UNIVERSITETET I AGDER

Can Norwegian Mutual Fund Managers Pick Stocks?

SUPERVISOR

Valeriy Zakamulin

MORTEN BLØRSTAD AND BJØRN OTTO BAKKEJORD

University of Agder, 2017

School of Business and Law

Faculty of Economic and Finance

Can Norwegian Mutual Fund Managers Pick Stocks?

Morten Blørstad and Bjørn Otto Bakkejord *

May 31, 2017

Master Thesis, MSc, Finance

Abstract

This thesis examines the performance and persistence of Norwegian mutual funds using a comprehensive dataset (surviving and non-surviving funds) over the period 1983-2015. We examine the performance of Norwegian mutual funds on aggregate level and individually, using the Carhart (1997) four-factor model as our performance model. We find that Norwegian mutual funds, on aggregate, do not produce significant abnormal risk-adjusted returns. When examining mutual funds individually, we apply a cross-section bootstrap to distinguish between skill and luck. The bootstrap is necessary for proper inference because the cross-section of alphas has non-normalities in the tails of the distribution due to idiosyncratic risk-taking and non-normalities in the individual fund's alpha distribution. We find no evidence of skills among the top performing funds. We do however find evidence of bad skill among the poorest performers. There is no evidence of persistence among top performers, but the performance of poor funds persists in short-term.

*Department of Economics and Business Administration, School of Management, Agder University , Serviceboks 422, 4604 Kristiansand, Norway,

Contents

1 Introduction

In this thesis, we evaluate the performance of Norwegian mutual fund managers. Within the subject of mutual fund performance there are the two key questions: i) Can actively managed mutual funds outperform their benchmarks, net of costs? and ii) do they do so persistently? The first question refers to whether actively managed mutual funds are able to outperform their benchmarks by creating risk-adjusted returns (i.e. abnormal returns) net of costs. Throughout this thesis, risk-adjusted returns and abnormal return is also referred as alpha.

An actively managed fund tries to outperform the benchmark by predicting the market or identify and exploit mispriced securities. Any successful attempt by a fund manager to outperform the market requires that the market is not efficient in the semi-strong form, as defined by the Efficient Market Hypotheses (EMH) of Fama (1970). If the market is fully efficient, it is not possible to create abnormal returns by exploiting mispricings, since prices reflect all available information. Therefore, any attempt of outperforming the market is a game of chance rather than skill. In contrast, Grossman and Stiglitz (1980) state the market is not always perfectly efficient, which gives a reason to believe that skilled managers are able to identify and exploit mispriced securities when the market is not fully efficient. Past research (see e.g. Jensen 1968; Carhart, 1997; Edelen, 1999) suggest mutual fund managers underperform and are not able to produce positive abnormal returns, net of costs, and argue that actively managed mutual funds do little besides collecting fees from the investors.

The second question refers to whether actively managed mutual funds are able to create abnormal returns persistently and for how long. Performance persistence is interesting from both an academic point of view as well as a practical viewpoint. From an academic point of view, persistence in performance is a violation of semi-strong form of the EMH and would support a rejection of this hypotheses. From the practical point of view, evidence of performance persistence suggests that investors can exploit past performers to earn risk-adjusted returns. Past literature (see e.g. Hendricks et. al. 1993; Brown and Goetzmann 1995; Carhart 1997; Bollen and Busse, 2005) shows little evidence of performance persistence of top performers. The literature does, however, indicate persistence in the performance among the poorest performers.

The challenge when evaluating mutual funds is to distinguish performances attributed to stock-picking skill from the performances due to luck. In the presence of multiple mutual funds, some funds perform well by chance, and some may perform well due to skill. The majority of past literature on mutual fund performance do not consider and model for the element of luck in performance outcomes. Arguably, literature on performance persistence do it to a large extent, but these models do not account for that luck can also persist. The first to explicitly model for luck without relying on parametric tests, was the paper by Kosowski, Timmermann, Wermers and White (2006). They found evidence of skilled fund managers in the US by implementing an innovative bootstrap method. They also found that the performance of the skilled managers persists, contradicting past studies based on parametric tests (e.g., Carhart (1997)).

These questions regarding performance and persistence are important for investors because they want to know if actively managed funds are worth their costs. If they are not able to create risk-adjusted abnormal returns, investors would be better off by investing in a low-cost passive index fund. Therefore, the purpose of our thesis is to answer the following questions: i) Can actively managed Norwegian mutual funds pick stocks in order to outperform their benchmark net of costs, and ii) do they do it persistently?

To be able to answer the first of these question, we apply the bootstrap method of Kosowski et. al. (2006) to the monthly net returns of Norwegian mutual funds in of our sample, using the Carhart (1997) four-factor model as the performance measurement model.[1] Our sample contains monthly net returns of 98 active Norwegian mutual funds during the period January 1983 to December 2015 and is controlled for survivorship bias.[2] The bootstrap is necessary because mutual funds exhibit non-normal properties which the bootstrap can provide reliable inference whereas traditional parametric inference cannot. We address the second of these questions by implementing a persistence test known as the recursive portfolio formation approach similar to Carhart (1997), but our statistical inference is based on the bootstrap rather than parametric t-tests.

Most of the research on mutual funds is done on US funds, and there are only a few comprehensive studies on Norwegian mutual fund. To our knowledge, it is only conducted two

[1]We have also applied several extensions of the bootstrap method of Kosowski et. al. (2006), to show that our results are robust to possible cross-sectional correlations residuals among Norwegian mutual fund or potential autocorrelations in mutual funds returns or in the factor returns

[2]Meaning, our sample includes all funds that existed during the sample period; this also includes liquidated and dead funds.

similar studies as to ours on Norwegian Mutual funds. Only one is published, the other one is a working paper.[3] The published paper by Gallefoss et. al. (2015) applied the bootstrap method of Kosowski et. al. (2006) to evaluate Norwegian mutual fund in period 2000-2010 using daily returns. They found that the performance of the top and bottom fund was not due to luck, but due to the manager's level of skills. They also found that the performance among the top and bottom funds persist for up to one year. The working paper by Sørensen (2009) found no evidence of skill among the top performers, but evidence of poor skill among the worst performers. Sørensen (2009) found no evidence of persistence among winners or losers.

Our thesis contributes to the existing literature by investigating the Norwegian mutual fund performance using the longest evaluation period up to date, covering almost the entire history of Norwegian mutual fund existence. Although the Norwegian economy is among the most developed in the world, there exist only a handful of comprehensive studies of Norwegian mutual fund, thus making it truly a subject of interest.

Our results indicate that the actively managed Norwegian mutual fund, on average do not possess sufficient skill to create abnormal returns to their investors net of costs. Evaluating fund performance individually, we find that top performing fund managers are lucky and not skilled, as they do not produce abnormal returns above what one would expect if such returns were determined by random chance. Interestingly, we find that underperforming fund managers create negative abnormal returns which cannot be explained by chance, which is evidence of poor skill. In short, from our bootstrap simulations, we find no evidence of superior fund managers but rather evidence of inferior fund managers. Furthermore, we have approximated the economic impact of the subgroups of funds that possess poor-skill. We find that inferior fund managers potentially destroy approximately 600 million NOK per year in investor wealth. Lastly, we find no evidence of performance persistence among the top performers, indicating that investor cannot earn abnormal risk-adjusted returns by exploiting past performance, which is also in line with the semi-strong form of the EMH. Among the worst performing fund, we find evidence of short-term persistence but disappears with increased time horizon, which is in line with the majority of previous studies (see e.g. Berk and Green, 2004; Bollen and Busse, 2005; Huij and Verbeek, 2007). The results of our analysis on Norwegian mutual fund suggest

[3]There is also a master thesis by Utseth and Sandvik (2015) which is similar to our study.

that Norwegian investors are best off by investing in a low-cost passive index fund.

The remainder of our thesis proceeds as follows. In Chapter 2 we will review the current and previous literature related to our subject matter. Chapter 3 contains a thorough review of the models and methods used in our analysis. Chapter 4 reviews the external data used in this paper and their properties. Chapter 5 provides the empirical results, including robustness tests of our bootstrap evidence of individual mutual fund. Chapter 6 concludes.

2 Literature Review

Throughout this chapter, we review past studies on performance and performance persistence of mutual funds. The purpose of this chapter is to inform the reader about the ideas and knowledge established on the subjects similar to ours, by reviewing the most relevant research on the topic of mutual fund performance evaluation. The chapter is structured as follows: We first examine the most relevant literature on mutual fund performance and then we assess the most important research on performance persistence in mutual funds.

2.1 Mutual Fund Performance

Markowitz (1952) created the field of portfolio theory. Here, among other things, the concept of diversification was developed. Later, Sharpe (1964) and Lintner (1965) developed the Capital Asset Pricing Model (CAPM) which is seen as the cornerstone of modern finance. Sharp (1966) was the first to evaluate the risk-adjusted performance of mutual funds by introducing the measure Sharpe Ratio (i.e. reward-to-variability). Utilizing the performance measurement Sharpe Ratio, he evaluated the returns of 34 open-end U.S mutual funds in the period 1945-1963 and found that 11 funds outperformed and 23 underperformed the benchmark, suggesting that the U.S capital market was highly efficient. His study concluded that investing in mutual funds was a bad investment.

Based on the CAPM, Jensen (1968) suggested another performance measurement to evaluate the risk-adjusted performance of mutual funds, the Jensen's alpha. The alpha is an estimate of the predictive ability of mutual fund managers and describes the abnormal return of a mutual fund. An actively managed mutual fund is supposed to produce positive alpha, and a passive fund is supposed to produce an alpha equal to zero. Jensen (1968) evaluated the returns, net of costs, of 115 U.S mutual funds for the period 1945-1964 using his alpha measure. He concluded that the 115 US mutual funds were on average unable to predict security prices well enough to outperform a "buy-the-market-and-hold" strategy. He also found little evidence that any individual fund outperformed the market index. Ippolito (1989) contradicted the findings of Jensen (1968). Using a sample of 143 US mutual funds for the period 1965-1984, Ippolito (1989) found that mutual funds, overall, outperformed the S&P500 index net of costs.

The paper of Roll (1977) criticized the CAPM and argued against the use of the CAPM proxy as a benchmark for performance since it presupposed complete knowledge of the true market portfolio's composition. Roll's critique led to the important issue of choosing an appropriate benchmark to evaluate abnormal performance. Research by Lehmann and Modest (1987), Grinblatt and Titman (1989), Connor and Korajczyk (1991) further addressed the issue of choosing an appropriate benchmark to evaluate performance. Lehmann and Modest (1987) studied the sensitivity of performance to the chosen benchmark. In their paper, they found that the performance was sensitive to the chosen benchmark and emphasized the need to use an appropriate benchmark that represents common factors determining the security returns.

Motivated by these papers and the issue of choosing an appropriate benchmark, Elton et. al. (1993) investigated the informational efficiency of U.S. mutual funds in the period 1965-84. The results of Elton et. al. (1993) contradicted the findings of Ippolito (1989). Elton et. al. (1993) argued that Ippolito's (1989) evidence of positive Jensen's alpha was due to the usage of an inappropriate benchmark. They showed that the reason why Ippolito (1989) found that mutual funds outperformed the S&P500 was because the funds included in his sample invested heavily in small stocks not listed in the S&P500 benchmark and that these stocks outperformed the S&P500 significantly during the period. When Elton et. al. (1993) accounted for the performance of non-S&P500 assets, they found that the positive Jensen's alphas became negative. Malkiel (1995) examined the returns of mutual funds with a sample comprised of all diversified U.S. mutual funds in existence over the period 1971-1991. Using the CAPM, he found that mutual funds underperformed the market, net of costs and gross of costs. However, Malkiel's (1995) results are, like many other previous studies, sensitive to the selection of a benchmark. The issue of choosing an appropriate benchmark lead to the development of multifactor models.

These multifactor models control for various anomalies in the equity market. The most well-known multifactor models are the three-factor model of Fama and French (1993) and Carhart's (1997) four-factor model. Fama and French (1992, 1993, 1996) augmented the single-factor model of Jensen's (1968) by including two new factors in addition to the market proxy, the value factor and the size factor. These factors captures the size and the value anomalies found empirically to operate in the market. To capture the momentum anomaly, Carhart (1997) added another factor to the three-factor model of Fama and French, the momentum factor of Jegadeesh and Titman (1993).

One of the first to use a multi-index for evaluating mutual fund performance is Gruber (1996), where he evaluated mutual funds in the period 1985-1994. He called the multi-index model a four-index model since it consisted of four factors. The four factors was the excess market return, the difference in return between a small cap portfolio and a large cap portfolio, the difference in return between a high growth portfolio and a value portfolio, and the excess return on a bond index. In his paper, he found that the average actively managed mutual fund underperformed compared to a set of indices, net of costs, by approximately 65 basis points. Adding back the average costs of 1.33%, Gruber (1996) stated that fund managers were able to outperform the market gross of costs. Those results indicated that mutual fund managers did have superior stock-picking skills, but not enough to cover the costs of investors.

Daniel et. al. (1997) evaluated the performance of U.S. mutual funds in the period 1973-1994. They examined whether managers had sufficient stock-picking skills to earn back a significant amount of the fees and expenses they generated. Specifically, Daniel et. al. (1997) examined whether funds' excess returns were attributed to Characteristic Selectivity and Characteristic Timing. They found, unlike the majority of previous studies, that some mutual funds, particularly aggressive-growth funds, exhibited stock selection abilities. They found, however, no evidence of market timing. They also observed that fund managers on average, beat a mechanical strategy, but only enough to earn back their average fees.

Edelen (1999) used the single factor model of Jensen (1968) and found significant negative alphas in his sample of 166 different U.S mutual funds. The average yearly alpha, net of costs, was equal to -1.63%, whereas the expenses ratio was 1.72%. This indicated that mutual funds did little beside collect fees. Edelen (1999) stated that the underperformance of mutual funds was due to expenses and not due to fund managers' inability to produce alpha.

Wermers (2000) evaluated the value of active management using a similar method as Daniel (1997). Wermers (2000) decomposed mutual funds into stock-picking skills, characteristic selectivity and timing of the stocks held, trading costs, and expenses. The decomposition of performance was based on stock holdings and the net returns of U.S. mutual funds in the period 1975-1994. He observed that funds held stocks that outperformed the market by 1.3% per year, but their net returns underperformed by -1 %. A difference of 2.3 % between the results. He found that 1.6 % of the difference was due to expenses and transaction costs, whereas the rest, 0.7 %, was due to the underperformance of nonstock holdings. Wermers (2000) evidence supports the value of active mutual fund management. Moskowizt (2000) questioned the results of Wermers (2000) based on characteristic selectivity. Specifically, he was skeptical of the benchmark used in the estimation. Moskowizt (2000) argued that the benchmark used contained certain stock characteristics that all funds avoid. For instance, funds tend to avoid extremely small, illiquid and risky firms. These types of firms did not perform particularly well over the sample period 1975-1994. Thus, the results of Wermers (2000) may have been overstated.

The literature reviewed so far indicates that there is little evidence that mutual funds, on average, creates value for the investor. 7 out of 9 studies found evidence of negative alphas, net of costs. This does however not imply that every individual mutual fund is not able to outperform its benchmark. According to the equilibrium model of Grossman and Stiglitz (1980), some mutual funds outperform, and some underperform from time to time. Meaning the markets are not always perfectly efficient and there exist temporarily mispriced securities. This raised the question whether the performance of individual funds is due to luck or stock-picking skill.

Using a new bootstrap statistical technique, the paper by Kosowski, Timmermann, Wermers and White (2006) address the question whether individual mutual fund performances are attributed to the manager's stock-picking skill or due to luck. They also stated the importance of the bootstrap inference, as it accounts for the complex non-normality in funds returns. Their bootstrap method uncovered findings that differ from past studies. Kosowski et. al. (2006) found that the performance of both the top and worst U.S. funds could not be explained by luck. Specifically, the top 10% of US funds had risk-adjusted returns that could not be explained by luck, and must be a result of the inherent skill of fund managers. The worst performers had risk-adjusted returns so poor that it could not be explained by (bad) luck. This lead to the conclusion that these funds returns are due to managers' bad skill.

Fama and French (2010) modified the bootstrap method of Kosowski et. al. (2006) and examined U.S. mutual funds for the period 1984-2006. They found that mutual funds on average created net returns that underperformed the CAPM, three-factor and four-factor model by about the costs. Investigating the fund individually, they found that few U.S mutual funds had the skills to cover their costs and expenses. Unlike Kosowski et. al. (2006), Fama and French (2010) found no evidence stock-picking skill among the top performers. They did agree with Kosowski et. al.'s (2006) findings regarding the worst performers, which both concluded to be due to bad skill.

2.2 Non-US Studies

Blake and Timmermann (1998) examined the performance of 2300 different U.K. mutual funds in the period 1972-1995. They found that the U.K. mutual fund on average produced a negative abnormal return of 1.8 percent. They also found some evidence that funds produced a small positive risk-adjusted return in the first year of the fund's existence. This effect dissipated and disappeared within the funds second year. Otten and Bams (2002) examined 506 mutual funds from France, Italy, Germany, UK and the Netherlands. Their sample was controlled for survivorship bias. They found, using the four-factor model of Carhart (1997), that on aggregate, French, Italian, Dutch and U.K funds outperformed the benchmark and provided positive alpha, gross of costs. German funds underperformed the benchmark, but not significantly. Deducting the costs, they found that only U.K. funds produced a significantly positive alpha net of costs. Cutbertson, Nitzche, and O'Sullivan (2007) examined UK mutual funds in the period 1975-2002 by implementing the bootstrap method of Kosowski et. al. (2006). Their results were similar to Kosowski et. al. (2006). Namely, there was evidence of stock-picking skills among the top performing funds and evidence of bad skill among the worst performing funds.

The coverage of mutual funds performance in Scandinavian countries is relatively limited compared to the U.S. studies. We review the following Scandinavian studies: Dahlquist, Engström, and Söderlind (2000, Sørensen (2009), Christensen (2013) and Gallefoss, Hansen, Haukass and Molnar (2015). Dahlquist, Engström, and Söderlind (2000) examined the performance and characteristics of Swedish mutual funds from 1993-1997. They found mixed results for regular equity funds, special equity funds, bond, and money market funds. They found that the performance of regular equity funds was somewhat superior. They concluded that there was some evidence suggesting that actively managed equity funds performed better than more passively managed funds.

Sørensen (2009) used the modified bootstrap method of Fama and French (2010) to do a comprehensive evaluation of the performance of Norwegian mutual funds, in the period 1982-2008. When he examined Norwegian mutual funds on aggregate, he found no statistically significant evidence of abnormal performance. When he looked at the funds individually; he found no evidence of skills among the best performers, only bad skill among the worst performers. He concluded that the best Norwegian mutual funds were just lucky while the worst ones exhibited bad skill.

After this, Christensen (2013) provided the first independent analysis of Danish mutual funds. He used the CAPM with the approach of Treynor and Mazuy (1966) to evaluate the performance of 71 Danish mutual funds, in the period 2001-2010. He found that on average funds underperformed the benchmark. Only five funds in his sample of 71 generated positive alphas. A whole 80% of funds in his sample underperformed.

Gallefoss, Hansen, Haukass, Molnar (2015) examined Norwegian mutual funds using daily data, controlled for survivorship bias, in the period 2000-2010. Using the bootstrap method of Kosowski et. al. (2006), they found statistically significant abnormal performances on both ends of the performance distribution. In other words, evidence of skill for the best performing funds and evidence of bad skill among the worst performing funds. Additionally, they found that Norwegian mutual fund underperforms at the aggregate level.

The master thesis of Utseth and Sandvik (2015) examined Norwegian mutual funds in the period 1983-2014. They implemented the bootstrap procedure of Kosowski et. al. (2006) and controlled the sample for survivorship bias. They found no evidence of superior fund managers, rather evidence indicating that fund managers performed worse than random chance would have produced, i.e. bad skill.

2.3 Performance Persistence

The studies reviewed so far show little to no evidence that actively managed mutual funds on aggregate outperform their benchmark. However, there still might be a chance that some individual fund managers occasionally outperform their benchmark and that the performance persists over a subsequent period. Hence, performance persistence is important from both an academic viewpoint as well as a practical viewpoint. From an academic viewpoint, evidence of performance persistence challenges the semi-strong form of the EMH. From the practical viewpoint, evidence of performance persistence suggests that investors can exploit past performers to earn risk-adjusted returns in future periods. Studies on persistence in mutual funds performance have a long history. They all agree on the relevance of performance persistence but disagree on whether performance persistence is present and how long the persistence lasts.

Hendricks, Patel, and Zeckhauser (1993) looked for persistence in US mutual funds in the period 1974-1988 using the recursive portfolio approach. They called the phenomenon of short-run persistence either "hot-hands" or "icy-hands." Hot-hands is when past good performers perform well in the near future. Icy-hands is the "evil" counterpart, persistence among the poorest performers. They found evidence for the icy-hands phenomena. Poor performers continued to be poor performers in the near future. They also found evidence of the hot-hands phenomenon. Both types of persistence were among growth funds. They did not use risk-adjusted returns which we consider a weakness. Brown and Goetzmann (1995) looked for persistence in U.S. mutual funds in the period 1976-1988. They found evidence of persistence among loser funds and winner funds. However, this effect was only present in 8 out of the 12 years. They said persistence was strongly dependent upon the time period of study. Carhart (1997) also looked for persistence and found evidence for persistence among the worst performers. He claimed that the hot-hands phenomenon found in Hendrick, Patel and Zackhauser (1993) could be explained by the momentum effect of Jegadeesh and Titman (1993). He attributed persistence to the four-factor loadings, the expense ratios and transaction costs of the funds, not the stock picking skills of the managers. In their study, Elton, Gruber and Blake (1996) used risk-adjusted returns and found evidence of short-run persistence among both winner and loser funds. Additionally, they found evidence of persistence over a longer time period, up to three years. They also stressed the importance of controlling the sample for survivorship bias when doing persistence studies.

Another study by Carpenter and Lynch (1999) looked at the effect of survivorship bias and look-ahead bias on performance persistence tests. They found that survivorship bias created spurious reversals in performance results, therefore failure to control for these biases distorted the results. They also examined the effect of attrition on persistence tests. The results of their tests reinforced the findings of earlier persistence tests. In our data and method, we have controlled for the biases mentioned.

Chen, Jegadeesh and Russ (2000) investigated whether high turnover funds outperformed low turnover funds in U.S. mutual funds. The theory was that skilled managers identify attractive investment opportunities more often than non-skilled managers. To be more specific, they investigated whether stocks held by high turnover funds outperformed stocks held by low turnover funds. The evidence pointed to a marginal ability by high turnover funds to outperform low turnover funds. They did not conclude that this difference in performance was big enough to outweigh the increased costs associated with increased turnover, whether the net result is positive. The authors speculated that this was the root of Carhart's (1997) negative relation between return and turnover. They also found weak evidence that stocks bought by funds outperform stocks newly sold, at least for the first year.

Berk and Green (2004) derived a model of active portfolio management. Using their model, they found evidence short-term persistence in net returns but the persistence effect disappeared with longer evaluation periods. They argued that the persistence in net returns was competed away by fund inflows. After this, Bollen and Busse (2005) investigated whether the ability to pick stocks persisted over time. They used daily data on U.S. mutual funds and the four-factor model of Carhart (1997). They found evidence of persistence over 3-month periods for the top decile of funds. Over longer periods the persistence dissipated and disappeared.

Kosowski et. al. (2006) found inferior and superior fund managers among U.S. mutual funds. They found that the performance of these funds persists. Thus, contradicting the findings of Carhart (1997). Kosowski et. al. (2006) used the similar test method as Carhart (1997), but made the inference based on the bootstrap rather than the parametric t-tests employed by Carhart (1997). Huij, Verbeek (2007) investigated whether there is short-term performance persistence among funds. The data they used consisted of 6400 U.S. funds, from 1984-2003. They concluded that past performance had some predictive power for future performance. More specifically, when they had ranking periods of 1 year, the top decile earned a statistically significant alpha of 0.26. This was short term and strongest one month after the ranking. After one month this positive alpha dissipated and gradually disappeared.

Next, Javier Vidal-García (2013) looked for persistence in the performance of European equity mutual funds in the period 1988-2010. He tested whether the persistence was related to investment style. He used a large sample consisting of data from six European countries and controlled for survivorship bias. He found strong evidence of persistence in benchmark-adjusted returns among European mutual funds. The inference was based on the bootstrap method of Kosowski et. al. (2006). He founds that the benchmark-adjusted returns persisted up to 36 months, making him conclude that mutual funds that performed well in the past are likely to do well in the future.

Among the studies on Scandinavian mutual funds, Dahlquist et. al. (2000) and Sørensen (2009) detected no evidence of persistence in the performance of Swedish mutual funds and Norwegian mutual funds, respectively. Using daily data, Gallefoss, Hansen, Haukass, Molnar (2015) found strong evidence of short-term persistence among the top performing fund and the worst performing fund. They found that the persistence lasted up to one year into the future.

3 Methodology

This Chapter presents the methods implemented in this study, starting with the method of estimating the risk-adjusted performance of Norwegian mutual fund, followed by the bootstrap method for separating skill from luck in the performances of the mutual funds and ending with the performance persistence methodology. The model we use is selected to get a good estimate of the funds' performance, accounting for known anomalies. The bootstrap method is chosen to evaluate the statistical significance of that performance correctly. This chapter is meant to give a walkthrough of the methods used in our analysis.

3.1 Model Selection

Following the traditional methods for fund performance evaluation, we compare a fund's return to the return implied by a factor model for returns. We want to isolate the return that is directly attributable to the individual fund manager's stock-picking skill or their lack of skill (i.e. bad-skill). To do that we have to explain and attribute the parts of a fund's return that are generated by the exposure to the risk factors. The return not explained by the factors is attributed to stock-picking skills. We isolate the return attributed to stock-picking skill, by running a time-series regression on the factor model. From the regression, an intercept (alpha) is obtained which is the risk-adjusted return attributed to stock-picking skills. The most common factor models are the single-factor model of Jensen (1968), the three-factor model of Fama and French (1993) and the Carhart's (1997) four-factor model. The single-factor model only controls for the market factor, and the return not explained by the market factor is estimated as stock-picking skills. This estimated may be misleading since it does not control for known market anomalies. The return generated by the anomalies must be controlled for since such return is not attributable to managers skill. One could employ a mechanical strategy to exploit the anomalies to gain positive risk-adjusted returns, but it does not indicate any level of skill in the stock selection. We use the four-factor model of Carhart (1997) since the model it controls for return generated by known anomalies in the market. The motivation of the choice of factor model will be discussed further in the following subsections.

3.1.1 Single-Factor Model

We start with the single-factor model (1), developed by Jensen (1968). It is the foundation for all risk-based performance measures. The model is based on the market equilibrium model known as Capital Asset Pricing Model (CAPM), developed by Sharpe (1964) and Lintner (1965). The CAPM describes the relationship between risk and returns for a given asset, based on its exposure to the market factor. Jensen (1968) modified the CAPM to add the alpha variable. The alpha measures the return which is above or below what the theoretical CAPM would predict. It is the abnormal return or the disequilibrium return of an asset. If the market is in equilibrium, the expected value of the alpha would be equal to zero. If the market is in not in equilibrium, positive and negative alpha may occur. A positive alpha (i.e., $\alpha > 0$) is interpreted as the return that exceeds the risk, meaning the fund performs better than what the model suggests. A negative alpha (i.e., $\alpha < 0$) means that the return of the fund performs worse than what the model suggests. The single factor model is presented below:

$$r_{i,t} = \alpha_i + \beta_i MKT + \varepsilon_{i,t} \tag{1}$$

where $r_{i,t}$ is the excess return of an asset i at month t and MKT is the market excess return. The error term, $\varepsilon_{i,t}$, is the specific risk of individual assets i and can be diversified away, leaving only the systematic risk or the market risk. The beta, β_i, tells us how much the asset will change in value when the market changes value. It's the sensitivity of the asset to the market. The beta indicates how much the asset is exposed to market risk.

In the single-factor model, the intercept, α_i, of fund i is the Jensen's alpha. The abnormal return, α_i is used as an indication of how well the fund i performs after accounting for the risk (risk-adjusted performance).

3.1.2 The Fama-French Three-factor Model

The single factor model by Jensen (1968) accounts only for the market factor. Evaluating the stock-picking skill using this model will not be adequate as it does not control for known anomalies in the stock market (i.e. some stock types is expected to perform better than others). It is important to control for these anomalies as fund managers tend to exploit them to generate positive alphas. Exploiting anomalies is not considered as stock-picking skills since anyone can do it. Studies on the behavior of expected stock returns lead to the development of multi-factor asset pricing models (e.g. Fama-French's three-factor; Carhart's four-factor). These models capture common risk factors in expected stock returns.

Fama and French (1993) augmented the single factor model to include two additional factors, the size factor, and the value factor. Earlier empirical research, (Stattman, 1980; Banz, 1981; Rosenberg, Reid and Lanstein, 1985), observed two main contradictions of the single factor model. The size effect and the value effect. These additional factors are meant to capture common risk factors in cross-section of returns and add further explanatory power to the model.

The size effect is related to the market capitalization of the stock and the anomaly that small capitalization stocks tend to give higher returns than predicted by their beta. Vice versa big capitalization stocks tend to give lower returns than predicted by their beta. The size effect is captured in the Small-Minus-Big or SMB factor. It is constructed as a portfolio with long positions in small market capitalization stocks and short positions in large market capitalization stocks. More accurately, the market is ranked on size, from big to small. The median size is then used to divide the market into two parts, one "small" and one "big." These two parts are then used in the hedging portfolio to create the factor.

The value anomaly is related to the book-to-market value of the stock. Stocks with high book-to-market ratio gives returns above what the CAPM predicts. High book-to-market stocks are known as growth stocks. The factor is constructed as a portfolio with long positions in the high B/M ratio stocks and short in low B/M ratio stocks. The market is sorted from high B/M to low B/M, then the highest 30% and the lowest 30% are put in a hedging portfolio, long in the high 30% and short in the low 30%. The value effect is captured in the high-minus-low or HML factor. The evidence from Fama and French (1992) indicates that book-to-market equity has more explanatory power than the size factor. The two factors are added to measure the mutual fund portfolio's exposure to these two classes of stock.

$$r_{i,t} = \alpha + \beta_{1i} MKT_t + \beta_{2i} SMB_t + \beta_{3i} HML_t + \varepsilon_{i,t} \tag{2}$$

3.1.3 The Carhart Four-factor Model

The three-factor model of Fama-French (1993) controls for the size and value anomalies, but it does not control for the momentum anomaly. To control for this momentum effect, Carhart (1997) augmented the three-factor model, by adding the one-year momentum factor of Jegadeesh and Titman (1993), PR1YR. The PR1YR factor is meant to capture the one-year momentum anomaly, discovered by Jegadeesh and Titman (1993). The momentum anomaly is that stocks that have risen (fallen) in value the previous period (within past year) have a tendency to rise (fall) further in the subsequent period. The momentum factor measures the portfolio's exposure to this anomaly. The factor itself is constructed as a portfolio with long positions in the 30% stocks with highest one-year lagged returns, and short in the 30% of stocks which have the lowest one-year lagged returns. For more detailed description of the risk factors, see Fama and French (1993) and Jegadeesh and Titman (1993).[4]The Four-factor model of Carhart (1997)

[4]For more detailed explanation of how the risk factors are constructed, see Kenneth French's data library (`http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`) and Jegadeesh and Titman (1993).

can be estimated as:

$$r_{i,t} = \alpha + \beta_{1i} MKT_t + \beta_{2i} SMB_t + \beta_{3i} HML_t + \beta_{4i} PR1YR_t + \varepsilon_{i,t} \tag{3}$$

We use the four-factor model of Carhart (1997) as our performance model. The four-factor model has the ability to control for the four most common non-diversifiable risk factors in stock returns. By controlling for these factors, we can correctly estimate the performance of mutual fund managers that are due to stock-picking skill.

3.2 Bootstrap

Traditional OLS inference is a parametric approach and depends upon the Gauss-Markov assumptions. One of this assumptions is that the residuals are normal distributed. Violation of this normality assumption makes the parametric inference unreliable. When analyzing mutual fund alphas, we find many properties that would lead to a rejection of the normality assumption, thus making the standard parametric test-statistics invalid. Kosowski et. al. (2006) pointed out several reasons why these properties may arise. First, the returns of individual stocks within the typical mutual fund tends to have a skewness and kurtosis statistically different from a normal distribution. Although the central limit theorem implies that a large equal-weighted portfolio of non-normal distributed stock will approach normality, fund managers generally hold large positions in a few stocks. Second, the returns of individual stocks tend to be auto-correlated and have heteroscedastic variance. Finally, mutual funds may implement dynamic strategies that involve adjusting their level of risk in response to the changes of risk in the overall market portfolio or response to their performance ranking to similar funds. Each of these regularities may contribute to non-normality of mutual fund alphas, which makes normality a poor approximation for the typical fund. Specifically, tests show that normality of residuals is rejected for 61% of the mutual fund in our sample. Additionally, the normality of returns is rejected for 74% of the mutual fund. [5]

We have now discussed reasons for non-normalities in individual mutual fund residuals and therefore in their alphas. These non-normalities in individual mutual funds implies non-

[5] The Jarque-Bera test is used to test for normality in the residuals and returns at the 5% significance level.

normalities in the cross-section of residuals as well as the cross-section of alphas. In addition to the non-normality in individual fund residuals, properties like cross-sectional differences in sample size (i.e., fund lives) and heterogeneous risk-taking across fund cause non-normalities in the cross-section. Meaning, non-normalities can occur in the cross-section of funds, even if individual mutual fund are normally distributed.[6] Evidence of one or more of these properties will lead to a complex and non-normal distributed cross-section of the funds' performance measure and must be evaluated with the bootstrap.

Given the bootstrap's great feature of not relying on any distribution assumptions, it can substantially improve the validity of inference about the performance of mutual funds. Bickel and Freedman (1984), Hall (1986), Horowitz (2003), Kosowski et. al. (2006), and Fama and French (2010), all argued that the bootstrap provides more accurate evaluation of the significance of alpha estimates. In Monte Carlo experiments conducted by Horowitz (2003), he demonstrates that the bootstrap can significantly reduce the difference between the true and nominal probability of correctly rejecting a given null hypothesis. For example, by identifying thick tails in individual fund returns, the bootstrap does not reject the abnormal performance as often compared to the standard parametric t-test.

We implement a bootstrap method to evaluate the performance of Norwegian mutual fund. The bootstrap is a nonparametric approach to statistical inference introduced by Efron (1979). The basic idea of bootstrap is to estimate the distribution of the statistic of interest in order to infer, instead of relying on parametric assumptions about the distribution like normality. The bootstrap estimates the distribution by resampling multiple times with replacement from the original sample and compute the statistic of each resample. In our paper, we use the bootstrap to estimate cross-sectional and individually distributions of mutual funds performance measure, $\hat{\alpha}$ (or $\hat{t}_\alpha$), where zero true performance is imposed by construction. Bootstrapping is necessary for proper inference about performance since mutual funds exhibit properties which make the parametric approach unreliable. These properties include non-normalities in individual funds

[6] As an example, consider a scenario where we have a sample of 1000 mutual funds, each existing 396 months, with normally distributed residuals, but with heterogeneous level of risk so that, across funds, residual variances vary uniformly between 0.5 and 1.5 (i.e., the mean variance is unity). Under these assumptions, the tails of the cross-sectional distribution of residuals and alphas will now be fatter than those of a normal distribution. Meaning, as we move further into the tails of the distributions, the probability of extreme outcomes does not fall very quickly compared to if it had been normally distributed, since high-risk funds more than compensate for the large drop in such extreme outcomes from low-risk funds (Kosowski et. al., 2006).

returns as well as non-normalities in the cross-section of funds alphas. How to implement the bootstrap is described in the next section.

3.2.1 Implementation of The Baseline Bootstrap Procedure

Following the bootstrap method developed by Kosowski et. al. (2006), which involves residual-only resampling under the null hypothesis of no abnormal return, the first step is to prepare for the bootstrap procedure. We use the Carhart's (1997) four-factor model to compute ordinary least squares (OLS)-estimates of alpha, factor loadings and residuals for fund i,

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_{1i}MKT_t + \hat{\beta}_{2i}SMB_t + \hat{\beta}_{3i}HML_t + \hat{\beta}_{4i}PR1YR_t + \hat{\varepsilon}^b_{i,t_\varepsilon} \tag{4}$$

saving the coefficients estimates, $\{\hat{\alpha}_i, \hat{\beta}_{1i}, \hat{\beta}_{2i}, \hat{\beta}_{3i}, \hat{\beta}_{4i}\}$, the estimated residuals, $\{\hat{\varepsilon}_{i,t}\}$, and the t-statistic of alpha, $t_{\hat{\alpha}_i}$ for fund i.

The second step is to draw a random sample with replacement from fund i's residuals to construct a pseudo-random time-series of resampled residuals that has the same length as the original sample of residuals, $\{\hat{\varepsilon}^b_{i,t_\varepsilon}, t_\varepsilon = s^b_{T_{i0},\ldots,T_{i1}}\}$, where T_{i0} and T_{i1} are the dates of the first and last monthly returns available for fund i and where b represents an index for the bootstrap iteration.

The next step is to use time-series of resample residuals combined with the fitted values from the first step (i.e. factor returns multiplied by the estimated beta coefficients) to construct a time-series of pseudo-monthly excess returns for fund i, where we impose the null hypothesis of zero true performance ($\alpha_i = 0$, or equivalently, $\hat{t}_{\hat{\alpha}_i} = 0$),

$$\{\tilde{r}^b_{i,t} = \underbrace{\beta_{1i}MKT_t + \hat{\beta}_{2i}SMB_t + \hat{\beta}_{3i}HML_t + \hat{\beta}_{4i}PR1YR_t}_{\text{Excess return with } \alpha = 0} + \underbrace{\tilde{\varepsilon}^b_{i,t_\varepsilon}}_{\text{Sampling Variation}} \} \tag{5}$$

The time-series of artificial returns in equation (5) is constructed to have a zero true performance. However, when the artificial returns, for a given bootstrap sample b, are regressed on the four-factor model of Carhart (1997), we may obtain a non-zero estimated alpha (and t-statistic) depending on the drawn residuals. For example, if an abnormally high number of positive (negative) residuals are drawn in a given bootstrap sample, b, a positive (negative) alpha (t-statistic) may result and is a result of sampling variation around the zero true performance (due to luck). For all bootstrap iterations, $b = 1,2,...,B$, the above steps are repeated across all funds to build the cross-section of bootstrapped alphas, $\tilde{\alpha}_i^b$, $i = 1,2,...,N$, or their t-statistics, $\tilde{t}_{\hat{\alpha}_i}$, $i = 1,2,...,N$. The bootstrapped alphas are then sorted from the lowest to the highest, $\tilde{\alpha}_{min}^b$ to $\tilde{\alpha}_{max}^b$, across all bootstrap iterations, resulting in a $B \times N$ matrix, where column 1 contains the lowest values of bootstrapped alphas and column N contains the highest values of bootstrapped alphas from the simulations. The columns of the matrix represent the luck distributions for each of ex-post ranked fund. For example, the last column, N, contains all the highest alpha values across all bootstrap iterations and is the luck distribution, $f(\tilde{\alpha}_{max}^b)$, for the top fund, $\hat{\alpha}_{max}$. This sorting procedure is also done for the t-statistic. By comparing the ex-post ranked fund alpha (or t-statistic) to its respectively luck distribution, we can separate skill from luck. For example, if we find for the top fund that the probability of the bootstrap generating alphas (or t-statistic) greater than the actual observed alpha (or $\hat{t}_{\hat{\alpha}_i}$) is less than our significance level of 5%, we conclude that performance is not due to sampling variation (i.e. luck), but rather superior manager skill.[7]

In the implementation, we consider two test-statistics, the estimated alpha, $\hat{\alpha}_i$, and its estimated t-statistic, $\hat{t}_{\hat{\alpha}_i}$. Of the two, alpha is the most widely used and measure the economic size of abnormal performance. However, Kosowski et. al. (2006), Busse et. al. (2010) and Fama and French (2010) argue for using the t-statistic of alpha rather than the estimated alphas. The reason why focusing on the t-statistic is more advantageous than alpha, is that the alpha tends to suffer from a lack of precision in the construction of confidence intervals. This imprecision is caused by the alpha estimates of short-lived funds or funds with a high level of risk-taking, as they will have a higher variance estimated alpha distribution. Thus these alphas tend to be spurious outliers in the cross-section of alphas. The t-statistic corrects for these spurious outliers. By rescaling the estimated alpha by its standard deviation, we get estimates which are adjusted for the difference in the precision of alpha. Thus the t-statistic can be seen as precision-adjusted estimates of alpha.

[7]The bootstrapped p-values for funds in located in the right-hand side of the distribution i calculated as the proportion of bootstrapped alpha (or $\hat{t}_{\hat{\alpha}}$) > actual alpha (or $\hat{t}_{\hat{\alpha}}$) at each percentile point. The bootstrapped p-values for funds in located in the left-hand side of the distribution i calculated as the proportion of bootstrapped alpha (or $\hat{t}_{\hat{\alpha}}$) < actual alpha (or $\hat{t}_{\hat{\alpha}}$) at each percentile point.

3.2.2 Bootstrap Extensions

We implement some different extensions of our baseline bootstrap procedure. Our baseline bootstrap procedure assumes independent residuals (i.e., no autocorrelation and no cross-correlation). In the first extension, we open up for the possibility of residual to be autocorrelated. The second extension, we address the issue that the factor returns may be autocorrelated. The third extension allows for the residuals to be cross-correlated. The final extension we implement the modified bootstrap procedure suggested by Fama and French (2010), which allows for correlation between the factor returns and residuals. In addition to these extensions, we test whether our results are robust to the choice of minimum data length requirement (i.e., look-ahead bias vs. imprecision). We go further into details in Section 5.3, Sensitivity Analysis.

3.3 Performance Persistence Methodology

As mentioned earlier, an investor wants to invest in a fund that not only generates positive alpha, but also does so consistently over time. A performance persistence test allows us to see if past performance is predictive of future performance. Testing for persistence is important from both an economic and an academic viewpoint. From an economic point of view, evidence of persistence among top performing funds suggests that investors would know which funds would perform well, based on prior performances. Investors would know ex-ante which funds would be worth buying. Performance persistence could be exploited to earn positive alpha, by simply buying the funds that recently performed best. From the academic viewpoint, evidence of performance persistence would challenge the semi-strong form of the EMH, supporting a rejection of this hypothesis.

If funds performance persist over time it is an indication of skill among the respective fund managers. If there is no type of skill we expect the funds performance to oscillate between good and bad performance over time, an indication that performance is random. The effect survivorship bias has on performance persistence tests are well documented in Carpenter and Lynch (1999), we have, however, controlled for this bias.

3.3.1 Implementation of Performance Persistence Test

The performance persistence test we perform is known as a recursive portfolio formation test (see, e.g., Hendricks et. al.,1993; Carhart, 1997). It involves sorting the funds into different quintile portfolios based on the funds' previous performance and then evaluate the risk-adjusted performance (i.e., alpha) of each quintile portfolio over a future holding period. We use the same sorting procedure as Hendricks et. al. (1993) and Carhart (1997), but rather than using a standard parametric t-test to evaluate the significance of the alpha, we use the bootstrap method described in Section 3.2. Given the properties of mutual funds, the bootstrap analysis will sharpen the p-values, but the alpha estimates will be unchanged compared to the parametric test.

The test uses two periods, a ranking period and a holding period. The ranking period is related to past performance, while the holding period is related to future performance. In the ranking period, which in our case lasts either a year or three years, we rank funds based on three different performance measures: Returns, the Carhart (1997) four-factor alpha estimates and the t-statistic of the four-factor alpha. We rank the funds on their performance, from best to worst. The funds are sorted according to the chosen performance measure, into quintile portfolios. Thus, we create a portfolio of the 20% best performing funds, all the way down to the 20% worst performing funds, resulting in five different 20% portfolios. In addition to these five portfolios, we have a spread portfolio which constructed by a long position in the top 20% portfolio and short in the worst 20% portfolio.

We then hold these portfolios over the holding period, which in our case lasts either a year or three months, before rebalancing them for the next ranking period. In the holding period, we estimate the alpha of these six different portfolios. We evaluate the performance persistence by looking at the alpha generated by the portfolios over the holding period. If any portfolio generates a statistically significant alpha, it indicates that there is persistence in the performance among the funds in that portfolio. We test the null hypothesis of zero true performance of the portfolios, rejection of the null is evidence of persistence. If the funds generate alpha significantly different from zero over the holding period, we have a statistically significant result that indicates performance persistence.

For example, with one-year ranking and holding periods, funds are first ranked on January 1, 1984, by their performance (return, alpha or t-alpha) over the period January 1983 to December 1983. The top 20% performing funds is put into a portfolio to construct an equal-weighted portfolio of the top 20% performing funds' monthly excess returns. Next holding period the portfolios are rebalanced to contain the top 20% of the ranked funds in the next ranking period. We repeat this procedure until the end of the sample period to construct a time-series composed of past year best performers.

4 Data

In this Chapter, the data used in our study to evaluate the performance of Norwegian mutual funds is presented. Throughout this section, details regarding the data and its providers will be reviewed.

4.1 Norwegian Mutual Funds

A mutual fund is a pooled investment vehicle, meaning that many individual investors pool their money and let a professional manager invest it on their behalf. The manager does not do it for free, of course. The expenses of the fund are retracted from the funds return before any investor gets his share. A mutual fund provides diversification, professional management, and liquidity for an investor.

The Norwegian Fund and Asset Management Association divides mutual funds into four distinctive classes: Norwegian stock/equity funds, Norwegian/international stock/equity funds, international stock/equity funds and sector stock/equity funds. To choose an appropriate benchmark, we only consider Norwegian stock/equity funds. A fund is defined as "Norwegian" if it invests at least 80% of its assets in domestic securities. Our paper uses a comprehensive data set comprised of monthly net returns of all 98 actively managed Norwegian mutual fund (surviving and non-surviving funds) during the period 1983-2015, covering almost the entire lifespan of Norwegian mutual fund.[8]

The monthly returns of a fund are computed using the fund's monthly Net Adjusted value (NAV). Daily NAVs are obtained from the TITLON database and converted into monthly NAVs

[8] A list of all mutual funds in our data set will be found in the Appendix

by using the last observed NAV in each month for all funds. [9] The NAV is the price that an investor pays to buy one share in the fund and can be compared to the price of a stock. The NAV is calculated in the following way:

$$NAV = \frac{(asset - liabilities)}{number\ of\ outstanding\ shares} \tag{6}$$

The liabilities contain all the fund's expenses, managerial fee and so on. Therefore, the returns we calculate are all net of fund costs. We use the NAV's to calculate the fund monthly returns net of cost as follows:

$$r_{i,t} = \frac{NAV_{i,t} - NAV_{i,t-1}}{NAV_{i,t-1}} \tag{7}$$

In Appendix, Table A.I provides an overview of the exact number of funds available at the end of each year, in addition to how many funds were born and liquidated each year, throughout the whole sample period. The table also shows the comparison between the returns of an equally weighted portfolio of all fund in our sample and the Oslo Børs Mutual Fund Index.

4.1.1 The Structure of Mutual Funds

All Norwegian mutual funds are all open-end domestic equity funds. Open-end means that there is no set on the number of fund shares available on the market and fund shares can be issued and redeemed at any given time. The Norwegian fund sector has grown significantly during its lifespan. Gjerde and Sættem (1991) report that the market value of the Norwegian mutual fund was 290 million NOK at the end of 1982.[10] During the period 1982 to 2015, the Norwegian mutual fund and Asset Management Association (VFF) reports that the market value of 290 million NOK has increased to 86.7 billion NOK.

Figure 1 below plots the number of Norwegian equity mutual funds available to investors each year at the Oslo Stock Exchange from 1983 to 2015. The figure shows that during the period 1990 to 1999, the number of Norwegian Mutual funds available grew rapidly, from 10

[9] `https://titlon.uit.no/`: TITLON is a database with financial data from Oslo Stock Exchange, developed jointly by a couple of universities in Norway and contains information on mutual funds, as well as stocks, options, warrants, futures, and bonds.

[10] Prior to 1982, there was only one single Norwegian mutual fund on Oslo Stock Exchange.

to 59. During the period 2000-2011, the number of funds is relatively stable, varying between 56 to 66. The last period, 2012-2015, seems to indicate a downward trend in the number of Norwegian mutual funds.

Figure 1: Number of equity mutual funds on the OSE

The figure plots the number of Norwegian equity mutual funds available to investors each year at the Oslo Stock Exchange from 1983 to 2015.

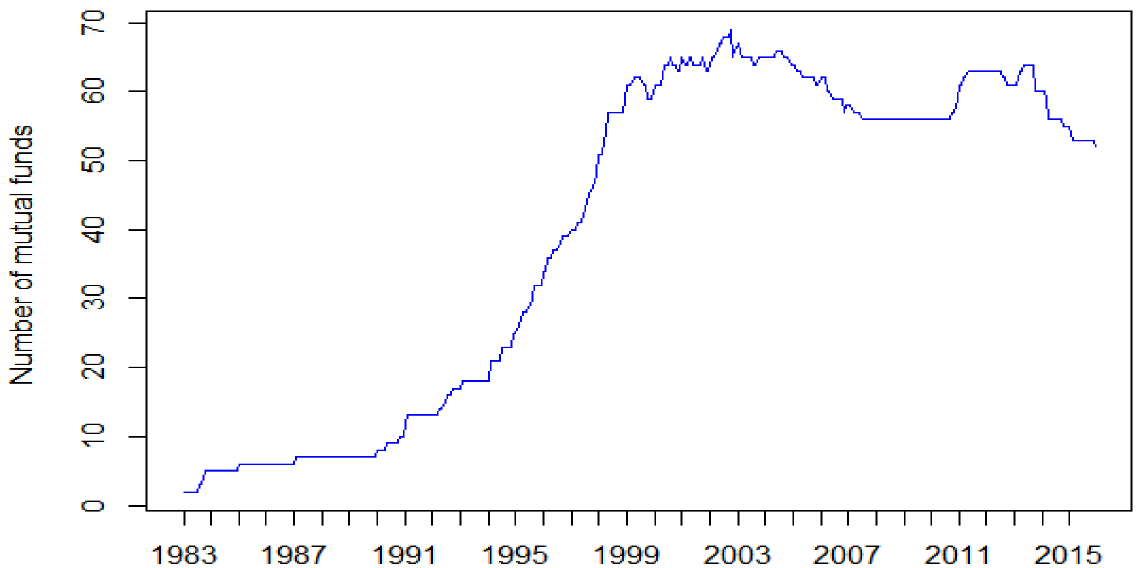


Table 1 reports some descriptive statistics on the Norwegian Mutual fund industry for the period 1994 to 2015. [11] Column three of Table 1, shows that the average number of customers per fund decreases gradually for each year in the period 1998 to 2015. The same trend is found in the last two columns, which show the Norwegian mutual fund market share, in percentage, of the total equity fund market, and of the total fund market. From 1994 to 2015, the Norwegian market share decreases from 92% to 19.9% of the total equity fund market, and 37% to 9.6% of the total fund market. VFF reports that international equity fund's market share of the total fund market increases from 3% in 1994 to 37% in 2015, which suggest that investors prefer the broader diversification provided by international equity funds instead of Norwegian mutual fund.

[11] We thank the VFF for providing us with the descriptive statistics in the Norwegian Mutual fund industry. The descriptive statistic only dates back to 1994, since VFF did not record data prior to 1994.

Column 4 of Table 1, shows the average size of Norwegian mutual fund, measured in Asset Under Management (AUM). The average fund size grew steadily from 215 million NOK in 2002 to 746 million NOK by the end of 2007, so to decrease drastically during 2008. In the end of 2008, the average AUM from 2007 was almost halved, from 746 million NOK to 359 million NOK. This drastic reduction in AUM is mainly due to significant drop in equity prices caused by the global financial crisis. The average AUM recovers quickly and by the end of 2009 the reported average AUM is 822 million NOK. From 2009, the average AUM increases steadily and by the end of 2015 the average fund size is 1112 million NOK.

Table 1
Descriptive Statistics of the Norwegian Mutual Funds Market

This table lists the statistics for Norwegian mutual funds in the period 1994 to 2015. Column two reports the number of funds in the sample each year. Column three reports the average number of customers per fund. Column four reports the average fund size measured in Asset Under Management. Column 6 reports the Norwegian mutual fund market share of the total equity fund market. Column 7 reports the Norwegian mutual fund market share of the total fund market. Fund size and flow data are reported in million NOK. The market shares are in percent.

Year	Number of funds included	Average number customers per fund	Average fund size	Average net inflow	% of total equity fund market	% of total fund market
1994	25	10,987	235	8	92.0	37.0
1995	32	9,689	220	8	91.9	33.5
1996	39	12,824	431	96	88.1	42.4
1997	48	14,858	604	140	80.1	47.8
1998	59	15,878	403	4	67.3	38.4
1999	59	14,255	573	7	46.1	30.8
2000	63	11,537	459	-22	38.3	24.6
2001	63	11,302	374	-11	37.0	20.7
2002	66	9,024	215	-11	37.1	15.8
2003	65	9,281	351	-1	35.9	17.3
2004	64	8,342	421	-52	31.8	16.8
2005	61	6,854	504	-61	26.2	14.1
2006	58	6,175	635	16	24.5	14.8
2007	56	6,726	746	-44	23.1	12.9
2008	56	6,571	359	-1	19.7	8.7
2009	56	6,874	822	153	24.8	13.9
2010	59	6,281	1063	60	26.6	15.6
2011	63	6,017	833	-18	24.6	12.5
2012	61	5,745	945	-10	24.5	12.2
2013	60	4,634	1087	-13	22.4	12.3
2014	55	4,138	1090	-25	20.9	10.2
2015	52	3,877	1112	-51	19.9	9.6

4.2 Interest Rate

In order to compute the abnormal performance of Norwegian mutual fund using the Carhart four-factor model described in Chapter 3, we need the funds' excess returns. We construct the excess returns by subtracting the risk-free interest rate from the return of each individual fund. In this paper, the proxy of the risk-free interest rate is collected from Ødegaards' database.[12] Ødegaard (2017) argues that the Norwegian Interbank Offered Rate (NIBOR) is the most appropriate proxy of the interest rate and has constructed a monthly risk-free rate from the one-month NIBOR rate. The overnight NIBOR rate is used as an approximation for the period 1983-1986 due to slightly messy interest rate data prior to 1986.[13] Since NIBOR is quoted in annualized terms, it must be changed into monthly terms. The monthly risk-free interest rate at time t is computed as:

$$r_{f,t} = (1 + NIBOR_t)^{1/12} - 1 \tag{8}$$

4.3 The Benchmark index

To measure the performance of mutual funds, we need to choose an appropriate benchmark index to represent the market returns. A natural choice is the Oslo Børs Mutual Fund Index (OSEFX) for which serves as the benchmark for most Norwegian mutual funds. The OSEFX is a capped version of the OSEBX and is adjusted to meet specific requirements, and to fulfill the rules set by the UCITS directives for regulating investments in mutual funds.[14] Norwegian mutual funds must invest in at least 16 different stocks, and the weight in any stock must not exceed 10%. The OSEFX takes these requirements into account. Our sample of mutual fund returns starts in January 1983, and since the OSEFX originated in December 1995, it cannot be used as the benchmark for the whole sample period. We, therefore, use the linked OSEFX which adjust the values of the total index (TOTX) by an adjustment-factor equal to the ratio between the total index and the OSEFX to create historic values before the official start of the OSEFX. The data for the Linked Oslo Børs Mutual Fund Index is obtained from the TITLON

[12] `http://finance.bi.no/~bernt/financial_data/ose_asset_pricing_data/index.html`

[13] The Figure in Appendix B shows a visual presentation of the development of the interest rate.

[14] UCITS: European Union's Undertakings for Collective Investment in Transferable Securities Directive

database for the period 1983 to 2015.

Table 2 below presents the descriptive statistics for the OSEFX returns in addition to total returns for an equal-weighted portfolio of all mutual fund returns, equal-weighted returns on dead funds and equal-weighted returns on surviving funds only. Panel A shows that the average mutual fund outperformed the benchmark over the whole sample period 1983-2015. This is however in terms of raw returns, not in risk-adjusted returns. Panel B shows that in the first half of the sample period the average mutual fund had a strong performance versus the benchmark index, OSEFX. Panel C shows that the average mutual fund performed slightly worse in the second half of the sample period.

Table 2
Summary Statistics of Benchmarks and Fund Returns

The table presents the descriptive statistics for the OSEFX returns, returns for an equal-weighted portfolio of all mutual fund returns, equal-weighted returns on non-surviving funds and equal-weighted returns on surviving fund. Column 1 - 2 shows the average annualized return, and standard deviation, Column 3-4 shows the monthly maximum and minimum return. Column 5-6 report the skewness and kurtosis of the portfolios. Panel A presents the whole sample period 1983-2015, Panel B and Panel C present the first (1983-1999) and second half (2000-2015) of the sample period.

	Average return	Standard deviation	Max	Min	Skewness	Kurtosis
Panel A: 1983:01 - 2015:12						
OSEAX	13.709	21.475	17.448	-27.435	-0.866	5.328
OSEFX	13.492	22.691	17.446	-27.433	-0.985	5.749
EW (ALL)	15.051	21.565	17.394	-25.214	-0.723	4.774
EW (DEAD)	13.349	21.684	21.632	-24.923	-0.728	4.908
EW (ALIVE)	16.078	21.707	19.900	-25.311	-0.670	4.741
Panel B: 1983:01 - 1999:12						
OSEAX	16.903	22.527	17.448	-27.435	-0.869	5.307
OSEFX	17.062	22.837	17.446	-27.433	-0.927	5.540
EW (ALL)	20.029	21.695	17.394	-22.842	-0.579	4.449
EW (DEAD)	18.537	21.653	21.632	-22.659	-0.585	4.714
EW (ALIVE)	21.172	22.067	19.900	-23.060	-0.496	4.372
Panel C: 2000:01 - 2015:12						
OSEAX	10.284	20.385	15.047	-23.934	-0.888	5.254
OSEFX	9.941	22.584	16.521	-27.166	-1.066	5.986
EW (ALL)	9.761	21.375	15.486	-25.214	-0.901	5.090
EW (DEAD)	7.807	21.660	15.065	-24.923	-0.891	5.065
EW (ALIVE)	10.665	21.263	15.626	-25.311	-0.904	5.116

4.4 Risk Factors

Ødegaard (2017) computed the risk factors for the Carhart (1997) four-factor model, based on empirical data for the Oslo stock exchange. He computed them using data from the time interval 1980-2016. The SMB, HML and PR1YR factors used in our computations are collected from this database. We explained what these factors capture and how they are computed in Section 3.1 on model selection.

Figure 2 plots time-series of the cumulative return on all risk factors MKT, SMB, HML and PR1YR over the whole sample period. The figure shows that all factors have positive accumulated return. The SMB and The PR1YR have the highest accumulated return, but quite different development. While SMB has had a stable upward trend over the whole sample period, the PR1YR is relatively flat from 1983-2013 for so to increase the last couple of years drastically.

Figure 2: Cumulative returns on factors

The figure plots the cumulative return on the factors in the Carhart's four-factor model from 1983 to 2015.

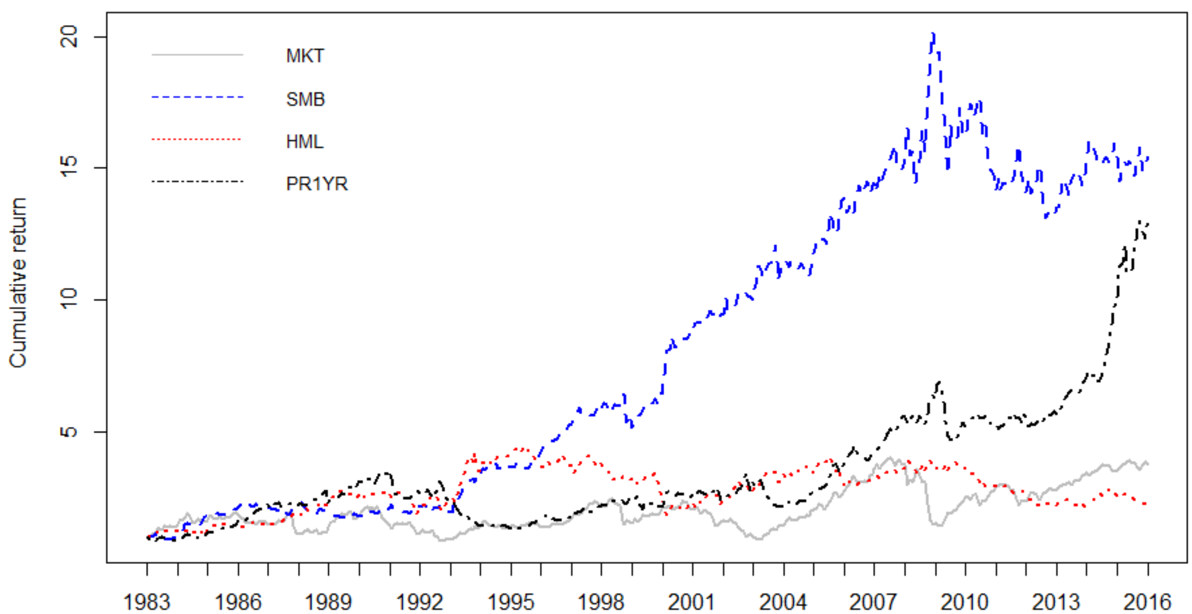


Table 3 below presents the descriptive statistics of these risk factors over the whole sample period, the first half and the second half of the sample period. Panel A and B show the average returns and standard deviations for each factor, over the three different periods. Additionally, we have included a correlation matrix in Panel C. As Panel A of Table 3 shows, the SMB and

PR1YR factor have the highest average return in over the whole sample period. The SMB factor generated especially high returns in the first half of sample period, 1983-1999, with an average return of 12.37% compared to the market factor with only 7.26%. While the PR1YR factor generates highest returns in the second half of the period, 2000-2015, with an average return of 10.05% compared to the market factor with only 6.18%. Panel B shows that the market factor has the highest volatility, measured by the standard deviation, and the SMB has the lowest volatility. The correlation matrix in Panel C shows the correlations between the factors. The correlations between factors are relatively small, mostly below |0.2|. On the other hand, SMB and the market factor has a correlation of -0.425.

Table 3
Descriptive Statistics of Factor Returns

This table presents the descriptive statistics of the risk factors of the whole, first half and second half of the sample period. Panel A and B shows the average return and the standard deviation of the risk factors, respectively. Panel C shows the correlation matrix between the factors. Average return and standard deviation are annualized and reported in percent.

	MKT	SMB	HML	PR1YR
Panel A: Average returns				
1983:01 - 2015:12	6.733	9.462	3.819	9.279
1983:01 - 1999:12	7.257	12.374	6.739	6.234
2000:01 - 2015:12	6.179	6.369	0.718	10.055
Panel B: Standard deviations				
1983:01 - 2015:12	22.782	15.222	16.998	17.109
1983:01 - 1999:12	22.904	16.615	18.416	17.819
2000:01 - 2015:12	22.711	13.573	15.346	16.537
Panel C: Correlation matrix				
MKT	1			
MKT.A	0.985			
SMB	-0.425	1		
HML	0.071	-0.132	1	
PR1YR	-0.152	0.14	-0.062	1

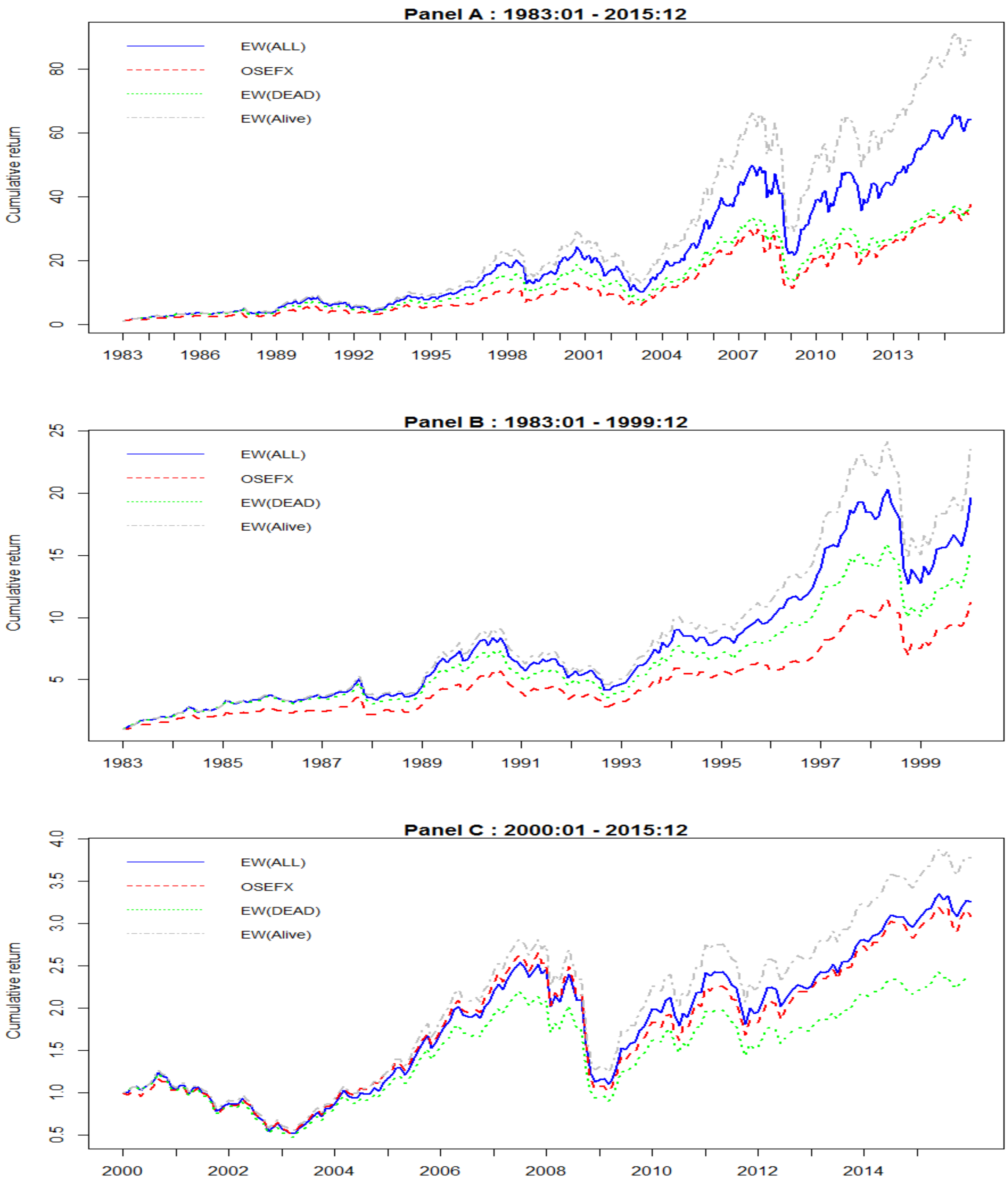
4.5 Potential Biases in Mutual Fund Returns

In this Section, we will discuss potential biases in mutual fund returns. One type of bias is the survivorship bias. This bias occurs if non-surviving funds are omitted from the sample. In other words, our sample is incomplete if we exclude the non-surviving and only include the surviving funds. If the sample is incomplete, then the results from that sample will also be incomplete. The incompleteness as a consequence of omitting the non-surviving will bias the results one way or another. Omitting losers bias our results and make the results for the industry look better than they in reality are. This is because non-surviving fund tends to perform worse than surviving funds. Elton, Gruber, and Blake (1996) claimed that mutual funds tend to disappear for two reasons, either their performance is too poor, or the management thinks that it no longer pays to maintain the fund. The last reason can also be attributed to the first: poor performance. Brown and Goetzmann (1995) found that the strongest predictors of fund attrition are poor track record, size, age and the fund's expense ratio. The strongest predictor among them for attrition is a poor track record. Carhart, Carpenter, Lynch, Musto (2000) concluded that fund disappear primarily because of poor multi-year performance, not poor one-year performance.

If we include only the surviving funds in our sample, we overstate the performance results and skew the results towards the positive. The non-surviving funds do not perform well enough for its continued existence to be justified. If we throw out bad performers and only keep the good ones, our sample is biased. The effect of the survivorship bias can be seen clearly when looking at the cumulative results of different portfolios.

Figure 3 show the cumulative returns of an equal weighted portfolio of all the funds, OSEFX, an equally weighted portfolio of the non-surviving funds and the surviving funds. Panel A is for the entire sample period, Panel B and C are subperiods. In panel A we can see that the portfolio of surviving funds has the highest cumulative return. The equal weighted portfolio of dead funds has a much lower cumulative return. The equal weighted returns of all the funds are in the middle of the dead and alive funds. It becomes clear how omitting losers bias the cumulative results positively, overstating the returns the fund earn for their investors.

Figure 3: Cumulative returns on equally weighted portfolios and the OSEFX
This figure plots the cumulative returns of the equally weighted portfolio for all, dead and alive mutual funds and the Oslo Børs Mutual Fund Index. Panel A depicts the whole sample period from 1983 to 2015. Panels B shows to the first half of the sample, 1983-1999. Panel C shows to the second half of the sample, 2000-2015



Another bias that may influence the results is the look-ahead bias. This bias might occur if one requires a fund to survive a minimum period. The choice of excluding rule is tradeoff between look-ahead bias and precision. We exclude fund that have less than 12 month of returns. This lead to no exclusion in our sample, since the shortest lived fund in our sample has 17 months of returns. Our exclusion rule only comes to play when evaluating the funds individual performance in subperiods. We do not exclude any fund to make sure that the data set free of survivorship bias, but the disadvantage of having a fund with only 17 observation may cause the regression estimate might be imprecise.

5 Empirical Results

5.1 The Performance of Aggregate Norwegian Mutual Fund

Table 4 below presents estimates for the equal weighted portfolios of all mutual fund returns in our sample based on the three different performance models. The intercept (alpha) in Table 4 indicates whether the mutual fund sector on average has produced abnormal risk-adjusted returns. Panel A shows that the alpha over the entire sample period for the equal-weighted portfolio based on the CAPM is 1.69 % yearly, but it is decreasing as we progress from CAPM to the Fama-French three-factor model (0.16%) and the Carhart four-factor model (-0.03%). However, the alphas are not statistically significant. Meaning the Norwegian mutual fund sector on average has not delivered statistically significant alpha given any of models. The drop from 1.69 % to 0.16% in the abnormal return of the average mutual fund can be explained by the significant and positive exposure to the SMB factor. This positive exposure indicates that the average mutual fund manager prefers small capitalization stocks. The average mutual fund has a negative exposure to the HML factor and positive exposure to the PR1YR factor. However, these coefficients are insignificant.

Panel B and Panel C show the performance of the Norwegian mutual fund sector on average in the first and second half of the sample period. The average fund produces a four-factor alpha of 0.509 % yearly in the first half and a four-factor alpha of -0.63 % in the second half. These estimated alphas are, however, not statistically significant. Comparing the performance of the average fund for the entire sample period with the performance in the first and second half, we find that the average fund performed best in the first half. This is consistent with our earlier findings.[15] Panel B and C shows that given the positive and significant exposure to the SMB factor, the average mutual fund heavily favored small capitalization stocks in the first and second half of sample period. There is some difference in the exposure to the HML factor in the two periods. In the second half of the period, the HML factor is significant at the 1% level. However, the coefficient of -0.044 is of no economic significance.

Table 4
Aggregate Fund Performance for Different Sample Periods

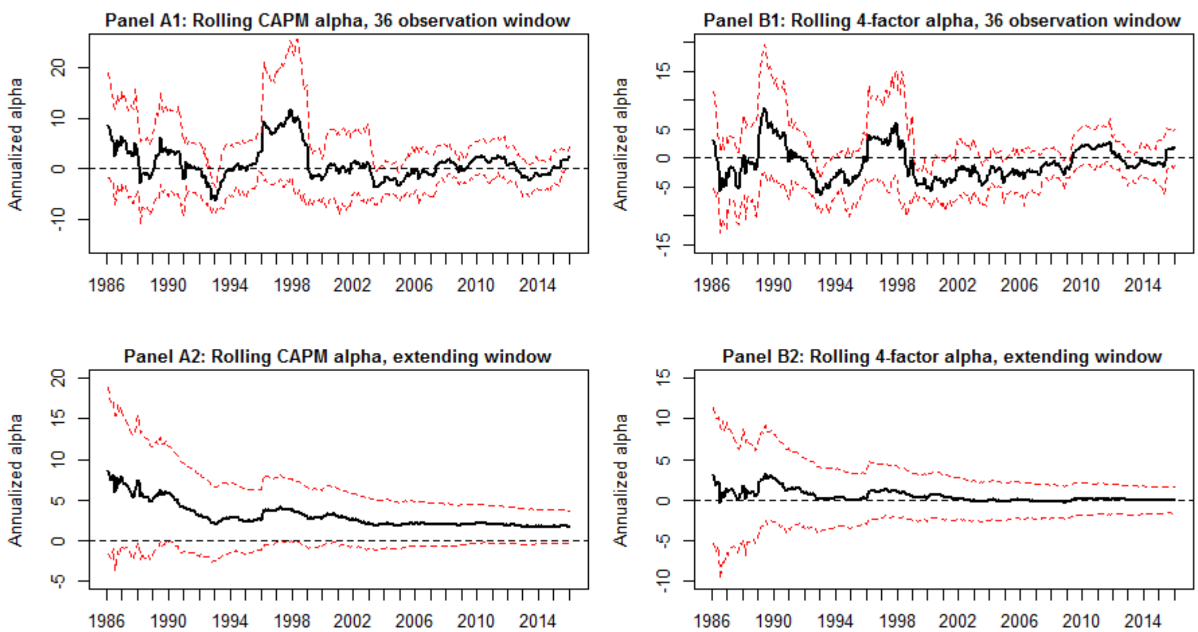
The table shows alphas, factor loadings and the adjusted R^2 for an equal weighted portfolio of mutual fund returns compared to CAPM, the Fama-French three-factor model, and the Carhart four-factor model. Numbers assigned stars, ***,**, and * indicates significance at the 1%, 5%, and 10% levels. The numbers in parentheses below the estimated coefficients are the t-statistics. The t-statistics is computed by using standard errors corrected for autocorrelation and heteroscedasticity following the procedure by Newey and West (1987) . Panel A reports the result for the whole sample period from 1983 to 2015. Panel B reports the result for the first half of the sample. Panel C reports the result for the second half of the sample.

Model	α	β_{MKT}	β_{SMB}	β_{HML}	β_{PR1YR}	R^2_{adj}
Panel A: 1983:01 - 2015:12						
CAPM	1.69	0.91^{***}				0.93
	(1.67)	(50.58)				
Fama-French	0.16	0.95^{***}	0.14^{***}	-0.03		0.94
	(0.19)	(56.34)	(7.06)	(-1.63)		
Carhart	-0.03	0.95^{***}	0.14^{***}	-0.03	0.02	0.94
	(-0.04)	(55.14)	(7.11)	(-1.57)	(1.1)	
Panel B: 1983:01 - 1999:12						
CAPM	2.902	0.888^{***}				0.893
	(1.56)	(32.49)				
Fama-French	0.786	0.926^{***}	0.149^{***}	-0.002		0.903
	(0.52)	(34.33)	(5.37)	(-0.08)		
Carhart	0.509	0.926^{***}	0.144^{***}	0.001	0.046	0.904
	(0.36)	(32.94)	(5.48)	(0.03)	(1.63)	
Panel C: 2000:01 - 2015:12						
CAPM	0.436	0.937^{***}				0.974
	(0.5)	(46.82)				
Fama-French	-0.662	0.976^{***}	0.139^{***}	-0.044^{***}		0.981
	(-0.85)	(52.71)	(7.54)	(-3.45)		
Carhart	-0.63	0.976^{***}	0.139^{***}	-0.044^{***}	-0.002	0.98
	(-0.809)	(58.064)	(7.816)	(-3.577)	(-0.138)	

[15] In Section 4.3, Table 2 reports that the equal weighted portfolio of all Norwegian mutual funds performed best in the first half of the sample period in terms of raw returns (i.e. before adjusting for risk).

Figure 4: Rolling and extending window estimations of the equally weighted portfolio

This figure plots estimated alphas of the equal weighted portfolio. The upper panels plot the rolling window estimates of alpha, where the window length is set to 36 months. The lower panels plot the extending window alpha estimates, where the window extends from 36 to 396 months. The left panels report CAPM estimates of alpha. The right panels report 4-factor estimates of alpha. The solid line shows the alpha estimates, whereas the dotted line is the Newey-West-corrected two standard errors bands. The sample period is 1983 to 2015. The alpha estimates are annualized and in percent.



To examine the development of the average Norwegian mutual fund, we estimate the alpha from rolling window and extending window regressions. The results are presented in Figure 4. The left panels compute alpha against the single-factor CAPM by Jensen (1968), whereas the right panels compute alpha against the Carhart's (1997) four-factor model. The upper panels report the rolling window estimates. From Panel A1, the CAPM estimates of alpha at the beginning of the sample period is positive with occasional drops below zero. From the middle of the sample period, the CAPM alpha stabilizes around zero. The rolling window estimates of 4-factor alpha, Panel B1, is negative throughout the sample period, with occasional spikes above the zero mark. The lower panels present the extending window alpha estimates. Panel A2 shows that the CAPM alpha is 8.6 % yearly at the start and it declines almost monotonic with time. The CAPM alpha is positive through the sample period, but insignificantly different from zero. In Panel B2, the 4-factor alpha is always lower than the CAPM alpha. It is also more stable than the CAPM alpha and closer to the zero mark. Same as the CAPM, the 4-factor alpha is never significantly different from zero.

The difference in the performance between the CAPM and the 4-factor model can be explained by the fund portfolio's exposure to the SMB, HML and PR1YR risk factors. The market risk explains most of the variation in the returns, but it does not account for the potential exposure to the other factors. For instance, Table 4 shows that the aggregate fund portfolio has a relatively high exposure to the SMB factor in the first half of the sample period. From section 4.4, Table 3 showed that the SMB performed particularly well during this period and may explain the difference between CAPM and 4-factor in the first half of the period. The average Norwegian mutual fund relatively high exposure to SMB factor throughout the entire sample period, which the 4-factor model capture and the CAPM does not. This indicates that the CAPM is a poor choice when evaluating the abnormal performance of Norwegian mutual funds since it fails to account for the exposure to the size, value and momentum factors. Based on the presented results, our conclusion is that actively managed Norwegian mutual funds, on average, do not possess sufficient skill to create statistically significant abnormal returns to their investors net of costs. Even if mutual fund managers on average do not have sufficient skill to produce significant alphas net of cost, one cannot rule out the possibility that there are some managers with superior or inferior skills. In the next subsection we will implement the bootstrap procedure of Kosowski et. al. examine the mutual fund's individual performance with the purpose of identifying superior and inferior fund managers.

5.2 Individual Funds - Distinguishing Skill from Luck

The results for the aggregate mutual fund's performance indicates that mutual fund managers as a group are not able to generate statistically significant positive alpha net of costs. However, this does not rule out the possibility that there are individual managers with skill. Perhaps this result occurs because the poor-skilled mutual funds are nullifying the effect of the skilled mutual funds. In any large sample, there will be funds that perform well and funds that perform poorly. The challenge is to separate the performances due to skill from the performances generated by chance. In this section, we apply a similar bootstrap approach to Kosowski et. al. (2006) with the purpose of separating skill from luck in individual mutual funds performance.

5.2.1 Bootstrap evidence

We apply the bootstrap procedure, described in Section 3.2, to build the distribution of these cross-sectional draws of alpha and their t-statistic. These drawn alphas (t-statistic) is a result of sampling variation. By examining these distributions, we can separate skill from luck. For example, consider the top (bottom) funds. If the bootstrap iterations generate far fewer extreme positive (negative) values of alpha (or t-statistic) compared to those observed in the actual data, we can conclude that the actual alpha is not a product of luck, but rather stock picking skills (or poor skill).

The result of the bootstrap procedure is presented in Table 5. The columns list the three bottom funds, bottom 5%, percentiles 10 to 90, top 5%, and the three top funds. Panel A shows the results when funds are ranked on their alpha. Panel B shows the results when funds are ranked on their t-statistic of alpha. Kosowski et. al. (2006), Busse et. al. (2010) and Fama and French (2010) argued for ranking funds on their t-statistics of alpha rather than the actual alpha estimates. By rescaling the alpha estimate by its standard deviation, we get estimates which are adjusted for the difference in the precision of alpha. The t-statistics is the precision-adjusted estimates of alpha. For completeness, we report the results of both rankings. Panel A shows that when ranked on the estimated alpha. The bottom funds' bootstrapped p-value of 0.02 is the probability that this fund generates an estimated alpha less than -22.83% per year, purely by chance. From the bootstrapped p-value of the bottom fund, we reject the null hypothesis that performance is due to luck at the 5% significance level. Which implies, that the managers of that fund have bad skills. The bootstrapped p-values of the bottom two, three, bottom 5% and the 10% decile fund suggest that estimated alpha of -17.07%, -14.52%, -10.49% and -7.22% per year is not generated purely by chance, suggesting bad skills among these funds' managers. The top fund's bootstrapped p-value of 0.16 is the probability that the top fund generates an estimated alpha of at least 12.91% per year, by chance. This p-value is not significant to reject the null hypothesis of the alpha being generated by luck. The bootstrapped p-values of the top two, three, top 5% and the 90% decile fund suggest that the estimated alpha of 6.33%, 5.25%, 3.45% and 1.72% per year is generated by chance, suggesting that the performances of these funds are a product of luck.

Table 5: Luck vs Skill

This table reports the results for the cross-section of the Norwegian mutual fund's performance measure for the whole sample period 1983-2015. Panel A shows the Norwegian mutual fund ranked on unconditional Four-Factor Model Alphas. The first and second row reports the OLS estimate of alphas and the cross-sectionally bootstrapped p-values of the alpha. The third row reports Parametric p-values of the t-statistic, which is based on standard critical values of t-statistic. Panel B shows the Norwegian mutual fund ranked on t-statistics unconditional Four-Factor Model Alphas. The first and second row reports the t-statistic of alphas and the cross-sectionally bootstrapped p-values of the t-statistic of alpha. The third row reports Parametric p-values of the t-statistic, which is based on standard critical values of t-statistic. The OLS estimated alphas are reported in percent and are annualized. The t-statistics of alpha are based on heteroskedasticity- and autocorrelation-consistent standard errors. The cross-sectionally bootstrapped p-value is based on the distribution of the best (worst) funds in 10,000 bootstrap resamples.

	Bottom	2nd	3rd	Bottom 5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	Top 5%	3rd.	2nd.	Top
Panel A: Fund Ranked on Four-Factor Model Alphas																	
Alpha	-22.83	-17.07	-14.52	-10.49	-7.22	-2.47	-1.90	-1.26	-0.57	0.18	0.57	0.99	1.72	3.45	5.25	6.33	12.91
Bootstrapped p-value	**0.02**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**	**0.00**	0.99	0.86	0.92	0.96	0.98	0.83	0.55	0.54	0.16
Parametric p-value	**0.01**	**0.00**	**0.01**	**0.00**	0.15	0.06	0.34	0.27	0.40	0.46	0.33	0.22	0.08	**0.02**	**0.03**	**0.01**	**0.00**
Panel B: Fund Ranked on t-Statistic Four-Factor Model Alphas																	
t-alpha	-4.05	-2.98	-2.92	-2.87	-2.07	-1.41	-1.12	-0.70	-0.30	0.10	0.28	0.61	1.16	1.99	2.35	2.50	3.35
Bootstrapped p-value	0.11	0.15	**0.04**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.98	0.88	0.97	0.95	0.81	0.13	0.25	0.38	0.27
Parametric p-value	**0.00**	**0.00**	**0.01**	**0.00**	**0.02**	0.08	0.13	0.24	0.38	0.46	0.39	0.27	0.12	**0.02**	**0.01**	**0.01**	**0.00**

Ranking the funds on their t-statistic rather than the alpha estimate is more favorable for the funds. Ranked on their t-statistic of alpha, Panel B reports that the evidence of bad skill among the worst performing funds is not as strong as in Panel A. The bootstrapped p-values for the bottom one and bottom two funds have increased from 0.02 and less than 0.01 to 0.11 and 0.15. We cannot reject the null hypothesis of performances being generated by luck. The top fund bootstrapped p-values has increased compared to Panel A, but for the top two, three and top 5% funds the p-values have decreased from 0.54, 0.55 and 0.83 to 0.38, 0.25 and 0.13. The changes in the bootstrapped p-values for these funds is not enough to change the conclusion. The funds in the right tail of the distribution continue to perform insignificantly different from what we would expect if performances were generated by luck.

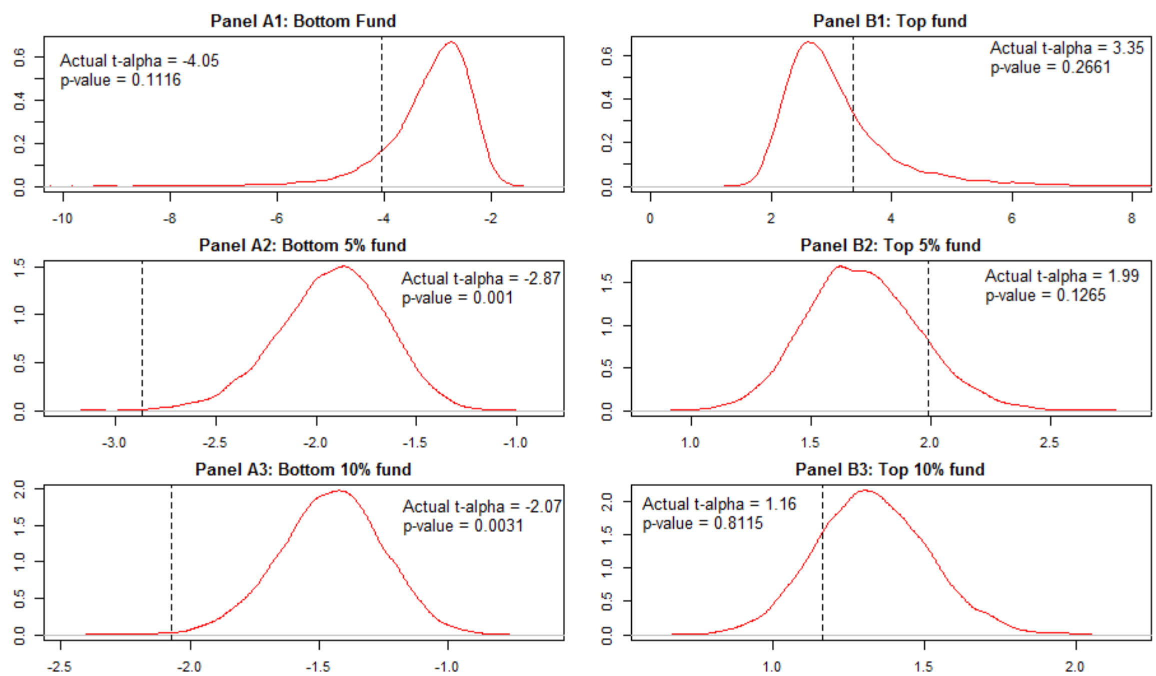
Note, that our inference about the skill of mutual fund managers is somewhat different with the bootstrap than a parametric approach. Comparing the bootstrap p-values with the parametric p-values, the importance of implementing the bootstrap procedure becomes clear. In Panel A and B, the cross-sectional bootstrapped p-value for the funds in the right tail is larger than their parametric p-values. Meaning, the bootstrap uncovers more probability mass in the right tail of the cross-section than expected under a parametric normal assumption. Looking at the funds in the left tail, ranked on t-statistic of alpha, their cross-sectional bootstrapped p-values are larger than their parametric p-values. The bootstrap uncovers more probability mass in the left tail of the cross-section than expected under the parametric normal assumption. More probability mass meaning fatter tails, which leads to over-rejection of the null hypothesis (without the bootstrap).

To explore further, Figure 5 shows the bootstrapped distributions of the t-statistic of alpha for funds at different percentile points in the cross-section using the Carhart's (1997) four-factor model. For example, Panel A1 shows that the distribution of the bottom fund is heavily left skewed and includes alpha t-statistics that vary from about -2 to less than -8 in the extreme cases. Using a 5% lower tail cut off point, the bootstrap does not reject the null hypothesis that the bottom fund (estimated) alpha t-statistic of -4.05 may be explained by luck, whereas the parametric t-tests do. The same difference between bootstrap and parametric inference is found in Panel B1. Panel B1 shows that the distribution of the top fund is heavily right-skewed and vary from about 2 to less than 8 in the extreme cases. Using a 5% upper tail cut off point, the bootstrap does not reject the null hypothesis that the top fund (estimated) alpha t-statistic of 3.35

may be explained by luck, whereas the parametric t-tests do.[16] These contradictions are due to the highly non-normal distribution of idiosyncratic risk across our bottom performing funds as well as our top performing funds. Both these cases demonstrate that standard parametric test statistics may provide misleading inference when evaluating funds in the extreme tails, and underline the importance of bootstrap for proper inference.

Figure 5: Estimated t-statistics of alpha vs bootstrapped t-statistics of alpha distribution for individual funds

The figure shows the bootstrapped distribution of the four-factor t-statistics of alpha (solid line) at different percentile points of cross-section. The dotted vertical line shows the actual (estimated) t-statistic of alpha relative to the bootstrapped distribution. Panels A1-A3 show funds in the left tail of the distribution and Panel B1-B3 show funds in the right tail of the distribution.



[16]As we move from very extreme tails and toward the center of the of the cross-sectional distribution (i.e., from Panels A1 to A3 and Panel B1 to B3), the alpha t-statistic distributions become more symmetric but continue to remain non-normal. Panel B2 and Panel B3 are cases where the rejection is so strong that both the bootstrap and the parametric t-test rejects the null.

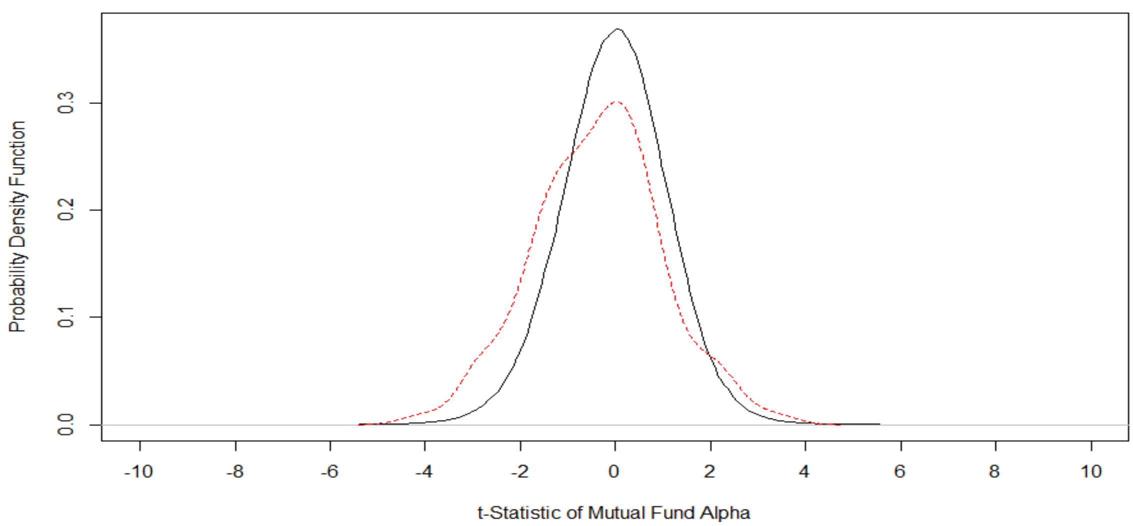
Figure 6 compares the cross-sectional distribution of the actual t-statistic of alpha with the bootstrapped distribution. The densities of these two are quite different in shape. Panel A, shows the probability distribution of the actual and the bootstrapped t-statistic. The distribution of the actual t-statistic has more probability mass in the tails, but less in the center than the bootstrap distribution. The differences in the tails are larger on the left side than the right side. We will also point out that the distribution of the actual t-statistic of alpha has several complex features such as "shoulders" in the tails. This further reinforces the importance of using the bootstrap for inference since it can capture the complex shape and measure the fat and thin tails of the distribution. Panel B shows that there is far more probability mass in the left tail of the actual distribution than the bootstrap distribution. This is an indication of overpopulation in the left tail. Panel B also shows that in the right tail the actual distribution occasionally lies below the bootstrap distribution, and it is much closer to the bootstrap distribution. In the extreme right tail, the two distributions are overlapping, indicating that the number of funds in the right tail of the actual distribution is due to luck.

To summarize, performance evaluation based on standard parametric inference suggests that there are both superior and inferior Norwegian mutual fund managers. Given that the distribution of alphas and t-statistic of alphas exhibit complex and non-normal properties, our bootstrap evidence contradicts the results based on the standard parametric approach. Our bootstrap results show no evidence of skilled Norwegian mutual fund managers. For all funds in the right tail of the distribution, we expect the t-statistic of alpha to be greater than the actual t-statistic of alpha too many times to reject the null hypothesis that the funds' performances can be explained by luck. Norwegian mutual fund managers in the right tail show no performance that cannot be explained by luck. We do however find evidence of inferior fund managers in the left tail that exhibit statistically significant negative t-statistic of alpha that cannot be explained by luck.

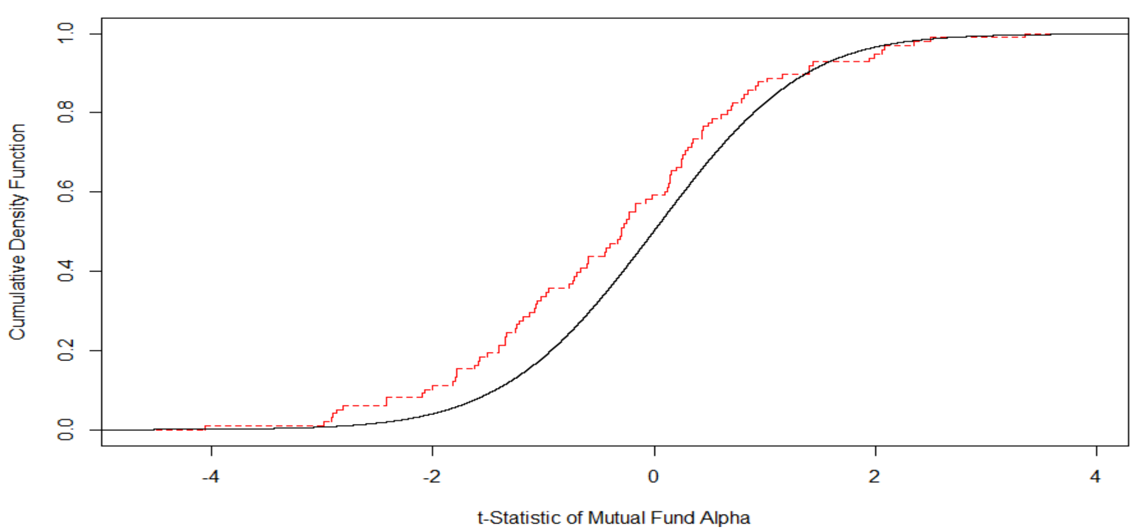
Figure 6: Estimated vs. bootstrap cross section of alpha t-statistics.

This figure plots the kernel density estimates of actual (dashed line) and bootstrapped (solid line) cross-sectional distribution of the t-statistic of mutual fund alphas. Panel A displays the kernel density estimate of the probability density function (PDF) of the distributions. Panel B shows the kernel density estimate of the cumulative density function (CDF) of the distributions. The t-statistic of alpha are computed using the four-factor model applied to all Norwegian mutual funds during the 1983 to 2015 period.

Panel A : PDF of actual (Dashed Line) and bootstrapped (Solid Line) Cross-sectional Distribution of t-statistic of Alphas



Panel B : CDF of actual (Dashed Line) and bootstrapped (Solid Line) Cross-sectional Distribution of t-statistic of Alphas



5.2.2 The Economic Impact

Following Kosowski et. al. (2006), we use the bootstrap distribution of alphas to calculate how many funds that are expected, by chance, to exceed a given level of performance. We then compare this number to the number of funds that we actually observe exceed this level of performance in our sample. Figure 7 Panel A plots the number of funds from the actual and from the bootstrapped distribution that performs above certain performance levels, while Panel B plots the number of funds from the actual and from the bootstrapped distribution that performs below certain performance levels.[17] For example, Panel A indicates that we expect five funds to have an alpha greater than 4% per year purely by chance, whereas we observe four funds achieve this. While in Panel B, we expect five funds to have an alpha less than -4% due to chance, whereas the observed number of funds are 14.

The result of this comparison, further strengthen the evidence of statistically insignificant abnormal performances in the right tail and statistically significant negative abnormal performances in the left tail of the distribution. For example, Panel A of Figure 7 indicates that among the fund managers with performances above 4% per year, none exhibit stock-picking talent sufficient to exceed their costs, it's simply due to luck. Panel B, on the other hand, indicates that among the fund managers performances below -4% per year, nine of 14 fund managers lack the stock-picking talent to cover their costs, while five fund managers are simply unlucky.

Further, we evaluate the overall potential economic impact of our findings by approximate the value added (or destroyed) by skilled (unskilled) managers. We estimate the skilled-based value added by managers as the difference between the value added by all funds (lucky and skilled funds) with a certain level of performance and the value added by funds that achieve the same performance level due to luck.[18]

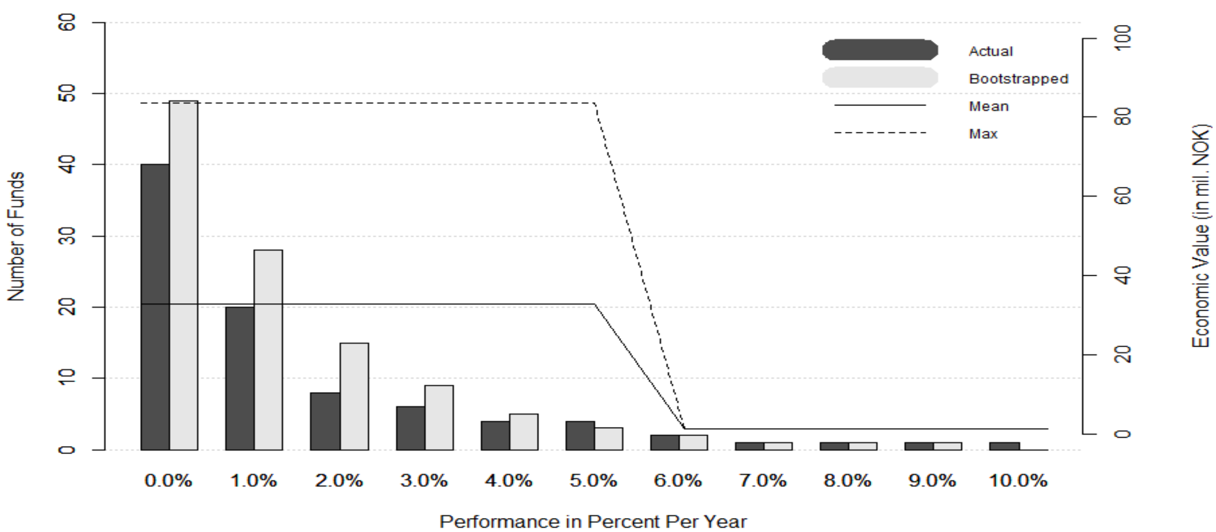
[17] The bootstrapped alphas are the average of each individual fund's luck distribution

[18] The skilled-based value added for a given 1% alpha interval is estimated as the average size of all fund in that interval multiplied by the difference between the total number of funds and number of lucky (unlucky) funds in the interval, multiplied by that level of alpha. As an example, We have 14 funds with estimated alphas less than -4% per year, while only five are expected to have such alphas by chance. The value destroyed by the additional nine fund are estimated by the average fund size of all 14 funds multiplied by -4%, multiplied by 9. The skilled-based value added of an alpha interval is cumulated for all alpha intervals above (in absolute term) and including a given interval. The solid lines in the figure present these estimated of skilled based value added. We have also included an upper bound which is the same computation only using the average size of the largest fund in each interval. The dotted lines present the maximum skilled-based value added. The average size of each fund is calculated as the average of its historical asset under management (AUM). The historical data on AUM provided by The Norwegian Fund and Asset Management Association (VFF) and covers the period January 1994 to December 2015.

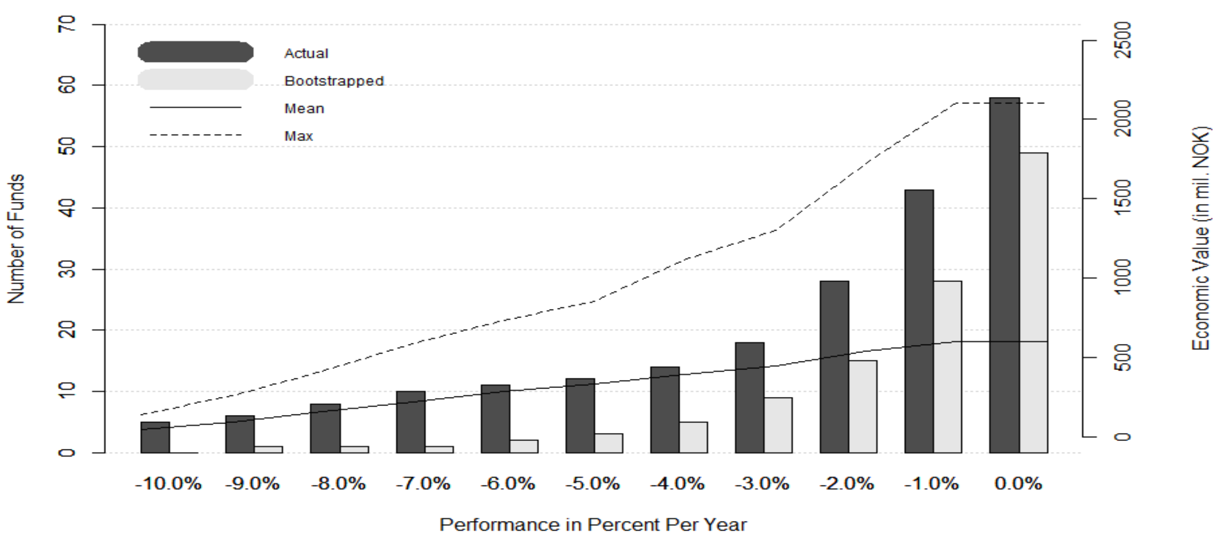
Figure 7: Cumulative economic value-added by funds above and below various alpha levels.

Panel A of the figure reports (as vertical bars) the number of funds from the actual and the bootstrapped cross-sectional distributions with performance above various four-factor alpha levels. Panel B the number of funds from the actual and the bootstrapped cross-sectional distributions with performance below various four-factor alpha levels. In Panel A, the solid (dashed) line show the mean (max) the cumulative economic value that an investor could potentially gain by investing in the number of funds exceeding the bootstrapped number of funds in all higher performance brackets. The solid line is based on the average total asset under management (AUM) in each performance bracket, whereas the dashed line based on the average of a subgroup of the largest funds' AUM. In Panel B, the solid (dashed) line show the mean and the max cumulative economic value that is potentially lost by the statistically significant underperformance of some funds.

Panel A : Cumulative Economic Value and Number of Funds from the Actual and Bootstrap Sample with Performance Above a Certain Value



Panel B : Cumulative Economic Value and Number of Funds from the Actual and Bootstrap Sample with Performance Below a Certain Value



Our findings indicate no evidence of skilled managers, we therefore in Panel A of Figure 7 plots the cumulative economic value a hypothetical investor could potentially gain by investing in the number of funds exceeding the bootstrapped number of funds in all higher performance brackets. However, Panel B shows the cumulative value destroyed below a certain performance level by fund managers poor skill. As Panel B of Figure 7 depicts, we estimate that approximately 600 million NOK per year is destroyed by truly underperforming funds inability to compensate for fees and trading costs.

We should point out, that many potential issues could complicate the results of our baseline bootstrap procedure and its interpretation. For instance, funds may have cross-sectionally correlated residuals. If so, our bootstrap could be biased as we have assumed residuals to be independent. We address these and other concerns in the next section.

5.3 Sensitivity Analysis

In this section, we conduct various tests to determine whether our bootstrap results in the last section are sensitive to changes in the bootstrap procedure.

A. Time-Series Dependence

Our bootstrap results assume that the return generating process and the error-term in the model are independent. As a robustness check, we allow for dependence and adopt the stationary bootstrap method suggested by Politis and Romano (1994), by resample the return residuals in data blocks, the dependence in the return residuals is kept.

The important question when implementing the stationary bootstrap method is how to determine the optimal block length. The paper by Hall, Horowitz, and Jing (1995) found that the optimal block length is determined by the asymptotic formula $l \sim T^{1/h}$, where $T =$ number of observation and $h = 3,4$ or 5. The authors found that the value of h depends on the context. For computing block bootstrap estimators of variance or bias, $h = 3$. For computing block bootstrap estimators of one-sided and two-sided distribution functions of the test statistic of interest, $h = 4$ and 5 respectively. Since our hypothesis tests are one-tail tests, $h = 4$. Another method of determining the optimal block length is the method proposed by Politis and White (2004) (see also the subsequent correction of the method by Patton, Politis, and White, 2009).

We run the stationary bootstrap for both optimal block length methods. We compare the results of these stationary bootstraps with our independent bootstrap results. In Appendix E, Table E.I shows that the stationary bootstrap method leads to almost identical bootstrapped p-values as the results from our baseline bootstrap procedure.

B. Residual and Factor Resampling

Next, we test if our findings could be affected by autocorrelation (persistence) in the factor returns. By implementing a bootstrap which resamples the regression residuals and the factor returns independently, we can investigate whether randomizing factor returns changes our results due to breaking the potential autocorrelation in factor returns.

We use the same draw across all funds when resampling factor returns. By doing so, the correlation between factor returns and all funds is preserved. The independently resampled factors and residuals for each bootstrap iteration b and fund i,

$$\left\{MKT_t^b, SMB_t^b, HML_t^b, PR1YR_t^b, t = \tau_{T_{i0}}^b, ..., \tau_{T_{i1}}^b\right\} and \left\{\hat{\varepsilon}_{i,t}^b, t = s_{T_{i0}}^b, ..., s_{T_{i1}}^b\right\} \tag{9}$$

is used to construct a time series of monthly returns for fund i, where we again impose the null hypothesis of zero true performance ($\hat{\alpha} = 0$, or $\hat{t}_{al\hat{p}ha} = 0$),

$$\{\tilde{r}_{i,t}^b = \beta_{1i} MKT_{t_F}^b + \beta_{2i} SMB_{t_F}^b + \beta_{3i} HML_{t_F}^b + \beta_{4i} PR1YR_{t_F}^b + \tilde{\varepsilon}_{i,t_\varepsilon}^b\}, \tag{10}$$

where $t_F = \tau_{T_{i0}}^b, ..., \tau_{T_{i1}}^b$ and $t_\varepsilon = s_{T_{i0}}^b, ..., s_{T_{i1}}^b$.

Next, we estimate, fund by fund, the four-factor model on the simulated draw of the monthly constructed returns, saving the alpha and t-statistic of alpha for each simulation b. Again, we find almost identical results (see AppendixE, Table E.II) as our baseline bootstrap procedure.

C. Cross-Sectional Bootstrap

As stated earlier, our baseline bootstrap procedure assumes no cross-correlation between fund residuals. In empirical tests, we find that the average residual correlation is 27.6% for the four-factor model. An explanation for why cross-correlation between fund residuals arises is that funds may buy, or otherwise own the same stocks at the same time. Addressing this issue may be especially important for the tails of the cross-section of alphas, as funds with very high (low) alphas may hold the same (or similar) stocks. We have therefore implemented an extension of the baseline bootstrap which allows for cross-correlation in residuals, with the goal of determining whether the cross-correlation between fund residuals bias our results. By drawing residuals, across all funds, during identical time periods, we are able to capture the potential effect of the cross-correlation in fund residuals.

The implementation of the cross-sectional bootstrap is almost the same as the baseline bootstrap. But rather than drawing a unique sequence of time periods, t_i for each of individual fund i, we draw a $(Tx1)$ vector of time periods from the set $t = 1,...,T$, creating a reindexed time sequence, $\tilde{T}_b$, which is used to resample residuals across all funds. This resampling method may result in some funds being allocated bootstrap entries from periods when the funds did not exist. Therefore a minimum number of observations are required for a fund to be included. We drop a fund if it does not have at least 12 observation during the reindexed time sequence. The results of the cross-sectional bootstrap are presented in Panel B of Table E.III in Appendix E. It shows that these results are almost identical to our baseline bootstrap procedure.

D. Joint Resampling of Fund return and Factor Returns

Our baseline bootstrap generates independent simulations for each fund, the drawback of this method is that our baseline does not account for correlation of estimated alphas for different funds that arise because the chosen benchmark model does not capture all the common variation in fund returns. The baseline bootstrap also misses any potential effects of correlated movement in the volatility of the factor returns and residuals. Fama and French (2010) states that failure to account for the joint distribution of fund returns, and of fund and factor returns, biases the results. To account for this, Fama and French (2010) argues for resampling the sample fund and factor returns jointly, instead of the resampling residuals-only method by Kosowski et. al. (2006). We have therefore implemented their joint resampling of fund returns and factor returns to determine whether this changes our results.

Following the method of Fama and French (2010), we subtract a fund's estimated alpha from its returns to construct returns that have an alpha equal to zero. From our original time index of 396 months, we draw 10,000 random samples with replacement creating reindexed time sequences which are used to resample all fund returns and factor returns jointly. Each of the resampled time-series of fund *i*'s returns is regressed on the resampled factor returns. We require a fund to have at least 12 observation during a reindexed time sequence to be included. We find that this approach gives almost identical results for both left-tail and right-tail funds as to our baseline approach. The results of this approach are reported in Panel C of Table E.III in Appendix E.

E. Length of Data Records

This test tests whether our bootstrap results are sensitive to the choice of minimum observation requirement. As Kosowski et al. (2006) point out, short-lived funds tend to generate more extreme alphas than long-lived funds, due to higher dispersion, leading to heteroskedasticity in the cross-section of alpha estimates, This effect can be reduced by imposing a larger number of observation as a minimum requirement. In addition, using the t-statistic of alpha instead of the alpha estimate is more favorable as it is less sensitive to these variance outliers. As discussed in section 3.5, minimum observation requirements may impose survivorship bias in our results.

To explore whether our results are sensitive to the choice of minimum observation requirement, we run our bootstrap procedure with a minimum observation requirement of 24, 36 and 60 months. The test results presented in Table E.V in Appendix E show that the bootstrap p-value for funds changes somewhat, but do not change the conclusion. Table E.V shows as the minimum observation requirement increases the more extreme values (positive and negative) get eliminated.

5.4 Baseline Bootstrap Tests for sub-periods

We apply the bootstrap procedure to the first half and second half of the sample period to examine whether the cross-sectional distribution of mutual fund performances changes over the sample period. Table 6 below reports our findings. Panel A reports the results for the first half of the sample period, January 1983 to December 1999. Panel B reports the results for the second half of the sample period, January 2000 to December 2015.[19]

Panel A shows that in the first half of the sample period, the performances of mutual funds in both tails of the distribution are not statistically significant at the 5% significance level; we cannot reject the null hypothesis that the performances are due to luck. Panel B, which reports the results for the second half of the sample period, indicates that there is evidence of inferior mutual fund performances in the left tail. Here we can reject the null hypothesis at the 5% significance level for the 3rd worst fund, the bottom 5% and the bottom 10%. The performances of these funds are not due to luck, but bad skill. In the right tail of the distribution, the performances are not statistically significant; we cannot reject the null hypothesis that performances are due to luck.

Figure 8 plots the cross-sectional distribution of the actual t-statistics of alphas versus the cross-sectional bootstrap distribution for the two subperiods. Panel A1 and Panel A2 plots the probability density functions (PDF) and the cumulative density functions (CDF) of the distributions for the first subperiod. From Panel A1, we observe that the actual distribution of t-statistics has more probability mass in the left tail and somewhat less probability mass in right tail compared to the bootstrapped distribution. Panel A2 shows that bootstrap distribution overlaps the actual distribution in the left tail and almost overlaps the actual distribution in the right tail. The visual presentation of the probability distributions for the first subperiod, suggests that there is no difference between the cross-sectional of actual and bootstrap distribution of t-statistics of alpha. Panel B1 and Panel B2 plots the PDF and the CDF of the distributions for the second subperiod. From Panel B1, we observe that the actual distribution of t-statistics has more probability mass in the left tail and right tail than the bootstrapped distribution. Panel B2 shows that bootstrap distribution is below the actual distribution in the left tail, but overlaps the actual distribution in the right tail. The visual presentation of the probability distributions for the second subperiod, suggests that there is difference between the cross-sectional of actual and bootstrap distribution of t-statistics of alpha.

[19] We find that the average cross-sectional correlation between residuals is somewhat higher in the first half than the second first half of the sample. Unreported tests show that for the first half of the sample, the results of our baseline bootstrap procedure are almost identical in the left tail of distribution as the cross-sectional bootstrap approach, but differ in the right tail of the distribution. For the second half of the sample period, our baseline bootstrap procedure and the cross-sectional bootstrap approach are almost identical in the tails. Thus, Panel A report the results of the cross-sectional bootstrap and Panel B report the results of the baseline bootstrap.

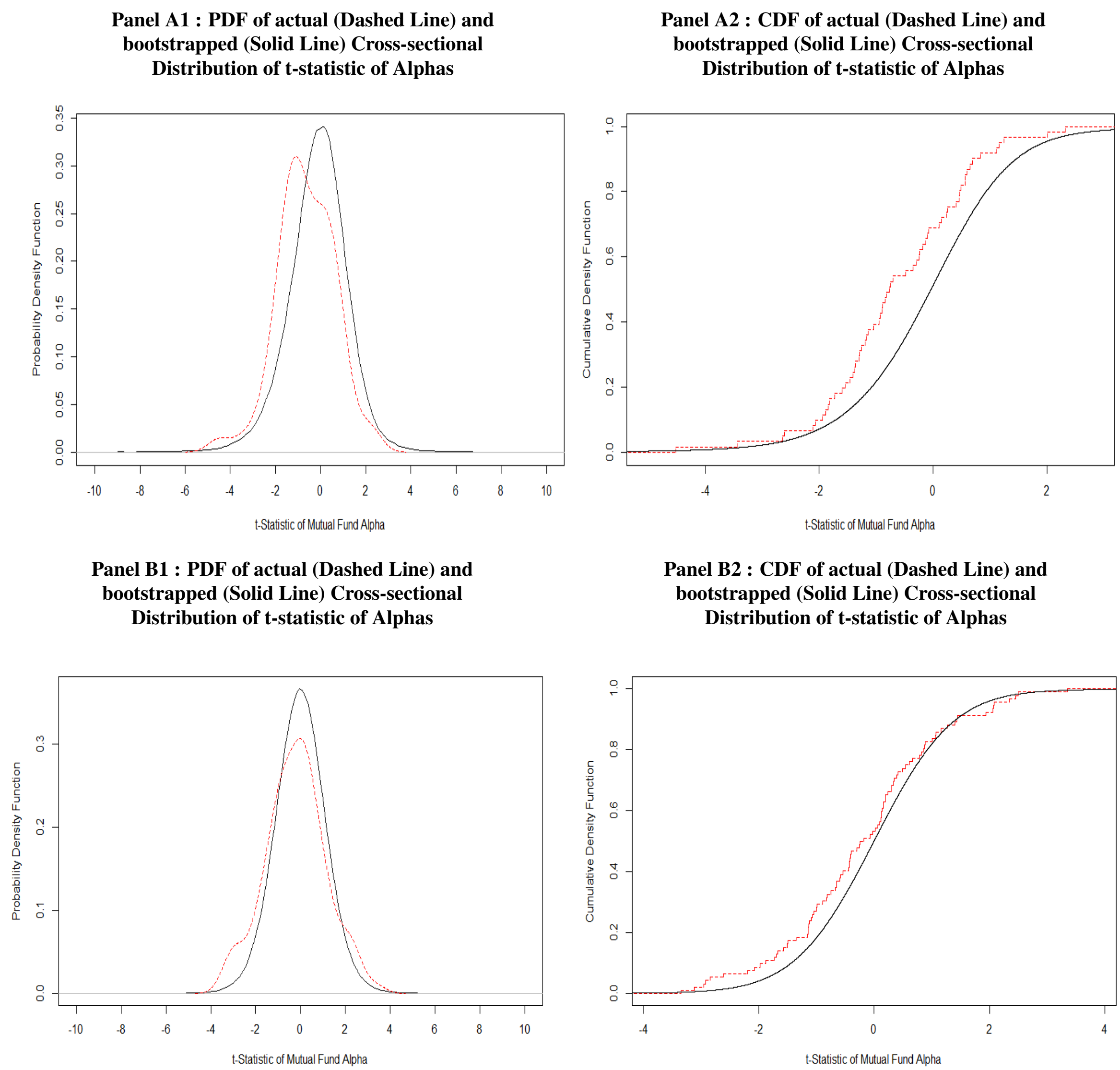
Table 6: Luck vs Skill for sub-periods

This table reports the results for the cross-section of the Norwegian mutual fund's performance measure for subperiod. Panel A shows the Norwegian mutual fund ranked on t-statistics Four-Factor Model Alphas for the subperiod 1983-1999. The first and second row reports the t-statistic of alphas and the cross-sectionally bootstrapped p-values of the t-alpha. The third row reports Parametric p-values of the t-statistic, which is based on standard critical values of t-statistic. Panel B shows the Norwegian mutual fund ranked on t-statistics Four-Factor Model Alphas. The first and second row reports the t-statistic of alphas and the cross-sectionally bootstrapped p-values of the t-statistic of alpha. The third row reports Parametric p-values of the t-statistic, which is based on standard critical values of t-statistic. The t-statistics of alpha are based on heteroskedasticity- and autocorrelation-consistent standard errors. The cross-sectionally bootstrapped p-value is based on the distribution of the best (worst) funds in 10,000 bootstrap resamples.

	Bottom	2nd	3rd	Bottom 5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	Top 5%	3rd.	2nd.	Top
Panel A: Fund Ranked on t-Statistic Four-Factor Model Alphas (1983-1999)																	
t-alpha	-4.52	-3.45	-2.64	-2.64	-2.07	-1.60	-1.30	-1.05	-0.82	-0.24	0.11	0.46	0.69	1.16	1.25	2.01	2.33
Bootstrapped p-value	0.17	0.19	0.27	0.27	0.25	0.18	0.14	0.11	0.90	0.79	0.82	0.74	0.82	0.75	0.81	0.51	0.57
Parametric p-value	**0.00**	**0.00**	**0.01**	**0.01**	**0.02**	0.06	0.10	0.16	0.21	0.40	0.46	0.32	0.25	0.13	0.11	**0.03**	**0.02**
Panel B: Fund Ranked on t-Statistic Four-Factor Model Alphas (2000-2015)																	
t-alpha	-3.35	-3.12	-2.95	-2.84	-1.98	-1.15	-0.89	-0.54	-0.25	0.13	0.34	0.85	1.43	2.07	2.46	2.50	3.35
Cross-sectionally Bootstrapped p-value	0.35	0.13	**0.04**	**0.00**	**0.02**	0.15	**0.02**	**0.04**	0.95	0.81	0.91	0.64	0.42	0.10	0.24	0.47	0.35
Parametric (standard) p-value	**0.00**	**0.00**	**0.00**	**0.00**	**0.03**	0.13	0.19	0.30	0.40	0.45	0.37	0.20	0.08	**0.02**	**0.01**	**0.01**	**0.00**

Figure 8: Estimated vs. bootstrap cross section of alpha t-statistics for subperiods.

This figure plots the kernel density estimates of actual (dashed line) and bootstrapped (solid line) cross-sectional distribution of the t-statistic of mutual fund alphas. Panel A1 shows the kernel density estimate of the probability density function (PDF) of the distributions for the first half of the period (1983-1999), and Panel B1 the kernel density estimate of the probability density function of the distributions for the second half of the period (2000-2015). Panel A2 (Panel B2) shows the kernel density estimate of the cumulative density function (CDF) for the first (second) half of the period.



All taken together, the bootstrap results for the subperiods show that the bottom funds' performances are due to bad luck in the first half, but due to bad skill in the second half. The results of the top performing funds show no evidence of superior skill in any of the sub-periods. Here we fail to reject the null hypothesis that the performances of the top funds are due to luck.

5.5 Bootstrap Performance Persistence Test

The results of our analysis in Section 5.2 suggests that the performance of the bottom fund managers is not a result of bad luck, but due to poor skill. This may imply that there is some level of persistence in the performance. We will also point out that even though the results of the bootstrap simulation showed no indication of superior fund performance among actively managed Norwegian mutual fund, evidence of performance persistence could mean that investors may have a reliable way of identifying funds that occasionally outperform the market and earn risk-adjusted returns. In this section will investigate the extent and duration of the persistence by implementing the method from Section 3.3.

We test for performance persistence in the quintile portfolios sorted on three different performance measures, raw returns, alpha and the t-statistic of alpha. The results are presented in Table 7, Table 8 and Table 9. We present the results in this order. We are mainly interested in the bottom and top 20% portfolios. In column 6 and 7 of the following tables the bootstrapped p-values of the right and left tail are shown. As you remember, we are testing the null hypothesis of an alpha equal to zero in the holding periods (i.e. no persistence).

Table 7 reports the result when funds are ranked on their past raw return. Panel A of Table 7 shows the results using a one year ranking and holding strategy. Panel B shows the results using a one year ranking and three months holding strategy. Panel C (Panel D) shows the results using a three year ranking and one year (three months) holding strategy. None of the portfolios generate a statistically significant alpha in the holding period, for any of the strategies. In other words we cannot reject the null hypothesis of an alpha equal to zero. The result implies that the investors cannot look at funds past raw returns to earn risk-adjusted returns in the future. Looking for performance persistence based on raw returns, we find no evidence of "hot hands" or "icy hands" in Norwegian mutual funds. These findings support the semi-strong form of the Efficient Market Hypothesis and are in line with previous research on persistence in Norwegian mutual using raw returns (see e.g. Sørensen, 2009).

Table 8 reports the result when funds are ranked on their past alpha estimates. In Panel A, we find no statistically significant results, using a one year ranking and holding strategy. It is, however, interesting to note that the top 20% portfolio generate a negative alpha of -1.24%. Panel B shows the results on a one year ranking and three month holding strategy. The statistically significant results we find using this strategy is with the bottom 20% portfolio indicated by the p-value of 0.046 in column 6. Hence, we reject the null hypothesis that the alpha is equal to zero. Using such a portfolio strategy yielded an alpha of -2.34%. This is evidence of short-term persistence among the worst performing funds, known as “icy hands”. The spread is also statistically significant with a p-value of 0.027. This spread portfolio strategy would yield an alpha return of 3.076%. However, a strategy exploiting this spread is not a realistic one with regards to the difficulty in shorting mutual funds and the transaction costs associated with constant rebalancing. In Panel C, the results for a 3-year ranking and 1-year holding strategy. Here, we have, again, statistically significant results for the bottom 20%, indicated by the p-value of 0.016, well below our significance level of 0.05. This portfolio yields an alpha return of -2.901%. Again, we find evidence for the "icy-hands" phenomenon. Panel D shows the results on a 3-year ranking and 1-month holding strategy. Again, we find statistically significant results for the bottom 20% portfolio. The bottom 20% portfolio has a p-value of 0.008, well below the significance level 0.05. The abnormal return generated by this portfolio is -2.195%.

In Table 9 we sort funds into quintile porfolios according to their past estimated t-statistic of alpha. In Panel A, there no statistically significant results for a one year ranking and holding strategy. Panel B shows that the the bottom 20% portfolio and the spread portfolio has significant results, using a one year ranking and three months holding strategy. The bottom 20% portfolio has a p-value of 0.015 and generates an alpha of -2.328%. The spread portfolio generates an alpha of 3.234. In panel C there is no significant results. In Panel D, we find statistically significant result for the bottom 20% portfolio using a three year ranking and three months holding strategy. The alpha generated by the bottom portfolio is -2.307%. Again we have evidence of short-persistence, up to three months.

To summarize, we find no evidence of short-term performance persistence when we rank according to raw returns. However, when we rank based on previous alpha and the t-statistic of alpha, we find statistically significant results. Kosowski et. al. (2006) argues that t-alphas are a better measure for performance persistence since it controls for the difference in risk taking among funds. We conclude that there is short-term persistence among the bottom 20% of funds. We found no evidence of persistence among the top performers, only the bottom 20% portfolios show short-term persistence in performance. We also find significant results with the constructed spread portfolio (e.i. winners minus loser), but exploiting this for economic gains are unrealistic for the average investor since it would require constant rebalancing and shorting of mutual funds. We were looking for evidence of the "hot hands" phenomena, but our evidence indicates the opposite, "icy hands." That is, short-term persistence in the results among the poorest performers. This means that investors cannot identify good performers based on previous performance to earn abnormal returns. Past winner funds do not stay winner funds and past loser funds continue to lose.

Table 7: Performance Persistence Across Quintiles Portfolios formed on lagged returns

The table reports the results of persistence tests when portfolios are formed on lagged returns. In Panel A, mutual funds are sorted into portfolios based on their average return over the prior year, rebalancing the portfolios every year. In Panel B, portfolios are formed based on their past one-year average return, and funds are held for one year before rebalancing. In Panel C, mutual funds are sorted into portfolios based on their past year average return over the prior 3 years, rebalancing the portfolios every year. In Panel D, mutual funds are sorted into portfolios based on their average return over the previous 3 years, rebalancing the portfolios every tree months. Column 1 through 4 report excess return, standard deviation, alpha and alpha t-statistic for each portfolio. Column 5 show the one-tailed parametric p-value of alpha. Column 6 and 7 reports bootstrapped p-value of t-statistic alpha. Column 6 reports the probability of that bootstrapped t-statistic of alpha is less than $(-|t(alpha)|)$, i.e. the left tail of the bootstrap distribution. Column 7 reports the probability of that bootstrapped t-statistic of alpha is greater than $(+(alpha))$, i.e. the right tail of the bootstrap distribution. The last 3 columns report the skewness, kurtosis and the p-values of a Jarque-Berra normality test. The reported excess returns, standard deviations, and alphas are annualized.

	Excess Return	Std. Dev.	Alpha	t-Stat of Alpha	One-tailed Parametric p-Value of Alpha	Bootstr. p-Value of t(Alpha) (Left Tail)	Bootstr. p-Value of t(Alpha) (Right Tail)	RMRF	SMB	HML	PR1YR	Adj.R^2	Skew	Kur	p-Value (JB-Test)
Panel A: 1-year ranking Periods, 1-year holding Period															
Top 20%	6.298	22.091	-0.698	-0.615	0.269	0.273	0.287	0.973	0.189	-0.056	0.036	0.893	-0.838	4.951	<0.01
2nd	5.67	21.351	-0.324	-0.29	0.386	0.415	0.364	0.935	0.103	-0.022	0.013	0.93	-0.839	4.85	<0.01
3nd	6.207	21.976	0.502	0.376	0.354	0.353	0.338	0.941	0.064	-0.036	0.018	0.908	-0.813	5.021	<0.01
4nd	5.957	22.093	0.038	0.039	0.484	0.45	0.512	0.964	0.113	-0.055	-0.009	0.925	-0.731	4.804	<0.01
Bot 20%	4.489	22.293	-2.016	-1.497	0.068	0.066	0.065	0.973	0.192	-0.027	-0.029	0.893	-0.756	4.958	<0.01
Spread	1.809	8.293	1.318	0.9	0.184	0.191	0.202	0	-0.003	-0.029	0.065	0.011	0.344	18.611	<0.01
Panel B: 1-year ranking, 3-month holding Period															
Top 20%	6.88	22.447	-0.214	-0.153	0.439	0.461	0.431	0.976	0.182	-0.06	0.057	0.893	-0.782	4.767	<0.01
2nd	6.679	22.185	0.361	0.276	0.391	0.421	0.376	0.953	0.119	-0.044	0.034	0.904	-0.782	4.745	<0.01
3nd	6.265	21.837	0.679	0.637	0.262	0.274	0.247	0.935	0.065	-0.026	0.005	0.935	-0.905	5.12	<0.01
4nd	6.13	22.2	0.542	0.479	0.316	0.298	0.303	0.95	0.095	-0.031	-0.031	0.928	-0.775	4.944	<0.01
Bot 20%	2.196	22.103	-2.319	-0.931	0.176	0.184	0.171	0.8	0.226	-0.054	-0.174	0.651	-0.786	4.964	<0.01
Spread	4.685	14.778	2.106	0.794	0.214	0.217	0.23	0.176	-0.044	-0.006	0.23	0.115	-0.707	13.89	<0.01
Panel C: 3-year ranking Periods, 1-year holding Period															
Top 20%	6.824	22.218	-0.352	-0.286	0.388	0.377	0.399	0.96	0.173	-0.098	0.029	0.897	-0.785	4.656	<0.01
2nd	6.598	22.152	0.109	0.095	0.462	0.484	0.443	0.954	0.092	-0.019	0	0.937	-0.788	4.801	<0.01
3nd	6.647	22.793	-0.03	-0.021	0.492	0.545	0.442	0.965	0.079	-0.011	0.028	0.905	-0.774	4.757	<0.01
4nd	5.978	22.176	-0.327	-0.262	0.397	0.392	0.413	0.948	0.081	0	-0.014	0.932	-0.943	5.31	<0.01
Bot 20%	4.961	22.13	-1.67	-1.483	0.069	0.072	0.061	0.954	0.136	0.013	-0.037	0.927	-0.854	4.879	<0.01
Spread	1.863	7.976	1.318	0.902	0.184	0.186	0.183	0.007	0.038	-0.111	0.066	0.08	0.254	15.296	<0.01
Panel D: 3-year ranking Period, 3-month holding Period															
Top 20%	6.643	22.537	-0.674	-0.464	0.321	0.32	0.304	0.963	0.203	-0.077	0.006	0.892	-0.705	4.458	<0.01
2nd	6.286	22.591	-0.021	-0.019	0.492	0.499	0.492	0.956	0.082	-0.018	-0.02	0.941	-0.884	4.842	<0.01
3nd	7.203	22.807	0.911	0.68	0.248	0.279	0.224	0.95	0.068	-0.027	0.005	0.911	-0.772	4.816	<0.01
4nd	5.364	22.348	-0.823	-0.787	0.216	0.24	0.22	0.944	0.066	-0.014	-0.009	0.943	-0.843	5.015	<0.01
Bot 20%	4.775	22.697	-2.081	-1.593	0.056	0.065	0.051	0.958	0.106	0.011	0.017	0.914	-0.855	4.693	<0.01
Spread	1.868	7.895	1.407	0.938	0.174	0.173	0.182	0.005	0.097	-0.088	-0.011	0.063	0.826	20.669	<0.01

Table 8: Performance Persistence Across Quintiles Portfolios formed on lagged Alpha

The table reports the results of persistence tests when portfolios are formed on lagged alpha. In Panel A, mutual funds are sorted into portfolios based on their four-factor model estimated alpha over the prior year, rebalancing the portfolios every year. In Panel B, portfolios are formed based on their past one-year alpha, and funds are held for one year before rebalancing. In Panel C, mutual funds are sorted into portfolios based on their past year alpha over the prior 3 years, rebalancing the portfolios every year. In Panel D, mutual funds are sorted into portfolios based on their alpha over the previous 3 years, rebalancing the portfolios every tree months. Column 1 through 4 report excess return, standard deviation, alpha and alpha t-statistic for each portfolio. Column 5 show the one-tailed parametric p-value of alpha. Column 6 and 7 reports bootstrapped p-value of t-statistic alpha. Column 6 reports the probability of that bootstrapped t-statistic of alpha is less than $(-|t(alpha)|)$, i.e. the left tail of the bootstrap distribution. Column 7 reports the probability of that bootstrapped t-statistic of alpha is greater than $(+(alpha))$, i.e. the right tail of the bootstrap distribution. The last 3 columns report the skewness, kurtosis and the p-values of a Jarque-Berra normality test. The reported excess returns, standard deviations, and alphas are annualized.

	Excess Return	Std. Dev.	Alpha	t-Stat of Alpha	One-tailed Parametric p-Value of Alpha	Bootstrap p-Value of t(Alpha) (Left Tail)	Bootstrap p-Value of t(Alpha) (Right Tail)	RMRF	SMB	HML	PR1YR	Adj.R^2	Skew	Kur	p-Value (JB-Test)
Panel A: 1-year ranking Periods, 1-year holding Period, ranked on alpha															
Top 20%	5.362	21.71	-1.247	-1.139	0.128	0.136	0.156	0.954	0.16	-0.039	0.024	0.904	-0.818	4.961	<0.01
2nd	5.646	21.251	-0.043	-0.039	0.484	0.488	0.481	0.926	0.084	-0.024	0.003	0.934	-0.837	4.891	<0.01
3nd	4.557	22.131	-1.392	-1.385	0.083	0.083	0.09	0.963	0.084	-0.03	0.011	0.93	-0.878	5.025	<0.01
4nd	6.399	22.131	0.589	0.45	0.326	0.347	0.311	0.951	0.092	-0.047	0.002	0.904	-0.732	4.85	<0.01
Bot 20%	6.265	22.538	-0.2	-0.152	0.44	0.435	0.437	0.983	0.188	-0.061	-0.025	0.894	-0.749	4.809	<0.01
Spread	-0.903	6.821	-1.047	-0.826	0.205	0.19	0.206	-0.03	-0.028	0.022	0.049	0.016	-1.099	12.661	<0.01
Panel B: 1-year ranking Periods, 3-month holding Period															
Top 20%	7.186	22.178	0.733	0.602	0.274	0.262	0.276	0.955	0.14	-0.053	0.033	0.899	-0.749	4.485	<0.01
2nd	5.266	21.29	-0.513	-0.524	0.3	0.327	0.299	0.921	0.094	-0.025	0.007	0.936	-0.807	4.619	<0.01
3nd	5.507	22.11	-0.145	-0.128	0.449	0.453	0.447	0.947	0.063	-0.027	0.007	0.935	-0.87	5.114	<0.01
4nd	5.774	22.221	-0.105	-0.1	0.46	0.49	0.436	0.952	0.094	-0.04	0.006	0.92	-0.921	5.53	<0.01
Bot 20%	4.498	23.036	-2.343	-1.794	**0.037**	**0.046**	**0.041**	1.001	0.228	-0.058	-0.033	0.896	-0.698	4.656	<0.01
Spread	2.688	7.988	3.076	1.984	**0.024**	**0.027**	**0.034**	-0.046	-0.088	0.005	0.066	0.032	-0.098	12.488	<0.01
Panel C: 3-year ranking Periods, 1-year holding Period															
Top 20%	5.52	21.901	-1.079	-0.95	0.171	0.164	0.204	0.936	0.128	-0.031	-0.003	0.901	-0.766	4.571	<0.01
2nd	6.224	21.788	0.034	0.031	0.488	0.5	0.475	0.934	0.065	-0.017	0.004	0.941	-0.9	5.016	<0.01
3nd	5.824	22.051	-0.406	-0.428	0.334	0.362	0.304	0.948	0.056	-0.027	0.011	0.949	-0.945	5.317	<0.01
4nd	8.787	22.514	2.187	1.439	0.076	0.119	0.064	0.951	0.088	-0.009	0.018	0.897	-0.801	4.668	<0.01
Bot 20%	4.758	22.886	-2.901	-2.29	**0.011**	**0.016**	**0.013**	1.002	0.226	-0.046	-0.014	0.908	-0.76	4.914	<0.01
Spread	0.763	7.515	1.823	1.47	0.071	0.075	0.081	-0.066	-0.099	0.015	0.011	0.034	-0.898	23.136	<0.01
Panel D: 3-year ranking Periods, 3-month holding Period, ranked on alpha															
Top 20%	5.557	22.549	-0.935	-0.801	0.212	0.237	0.234	0.943	0.095	-0.029	0.009	0.902	-0.834	4.683	<0.01
2nd	7.152	21.885	1.022	0.845	0.199	0.225	0.193	0.924	0.06	-0.024	0.012	0.938	-0.852	4.722	<0.01
3nd	7.047	22.745	0.71	0.5	0.309	0.383	0.292	0.952	0.075	-0.05	0.013	0.913	-0.736	4.76	<0.01
4nd	6.778	22.479	0.304	0.267	0.395	0.388	0.397	0.952	0.082	-0.006	0.003	0.935	-0.919	5.053	<0.01
Bot 20%	4.657	23.324	-3.035	-2.195	**0.014**	**0.008**	**0.012**	1.004	0.227	-0.034	-0.025	0.905	-0.744	4.692	<0.01
Spread	0.9	7.991	2.1	1.396	0.082	0.079	0.095	-0.061	-0.132	0.004	0.034	0.044	-0.745	19.348	<0.01

Table 9: Performance Persistence Across Quintiles Portfolios formed on lagged t-statistic of Alpha

The table reports the results of persistence tests when portfolios are formed on lagged t-statistic of Alpha. In Panel A, mutual funds are sorted into portfolios based on their t-statistic of alpha over the prior year, rebalancing the portfolios every year. In Panel B, portfolios are formed based on their past one-year t-statistic of alpha, and funds are held for one year before rebalancing. In Panel C, mutual funds are sorted into portfolios based on their past year t-statistic of alpha over the prior 3 years, rebalancing the portfolios every year. In Panel D, mutual funds are sorted into portfolios based on their t-statistic of alpha over the previous 3 years, rebalancing the portfolios every tree months. Column 1 through 4 report excess return, standard deviation, alpha and alpha t-statistic for each portfolio. Column 5 show the one-tailed parametric p-value of alpha. Column 6 and 7 reports bootstrapped p-value of t-statistic alpha. Column 6 reports the probability of that bootstrapped t-statistic of alpha is less than $(-|t(alpha)|)$, i.e. the left tail of the bootstrap distribution. Column 7 reports the probability of that bootstrapped t-statistic of alpha is greater than $(+(alpha))$, i.e. the right tail of the bootstrap distribution. The last 3 columns report the skewness, kurtosis and the p-values of a Jarque-Berra normality test. The reported excess returns, standard deviations, and alphas are annualized.

	Excess Return	Std. Dev.	Alpha	t-Stat of Alpha	One-tailed Parametric p-Value of Alpha	Bootstrap p-Value of t(Alpha) (Left Tail)	Bootstrap p-Value of t(Alpha) (Right Tail)	RMRF	SMB	HML	PR1YR	Adj.R^2	Skew	Kur	p-Value (JB-Test)
Panel A: 1-year ranking Periods, 1-year holding Period, ranked on t-alpha															
Top 20%	5.675	21.634	-0.864	-0.767	0.222	0.222	0.238	0.951	0.148	-0.041	0.03	0.908	-0.757	4.688	<0.01
2nd	4.758	21.859	-1.492	-1.381	0.084	0.099	0.082	0.959	0.115	-0.029	0.019	0.926	-0.886	5.299	<0.01
3nd	6.565	22.316	0.635	0.444	0.329	0.334	0.303	0.959	0.119	-0.066	-0.006	0.893	-0.729	4.729	<0.01
4nd	5.485	22.178	-0.448	-0.449	0.327	0.345	0.35	0.964	0.111	-0.029	-0.015	0.922	-0.918	5.308	<0.01
Bot 20%	5.754	21.874	-0.234	-0.199	0.421	0.411	0.431	0.953	0.126	-0.031	-0.015	0.917	-0.834	5.043	<0.01
Spread	-0.079	6.224	-0.63	-0.549	0.292	0.276	0.31	-0.002	0.022	-0.01	0.045	0.011	-1.257	17.806	<0.01
Panel B: 1-year ranking Periods, 3-month holding Period															
Top 20%	7.176	22.056	0.791	0.673	0.251	0.237	0.265	0.951	0.136	-0.054	0.031	0.904	-0.771	4.542	<0.01
2nd	5.941	21.942	-0.297	-0.264	0.396	0.399	0.388	0.951	0.118	-0.037	0.025	0.925	-0.69	4.557	<0.01
3nd	5.807	22.14	-0.044	-0.037	0.485	0.516	0.448	0.954	0.114	-0.055	-0.011	0.926	-0.814	4.909	<0.01
4nd	4.94	22.24	-0.774	-0.678	0.249	0.262	0.25	0.949	0.096	-0.038	-0.013	0.917	-0.93	5.48	<0.01
Bot 20%	4.022	22.414	-2.442	-2.328	**0.01**	**0.015**	**0.014**	0.972	0.153	-0.022	-0.003	0.916	-0.844	4.812	<0.01
Spread	3.154	7.422	3.234	2.34	**0.01**	**0.012**	**0.011**	-0.021	-0.017	-0.032	0.034	0.005	0.011	15.985	<0.01
Panel C: 3-year ranking Periods, 1-year holding Period															
Top 20%	6.415	21.711	-0.212	-0.205	0.419	0.4	0.424	0.943	0.109	-0.039	0.017	0.937	-0.822	4.846	<0.01
2nd	5.931	22.61	-0.204	-0.158	0.437	0.428	0.461	0.937	0.095	-0.04	-0.028	0.871	-0.92	5.275	<0.01
3nd	7.139	22.396	0.633	0.441	0.33	0.365	0.316	0.95	0.087	-0.051	0.022	0.903	-0.8	4.801	<0.01
4nd	5.951	23.054	-1.332	-1.138	0.128	0.132	0.134	1.002	0.131	0.002	0.019	0.931	-0.887	5.018	<0.01
Bot 20%	5.319	21.931	-1.3	-1.053	0.147	0.148	0.134	0.948	0.137	-0.017	-0.025	0.925	-0.9	5.163	<0.01
Spread	1.096	5.462	1.089	1.137	0.128	0.136	0.129	-0.005	-0.029	-0.022	0.042	0.014	-0.539	5.7	<0.01
Panel D: 3-year ranking Periods, 3-month holding Period															
Top 20%	6.178	22.232	-0.4	-0.347	0.364	0.376	0.351	0.947	0.096	-0.026	0.016	0.933	-0.834	4.776	<0.01
2nd	7.183	22.96	1.083	0.834	0.202	0.161	0.253	0.936	0.067	-0.06	0.006	0.872	-0.953	5.14	<0.01
3nd	6.704	22.517	0.237	0.162	0.436	0.459	0.408	0.943	0.091	-0.03	0.011	0.906	-0.787	4.777	<0.01
4nd	6.024	22.88	-0.68	-0.578	0.282	0.269	0.286	0.97	0.105	-0.009	-0.005	0.928	-0.801	4.793	<0.01
Bot 20%	4.805	22.665	-2.307	-1.9	**0.029**	**0.026**	**0.032**	0.974	0.175	-0.009	-0.032	0.926	-0.901	4.936	<0.01
Spread	1.373	6.304	1.907	1.323	0.093	0.099	0.093	-0.028	-0.079	-0.017	0.049	0.032	1.714	18.052	<0.01

6 Conclusion

This thesis uses a data set on Norwegian mutual funds over the period 1983-2015. Our data set contains 98 different actively managed mutual funds with monthly net returns and is free of survivorship bias. We examine the performance of individual funds as well as the performance on the aggregate level, using the Carhart's (1997) four-factor model as our performance model. We apply a bootstrap method similar to Kosowski et. al. (2006) to distinguish skill from luck among the individual funds. In addition to the ability to distinguish skill from luck, the bootstrap is necessary for proper inference, since the cross-sectional distribution of fund alphas exhibit complex non-normalities in the tails. Heterogeneous risk-taking causes these non-normalities among funds, as well as non-normalities in the individual fund's distribution of alpha.

We find that the average Norwegian mutual fund provides an alpha that is indistinguishable from zero, which suggest that if the average mutual fund manager does exhibit stock-picking skills, the return generated by skill is charged as fees. When examining individual fund performance, we find no evidence of skill among top performing funds located in the right tail of the alpha distribution, but rather evidence of poor skill among worst performing funds in the left tail. We reject the hypothesis that poor performance among the worst performing funds is caused by bad luck alone. This paper also states the importance of the bootstrap, as inference using the bootstrap deviates significantly from standard parametric tests. Our results indicate that it is difficult for an investor to earn abnormal risk-adjusted returns by investing in actively managed Norwegian mutual funds.

To evaluate persistence, we used a recursive portfolio formation test. To obtain a comprehensive examination of performance persistence, we have used three performance measures, raw returns, alpha, and t-statistic of alpha. The results of the persistence tests showed no evidence of persistence among the top performing funds, indicating that investor cannot earn abnormal risk-adjusted returns by exploiting past performance, which is also in line with the semi-strong form of the EMH. Among the worst performing funds, we find evidence of short-term persistence, but it disappears with increased time horizon. Ranked on t-statistic of alpha, the performance of the worst funds persists for three months. Ranked on alpha, the performance of the worst funds persists up to one year. Additionally, a strategy of buying winners and selling losers (i.e. the spread portfolio), show statistically significant persistence up to 3 months, but exploiting this strategy for economic gain is not realistic since it would require constant rebalancing, incurring large transactions costs.

Our findings are in agreement with previous research on mutual fund performance. Our empirical evidence would support the semi-strong form of EMH. Active management does not seem justified from the investor's point of view. Investors would be better off by investing in a low-cost passive index fund.

References

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

BERK, J. B. & R. C. GREEN (2004): "Mutual Fund Flows and Performance in Rational Markets", *Journal of Political Economy*, 112(6), 1269–1295.

BICKEL, P. J. & D. A. FREEDMAN (1984): "Asymptotic Normality and the Bootstrap in Stratified Sampling", *Annals of Statistics*, 12(2), 470–482.

BLAKE, D. & A. TIMMERMANN (1998): "Mutual Fund Performance: Evidence from the UK", *European Finance Review*, 2(1995), 57–77.

BOLLEN, N. P. B. & J. A. BUSSE (2005): "Short-Term Persistence in Mutual Fund Performance", *Review of Financial Studies*, 18(2), 569–597.

BROWN, S. J. & W. N. GOETZMANN (1995): "Performance Persistence", *Journal of Finance*, 50(2), 679–698.

BUSSE, J. A., A. GOYAL & S. WAHAL (2010): "Performance and Persistence in Institutional Investment Management", *Journal of Finance*, 65(2), 765–790.

CARHART, M. M. (1997): "On Persistence in Mutual Fund Performance", *Journal of Finance*, 52(1), 57–82.

CARPENTER, J. N. & A. W. LYNCH (1999): "Survivorship Bias and Attrition Effects in Measures of Performance Persistence", *Journal of Financial Economics*, 54(3), 337–374.

CHEN, H.-L., N. JEGADEESH & R. WERMERS (2000): "The Value of Active Mutual Fund Management : An Examination of the Stockholdings and Trades of Fund Managers", *Journal of Financial and Quantitative Analysis*, 35(3), 343–368.

CHRISTENSEN, M. (2013): "Danish Mutual Fund Performance", *Applied Economics Letters*, 20(8), 818–820.

CONNOR, G. & R. A. KORAJCZYK (1991): "The Attributes , Behavior and Performance of U. S. Mutual Funds", *Review of Quantitative Finance and Accounting*, 1(1), 5–26.

CUTHBERTSON, K., D. NITZSCHE & N. O'SULLIVAN (2008): "UK Mutual Fund Performance: Skill or Luck?", *Journal of Empirical Finance*, 15(4), 613–634.

DAHLQUIST, M., S. ENGSTRÖM & P. SÖDERLIND (2000): "Performance and Characteristics of Swedish Mutual Funds", *Journal of Financial and Quantitative Analysis*, 35(3), 409–423.

DANIEL, K., M. GRINBLATT, S. TITMAN & R. WERMERS (1997): "Measuring Mutual Fund Performance with Characteristic-Based Benchmarks", .

EDELEN, R. M. (1999): "Investor Flows and the Assessed Performance of Open-End Mutual Funds", *Journal of Financial Economics*, 53(3), 439–466.

EFRON, B. (1979): "Bootstrap Methods: Another look at the jackknifing", *Annals of Statistics*, 40(1), 1–26.

ELTON, E. J., M. J. GRUBER & C. R. BLAKE (1996): "The Persistence of Risk-Adjusted Mutual Fund Performance", *Journal of Finance*, 69(2), 133–157.

ELTON, E. J., M. J. GRUBER, S. DAS & M. HLAVKA (1993): "Efficiency with Costly Information: A Reinterpretation of Evidence from Managed Portfolios", *Review of Financial Studies,*, 6(1), 1–22.

FAMA, E. F. (1970): "Efficient Capital Markets: a Review of Theory and Empirical Work", *Journal of Finance*, 25(2), 383–417.

FAMA, E. F. & K. R. FRENCH (1992): "The Cross-Section of Expected Stock Returns", *Journal of Finance*, 47(2), 427–465.

——— (1993): "Common Risk Factors in The Returns on Stocks and Bonds", *Journal of Financial Economics*, 33(1), 3–56.

——— (1996): "Multifactor Explanations of Asset Pricing Anomalies", *Journal of Finance*, LI(1), 55–84.

——— (2010): "Luck versus Skill in the Cross-Section of Mutual Fund Returns", *Journal of Finance*, 65(5), 1915–1947.

GALLEFOSS, K., H. H. HANSEN, E. S. HAUKAAS & P. MOLNÀR (2015): "What Daily Data Can Tell Us About Mutual Funds: Evidence from Norway", *Journal of Banking and Finance*, 55(June), 117–129.

GJERDE, Ø. & F. SÆTTEM (1991): "Performance Evaluation of Norwegian Mutual Funds", *Scandinavian Journal of Management*, 7(4), 297–307.

GRINBLATT, M. & S. TITMAN (1989): "Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings", .

GROSSMAN, S. J. & J. E. STIGLITZ (1980): "On the Impossibility of Informationally Efficient Markets", *American Economic Review,*, 70(3), 393–408.

GRUBER, M. J. (1996): "Another Puzzle: The Growth in Actively Managed Mutual Funds", *Journal of Finance*, 51(3), 783–810.

HALL, P. (1986): "On the Number of Bootstrap Simulations", *Annals of Statistics*, 14(4), 1453–1462.

HALL, P., J. L. HOROWITZ & B.-Y. JING (1995): "On Blocking Rules for the Bootstrap with Dependent Data", *Biometrika*, 82(3), 561–574.

HENDRICKS, D., J. PATEL & Z. RICHARD (1993): "Hot Hands in Mutual Funds : Short-Run Persistence of Relative Performance , 1974-1988", *Journal of Finance*, 48(1), 93–130.

HOROWITZ, B. Y. J. L. (2003): "Bootstrap Methods for Markov Processes", *The Econometric Society*, 71(4), 1049–1082.

HUIJ, J. & M. VERBEEK (2007): "Cross-Sectional Learning and Short-Run Persistence in Mutual Fund Performance", *Journal of Banking and Finance*, 31(3), 973–997.

IPPOLITO, R. A. (1989): "Efficiency With Costly Information: A Study of Mutual Fund Performance, 1965–1984", *Quarterly Journal of Economics*, 104(1), 1–23.

JEGADEESH, N. & SHERIDAN TITMAN (1993): “Returns to Buying Winners and Selling Losers : Implications for Stock Market Efficiency”, *Journal of Finance*, 48(1), 65–91.

JENSEN, M. C. (1968): “The Performance of Mutual Funds in the Period 1945-1964.”, *Journal of Finance*, 23(2), 389–416.

KOSOWSKI, R., A. TIMMERMANN, R. WERMERS & H. WHITE (2006): “Can Mutual Fund " Stars " Really Pick Stocks ? New Evidence from a Bootstrap Analysis”, *Journal of Finance*, 61(6), 2551–2595.

LEHMANN, B. N. & D. M. MODEST (1987): “Mutual Fund Performance Evaluation : A Comparison of Benchmarks and Benchmark Comparisons”, *Journal of Finance*, 42(2), 233–265.

LINTNER, J. (1965): “The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets”, *Review of Economics and Statistics*, 47(1), 13–37.

MALKIEL, B. G. (1995): “Returns from Investing in Equity Mutual Funds 1971 to 1991”, *Journal of Finance*, 50(2), 549–572.

MARKOWITZ, H. (1952): “Portfolio Selection”, *Journal of Finance*, 46(1), 409–419.

MOSKOWITZ, T. J. (2000): “Discussion”, *Journal of Finance*, 55(4), 1695–1703.

NEWEY, W. K. & K. D. WEST (1987): “A Simple , Positive Semi-Definite , Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”, *Econometrica*, 55(3), 703–708.

ØDEGAARD, B. A. (2017): “Empirics of the Oslo Stock Exchange:. Asset pricing results. 1980–2016.”, .

OTTEN, R. & D. BAMS (2002): “European Mutual Fund Performance”, *European Financial Management*, 8(1), 75–101.

PATTON, A., D. N. POLITIS & H. WHITE (2009): “Correction to "Automatic block-Length Selection for the Dependent bootstrap" by D. Politis and H. White”, *Econometric Reviews*, 28(4), 372–375.

POLITIS, D. N. & J. P. ROMANO (1994): "The Stationary Bootstrap", *Journal of the American Statistical Association*, 89(428), 1303–1313.

POLITIS, D. N. & H. WHITE (2004): "Automatic Block-Length Selection for the Dependent Bootstrap", *Econometric Reviews*, 23(1), 53–70.

ROLL, R. (1977): "A Critique of the Asset Pricing Theory's Tests Part I: On Past and Potential Testability of the Theory", *Journal of Financial Economics*, 4(2), 129–176.

ROSENBERG, B., K. REID & R. LANSTEIN (1985): "Persuasive evidence of market inefficiency", *Journal of Portfolio Management*, 11(3), 9–16.

SHARPE, W. (1964): "Capital Asset Pricing: A Theory of Equilibrium under Conditions of Risk", *Journal of Finance*, 32(4), 425–442.

SHARPE, W. F. (1966): "Mutual Fund Performance", *Journal of Business*, 39(1), 119–138.

SØRENSEN, L. Q. (2009): "Mutual Fund Performance at the Oslo Stock Exchange", .

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

TREYNOR, J. L. & K. K. MAZUY (1966): "Can Mutual Funds Outguess the Market?", *Harvard Business Review*, 44(4), 131–136.

UTSETH, L. P. & K. B. SANDVIK (2015): "Performance , Persistence and Business Cycle Asymmetries in Norwegian Mutual Fund Returns Do mutual funds perform when it matters the most?", .

VIDAL-GARCÍA, J. (2013): "The Persistence of European Mutual Fund Performance", *Research in International Business and Finance*, 28(1), 45–67.

WERMERS, R. (2000): "Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses", *Journal of Finance*, 55(4), 1655–1695.

Appendix A

Table A.I
Mutual Fund Database Summary Statistics

The Table presents summary statistics of our data set of mutual funds. Column 1 reports the number of funds available for the investors as of year-end each year. Column 2 reports the number of funds that are born each year and column 3 report the number of funds that are liquidated each year. Column 4 reports the return on an equally-weighted portfolio consisting of all funds each year and column 5 reports the return on the OSEFX benchmark. Returns are in percent and annualized.

	Number of funds			Returns	
Year	Year-end	Born	Liquidated	EW (All)	OSEFX
2015	52	0	3	7.328	7.143
2014	55	1	6	8.231	5.869
2013	60	3	4	22.222	22.342
2012	61	0	2	15.792	20.011
2011	63	4	0	-19.622	-18.914
2010	59	3	0	21.969	23.117
2009	56	0	0	57.513	56.587
2008	56	0	0	-65.115	-71.697
2007	56	0	2	11.441	10.611
2006	58	3	6	26.905	29.766
2005	61	1	4	40.000	35.464
2004	64	1	2	32.679	34.799
2003	65	2	3	44.994	46.588
2002	66	7	4	-38.815	-35.600
2001	63	4	4	-14.062	-14.195
2000	63	7	3	4.715	3.591
1999	59	3	3	45.830	42.462
1998	59	11	0	-32.431	-24.722
1997	48	9	0	30.083	31.030
1996	39	7	0	35.733	29.187
1995	32	7	0	16.645	1.611
1994	25	7	0	5.172	8.718
1993	18	1	0	58.404	52.835
1992	17	4	0	-12.583	-7.108
1991	13	3	0	-8.022	-6.856
1990	10	3	0	-16.035	-11.606
1989	7	0	0	52.072	46.903
1988	7	0	0	25.625	33.884
1987	7	1	0	3.727	-4.852
1986	6	0	0	-3.114	-8.195
1985	6	1	0	23.308	28.676
1984	5	0	0	38.984	24.494
1983	5	3	0	77.096	52.169

Appendix B

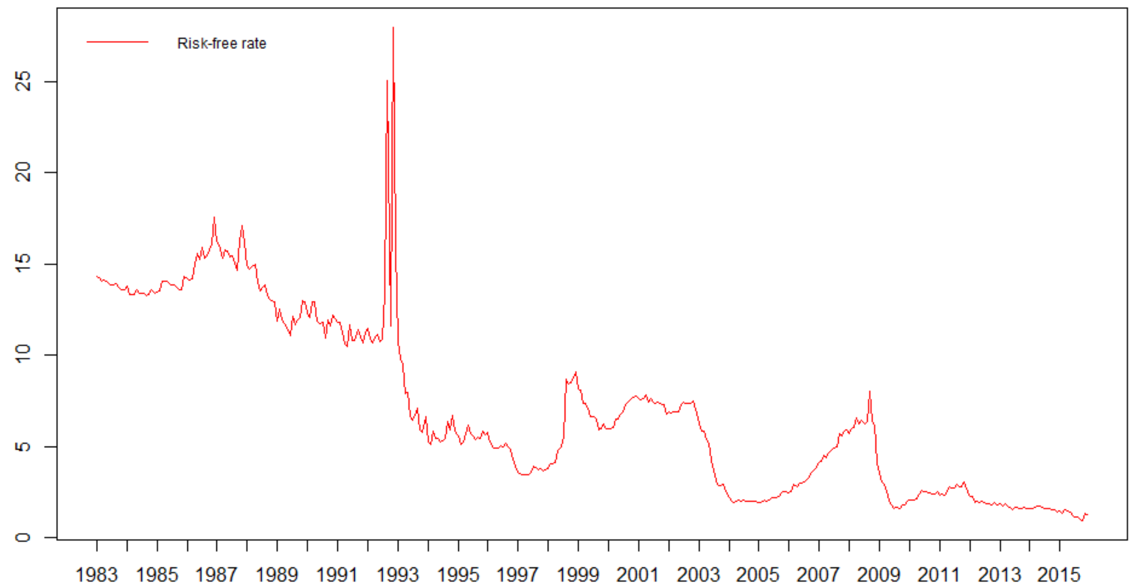


Figure B.I
Monthly risk-free interest rate from January 1983 to December 2015.

This figure plots the annualized one-month risk-free interest rate from 1983 to 2015. Prior to 1986 the overnight NIBOR is used. The period 1986 to 2015 the one-month NIBOR rate is used.

Appendix C

Table C.I
Descriptive Statistics on Individual Mutual Funds (1/2)

The table shows descriptive statistics on all individual Norwegian mutual funds in our mutual fund database. Column 1 reports the number of monthly return observations in each fund, whereas yearly excess returns, skewness, and kurtosis are reported in Columns 2 - 4. The estimates are constructed by running regressions on each individual fund from 1983 to 2015

	Obs	Mean	Standard deviation	Skewness	Kurtosis	Maximum	Minimum
ABIF Norge ++	55	3.427	0.069	-0.334	2.539	0.133	-0.168
Alfred Berg Aksjef Norge	113	3.700	0.062	-0.783	4.897	0.126	-0.257
Alfred Berg Aksjespar	105	2.764	0.067	-0.893	5.218	0.129	-0.287
Alfred Berg Aktiv II	180	4.815	0.074	-0.612	4.181	0.174	-0.280
Alfred Berg Aktiv	240	10.021	0.070	-0.833	5.150	0.206	-0.277
Alfred Berg Gambak	299	11.432	0.070	-0.495	5.440	0.280	-0.280
Alfred Berg Humanfond	192	6.599	0.064	-0.937	5.357	0.159	-0.265
Alfred Berg N. Pensjon	51	7.314	0.061	-1.370	7.528	0.112	-0.255
Alfred Berg Norge +	195	7.075	0.069	-0.992	5.236	0.170	-0.276
Alfred Berg Norge Classic	302	6.023	0.065	-1.047	5.591	0.169	-0.277
Alfred Berg Norge Etisk	144	8.792	0.070	-1.073	5.601	0.165	-0.285
Alfred Berg Norge Inst	23	13.881	0.023	-0.846	3.481	0.044	-0.051
Alfred Berg Vekst	70	2.161	0.078	-0.493	4.807	0.190	-0.285
Arctic Norwegian Equities Class A	60	7.853	0.036	-0.516	4.154	0.093	-0.095
Arctic Norwegian Equities Class B	61	10.183	0.038	-0.338	4.121	0.096	-0.094
Arctic Norwegian Equities Class D	34	12.803	0.025	-0.452	3.948	0.070	-0.061
Arctic Norwegian Equities Class I	61	10.181	0.037	-0.357	4.104	0.094	-0.094
Atlas Norge	214	7.201	0.076	-0.123	5.868	0.364	-0.259
Banco Norge	36	8.831	0.072	-0.340	2.622	0.137	-0.177
Carnegie Aksje Norge	245	10.441	0.064	-0.916	5.658	0.193	-0.282
Danske Invest Aktiv Formuesf. A	19	15.278	0.047	-0.898	3.462	0.074	-0.109
Danske Invest Norge 2	263	7.774	0.062	-1.028	6.199	0.147	-0.302
Danske Invest Norge Aksj. Inst 1	188	8.825	0.062	-0.908	5.085	0.153	-0.235
Danske Invest Norge Aksj. Inst 2	109	7.381	0.062	-1.096	6.121	0.148	-0.234
Danske Invest Norge I	263	7.130	0.062	-1.024	6.147	0.146	-0.295
Danske Invest Norge Vekst	263	9.750	0.068	0.318	9.008	0.413	-0.263
Delphi Norge	258	11.471	0.072	-0.555	4.704	0.225	-0.255
Delphi Vekst	191	6.684	0.076	-0.362	3.992	0.251	-0.237
DNB Norge (Avanse I)	385	6.554	0.063	-0.808	5.100	0.205	-0.269
DNB Norge (Avanse II)	285	5.231	0.063	-0.992	5.480	0.159	-0.269
DNB Norge (I)	287	3.964	0.063	-0.791	4.525	0.156	-0.248
DNB Norge (III)	238	7.460	0.062	-0.891	5.118	0.157	-0.248
DNB Norge (IV)	157	12.171	0.062	-0.908	5.520	0.158	-0.249
DNB Norge Selektiv (II)	168	9.569	0.063	-0.822	5.081	0.167	-0.244
DNB Norge Selektiv (III)	258	7.172	0.061	-0.834	5.017	0.168	-0.247
DNB Norge Selektiv	236	7.556	0.066	-0.731	4.639	0.166	-0.247
DNB Norge	244	5.806	0.062	-0.858	5.124	0.156	-0.248
DnB Real-Vekst	154	-3.024	0.061	-0.671	4.038	0.127	-0.231
DNB SMB	177	10.338	0.075	-0.473	3.820	0.173	-0.271
Eika Norge	147	14.657	0.061	-1.115	6.702	0.182	-0.256
Eika SMB	185	4.091	0.068	-0.675	4.272	0.168	-0.236
Fokus Barnespar	30	-8.082	0.081	-1.169	6.065	0.121	-0.288
Fondsfinans Aktiv II	47	-5.307	0.068	-0.252	2.907	0.139	-0.171
Fondsfinans Norge	156	14.951	0.063	-0.829	5.649	0.161	-0.264

Table C.II
Descriptive Statistics on Individual Mutual Funds (2/2)

	Obs	Mean	Standard deviation	Skewness	Kurtosis	Maximum	Minimum
FORTE Norge	58	1.540	0.044	0.209	4.677	0.143	-0.118
FORTE Trønder	32	12.823	0.029	0.520	3.526	0.094	-0.041
GAMBAK Oppkjøp	17	-3.403	0.058	0.343	3.156	0.135	-0.098
GJENSIDIGE AksjeSpar	150	1.874	0.066	-0.977	5.236	0.157	-0.280
GJENSIDIGE Invest	102	9.073	0.060	-0.883	5.454	0.129	-0.219
Globus Aktiv	87	8.159	0.087	-0.298	3.177	0.231	-0.232
Globus Norge II	94	5.241	0.085	-0.255	3.234	0.227	-0.235
Globus Norge	102	0.577	0.086	-0.334	3.244	0.219	-0.241
Handelsbanken Norge	250	9.359	0.064	-1.182	6.612	0.176	-0.295
Holberg Norge	180	7.030	0.062	-0.498	4.283	0.157	-0.245
K-IPA Aksjefond	35	7.242	0.068	-1.058	5.195	0.118	-0.225
KLP Aksjeinvest	95	-0.833	0.062	-0.789	4.658	0.145	-0.229
KLP AksjeNorge	201	8.197	0.064	-0.908	5.809	0.174	-0.304
Landkreditt Norge	115	4.352	0.060	-0.760	5.214	0.169	-0.212
Landkreditt Utbytte	34	13.157	0.022	-1.108	4.227	0.046	-0.048
NB-Aksjefond	205	5.777	0.066	-0.955	5.134	0.180	-0.254
Nordea Avkastning	387	5.907	0.066	-0.843	5.281	0.194	-0.287
Nordea Barnespar	46	-9.088	0.062	-0.361	2.625	0.109	-0.169
Nordea Kapital II	82	8.956	0.067	-0.466	2.771	0.131	-0.181
Nordea Kapital III	68	7.817	0.069	-0.551	2.669	0.131	-0.181
Nordea Kapital	249	9.154	0.063	-1.011	5.525	0.165	-0.264
Nordea Norge Pluss	56	5.770	0.041	-0.407	4.225	0.119	-0.113
Nordea Norge Verdi	238	8.564	0.059	-0.856	5.219	0.150	-0.251
Nordea SMB II	68	-20.529	0.078	0.166	3.037	0.182	-0.197
Nordea SMB	212	2.692	0.069	-0.252	3.541	0.178	-0.239
Nordea Vekst	378	4.304	0.066	-0.840	4.846	0.186	-0.269
ODIN Norge II	138	9.297	0.057	-1.013	6.193	0.134	-0.246
ODIN Norge	282	11.845	0.064	-0.471	4.962	0.220	-0.247
Orkla Finans 30	160	13.780	0.064	-0.741	4.519	0.142	-0.269
Pareto Aksje Norge	171	10.895	0.058	-0.877	6.236	0.159	-0.267
Pareto Aktiv	159	10.280	0.056	-0.929	7.133	0.157	-0.267
Pareto Investment Fund A	372	7.347	0.066	-0.965	5.812	0.186	-0.296
Pareto verdi	120	2.742	0.056	-1.200	7.669	0.138	-0.266
PLUSS Aksje (Fondsforval)	228	7.969	0.064	-0.742	4.879	0.172	-0.262
PLUSS Markedsverdi (Fondsforv)	251	8.064	0.060	-0.982	5.761	0.158	-0.257
Postbanken Aksjevekst	95	1.825	0.070	-0.402	3.098	0.144	-0.204
RF-Plussfond	52	15.209	0.073	-0.431	2.515	0.143	-0.176
RF Aksjefond	114	6.037	0.063	-0.772	4.297	0.131	-0.245
SEB Norge LU	66	-12.205	0.074	-0.657	4.199	0.149	-0.268
Skandia Horisont	95	5.624	0.065	-0.767	4.073	0.157	-0.222
Skandia SMB Norge	95	-6.010	0.070	-0.994	5.345	0.132	-0.279
Storebrand Aksje Innland	233	7.672	0.063	-1.012	5.539	0.152	-0.271
Storebrand AksjeSpar	413	3.775	0.057	-0.608	4.996	0.188	-0.262
Storebrand Norge A	41	23.053	0.071	-0.627	2.942	0.144	-0.178
Storebrand Norge I	188	7.480	0.066	-0.965	5.454	0.147	-0.292
Storebrand Norge Institusjon	37	4.231	0.043	-0.477	3.612	0.097	-0.099
Storebrand Norge	387	7.027	0.066	-0.801	4.912	0.176	-0.295
Storebrand Optima Norge	180	8.719	0.066	-0.990	5.473	0.144	-0.299
Storebrand Vekst	279	11.922	0.074	-0.040	6.220	0.362	-0.307
Storebrand Verdi	216	7.075	0.063	-0.926	5.504	0.133	-0.272
Swedbank Generator	63	13.272	0.051	-0.592	3.639	0.116	-0.146
Terra Norge	185	5.092	0.072	-0.762	4.437	0.186	-0.268
Terra Vekst	127	-0.698	0.086	1.082	10.831	0.488	-0.238
VåR Aksjefond	37	0.426	0.072	-1.168	6.340	0.109	-0.268

Appendix D

Table D.I
Individual Mutual Fund Characteristics (1/2)

The table shows individual fund characteristics on all Norwegian funds in our mutual fund database. Columns 1 - 6 report alpha, alpha t-statistic, and individual factor loadings on the risk factors used in the unconditional four-factor model of Carhart (1997), MKT, SMB, HML and PR1YR, respectively. Column 7 reports the R-squared. The estimates are constructed by running regressions on each individual fund from 1983 to 2015. Alphas are reported in percent per year.

Mutual Fund	α	t_α	β_{MKT}	β_{SMB}	β_{HML}	β_{PR1YR}	R^2_{adj}
ABIF Norge ++	1.142	0.809	0.973	0.038	-0.049	0.037	0.983
Alfred Berg Aksjef Norge	-2.221	-1.224	0.958	0.098	-0.003	-0.025	0.926
Alfred Berg Aksjespar	-4.400	-1.257	1.036	0.127	-0.002	0.035	0.890
Alfred Berg Aktiv II	-3.016	-1.242	1.030	0.353	-0.143	0.051	0.884
Alfred Berg Aktiv	-2.408	-1.095	1.095	0.352	-0.143	0.093	0.901
Alfred Berg Gambak	-0.454	-0.211	1.072	0.402	-0.236	0.171	0.822
Alfred Berg Humanfond	0.180	0.118	0.967	0.062	-0.075	0.008	0.950
Alfred Berg N. Pensjon	-2.273	-0.661	0.966	0.098	-0.114	0.026	0.863
Alfred Berg Norge +	0.233	0.249	0.997	0.085	-0.035	0.040	0.981
Alfred Berg Norge Classic	-1.113	-0.952	1.013	0.078	0.006	0.057	0.959
Alfred Berg Norge Etisk	-0.510	-0.443	0.994	0.047	-0.071	0.001	0.979
Alfred Berg Norge Inst	6.332	2.466	0.891	0.136	-0.002	0.058	0.888
Alfred Berg Vekst	-9.251	-2.133	1.087	0.361	-0.144	0.141	0.789
Arctic Norwegian Equities Class A	-0.386	-0.193	0.923	0.116	-0.040	0.143	0.872
Arctic Norwegian Equities Class B	0.275	0.140	0.912	0.112	-0.039	0.141	0.885
Arctic Norwegian Equities Class D	0.951	0.262	0.853	0.067	-0.031	0.150	0.763
Arctic Norwegian Equities Class I	0.307	0.156	0.906	0.112	-0.036	0.143	0.883
Atlas Norge	-1.613	-0.804	1.080	0.219	-0.250	0.066	0.877
Banco Norge	-2.020	-1.783	0.980	0.176	-0.164	-0.033	0.967
Carnegie Aksje Norge	1.724	1.399	0.955	0.033	-0.127	0.112	0.935
Danske Invest Aktiv Formuesf. A	-22.834	-3.097	0.612	0.162	0.603	0.637	0.688
Danske Invest Norge 2	1.238	0.953	0.961	0.072	-0.002	-0.027	0.928
Danske Invest Norge Aksj. Inst 1	2.473	2.377	0.931	0.015	0.013	0.002	0.971
Danske Invest Norge Aksj. Inst 2	3.447	2.197	0.938	0.046	0.024	0.013	0.966
Danske Invest Norge I	0.573	0.462	0.960	0.076	-0.002	-0.028	0.927
Danske Invest Norge Vekst	-1.822	-0.645	1.012	0.470	-0.233	0.090	0.765
Delphi Norge	-0.503	-0.318	1.121	0.389	-0.195	0.044	0.867
Delphi Vekst	-2.119	-1.081	1.048	0.407	-0.252	0.018	0.858
DNB Norge (Avanse I)	-0.074	-0.076	0.910	0.060	-0.005	-0.006	0.926
DNB Norge (Avanse II)	-1.522	-1.769	0.947	0.035	-0.017	0.012	0.966
DNB Norge (I)	-1.066	-1.441	0.931	0.010	-0.009	0.038	0.961
DNB Norge (III)	-0.116	-0.161	0.949	0.006	0.000	0.015	0.978
DNB Norge (IV)	0.615	0.587	0.938	-0.024	0.004	0.039	0.977
DNB Norge Selektiv (II)	1.004	0.853	0.931	-0.023	0.028	0.034	0.967
DNB Norge Selektiv (III)	-0.976	-0.738	0.961	0.095	-0.032	0.026	0.944
DNB Norge Selektiv	0.363	0.249	0.985	0.061	-0.019	-0.048	0.934
DNB Norge	-1.013	-1.187	0.937	-0.017	0.002	0.020	0.970
DnB Real-Vekst	-2.910	-1.772	0.932	0.104	-0.005	0.007	0.947
DNB SMB	0.579	0.197	1.151	0.566	-0.081	-0.109	0.838
Eika Norge	3.227	1.978	1.016	0.221	0.050	-0.061	0.953
Eika SMB	-1.258	-0.622	0.950	0.232	0.008	-0.108	0.899
Fokus Barnespar	-14.523	-1.542	0.972	0.140	-0.003	-0.282	0.858
Fondsfinans Aktiv II	-1.646	-0.553	0.888	0.038	-0.011	-0.026	0.907
Fondsfinans Norge	2.875	1.403	0.968	0.122	-0.033	-0.067	0.921

Table D.II
Individual Mutual Fund Characteristics (2/2)

Mutual Fund	α	t_α	β_{MKT}	β_{SMB}	β_{HML}	β_{PR1YR}	R^2_{adj}
FORTE Norge	-1.897	-0.459	0.895	-0.092	-0.011	-0.070	0.746
FORTE Trønder	12.910	3.182	0.560	-0.085	-0.106	-0.179	0.543
GAMBAK Oppkjøp	-7.218	-0.988	0.522	0.102	-0.365	0.320	0.834
GJENSIDIGE AksjeSpar	-3.381	-2.074	0.934	0.055	0.024	0.011	0.915
GJENSIDIGE Invest	-4.517	-1.575	0.961	0.193	0.104	0.048	0.907
Globus Aktiv	-7.887	-2.167	1.178	0.268	-0.149	-0.220	0.873
Globus Norge II	-10.489	-3.055	1.185	0.293	-0.149	-0.224	0.875
Globus Norge	-8.492	-2.466	1.136	0.312	-0.157	-0.230	0.883
Handelsbanken Norge	0.389	0.273	0.995	0.080	-0.038	0.068	0.941
Holberg Norge	-0.568	-0.249	0.963	0.305	-0.045	-0.031	0.895
K-IPA Aksjefond	1.028	0.237	0.968	0.199	0.051	0.024	0.893
KLP Aksjeinvest	-2.147	-1.056	0.891	0.058	-0.037	-0.001	0.912
KLP AksjeNorge	0.691	0.785	0.975	0.042	0.002	-0.004	0.963
Landkreditt Norge	0.923	0.369	0.922	0.162	0.052	-0.085	0.902
Landkreditt Utbytte	5.122	1.332	0.706	0.250	0.136	0.053	0.587
NB-Aksjefond	-1.621	-1.554	0.956	0.116	0.025	-0.054	0.960
Nordea Avkastning	-0.256	-0.266	0.938	0.080	-0.019	0.004	0.831
Nordea Barnespar	-3.907	-1.262	0.931	0.069	-0.026	0.041	0.932
Nordea Kapital II	-1.760	-1.235	1.009	0.058	-0.039	0.015	0.974
Nordea Kapital III	-1.700	-1.059	0.995	0.037	-0.030	-0.018	0.981
Nordea Kapital	0.645	0.712	0.974	0.079	-0.040	0.021	0.958
Nordea Norge Pluss	-0.044	-0.019	1.021	0.102	-0.007	-0.027	0.930
Nordea Norge Verdi	0.416	0.296	0.922	0.230	0.000	-0.058	0.922
Nordea SMB II	-17.067	-3.438	0.952	0.573	-0.123	-0.016	0.759
Nordea SMB	-6.561	-2.826	1.027	0.548	-0.042	-0.059	0.819
Nordea Vekst	-2.130	-1.816	0.967	0.085	-0.015	-0.003	0.909
ODIN Norge II	-1.320	-0.581	0.941	0.340	0.028	0.000	0.856
ODIN Norge	1.081	0.454	0.968	0.342	0.102	-0.028	0.806
Orkla Finans 30	-0.681	-0.322	0.997	0.184	-0.008	0.012	0.909
Pareto Aksje Norge	1.058	0.470	0.897	0.228	0.069	0.029	0.843
Pareto Aktiv	-2.354	-1.023	0.874	0.238	0.020	0.049	0.831
Pareto Investment Fund A	0.986	0.795	0.984	0.110	-0.044	0.025	0.896
Pareto verdi	-0.825	-0.501	0.872	0.246	-0.027	-0.148	0.861
PLUSS Aksje (Fondsforval)	1.251	0.996	0.943	0.051	-0.019	-0.014	0.926
PLUSS Markedsverdi (Fondsforv)	1.742	2.029	0.909	-0.051	0.010	0.008	0.965
Postbanken Aksjevekst	-3.444	-1.792	0.954	0.088	-0.151	-0.020	0.931
RF-Plussfond	-5.154	-1.580	1.065	0.176	-0.168	-0.015	0.938
RF Aksjefond	-1.534	-1.413	0.938	0.049	-0.007	-0.018	0.968
SEB Norge LU	-0.699	-0.264	1.036	0.057	-0.050	0.044	0.950
Skandia Horisont	1.978	0.571	0.964	0.243	-0.117	0.063	0.840
Skandia SMB Norge	-11.445	-3.140	0.989	0.476	-0.139	-0.058	0.824
Storebrand Aksje Innland	-0.236	-0.342	0.959	0.014	-0.002	0.041	0.978
Storebrand AksjeSpar	-2.475	-1.605	0.824	0.134	0.038	0.009	0.824
Storebrand Norge A	0.175	0.103	1.027	0.018	-0.055	-0.003	0.980
Storebrand Norge I	0.663	0.612	0.989	0.056	-0.028	-0.004	0.967
Storebrand Norge Institusjon	0.326	0.161	0.907	-0.031	0.028	-0.008	0.966
Storebrand Norge	0.577	0.403	0.965	0.062	-0.004	0.031	0.894
Storebrand Optima Norge	1.636	1.187	0.989	0.076	-0.029	0.001	0.946
Storebrand Vekst	0.420	0.146	1.005	0.315	-0.399	0.029	0.740
Storebrand Verdi	1.148	0.936	0.926	-0.025	0.158	0.090	0.925
Swedbank Generator	5.248	1.793	1.138	0.134	0.089	-0.066	0.802
Terra Norge	-1.115	-0.695	1.004	0.158	-0.124	0.010	0.944
Terra Vekst	-8.047	-2.778	0.950	0.335	-0.300	0.031	0.844
VÅR Aksjefond	-0.883	-0.293	1.068	0.184	0.170	-0.023	0.941

Appendix E

Table E.I
Sensitive Analysis of the Cross Section of Mutual Fund Alphas to Time-Series Dependence

This table reports the results for the cross-section of the Norwegian mutual fund's performance measure for the whole sample period 1983-2015 adopting the stationary bootstrap with different block lengths. Panel A shows results with a block length of one monthly return. Panel B, Panel C and Panel D show the results with a block length of 2, 5 and 10 monthly returns. The first and second row reports the OLS estimate of alphas and the cross-sectionally bootstrapped p-values of the t-statistic of alpha. The third row reports Parametric p-values of the t-statistic, which is based on standard critical values of t-statistic. The t-statistics of alpha are based on hetroskedasticity- and autocorrelation-consistent standard errors. The cross-sectionally bootstrapped p-value is based on the distribution of the best (worst) funds in 10,000 bootstrap resamples.

	Bottom	2nd	3rd	Bottom 5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	Top 5%	3rd.	2nd.	Top
Panel A: Block Length of One Monthly Return																	
t-alpha	-4.05	-2.98	-2.92	-2.87	-2.07	-1.41	-1.12	-0.70	-0.30	0.10	0.28	0.61	1.16	1.99	2.35	2.50	3.35
Bootstrapped p-value	0.11	0.15	**0.04**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.98	0.88	0.97	0.95	0.81	0.13	0.25	0.38	0.27
Parametric p-value	**0.00**	**0.00**	**0.01**	**0.00**	**0.02**	0.08	0.13	0.24	0.38	0.46	0.39	0.27	0.12	**0.02**	**0.01**	**0.01**	**0.00**
Panel B: Average Block Length of $T^{1/5}$ Monthly Returns																	
t-alpha	-4.05	-2.98	-2.92	-2.87	-2.07	-1.41	-1.12	-0.70	-0.30	0.10	0.28	0.61	1.16	1.99	2.35	2.50	3.35
Bootstrapped p-value	0.10	0.15	**0.04**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.98	0.88	0.97	0.94	0.81	0.14	0.26	0.37	0.25
Parametric p-value	**0.00**	**0.00**	**0.01**	**0.00**	**0.02**	0.08	0.13	0.24	0.38	0.46	0.39	0.27	0.12	**0.02**	**0.01**	**0.01**	**0.00**
Panel C: Average Block Length of $T^{1/4}$ Monthly Returns																	
t-alpha	-4.05	-2.98	-2.92	-2.87	-2.07	-1.41	-1.12	-0.70	-0.30	0.10	0.28	0.61	1.16	1.99	2.35	2.50	3.35
Bootstrapped p-value	0.13	0.18	**0.04**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.98	0.88	0.97	0.95	0.81	0.12	0.24	0.36	0.28
Parametric p-value	**0.00**	**0.00**	**0.01**	**0.00**	**0.02**	0.08	0.13	0.24	0.38	0.46	0.39	0.27	0.12	**0.02**	**0.01**	**0.01**	**0.00**
Panel D: Average Block Length determined by method of Politis and White (2004)																	
t-alpha	-4.05	-2.98	-2.92	-2.87	-2.07	-1.41	-1.12	-0.70	-0.30	0.10	0.28	0.61	1.16	1.99	2.35	2.50	3.35
Bootstrapped p-value	0.12	0.16	0.06	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.97	0.87	0.98	0.95	0.81	0.12	0.26	0.40	0.27
Parametric p-value	**0.00**	**0.00**	**0.01**	**0.00**	**0.02**	0.08	0.13	0.24	0.38	0.46	0.39	0.27	0.12	**0.02**	**0.01**	**0.01**	**0.00**

Table E.II
Residual and Factorial Resampling

This table reports the results for the cross-section of the Norwegian mutual fund's performance measure for the whole sample period 1983-2015 adopting the stationary bootstrap with different block lengths. Panel A shows our result of our baseline bootstap. Panel B shows results Residual and Factorial Resampling. The first and second row reports the OLS estimate of alphas and the cross-sectionally bootstrapped p-values of the t-statistic of alpha. The third row reports Parametric p-values of the t-statistic, which is based on standard critical values of t-statistic. The t-statistics of alpha are based on hetroskedasticity- and autocorrelation-consistent standard errors. The cross-sectionally bootstrapped p-value is based on the distribution of the best (worst) funds in 10,000 bootstrap resamples.

	Bottom	2nd	3rd	Bottom 5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	Top 5%	3rd.	2nd.	Top
Panel A: Our Baseline Bootstrap																	
t-alpha	-4.05	-2.98	-2.92	-2.87	-2.07	-1.41	-1.12	-0.70	-0.30	0.10	0.28	0.61	1.16	1.99	2.35	2.50	3.35
Bootstrapped p-value	0.11	0.15	**0.04**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.98	0.88	0.97	0.95	0.81	0.13	0.25	0.38	0.27
Parametric p-value	**0.00**	**0.00**	**0.01**	**0.00**	**0.02**	0.08	0.13	0.24	0.38	0.46	0.39	0.27	0.12	**0.02**	**0.01**	**0.01**	**0.00**
Panel B: Residual and Factorial Resampling																	
t-alpha	-4.05	-2.98	-2.92	-2.87	-2.07	-1.41	-1.12	-0.70	-0.30	0.10	0.28	0.61	1.16	1.99	2.35	2.50	3.35
Bootstrapped p-value	0.13	0.17	**0.04**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.98	0.89	0.97	0.96	0.81	0.12	0.24	0.37	0.26
Parametric p-value	**0.00**	**0.00**	**0.01**	**0.00**	**0.02**	0.08	0.13	0.24	0.38	0.46	0.39	0.27	0.12	**0.02**	**0.01**	**0.01**	**0.00**

Table E.III
Cross-sectional Bootstrap and Joint Resampling of Fund Returns and Factor Returns

This table reports the results for the cross-section of the Norwegian mutual fund's performance measure for the whole sample period 1983-2015 adopting the stationary bootstrap with different block lengths. Panel A shows results baseline bootstrap. Panel B reports the result of test for cross-section correlation in residuals. Panel C reports the result of the Joint Resampling of Fund Returns and Factor Returns procedure. The first and second row reports the OLS estimate of alphas and the cross-sectionally bootstrapped p-values of the t-statistic of alpha. The third row reports Parametric p-values of the t-statistic, which is based on standard critical values of t-statistic. The t-statistics of alpha are based on hetroskedasticity- and autocorrelation-consistent standard errors. The cross-sectionally bootstrapped p-value is based on the distribution of the best (worst) funds in 10,000 bootstrap resamples.

	Bottom	2nd	3rd	4th	5th	6th	7th	8th	9th	Bottom 10th	Top 10th	9th	8th	7th	6th	5th	4th	3rd.	2nd.	Top
Panel A: Our Baseline Bootstrap																				
t-alpha	-4.05	-2.98	-2.92	-2.91	-2.87	-2.81	-2.42	-2.41	-2.10	-2.07	1.16	1.40	1.40	1.43	1.94	1.99	2.08	2.35	2.50	3.35
Bootstrapped p-value	0.11	0.15	**0.04**	**0.01**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**	0.81	0.48	0.60	0.69	0.08	0.13	0.32	0.25	0.38	0.27
Parametric p-value	**0.00**	**0.00**	**0.01**	**0.00**	**0.00**	**0.00**	**0.01**	**0.01**	**0.02**	**0.02**	0.12	0.09	0.08	0.08	0.03	**0.02**	**0.02**	**0.01**	**0.01**	**0.00**
Panel B: Cross-Sectional Bootstrap																				
t-alpha	-4.05	-2.98	-2.92	-2.91	-2.87	-2.81	-2.42	-2.41	-2.10	-2.07	1.16	1.40	1.40	1.43	1.94	1.99	2.08	2.35	2.50	3.35
Bootstrapped p-value	0.13	0.18	0.08	**0.04**	**0.02**	**0.02**	**0.05**	**0.04**	0.09	0.07	0.57	0.40	0.46	0.49	0.16	0.20	0.30	0.25	0.33	0.27
Parametric p-value	**0.00**	**0.00**	**0.01**	**0.00**	**0.00**	**0.00**	**0.01**	**0.01**	**0.02**	**0.02**	0.12	0.09	0.08	0.08	0.03	**0.02**	**0.02**	**0.01**	**0.01**	**0.00**
Panel C: Joint Resampling of Fund Returns and Factor Returns																				
t-alpha	-4.05	-2.98	-2.92	-2.91	-2.87	-2.81	-2.42	-2.41	-2.10	-2.07	1.16	1.40	1.40	1.43	1.94	1.99	2.08	2.35	2.50	3.35
Bootstrapped p-value	0.16	0.19	0.10	**0.04**	**0.03**	**0.02**	0.06	**0.04**	0.10	0.08	0.53	0.38	0.43	0.46	0.16	0.19	0.29	0.25	0.33	0.30
Parametric p-value	**0.00**	**0.00**	**0.01**	**0.00**	**0.00**	**0.00**	**0.01**	**0.01**	**0.02**	**0.02**	0.12	0.09	0.08	0.08	**0.03**	**0.02**	**0.02**	**0.01**	**0.01**	**0.00**

Table E.V
Sensitive Analysis of the Cross Section of Mutual Fund Alpha to Minimum Number of Observations

This table reports the results for the cross-section of the Norwegian mutual fund's performance measure for the whole sample period 1983-2015 with different minimum return requirements. Panel A shows results with 12 month minimum observation requirement. Panel B, Panel C and Panel D show is the results with 24, 36 and 60 month minimum observation requirement. The first and second row reports the OLS estimate of alphas and the cross-sectionally bootstrapped p-values of the alpha. The third row reports Parametric p-values of the t-statistic, which is based on standard critical values of t-statistic. The t-statistics of alpha are based on hetroskedasticity- and autocorrelation-consistent standard errors. The cross-sectionally bootstrapped p-value is based on the distribution of the best (worst) funds in 10,000 bootstrap resamples.

	Bottom	2nd	3rd	Bottom 5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	Top 5%	3rd.	2nd.	Top
Panel A: The whole sample																	
t-alpha	-4.05	-2.98	-2.92	-2.87	-2.07	-1.41	-1.12	-0.70	-0.30	0.10	0.28	0.61	1.16	1.99	2.35	2.50	3.35
Bootstrapped p-value	0.11	0.15	**0.04**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.98	0.88	0.97	0.95	0.81	0.13	0.25	0.38	0.27
Parametric p-value	**0.00**	**0.00**	**0.01**	**0.00**	**0.02**	0.08	0.13	0.24	0.38	0.46	0.39	0.27	0.12	**0.02**	**0.01**	**0.01**	**0.00**
Panel B: Minimum of 24 Observations Per Fund																	
t-alpha	-4.05	-2.98	-2.91	-2.81	-2.00	-1.41	-1.12	-0.67	-0.29	0.10	0.26	0.61	1.16	1.94	2.08	2.35	3.35
Bootstrapped p-value	0.07	0.11	**0.03**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	0.98	0.86	0.97	0.95	0.84	0.12	0.48	0.45	0.18
Parametric p-value	**0.00**	**0.00**	**0.00**	**0.00**	**0.02**	0.08	0.13	0.25	0.39	0.46	0.40	0.27	0.12	**0.03**	**0.02**	**0.01**	**0.00**
Panel C: Minimum of 36 Observations Per Fund																	
t-alpha	-4.05	-2.98	-2.91	-2.87	-2.00	-1.41	-1.12	-0.70	-0.31	-0.07	0.25	0.52	0.94	1.94	2.06	2.08	2.35
Bootstrapped p-value	0.06	0.09	**0.02**	**0.00**	**0.01**	**0.00**	**0.00**	**0.00**	**0.98**	0.99	0.98	0.98	0.98	0.19	0.42	0.67	0.73
Parametric p-value	**0.00**	**0.00**	**0.00**	**0.00**	**0.02**	0.08	0.13	0.24	0.38	0.47	0.40	0.30	0.17	**0.03**	**0.02**	**0.02**	**0.01**
Panel D: Minimum of 60 Observations Per Fund																	
t-alpha	-4.05	-2.98	-2.91	-2.87	-2.07	-1.50	-1.12	-0.72	-0.29	-0.07	0.28	0.61	1.02	1.94	2.06	2.08	2.35
Bootstrapped p-value	**0.02**	0.06	**0.01**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.97	0.98	0.93	0.92	0.91	0.10	0.28	0.54	0.62
Parametric p-value	**0.00**	**0.00**	**0.00**	**0.00**	**0.02**	**0.07**	0.13	0.24	0.39	0.47	0.39	0.27	0.15	**0.03**	**0.02**	**0.02**	**0.01**

Reflection note.

In this paper, we looked at the performance of Norwegian mutual fund managers. When evaluating mutual fund performance there are two key topics. The first topic is whether fund managers are able to create abnormal risk adjusted returns, net of costs. According to the efficient market hypothesis of Fama (1970), it is not possible to create abnormal risk adjusted returns or beat the market. If one does, it would be purely by chance because market prices reflect all available information. Since prices reflect all available information it is impossible for investors to identify underpriced stocks. This theory is disputed among professionals who points to individuals like Warren Buffet, who has had massive success in the markets. Some fund also perform well, the question is however if the performance is due to luck or genuine skill. Earlier research suggests managers underperform and are not able to produce positive risk adjusted abnormal returns.

The second topic is whether such abnormal returns persist and for how long. If there is persistence in performance investors would be able to, ex ante, identify which funds would perform well in the future. It would seem however that identifying good performers ex ante is difficult, if not impossible. Earlier research shows little evidence of performance persistence of positive risk-adjusted returns. The research does however indicate persistence in the performance among the poorest performers.

These issues are important for investors because they want to know if actively managed funds are worth their costs. If they are not able to create risk adjusted abnormal returns investors would be better off by investing in a low-cost passive index fund. The topic is important from an economic perspective considering the amount of assets under management in Norwegian mutual funds. Sub-optimal placement of assets would lead to massive alternative costs and efficiency loss. This is a very important issue for both individuals and society. In tests we have done we found that the alternative costs for the funds in our sample amount to 600 million kroner a year. The topic is also important from an academic point of view since evidence of skilled performance and persistence in performance would support a rejection of the efficient market hypothesis.

Our problem statement is “Can Norwegian mutual fund managers pick stocks?”
We want to know which way the evidence points, is the performance of managers due to simple luck or do they truly have skill in stock selection. To be able to answer our problem statement we have two tests. The first one is the bootstrap performance test, the second is the persistence test. The bootstrap method to evaluate the risk-adjusted performance of funds, encapsulated in the alpha coefficient. The persistence test to see if the alphas are continuous over time. In good funds, we want not only positive alpha, we also want it consistently.

Our results tell us that managers do not produce risk-adjusted returns above what one would expect to find if such returns were determined by random chance. Managers do not show significant predictive ability to outperform the market factors and benchmark. Rather than evidence for skill, we find evidence for “bad-skill”. Bad-skill is the antithesis of skill, it is anti-skill. Managers at the negative end of the performance spectrum cannot contribute their poor performance due to bad luck or chance, their poor results must be attributed to their lack skills.

In our persistence test, we find that there is no persistence among the top performers. There are however statistically significant results among the poorest 20% performers of funds. Among them the poor results tend to persist, a further proof of bad skill. Poor performers tend to stay poor performers. So, all in all we do not find results telling us that Norwegian mutual fund managers have good skills. We only find evidence of bad skills, the skills that produce bad results.

Regarding innovation, there has been a lot of talk about new fin-tech the last few years. Fin-tech is short for financial technology. For example, robots (computer programs) who can pick and buy stocks automatically, to achieve positive returns. A “smart Beta” fund is a new type of robot fund emerging which, for example, exploits known market anomalies such as the value factor and size factors. Robots such as these would be cheaper than a human being, therefore such technology could decrease the costs of the fund and possibly give this new margin back to investors. The alpha generated, net of costs, could be bigger with such technology. Also, advanced computer programs could create cheaper index-funds. They could automatically rebalance the portfolio to mimic the market, possibly a lot faster and cheaper than a human being could.

When it comes to ethical challenges, there are some things that could be discussed. Fund products are complicated and one could wonder whether all customers understand the products they are buying. Would investors buy actively managed mutual funds if they knew that fund managers exhibit no skill in stock selection? Investors should know beforehand what they are buying and the risks associated with fund products. Funds themselves should explain the products, but investors themselves have the main responsibility to educate themselves and know what they spend their money on. In one of our tests we find that the possible returns lost due to unskilled managers amounts to 600 million NOK a year. This is quite the alternative cost. However, even if funds fail to provide sufficient alphas, funds provide other services that are valuable, such as diversification and professional management. Even if Norwegian mutual funds fail to provide sufficient alphas, they are still relatively secure products that provide, mostly, positive raw returns. The returns investors get are bigger than the rate they get in a normal bank account.

Regarding internationalization, the markets are by definition very international. However, we might see more consolidation in the sector across national lines. Big actors like Vanguard, might expand and claim a bigger part of the passive index market in other countries.