.01	-0.06	0.04	0.94	
Danske Invest Norge II	0.67	1.04	0.07	-0.03	-0.02	0.93	
Danske Invest Norge I	-0.08	1.04	0.07	-0.03	-0.02	0.93	
Danske Invest Norge Vekst	3.62	1.05	0.16	-0.01	-0.08	0.86	
Danske Invest Norge Aksj. Inst 1	0.75	1.03	0.05	-0.04	-0.01	0.94	
Danske Invest Norge Aksj. Inst 2	0.65	1.01	0.03	-0.05	0.00	0.93	
DNB Norge (Avanse I)	-1.60	1.10	0.08	0.00	-0.08	0.97	
DNB Norge (I)	-0.65	1.02	0.03	-0.02	0.00	0.99	
DNB Norge (Avanse II)	-1.44	1.10	0.07	-0.02	-0.08	0.97	
DNB Norge Selektiv E	0.37	1.08	0.06	-0.04	-0.04	0.93	
DNB Norge	-1.25	1.02	0.04	-0.04	-0.04	0.96	
DNB Norge (III)	-0.58	1.02	0.04	-0.04	-0.04	0.96	
DNB SMB A	1.46	1.17	0.48	-0.01	-0.19	0.66	
NB-Aksjefond	-2.59	1.20	0.33	0.08	-0.12	0.89	
Eika SMB	1.92	1.14	0.47	0.13	-0.27	0.79	
Terra Norge	0.44	1.22	0.35	0.08	-0.15	0.92	
Eika Norge	-2.28	1.11	0.22	0.02	-0.06	0.87	
FIRST Generator S	-6.52	1.64	0.49	0.06	-0.04	0.68	
Fondsfinans Norge	1.90	1.15	0.25	0.02	-0.14	0.85	
PLUSS Markedsverdi (Fondsforvaltning)	0.32	1.01	-0.01	-0.07	0.00	0.96	
PLUSS Aksje (Fondsforvaltning)	-0.17	0.97	-0.02	-0.09	0.01	0.93	
FORTE Norge	4.28	0.97	0.21	-0.01	-0.02	0.56	
FORTE Trønder	-2.90	1.14	0.10	0.12	0.24	0.54	
Handelsbanken Norge	-0.14	1.10	0.18	-0.03	0.03	0.79	
Holberg Norge A	-1.97	1.08	0.30	-0.01	-0.04	0.82	
KLP AksjeNorge	-0.01	1.11	0.11	-0.02	-0.07	0.94	
Landkreditt Norge	-3.33	1.11	0.24	-0.07	-0.06	0.83	
Landkreditt Utbytte A	2.28	0.59	0.08	0.03	0.10	0.40	

Table A4: Individual Fund Regressions (2/2)

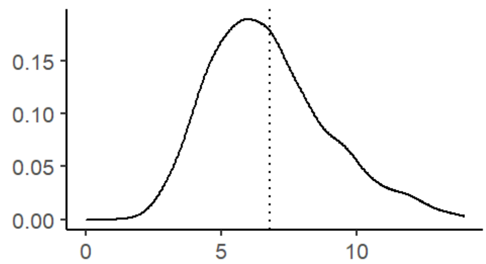
The table shows output from individual fund regressions estimated with the FFC4 performance evaluation model using monthly. Regressions are estimated using all observations for each fund available during our sample period from 2009 through 2019. Column 1 refers to the number of observations for each fund and column 2 shows the annualized alpha estimated from these observations. Columns 3-6 show the loading on the MKT, SML, HML and PR1YR factors, while column 7 report the adjusted R-squared.

	N	α	β_{MKT}	β_{SMB}	β_{HML}	β_{PR1YR}	$R^2_{Adj.}$
Nordea Vekst	-0.31	1.10	0.08	-0.06	-0.03	0.96	
Nordea Avkastning	0.27	1.10	0.12	-0.04	-0.02	0.95	
Nordea Kapital	0.70	1.07	0.09	-0.04	-0.03	0.95	
Nordea Norge Verdi	4.05	0.91	0.19	0.08	-0.09	0.83	
Nordea SMB	-8.46	1.17	0.65	-0.03	-0.28	0.76	
Nordea Norge Pluss	-0.01	1.06	0.25	0.02	0.01	0.84	
ODIN Norge C	-1.87	0.95	0.21	0.02	-0.05	0.79	
Pareto Investment Fund A	0.13	1.12	0.10	-0.09	0.04	0.81	
Pareto Aksje Norge A	0.00	0.93	0.24	0.02	-0.10	0.79	
Storebrand Norge	0.29	1.04	0.04	-0.02	0.03	0.92	
Storebrand Vekst	7.12	1.00	0.19	-0.13	-0.14	0.62	
Delphi Norge	-0.85	1.14	0.23	-0.04	0.07	0.84	
Storebrand Aksje Innland	-0.38	0.99	0.03	0.02	-0.01	0.97	
Delphi Vekst	-3.01	1.13	0.38	-0.01	-0.03	0.83	
Storebrand Verdi A	-0.25	0.95	0.02	0.12	-0.01	0.95	
Storebrand Norge I	0.09	1.01	0.07	0.05	-0.01	0.96	
Storebrand Optima Norge	0.29	1.02	0.10	0.05	-0.03	0.87	

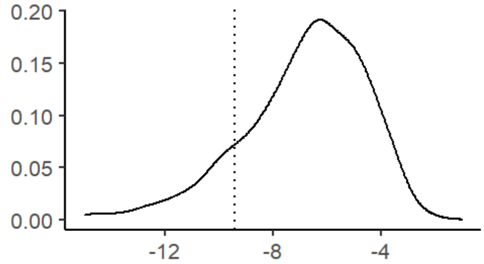
Figure A2: Estimated Alphas vs. Bootstrapped Distributions of Alpha for Individual Funds

The figure displays kernel density estimates of the bootstrapped distributions of alpha (solid line). To produce the distributions we follow the methodology of Kosowski et al. (2006) as described in section 4.2, using alpha as ranking measure and the CAPM for performance evaluation. The x-axis shows annualized alphas in percent and the y-axis shows the kernel density estimate. The vertical dotted line represents the actual (estimated) fund alpha. The top (bottom) fund refers to the fund with highest (lowest) alpha during our sample period (2009-2019). Similarly, the top (bottom) 3 funds refers to an equally weighted portfolio of the 3 funds with highest (lowest) alpha during the same period.

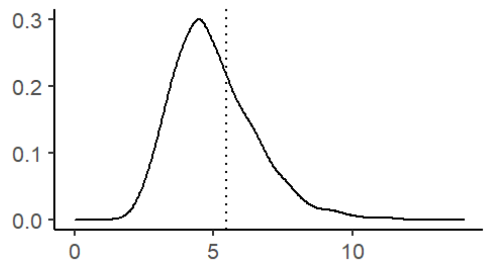
(a) Top Fund



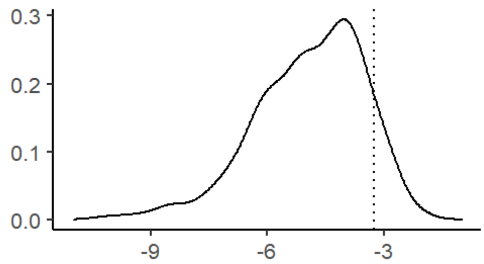
(b) Bottom Fund



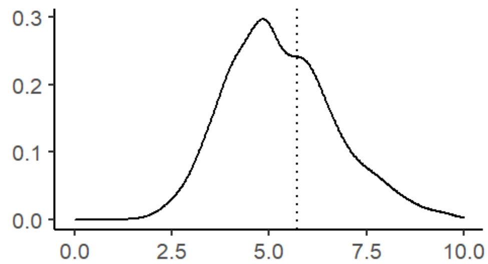
(c) Second Best



(d) Second Worst



(e) Top 3 Funds



(f) Bottom 3 Funds

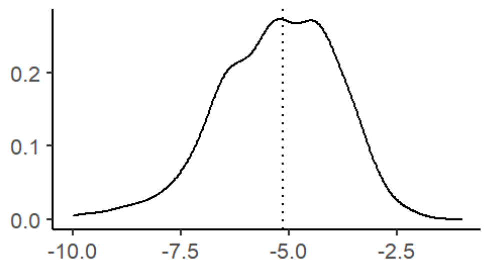


Table A5: Alphas of Optimal Portfolios Estimated using Daily Returns. The table presents annualized percentage alphas for optimal portfolios calculated using funds in the 75th percentile of alpha during the previous year. The 'Formation Model' is the model used to calculate the prior year alphas and residuals, which are used to identify the 75th percentile and to calculate the optimal weights. The 'Evaluation Model' is the model used to measure the alphas of the resulting optimal portfolios. The set of performance models used for formation and evaluation is the CAPM, the Fama-French 3-Factor model and the Carhart 4-factor model. Alpha is estimated using daily returns over the time period Jan. 2009 through Dec. 2019. Alpha t-statistics is shown in the brackets below each measure of alpha.

		Formation Model		
		CAPM	FF 3-Factor	Carhart 4-Factor
Evaluation Model	CAPM	2.57	3.04	2.74
		[1.98]	[2.45]	[2.25]
	FF 3-Factor	1.94	2.49	2.19
		[1,53]	[2.06]	[1.84]
	Carhart 4-Factor	1.72	2.29	2.07
		[1.36]	[1.89]	[1.74]

Table A6: Alphas of Equally Weighted Portfolios Estimated using Daily Returns.

The table presents annualized percentage alphas for optimal portfolios calculated using funds in the 75th percentile of alpha during the previous year. The 'Formation Model' is the model used to calculate the prior year alphas and residuals, which are used to identify the 75th percentile and to calculate the optimal weights. The 'Evaluation Model' is the model used to measure the alphas of the resulting optimal portfolios. The set of performance models used for formation and evaluation is the CAPM, the Fama-French 3-Factor model and the Carhart 4-factor model. Alpha is estimated using daily returns over the time period Jan. 2009 through Dec. 2019. Alpha t-statistics is shown in the brackets below each measure of alpha.

		Formation Model		
		CAPM	FF 3-Factor	Carhart 4-Factor
Evaluation Model	CAPM	2.85	3.30	3.10
		[2.09]	[2.53]	[2.42]
	FF 3-Factor	1.96	2.49	2.30
		[1.50]	[1.99]	[1.85]
	Carhart 4-Factor	1.79	2.32	2.18
		[1.37]	[1.84]	[1.76]

Table A7: Performance of Portfolios formed using Different Top Alpha Percentiles

This table shows annualized alphas for optimal (O) and equally weighted (EW) portfolios for four different percentiles of top alpha. Formation model is the model used to estimate the alpha and tracking error used in the Treynor-Black portfolio optimization. Evaluation model is the model used to estimate the alphas of the optimal and EW portfolios. Alpha is estimated using daily returns over the period from Jan. 2009 through Dec. 2019. In Panel A all funds with a full year of daily returns are included in portfolios. In Panel B-D, funds not in the top 50, 25 and 10 percent, respectively, are given a weight of zero before calculating optimal weights. Only funds with a positive weight in the optimal portfolio is included in the EW portfolio.

(a) Panel A: Top 100 of alpha

		Evaluation Model					
		CAPM O	CAPM EW	FF3 O	FF3 EW	FF4 O	FF4 EW
Formation Model	CAPM	1.74	1.81	2.10	1.96	2.14	2.00
		[1.63]	[1.68]	[2.01]	[1.84]	[2.04]	[1.87]
	FF 3-Factor	1.26	1.14	1.67	1.33	1.69	1.37
		[1.20]	[1.09]	[1.63]	[1.31]	[1.65]	[1.33]
	Carhart 4-Factor	1.04	1.00	1.45	1.12	1.50	1.21
		[1.00]	[0.96]	[1.42]	[1.13]	[1.47]	[1.17]

(b) Panel B: Top 50 percent of alpha

		Evaluation Model					
		CAPM O	CAPM EW	FF3 O	FF3 EW	FF4 O	FF4 EW
Formation Model	CAPM	1.93	2.18	2.21	2.18	2.26	2.23
		[1.72]	[1.87]	[2.03]	[1.91]	[2.09]	[1.98]
	FF 3-Factor	1.40	1.44	1.74	1.49	1.78	1.55
		[1.29]	[1.28]	[1.64]	[1.36]	[1.69]	[1.42]
	Carhart 4-Factor	1.19	1.31	1.54	1.33	1.63	1.43
		[1.09]	[1.16]	[1.45]	[1.21]	[1.54]	[1.31]

(c) Panel C: top 25 of alpha

		Evaluation Model					
		CAPM O	CAPM EW	FF3 O	FF3 EW	FF4 O	FF4 EW
Formation Model	CAPM	2.57	2.85	3.04	3.30	2.74	3.10
		[1.98]	[2.09]	[2.45]	[2.53]	[2.25]	[2.42]
	FF 3-Factor	1.94	1.96	2.49	2.51	2.19	2.30
		[1.53]	[1.50]	[2.06]	[1.99]	[1.84]	[1.85]
	Carhart 4-Factor	1.72	1.79	2.29	2.32	2.07	2.18
		[1.36]	[1.37]	[1.89]	[1.84]	[1.74]	[1.76]

(d) Panel D: top 10 percent of alpha

		Evaluation Model					
		CAPM O	CAPM EW	FF3 O	FF3 EW	FF4 O	FF4 EW
Formation Model	CAPM	3.51	4.01	3.77	4.36	3.31	3.74
		[2.04]	[2.28]	[2.28]	[2.57]	[2.08]	[2.29]
	FF 3-Factor	2.66	2.93	3.02	3.37	2.56	2.76
		[1.59]	[1.73]	[1.87]	[2.05]	[1.64]	[1.74]
	Carhart 4-Factor	2.34	2.59	2.76	3.02	2.49	2.64
		[1.40]	[1.53]	[1.71]	[1.84]	[1.59]	[1.66]

Table A8: Long Run Performance of Optimal and Equal Weight Portfolios formed using an Estimation Window of 6 months.

The table displays annualized alphas (in percent) for the optimal and equal weight portfolios formed using lagged weights. The weights are calculated using the Fama-French 3-Factor model and only funds in the 75th percentile of top alpha during the prior 6 months is included. Alphas are estimated using daily returns and three different evaluation models, which include the CAPM, the Fama-French 3-Factor model and the Carhart 4-factor model. For each set of weights, alpha is calculated in the first month after portfolio formation (t+1) and in each of the next 12 months (t+2 through t+13). The time period is Jan. 2010 through Dec. 2019. Alpha t-statistics are shown in the brackets below each measure of alpha.

	Optimal			Equal Weight		
	CAPM	FF3	FF4	CAPM	FF3	FF4
t+1	2.89	1.81	0.29	2.91	1.65	0.18
	(2.21)	(1.44)	(0.22)	(2.02)	(1.19)	(0.12)
t+2	1.52	0.38	-1.21	1.67	0.42	-1.09
	(1.16)	(0.30)	(-0.89)	(1.17)	(0.30)	(-0.73)
t+3	0.87	0.03	-1.47	0.74	-0.37	-1.84
	(0.70)	(0.02)	(-1.11)	(0.53)	(-0.28)	(-1.25)
t+4	1.28	0.37	-1.29	0.92	-0.27	-1.59
	(0.98)	(0.28)	(-0.92)	(0.65)	(-0.20)	(-1.06)
t+5	1.11	-0.06	-1.32	1.17	-0.17	-1.52
	(0.84)	(0.04)	(-0.94)	(0.79)	(-0.12)	(-0.99)
t+6	0.73	-0.33	-1.94	0.71	-0.69	-2.14
	(0.55)	(-0.26)	(-1.39)	(0.47)	(-0.49)	(-1.38)
t+7	0.38	-0.58	-2.03	0.11	-1.11	-2.45
	(0.27)	(-0.42)	(-1.36)	(0.08)	(-0.77)	(-1.56)
t+8	0.70	-0.29	-1.32	0.35	-0.86	-1.57
	(0.50)	(-0.22)	(-0.87)	(0.24)	(-0.59)	(-0.98)
t+9	0.33	-0.75	-1.83	0.33	-0.94	-1.59
	(0.24)	(-0.55)	(-1.23)	(0.21)	(-0.63)	(-0.96)
t+10	-0.93	-1.90	-3.14	-0.98	-2.30	-2.98
	(-0.70)	(-1.46)	(-2.21)	(-0.66)	(-1.61)	(-1.88)
t+11	-0.77	-1.72	-2.36	-1.04	-2.47	-2.59
	(-0.59)	(-1.33)	(-1.67)	(-0.71)	(-1.75)	(-1.66)
t+12	-0.92	-1.58	-2.05	-1.20	-2.21	- 2.21
	(-0.72)	(-1.20)	(-1.42)	(-0.86)	(-1.59)	(-1.43)
t+13	-0.47	-1.27	-1.78	-0.62	-1.64	-1.52
	(-0.36)	(-1.01)	(-1.28)	(-0.45)	(-1.20)	(-1.01)

Table A9: Long Run Performance of Optimal and Equal Weight Portfolios formed using an Estimation Window of 24 Months.

The table displays annualized alphas (in percent) for the optimal and equal weight portfolios formed using lagged weights. The weights are calculated using the Fama-French 3-Factor model and only funds in the 75th percentile of top alpha during the prior 24 months is included. Alphas are estimated using daily returns and three different evaluation models, which include the CAPM, the Fama-French 3-Factor model and the Carhart 4-factor model. For each set of weights, alpha is calculated in the first month after portfolio formation (t+1) and in each of the next 12 months (t+2 through t+13). The time period is Jan. 2010 through Dec. 2019. Alpha t-statistics are shown in the brackets below each measure of alpha.

	Optimal			Equal Weight		
	CAPM	FF3	FF4	CAPM	FF3	FF4
t+1	1.80	0.88	-0.73	1.68	0.59	-0.65
	(1.37)	(0.42)	(-0.79)	(1.26)	(0.45)	(-0.46)
t+2	1.44	0.55	-1.09	1.13	-0.04	-1.27
	(1.12)	(0.42)	(-0.78)	(0.82)	(-0.03)	(-0.87)
t+3	1.22	0.33	-1.29	0.0.98	-0.30	-1.58
	(0.93)	(0.25)	(-0.92)	(0.72)	(-0.23)	(-1.10)
t+4	0.74	-0.06	-1.68	0.56	-0.61	-1.81
	(0.57)	(-0.04)	(-1.21)	(0.42)	(-0.47)	(-1.28)
t+5	0.76	-0.09	-1.57	0.90	-0.28	-1.54
	(0.59)	(-0.07)	(-1.12)	(0.67)	(-0.21)	(-1.08)
t+6	0.85	-0.00	-1.33	1.01	-0.23	-1.34
	(0.66)	(-0.00)	(-0.95)	(0.71)	(-0.17)	(-0.89)
t+7	0.85	0.04	-1.20	0.59	-0.64	-1.65
	(0.65)	(0.29)	(-0.84)	(0.40)	(-0.44)	(-1.06)
t+8	1.08	0.19	-0.95	0.87	-0.45	-1.38
	(0.79)	(0.13)	(-0.64)	(0.59)	(-0.32)	(-0.88)
t+9	0.98	0.02	-0.78	0.65	-0.53	-1.28
	(0.73)	(0.01)	(-0.53)	(0.45)	(-0.46)	(-0.83)
t+10	0.95	-0.06	-0.85	0.84	-0.54	-1.14
	(0.68)	(-0.04)	(-0.56)	(0.57)	(-0.37)	(-0.72)
t+11	0.67	-0.36	-0.98	0.82	-0.48	-0.84
	(0.49)	(-0.27)	(-0.66)	(0.56)	(-0.33)	(-0.53)
t+12	1.02	0.09	-0.47	0.84	-0.34	-0.68
	(0.76)	(0.07)	(-0.32)	(0.58)	(-0.24)	(-0.43)
t+13	0.35	0.54	0.05	0.52	0.36	-0.00
	(-0.23)	(0.40)	(0.03)	(0.41)	(0.26)	(0.00)