Mathias Tidemann Krafft

Petter Bringedal

-Master Thesis-

Distinguishing between skill and luck in the returns of Norwegian mutual funds

Hand-in date:

10.08.2016

Supervisor:

Kjell Jørgensen

Campus:

BI Norwegian Business School, Oslo

Programme:

Master of Science in Business, Major in Finance

Examination code and name:

GRA 19003 Master Thesis

Abstract

Throughout this thesis we examine the risk adjusted performance of all actively managed Norwegian equity mutual funds, using a comprehensive dataset free of survivorship bias, spanning the period 1983 to 2015. We utilize a bootstrapping methodology which enables us to distinguish between skill and luck in the cross sectional distribution of mutual fund α estimates. A methodology of injecting alpha in the bootstrapping regressions is pursued to estimate the features of true α (defined as the skill to cover fees). After adjusting for luck, we find evidence that the top 5% of funds exhibit skills to earn 3.3% or more in annual alpha above the fees they charge, whereas the bottom 5% destroy at least 3.7% per year.

Acknowledgements

We would especially like to thank our supervisor, professor *Kjell Jørgensen* at BI Norwegian Business School, for valuable input and support throughout the process of writing this thesis.

Additionally, not in order of importance, we thank: *Lars Qvigstad Sørensen* for providing us with a list of tickers for the funds used in his research; *Bernt Arne Ødegaard* for making available and allowing us to use his constructed factor and interest rate returns, and for providing access to constructed fund returns from the OBI database; *Børsprosjektet NHH* for helping us with acquiring an additional subset of the fund returns used; *Verdipapirfondenes Forening (VFF)* for giving us descriptive statistics for the Norwegian mutual fund market; *Matteo Ottaviani and MathWorks* for allowing us to use MATLAB for free in our research; and finally *Bruno Gerard* for helpful comments.

Table of contents

1. Introduction

Throughout this master thesis, we pursue to analyse the performance of actively managed Norwegian equity mutual funds, listed on Oslo Stock Exchange in the period from 1983 to 2015. The funds included all have a mandate of investing solely in Norwegian equity, which eases the choice of an appropriate benchmark. In the performance evaluation of mutual funds, we expect to identify a cross-sectional distribution alpha ranging from low to high. The main goal of this paper is to disclose whether good performance of some mutual funds can be attributed to skill or if it is most likely just due to luck, and likewise if bad returns are due to a lack of skills or simply due to bad luck.

According to the Norwegian Fund and Asset Management Association *(VFF)* the capital inflows to mutual funds have increased in recent years, and reached an all-time high in the first half of 2015. Many investors perceive mutual funds to be an attractive investment alternative, especially for the time being with historically low interest rates on bank deposits (Figure A1). We hope that our research on actively managed mutual funds could further contribute to the investors' assessment between active and passive management.

A similar study as ours has been conducted by Sørensen (2009). We lengthen the time period covered up until 2015, and believe this to be a contribution, as it covers the post-financial crisis period. Another difference from that of Sørensen, is that we have used a dataset containing a sub-set of the funds, excluding all passively managed funds. Fama and French (2010) also does this "to focus better on the performance of active managers", and we believe it may increase the power of our analysis. We also use an additional methodology of Fama and French (2010) which involves injecting alpha into bootstrapping simulations. To our knowledge, this is the first paper that seeks to estimate the distribution of true α for the Norwegian mutual fund sector using this methodology.

In order to address our research issue, we compare the historical distribution of actual $t(\alpha)$ with 'luck distribution' obtained from 10 000 bootstrap simulations. The simulation samples are drawn from a return population which by construction have a true α equal to zero. This simulated distribution can be interpreted as a luck distribution, where abnormal returns are attributed to luck only. A comparison of the distributions of actual $t(\alpha)$ and simulation distribution enables us to disentangle skill from luck. This approach will hence reveal the potential existence of skills, both positive and negative.

The empirical results show evidence of both inferior and superior fund performance, in the left and right tails of the distribution, respectively. The overall distribution of actual $t(\alpha)$ seem to be shifted to the left of the simulation distribution, suggesting that fund managers as a group do not possess sufficient skills to fully cover for the fees imposed on customers. Furthermore, the data suggest a relationship between performance and shutdowns. For example, the 20 worst performing funds have all ended their operations, whereas the 20 best performers still operate.

The rest of the paper is organised as follows: Section 2 provides background information of the mutual fund industry in general and in Norway, and a review of previous research literature on similar topics. Section 3 provides relevant background theory for our study, as well the hypothesis we pursue to test. Section 4 covers the methodology used, and provides a step-by-step explanation of the statistical approach used. Section 5 provides a description of how all the data, factors and fund returns, was collected and treated. Section 6 gives the empirical results, whereas section 7 concludes.

2. Background and literature

Background

In this section we will provide a brief description of mutual funds in general, as well as of the Norwegian mutual fund market.

Mutual funds in general

A mutual fund is an investment vehicle which pools money from several types of investors and can hence make large purchases in a variety of securities or assets (e.g. stocks, bonds, money market instruments and real-estate) or a combination of these.

The aggregate holding of a fund's investments constitutes its portfolio (SEC). An investor's fraction of shares in the fund represents the proportionate claim on these assets, as well as the income these assets might generate. Investors purchase mutual fund shares directly from the fund, instead of in a secondary market. The price equals the net asset value (NAV) per share, plus potential front-end costs. During the period of ownership in a mutual fund, investors are obliged to pay management fees, and at the time of redemption some mutual funds charge additional back-end costs as well.

Mutual funds can be subdivided into two main types, namely passive and active. A passive, or indexed, mutual fund seeks to replicate the components of a market index, such as for instance Oslo Børs Benchmark Index (OSEBX). This index is rebalanced every half a year, which is associated with relatively low transaction costs. An active mutual fund seeks to reap profits mainly from exploitation of mispricing in the market. Perceptions of mispricing tend to change quite frequently, which in turns lead to frequent trading, hence the term *active* (Sharpe 1991). Consequently, active mutual funds are subject to higher costs originating from transaction and research activity.

Investing in a mutual fund provides several advantages for investors. The arguably most important one is the benefit from diversification. Most mutual funds are restricted by legislation in terms of diversification, and European mutual funds are subject to the *Undertakings for Collective Investment in Transferable Securities* (UCITS). The most common restriction imposed by this directive is the so-called 5/10/40-rule. This rule restricts a fund to invest a maximum of 10% of its net assets in one single security, and investments greater than 5% in a single security issuer, must not exceed 40% of the whole portfolio-value when these investments are added up. Such legislation provides diversification which can be difficult to obtain otherwise for small investors, especially if the amount to invest is modest.

Furthermore, an additional advantage of investing in a mutual fund is the access to professional money managers and researchers. Many small investors do neither have the time nor knowledge to select and monitor the performance of the available securities. Hence, many find it tempting to outsource this task to professionals.

On the flip-side, mutual funds also have its disadvantages. According to SEC, the greatest one might be the need to pay management fees and costs, regardless of fund performance. Moreover, another disadvantage could be the lack of control. As mentioned, an advantage with mutual funds is that investors let professionals manage their funds, but at the same time it could be difficult or even impossible for an investor to influence in decisions of which securities the fund should acquire or sell. In addition, it is typically difficult to ascertain the exact composition of the fund's portfolio at any given time, and hence the value of ones ownership or level of risk. Unlike an investment in the stock market, the mutual fund investors can not know the exact value of their investment from second to second, it is hence embedded inertia in the price. When an investor decides to sell his shares, he might have to wait for several hours after the order is placed to obtain the exact value.

The Norwegian mutual fund market

According to Gjerde and Sættem (1991) there was only one mutual fund listed on Oslo Stock Exchange prior to 1982. In the same year, a scheme with tax rebate concerning mutual fund investments was introduced in Norway, which subsequently led to a sharp increase in the number of mutual funds. In 1990 the total market value of mutual funds in Norway amounted to 5.5% of the total market value of all stocks listed on the Oslo Stock Exchange (Gjerde and Sættem 1991). Figure A2 in the appendix illustrates the evolution of the total number of actively managed Norwegian mutual funds. The figure shows an almost monotonic increase from 1983, reaching a maximum of 66 funds operating simultaneously in 2002. In 2015 there was a total of 78 Norwegian mutual funds, and we estimate that 57 of these were actively managed.

Table A1 in the appendix shows descriptive statistics for the Norwegian mutual fund market in the period from 1995 to 2015. The most remarkable observation from this table is in our opinion the third left column. This column displays the proportion of mutual funds which solely invests in Norwegian shares, relative to the total Norwegian fund market. In 1995, this fraction amounted to approximately 92 per cent. Subsequently, it has been in a steady decline, coinciding with increased globalization of financial markets easing the access to foreign stock markets. The fraction reached a minimum in 2008 at 19.7 per cent, and now seems to have stabilized at the lower half of the twenties, in terms of percent. This is quite close to the guidelines from VFF (2015). They recommend the average private investor to hold as much as 75% of savings in foreign shares, in order to secure diversification. Assets under management have had a yearly average growth rate of 9.2% which includes net inflow. Adjusted for net inflows, the Norwegian mutual fund industry had an average yearly organic growth in assets under management amounting to 8.5%. The total number of mutual funds with a Norwegian mandate has been quite stable over the timespan, averaging at 73. The number of funds included in our dataset averages at 56 per year, which is smaller due to the exclusion of passive index funds.

Literature Review

Literature on how to measure the performance of mutual funds has existed for decades, and academia has introduced a large variety of suggestions for measures, models and procedure for this purpose. Estimating whether the good/bad performance of the mutual fund is due to skill/incompetence or luck/misfortune, however, is a relatively new topic. Kosowski et al. (2006) were the first to utilize a bootstrap approach in order to distinguish skill from luck in mutual fund performance. They analysed the U.S. domestic equity mutual fund industry from 1975 to 2002, in order to investigate if mutual fund “stars” do possess a stock-picking skill. Their findings reveal that most fund managers do not provide good enough returns to more than cover their costs, but that a few fund managers actually do have a superior true alpha.

A relatively similar approach was put forward by Cuthbertson, Nitzsche and O'Sullivan (2008) on UK equity mutual funds. This study revealed stock picking ability in quite a small group among the best performing funds. In the other end of the scale, the study disclosed that persistent poor performance could not be attributed to bad luck only, but to some degree of “bad skill” as well.

Fama and French (2010) conducted a study of “luck versus skill” among mutual fund managers in the U.S. from 1984 – 2006, using a slightly modified bootstrap procedure than the one presented by Kosowski et al. (2006). The main difference in approach is that Fama and French jointly sample fund returns, compared to Kosowski et al. who simulate for each fund independently. Fama and French conducted the analysis on both gross and net returns. Using gross returns, they revealed evidence of both inferior and superior performance (nonzero true α), whereas the results using net returns were more devastating showing that only a few funds managed to produce expected returns sufficient to cover costs. Hence, this study finds less evidence of skill than Kosowski et al. (2006).

Other studies have also investigated fund manager performance in the Norwegian market. One study in especially in close relation to ours is "Mutual Fund Performance at the Oslo Stock Exchange" by Sørensen (2009). This study distinguishes luck from skill, based on the previously mentioned methodology presented by Fama and French (their article was published in 2010, but they presented the methodology and much of the results in a working paper previous to this), finding no significant evidence of superior performance among listed Norwegian mutual funds in the period of 1982 – 2008. Furthermore, Sørensen points out that there is no persistence in the performance of either top or bottom fund managers. In our research, we update the results of Sørensen by extending the period up until 2015. Additionally, we differentiate from his study by excluding passive funds from our dataset, in order to make a better analysis of active management. Furthermore, we conduct an analysis of the features of true alpha by simulating with injected alpha, in line with the methodology of Fama and French (2010).

3. Theory / Hypothesis

In the following section we will explain and discuss some of the most relevant theories, as well as the implications they make, with respect to our research.

The Efficient Market Hypothesis

The Efficient Market Hypothesis (EMH) was first introduced by Eugene Fama (1970). Jensen (1978) claims that this is the proposition in economics with most solid empirical evidence. According to Fama (1970), in an efficient market the prices will always “fully reflect” all available information. Hence, thorough research and analysis aiming to reveal mispriced securities or active investment strategies in general, will be in vain if markets are fully efficient. Fama presents three different levels of efficiency, depending on the degree of information incorporated in market prices.

First, the *weak form efficiency* states that prevailing share prices reflects all available information with respect to historical trading data, such as prices and volumes. In presence of weak form efficiency, technical analysis seeking to reveal price patterns will be a waste, as future prices are completely independent of past developments.

Next, *semi-strong form efficiency* comprises the weak form, and in addition that prices reflect all relevant public information. If the requirements of semi-strong form efficiency are fulfilled, neither technical nor fundamental analyses based on public information will enable traders to outperform the market (Dimson and Mussavian 1998).

Finally, under *strong form efficiency*, share prices reflect all information regarding the company. Unlike semi-strong form, the strong form does not pose a restriction that the information is publicly available, thus accounting for information possessed by insiders. This is the most extreme form of EMH, and probably more hypothetical than realistic. In developed financial markets, it is common to observe large share price movements in response to announcements of unexpected information regarding a specific company, which violates the strong form hypothesis.

Of the three abovementioned forms, the semi-strong form is the most likely to apply for stock markets such as the Norwegian (Koller, Goedhart and Wessels 2010).

Identification of investors or fund managers who persistently achieves to outperform the market, does not serve as evidence against the existence of efficient markets. First, as taught in basic statistic courses, failing to reject a hypothesis does not imply acceptance of the same hypothesis. Additionally, tests of market efficiency will face a joint hypothesis problem since one will test if the market is efficient given a specific asset-pricing model, i.e. a simultaneous test that the market is efficient and the model is correct. Hence, disclosing deviations from EMH could rather be viewed as indication that models used to predict equilibrium return are flawed (Summers 1986).

Equilibrium Accounting

The participants in the market can be separated into two subgroups, consisting of active and passive investors. A passive investor holds a portfolio consisting of all the shares in the market (Sharpe 1991). Each security in a passive portfolio is held in the same fraction as this particular security's part in the market as a whole. Thus, if a security constitutes one per cent of the total value of the market, a passive investor will invest one per cent of his or her funds in this particular security. Passive investing can be seen as a buy and hold strategy, where rebalancing is only needed after particular events such as initial or seasoned public offerings, share buybacks and changes is the index composition.

Active management is based on perceptions of under-priced shares, and not what fraction an individual share constitutes in a given index. Active fund managers attempt to outperform the market. Active investors' assumptions of mispricing tend to change quite frequently, leading to a need for active rebalancing of the holdings (ibid).

For any given period, the market return will equal the value weighted return from all securities that comprise the market. This will equal the gross return acquired by truly passive investors, gross of fees and transaction costs. Following from the previously mentioned condition that passive and active investors constitute the whole market, the market return is a weighted average of the returns from the two subgroups. As Sharpe (1991) points out, this implies that the average return achieved by active investors in the same period must equal the return gained by passive investors, referred to as *equilibrium accounting*.

As mentioned, active investing requires buying and selling securities more frequently compared to passive investing. This activity generates more transaction costs. Additionally, active fund managers charge higher fees to fund their research to find mispriced securities. For example, Norwegian equity mutual funds charge an average of 1.4 per cent of the customers' holdings in management fees per year (Strøm 2014), and fees can be much higher, especially for alternative investments such as hedge funds, funds of funds and private equity. For passive managers these fees are usually much lower, reflecting the simple buy and hold strategy and the smaller proportion of resources required to operate a passive fund.

From this, it follows that on average, the return net of fees and costs provided to investors from active investment management must be lower than the return provided by passive investments. Hence, active investors participate in a negative sum game, and the ones who receive excess returns, must do so at the expense of other active investors (Fama and French 2010). This theory does not exclude the possibility that some managers are able to persistently *beat the market,* but they do so at the expense of other active investors.

Hypothesis

In the remainder of this thesis, we will investigate the risk adjusted performance of Norwegian Mutual Funds and distinguish whether the performance is attributable to skill or luck. We will do this using an overall economic hypothesis as follows:

> *H_0: Managers of mutual funds do not possess skill (positive or negative), and the cross sectional distribution of mutual fund alphas is due to luck only*
>
> *H_1: Managers of mutual funds are endowed with different levels of skill (positive or negative), and the cross sectional distribution of mutual fund alphas is due to a combination of skill and luck.*

Although we do believe the skill of mutual fund managers to be a factor influencing mutual fund returns, we expect luck to be the major determinant of mutual fund alpha, and do not expect to find significant evidence of positive/negative skill in our data.

4. Methodology

In our tests we will utilize the risk adjusted performance measure referred to as α, and its t-statistic t(α) to measure the performance of Norwegian mutual funds. This is combined with a bootstrap procedure, which allows us to compare actual fund performance with a 'luck distribution'. In this section we first look at models and the regression framework used, before explaining the bootstrapping procedure in detail. Finally, we briefly explain how the simulated results of the bootstrap can be compared to actual results, in order to draw inference on mutual fund performance.

Model specification and regression framework

The models we consider in our regressions are unconditional factor models, of which on a general form, can be specified as follows:

$$r_{i,t}^{e} = r_{i,t} - r_{f,t} = a_i + \sum_{j=1}^{K} \beta_{i,j} * f_{j,t} + \varepsilon_{i,t} \qquad (1)$$

where $r_{i,t}^{e} = r_{i,t} - r_{f,t}$ is the asset (e.g. mutual fund) risk premium, $r_{i,t}$ is the return of an asset with index number i, between time $t-1$ and t, $r_{f,t}$ is the risk free rate, a_i is the asset excess return (or mispricing), K is the numbers of risk factors, $\beta_{i,j}$ is asset i's loading to risk factor j, $f_{j,t}$ is the value of risk factor j at time t and $\varepsilon_{i,t}$ are the residuals.

The simplest of such forms includes only one factor, namely the model developed by Jensen (1968), based on the CAPM. Here, asset risk premiums $(r_{i,t} - r_{f,t})$ are linear functions of the market risk premium $(r_{m,t} - r_{f,t})$ and the systematic risk of the asset (β_i), where $r_{m,t}$ is the market return between time $t-1$ and t, as follows:

$$r_{i,t}^{e} = r_{i,t} - r_{f,t} = a_i + \beta_i * (r_{m,t} - r_{f,t}) + \varepsilon_{i,t} \qquad (2)$$

The a_i in this equation is referred to as the *Jensen's Alpha* of the asset (in this case mutual fund), and is a commonly used performance measure. The model is extendable in numerous ways, mainly by including additional factors as independent variables, but also for example by allowing for time-varying coefficients. As a performance measure, the constant term remains the main focus. When the expected value of the α_i-term is zero, superior/inferior stock-picking ability possessed by an individual fund manager will be reflected in a statistically significant nonzero α-value.

In our brief initial tests on an EW portfolio (see Table 2 in section 6) we considered five different models combining the factors $r_{i,t}^e$, SMB_t, HML_t, UMD_t and LIQ_t in various ways, including but not limited to three common model specifications; the Jensen (1968) 1-factor model, the Fama and French (1993) 3-factor, and Fama and French's variation of the Carhart (1997) 4-factor model. Each of the different factors are described in detail in the 'Factor construction' part under section 5.

We estimate the models using basic OLS regression. Standard errors are corrected for autocorrelation and heteroscedasticity using the Newey and West (1987) procedure, as signs of autocorrelation and heteroscedasticity are evident in some of our data. For consistency, the procedure is used in all regressions.

Model selection

We focus on, report results for and discuss two different model specifications in the remainder of the thesis. The first is the widely used 3-factor model developed by Fama and French (1993), specified as:

$$r_{i,t}^e = r_{i,t} - r_{f,t} = a_i + \beta_{i,r_m^e} * r_{m,t}^e + \beta_{i,SMB} * SMB_t + \beta_{i,HML} * HML_t + \varepsilon_{i,t} \quad (3)$$

Here, the excess return on the market portfolio $r_{m,t}^e$, the returns of a size portfolio SMB_t, and the return of a value portfolio HML_t, are the three explanatory variables. This model is the industry norm, and allows for comparisons to important research on the same topic. It is the main model used in the paper 'Luck versus Skill in the Cross-Section of Mutual Fund Returns' by Fama and French (2010), which first introduced the form of the bootstrap procedures which we utilize. Sørensen (2009) also uses it in the paper 'Mutual Fund Performance at the Oslo Stock Exchange', which is the paper we mainly seek to produce an extension of. He utilizes many of the same procedures as we do, and we have constructed our dataset in a similar manner as him. The main difference is that we have extended the end date of the period by 7 years; he has data for the period 1982-2008, while we use data from 1983 up to and including 2015. In addition to this, we have excluded passive funds from our dataset, and focused solely on actively managed funds (although some funds may be 'closet index funds', i.e. they are reported as active, but are in fact index-tracking and more like a passive fund).

The second model we use is motivated by the findings of Næs, Skjeltorp and Ødegaard (2009). The main results from their analysis is that "the return at the OSE can be explained reasonably well by a multi-factor model consisting of the market index, a size index, and a liquidity index", and they exclude other factors such as value- and momentum indices in explaining market returns. We find it interesting to compare results using this model to the results using the classic Fama-French 3-factor model. The second model specification is:

$$r_{i,t}^{e} = r_{i,t} - r_{f,t} = a_i + \beta_{i,r_m^e} * r_{m,t}^{e} + \beta_{i,SMB} * SMB_t + \beta_{i,LIQ} * LIQ_t + \varepsilon_{i,t} \qquad (4)$$

where LIQ_t is the return of a liquidity portfolio LIQ_t. In the section 'Extension of the bootstrap procedure: Injecting alpha' where we seek to estimate the distribution of true alpha (i.e. alpha excluding luck, or skill above fees), we do not have comparable results for the Norwegian Mutual Fund market (we believe we are the first to utilize this procedure on this data), and choose to use the second model specification (equation 4), under the assumption that the results of Næs, Skjeltorp and Ødegaard (2009) still hold.

The bootstrap procedure: five steps

As mentioned, we will utilize a bootstrapping procedure for distinguishing skill from luck in the performance of Norwegian mutual funds. We will follow a procedure in line with that of Fama and French (2010). It is a modification to the bootstrapping procedure introduced by Kosowski et al. (2006), who were the first to use bootstrapping for the purpose of distinguishing skill from luck in the performance of mutual funds. It implies simultaneously simulating residuals and factor returns for all funds, instead of only residuals for just one fund at a time, in order to preserve the cross-correlation of fund returns (Fama and French 2010). The procedure can be broken down in five steps, each presented below.

The first step of the procedure is to estimate regular benchmark regression models, one regression for each fund. The observed historical returns of each individual fund are regressed against the returns of a specified set of risk factors over the corresponding period which the fund is present in the data. For each fund, we save the estimate of actual alpha, a_i, and its corresponding t-statistic, $t(a_i)$, the estimates of coefficients for risk factor exposure and a vector for residuals ε_i.

The second step is to produce a set of S number of simulation runs (e.g. 10 000 runs). The set of simulation runs is always the same, for each and every fund, and irrespective of which model specification we use. This is, as mentioned above, in order to preserve the cross-correlation of fund returns and comparability between models. The procedure used to produce a simulation run is described below (the complete algorithms used to produce all runs can be found in the appendix). First, we draw a $(T\ x\ 1)$-dimension vector from the uniform distribution, where T is equal to the number of periods in the data set ($T = 396$ in our main model, $12\ months * 33\ years$, from the first observation of the initial funds in January 1983, to the last data point in December 2015). Note that no fund will have return histories for the whole period, as all funds initiated later than the first and/or terminated before the last observation in the period. We then multiply the matrix with the scalar T and round up to nearest integer. This yields a vector of time indices, randomly drawn with equal probability and with replacement, from the set of available points in time:

$$\tilde{T}_s = roundup(T * \{U_t(0,1)\}_{t=1}^{T}) \tag{5}$$

In the third step, we use the simulations runs, which consist of simulated time indices, to construct new series of alpha free fund returns and new series of risk factor returns, as follows: For each $\tilde{T}_s$, we construct a new series of risk factor returns, $F(\widetilde{T_s})$, with dimensions $(T\ x\ K)$, where K is the number of factors. The returns are "picked" from the populations according to the drawn time indices of the simulation run. The same is done to construct a matrix of $(T\ x\ N)$ residuals $E(\widetilde{T_s})$, where N is the number funds in our sample. Each $(T\ x\ 1)$ column-vector $\varepsilon_i(\widetilde{T_s})$, now consist of drawn residuals from the original regression model of one fund from the first step. However, as funds are not present the whole period (all funds have less than 396 observations), some draws yield a blank (represented by NaN in our data), and the number of returns for one fund varies between simulations. We use a cut-off off at least 15 returns for a simulation to be valid and included in the results. Then, using the saved coefficients and simulated time series of risk factors and residuals, we construct new fund excess returns, but leave out alpha (a_i), so that all returns have zero alpha by construction:

$$r^e_{i,\tilde{t}_s} = \sum_{j=1}^{K} \beta_{i,j} * f_{j,\tilde{t}_s} + \varepsilon_{i,\tilde{t}_s} \quad (6)$$

The fourth step is to run regressions in the same manner as we did with actual fund returns in the first step, but now with the constructed excess fund returns of the simulation as dependent variables, and the corresponding set of risk factor returns as explanatory variables.

Finally, we use different $\tilde{T}_s$ from step two and repeat the processes in step three and four S times to produce a set of S simulated alpha-estimates with corresponding t-statistics for each of the N funds.

The fifth step is to in various ways compare estimated $t(a)$ based on actual historical returns with the estimated $t(a)$ from all the simulations. In accordance with previous research, we focus our analysis on the $t(a)$-estimates because this incorporates the precision with which the α is measured (Fama and French 2010). For completeness, we make the same calculations for the alphas as well, and report results in tables together with those for $t(a)$. A detailed explanation on how the numbers are compared is presented below.

Comparing regression results on historical vs. simulated returns

For both the benchmark regression performed in step one and for each individual simulation run, we separately sort the N fund alphas and t-stats. Different ranks/percentiles of actual alphas/t-stats are then compared with the corresponding ranks/percentiles of the S simulations. For example, we compare the $t(a)$ of the fund that performed the best in our benchmark regression, with the best $t(a)$ of each of the

S simulations, where alpha is completely due to luck. Similarly, we can compare how the worst, the 5th best or the 10th percentile-fund actually performed with how well they perform in each of the simulations.

A simple and intuitive way of comparing the numbers is to look at the average over all simulations (for a given rank/percentile) of $t(a)$ vs. the actual estimates (for the corresponding rank/percentile). This is done in a qualitative manner in order to gain some insight and perspective on the simulations and the relative levels of performance of the funds. The average of the simulations provides a measure of how well the given rank/percentile should perform gross of fees when there is no presence of skill; i.e. all performance is attributable to luck only.

Another, more useful and powerful comparison, is to measure the fraction of times which the simulated $t(a)$ is either larger or smaller than the actual number for the given rank/percentile. The fractions can be interpreted as p-values, and they allow us to more formally measure whether the actual performance is extreme compared to (not just different from), the performance in the simulations. For example, if a low fraction of simulation runs produces $t(a)$s in the left tail lower than the estimates from actual fund returns (or equivalently that a high fraction of simulations produces alphas/t-stats higher than the actual results), we can infer that some managers lack the skill to cover fees and trading costs. The lower this fraction is, the more confident we are in the existence of negative skill. The opposite is true for the right tail; if a high fraction of simulation runs produces alphas/t-stats lower than the actual estimates we infer that some managers are more than skilled enough to cover fees and trading cost.

Extension of the bootstrap procedure: Injecting alpha

In this section, we continue to follow the methodology of Fama and French (2010). Here, the full bootstrap simulation procedure above is repeated several times, but this time with random values of α injected into the new constructed fund returns of step three, varying the standard deviation of α for each repetition. The results from these repetitions are then compared to the actual historical results in order to estimate the tail distribution of true α. As Fama and French (2010) point out: the new simulated numbers allow us to examine (i) which levels of α is necessary to reproduce the t(α) estimates for actual fund returns, and (ii) levels of α too extreme to be consistent with t(α) estimates for actual fund returns.

The procedure for bootstrapping with injecting alpha is as follows. Overall, the same 5-step procedure outlined in the previous sections is still used: alpha is injected by altering the third step of the procedure, while the other steps are left unchanged. In three, instead of leaving out alpha and constructing a 'luck distribution' as we did previously, equation (6) now becomes:

$$r^e_{i,\tilde{t}_s} = \frac{\alpha_{i,s}}{12} * s_i + \sum_{j=1}^{K} \beta_{i,j} * f_{j,\tilde{t}_s} + \varepsilon_{i,\tilde{t}_s} \qquad (7)$$

where $\alpha_{i,s}$ is the annual alpha, a random number drawn from the normal distribution, individually and independently drawn for each fund and for each simulation (and constant over time), with mean equal to zero and standard deviation equal to σ (returns are per month and σ is the average injected annual standard deviation of alpha). And, s_i is a scalar adjusting for the individual funds different levels of diversification, defined as follows:

$$s_i = \frac{SE(\varepsilon_i)}{\left(\frac{\sum_{i=1}^{N} SE(\varepsilon_i)}{N}\right)} \qquad (8)$$

where $SE(\varepsilon_i)$ is the standard error of the residuals of the initial benchmark model regression for fund i and N is the number of funds included, such that the denominator becomes the average standard error of the residuals and s_i becomes a scalar that decrease with diversification. The term s_i is included because: "it seems reasonable that more diversified funds have less leeway to generate true α" (Fama and French 2010). The implication is that relatively more diversified funds with low standard errors of residuals compared to the average and lower s_i (than 1), will have their absolute values of total injected alpha effectively scaled down, while less diversified funds with high standard errors of residuals compared to the average and higher s_i (than 1), will have their absolute values of total injected alpha effectively scaled up.

The overall 5-step bootstrap procedure, with adjustments to step three made according to the above, is repeated several times, for different values of σ (average injected annual standard deviation of alpha).

Two different techniques are then used to find (i) the levels of α necessary to reproduce the t(α) estimates for actual fund returns, or 'likely levels of performance'. The first technique (see grey markings in the left panel of Table 5 under section 6 for illustration) comprises looking for the value of σ which gives average ranks/percentiles of simulations equal to those of actual fund returns. The second technique (see grey markings in the right panel of Table 5 under section 6) comprises looking for the value of σ which gives median ranks/percentiles of simulations equal to those of actual fund returns (we look for the value of σ where 50% of simulated ranks/percentiles are smaller than corresponding actual ranks/percentiles).

In order to find (ii) levels of α too extreme to be consistent with t(α) estimates for actual fund returns, or 'unlikely levels of performance', a technique similar to the second technique above is used. We construct the equivalent of a confidence interval, where we accept a 20% likelihood of setting a lower band that is too high and a 20% likelihood of setting an upper band that is too low (Fama and French (2010). We use the same thresholds as Fama and French (2010), which they consider to: "imply a narrower range than we would have with standard significance levels, but they are reasonable if our goal is to provide perspective on likely values of σ." In the left tail, the value of σ where 20% of simulated ranks/percentiles are smaller than corresponding actual ranks/percentiles marks the *lower* bound, while the value of σ where 80% of simulated ranks/percentiles are smaller than corresponding actual ranks/percentiles marks the *upper* bound. The converse is true for the right tail. Here, the value of σ where 20% of simulated ranks/percentiles are *smaller* than corresponding actual ranks/percentiles marks the *upper* bound, while the value of σ where 80% of simulated ranks/percentiles are smaller than corresponding actual ranks/percentiles marks the *lower* bound.

One very important difference in our analysis compared to that of Fama and French (2010) is that all our actual fund returns are net of fees, while they have access to gross fund returns. This becomes important in this section, when analysing and trying to estimate the true distribution of alpha. Due to (or in spite of) the nature of the data we have, we make one important simplification; we use the assumption that each fund is endowed with annual *net* return alpha drawn from a normal distribution with mean equal to zero and standard deviation per year equal to σ, whereas Fama and French (2010) make the same assumption for annual *gross* return alpha. This is a caveat, as we do not actually expect net return alpha to be symmetric around zero. A more reasonable assumption would have been that each fund is endowed with annual *net* return alpha drawn from a normal distribution with mean equal to minus its annual fee and standard deviation per year equal to σ. However, we do not have access to the level of fees for each fund (or how this has varied over time). And if we did, we would be able to construct gross fund returns and estimate corresponding actual gross α and t(α), avoiding this issue entirely. Another possibility would be to assume constant fees over time, and for the funds we do not have data on management fees, assume for example a fee equal to the average of the fees for the funds that we do have data, but here also the preferred procedure would be to construct gross fund returns, not to subtract fees in the simulation. In general, throughout the thesis, we have chosen not to construct estimates of gross returns, but rather to work with the precise net returns we have at hand. We do not go away from this in the current section, but are aware of the potential significance of management fees in the results and keep this in mind in our discussions and results.

5. Data

Mutual fund returns

In order to construct an as comprehensive data set as possible, we have searched through several sources of information. As students of BI Norwegian Business School we were granted access to a OBI (*Oslo Børs Informasjon)* database. Through this portal we were able to retrieve names and tickers of the Norwegian mutual funds that were still running in 2015, together with their return series. For defunct funds, however, the same data could only be retrieved individually contingent on first knowing the specific funds' tickers. As one of our main goals has been to extend the research of Lars Qvigstad Sørensen, constructing a dataset free of survivorship bias was paramount. Fortunately, Sørensen was willing to provide us with a list of the funds he used in his research. This included names and tickers of all Norwegian mutual funds that had existed between 1982 up until 2008. The remaining funds which we would now potentially miss, would be funds which initiated after 2008 and shut down before 2015. Through a request to Oslo Stock Exchange we were informed that there was one such mutual fund, namely *Storebrand Norge Institusjon*, which operated from 2010 until 2014. Through the above steps, and with the inclusion of this last fund, we were able to construct an extensive list containing every Norwegian mutual fund present from 1982 up until 2015. Using this list of tickers, we have retrieved monthly returns for each fund from the OBI database, and constructed the dataset.

We were able to obtain access to a similar database from *Børsprosjektet* at the Norwegian School of Economics (*NHH*), also based on data from OBI. This contained return histories of all Norwegian funds, global and domestic, operating over our desired timespan, but here as well retrieving returns was contingent on first knowing funds' tickers or names. By comparing funds individually, we found that all corresponding data points were consistent between the two databases when they existed in both of them, but disclosed minor discrepancies with respect to the starting date and length of time series of some funds, mostly that histories from *Børsprosjektet NHH* initiated somewhat earlier than those from the OBI database. We have consistently used the source that provided the longest return history.

Given our research focus on active management, we pursued with investigation of each individual fund in order to exclude passively managed funds. We withdrew every fund containing any variety of the word *index*, as well as searching up each fund's investment strategy in order to disclose passively managed funds without index in its name or ticker. For a small group of funds that closed several years ago, we were not able to obtain reliable information concerning investment strategy. Hence, we are aware that we might have unintentionally included a few passive funds. Although this may reduce the power of our conclusions slightly, we are confident that a possible wrongful inclusion of a few passive funds will not make a large impact on our analysis. Most of our research is mainly concerned around the tails of the alpha-distribution, and the passive funds would most likely exhibit an alpha close to zero and hence place them in the middle of the distribution.

We were not able to get reliable data for the OSEAX (previously named *Totalindeksen*) before 1983, and thus chose to begin our analysis from January 1983. This means that we excluded one year of returns in 1982 for the two funds that were in operation this early, but this should not have any major impact for our analysis.

Our final data set, ranging over the period from January 1983 to December 2015, contains the return series for 101 actively managed mutual funds, comprising a total of 15 408 observations of monthly returns, corresponding to an average presence for each fund of 153 months.

Survivorship bias in mutual fund returns

Motivated by the findings of Sørensen (2009) we test for the existence of a survivorship bias among Norwegian mutual funds. Sørensen found evidence of a survivorship bias existing among Norwegian funds in the time period 1982 – 2008. We run the same tests as Sørensen to verify the results and also test the period from 2008 up until 2015. Moreover, unlike Sørensen, we have decided to exclude all passively managed funds.

Several of the funds in our data have ceased to exist at some point in time prior to the end date. A survivorship bias is believed to arise if one excludes defunct funds, assuming poor performing funds are more likely to be closed down, whereas good performers tend to continue its operations, i.e. funds do not exit the dataset randomly. Hence, by excluding the mutual funds which are not active at the end date, one will run the risk of ending up with a dataset containing all of the well performing funds, while the bad performers are not considered. When conducting a study of the overall performance of the whole mutual fund industry, it is crucial to use an as unbiased dataset as possible. This serves as the most predominant reason as of why we chose to include such a wide range of mutual funds in terms of length and period of existence. Studies such as Brown et al. (1992) found the survivorship bias to be significant among U.S mutual funds, and argues that neglecting this fact would give rise to false inferences.

Figure A3 in the appendix shows the cumulative return on 1 NOK invested in an equal-weighted portfolio comprising all funds, compared to a 1 NOK investment in an equal-weighted portfolio only consisting of funds that were alive at the end of 2015. The data for this plot starts in the second half of 1983, since none of the surviving funds existed prior to this date. After the first month of 1985, the cumulative return on the portfolio consisting solely of surviving funds are everywhere above the return of the portfolio in which defunct funds are included. A difference among the two groups seems quite conspicuous, which insinuates the existence of a survivorship bias.

In order to formally test for the existence of such a bias, we chose to divide the funds into two sub-samples, namely defunct and active, and conduct a two-sample t-test. The first sample is the excess return of all the funds which ceased to exist at some point in time between 1983 and 2015, the second sample consists of excess return of all the funds still active at the end of 2015. We exclude the last month of 2015, since the last fund to close down did so in November 2015. 54 of the 101 funds considered were still active at the end of 2015, while 47 had ended their operations. We state a null hypothesis that the mean of the two samples is equal, and the observations are random draws. The t-statistic for this test is

$$t = \frac{\mu_d - \mu_a}{s_{d,a} \cdot \sqrt{\frac{1}{n_d} + \frac{1}{n_a}}} \sim t(n_d + n_a - 2) \tag{9}$$

Where n_d and n_a is the number of observations in the sample of defunct and active funds respectively, $\mu_d - \mu_a$ is the difference between the two means and $s_{d,a}$ is the pooled standard deviation computed as follows

$$s_{d,a} = \sqrt{\frac{(n_d - 1)s_d^2 + (n_a - 1)s_a^2}{n_d + n_a - 2}} \tag{10}$$

Table A2 shows the difference in means for the whole sample period to be -0.23% per month, or -2.79% annualized, with a corresponding t-statistic of -6.14, confirming the impression from Figure A3. This leads us to reject the null hypothesis that the two means for defunct and extant funds can be assumed to be equal. Furthermore, the table shows the results for the same test carried out with varying timespan, yielding the same conclusion. The difference in average returns is quite similar regardless of start date for the sample, and the corresponding t-statistic is highly significant. The last subsample excludes the period with financial crisis, which helps to explain the reduced volatility. This would, ceteris paribus, increase the t-statistic, but the relatively low number of observations makes the t-statistic decline, even though still at a highly significant level. The difference in excess returns between the surviving funds and the entire sample is 0.086% per month, or 1.03% annualized. This is consistent, though slightly higher, than the findings of Brown and Goetzmann (1995) and Dahlquist, Engström and Söderlind (2000) who finds this difference to be 0.8% for U.S mutual funds and 0.7% for Swedish mutual funds per year, respectively.

Table A3 shows the equal weighted average for both total and excess returns calculated for each year in our dataset, excluding the last month of 2015 due to comparability issues. The first sample contains all mutual funds in the data set, the following two are subsamples containing only funds that were still operating in December of 2015 and a sample of mutual funds which had closed down prior to this date, respectively. The table shows that the sample containing only extant mutual funds had superior excess returns compared to the defunct mutual funds in 28 of the 33 years considered. Moreover, after 2003 the sample of surviving funds outperforms the sample containing all mutual funds every single year up until 2015, whereas the sample of defunct funds underperforms relative to the total over the same time span.

This seems to propose that the complete dataset can be subdivided into two groups; one with the top performers and one with relatively bad performers, where mutual funds still alive tend to belong in the group of top performers while defunct funds tend to belong in the bad performing group. We, as did Sørensen, conclude that omitting defunct funds would result in a bias, which justifies the extra work of obtain a complete dataset free from survivorship bias.

Factors on the Norwegian Market

As the market factor, we use a combination of two indices, initially the Oslo Børs All Share Index (OSEAX) (1983 to 1995) and from when it is available, the Oslo Børs

Mutual Fund Index (OSEFX) (1996-2015). The size, value and momentum factors used in this paper are based on the methodology of Fama and French (1998), whereas the liquidity portfolio is based on the approach in Næs, Skjeltorp and Ødegaard (2009). All factors are obtained from Bernt Arne Ødegaard's online resources using Norwegian data (Ødegaard 2016). See the *Factor construction*-section below for further explanations regarding methodology on how factors are constructed.

Factor construction

Market return

There exist several indexes which could serve as proxies for the return on the Norwegian market. The most commonly used Norwegian index is arguably Oslo Børs Benchmark Index (OSEBX), which is an investible index, composed of the most traded shares. This index would serve as an appropriate benchmark when evaluating individual investors, but could be perceived as unfair when assessing mutual fund performance. As discussed in Section 2, Norwegian mutual funds are subject to legislation forcing diversification, which is not the case for OSEBX. Oslo Børs Mutual Fund Index (OSEFX) has historical returns from January 1996, and is constructed to comply with legislation concerning mutual funds (OSE 2016) . This is hence a common benchmark used for mutual funds. In this paper we have consistently used this index as a proxy for market return for the time period the index has existed. Since our data starts in 1983 we need an additional index to serve as the market portfolio from 1983 up until 1995. Inspired by methodology in similar studies as ours, especially by Sørensen (2009), we decided to use the Oslo Børs All Share Index (OSEAX) for this period, combined with OSEFX for the period 1996 to 2015. The advantage with OSEAX is that it has reliable data ranging all the way back to 1983. On the flipside, the index consists of small illiquid shares which would incur considerable transaction costs and share price movements in an attempt to replicate the index.

Pricing factors

The Asset Pricing Model of Sharpe (1964) Lintner (1965) and Black (1972) long served as the most important model for explaining asset returns in relation to risk.

The CAPM-model assumes that the market portfolio is mean-variance efficient, as described by Markowitz (1959). Under this assumption, the CAPM-model predicts expected return of any security as a positive linear relationship with a slope (β) equal to the specific security's exposure to market risk.

The preciseness of CAPM relies on the assumption that market βs adequately describes the cross sectional differences in the distribution of expected returns. This assumption has in subsequent years been relaxed by including additional factors. The most influential augmentation of CAPM is possibly the Fama and French (1993) three-factor model, which extends CAPM by including two factors in addition to the market risk premium. Both factors are constructed as zero investment portfolios, using publicly available information at the time of construction. In order to construct the factors, companies are sorted into three book-to-market value of equity (B/M) portfolios, namely *high, medium* and *low* using the 30^{th} and 70^{th} percentile as breakpoints. In each of the B/M-categories, companies are classified as either *small* or *big*, using the median company as cut-off point. This generates a three-by-two matrix consisting of the following portfolios: (SH, SM, SL, BH, BM, BL).

The SMB-factor is based on the results from Banz (1981), who through an empirical study revealed that smaller firms, in terms of market value, on average had higher risk adjusted returns than larger companies, commonly referred to as the 'size effect'. The SMB-factor is constructed in order to capture this effect, and is constructed as follows:

$$SMB = \frac{1}{3}(SH + SM + SL) - \frac{1}{3}(BH + BM + BL) \quad (11)$$

The portfolio is a zero-sum investment which takes a long position in an equal-weighted average of the small companies, and a short position in an equal-weighted average of the big companies.

Moreover, the HML-factor *(high minus low)* is based on findings by Bhandari (1988), Stattman (1980) and Rosenberg, Reid and Lanstein (1985) who found a positive relationship between stock returns and the previously described B/M-ratio. The HML factor is constructed as follows:

$$HML = \frac{1}{2}(SH + BH) - \frac{1}{2}(SL + BL) \quad (12)$$

The factor mimics a portfolio which is long in stocks with high book-to-market ratio (value stocks), and short in stocks with low book-to-market ratio (growth stocks).

The three-factor model is frequently augmented with a fourth factor in order to capture the effect described by Jegadeesh and Titman (1993). In an empirical study on the U.S. stock market they discovered that a strategy named *momentum* which consists of buying stocks that have performed well in the recent past combined with selling stocks that have underperformed in the same period. The authors showed that this strategy had provided excess returns. Carhart (1997) four-factor model includes a momentum factor *PR1YR*, as well as the three previously mentioned factors. *PR1YR* is constructed as the equal weighted average of the companies with the top 30 per cent return in the past eleven months, lagged one month, minus the bottom 30 per cent companies in the same time-period.

In our analysis we chose to make use of the momentum factor proposed by Fama and French, namely *UMD* (*up minus down)*. *UMD* is quite similar to Carhart's *PR1YR*, slightly modified in order to remove any prevailing size-effect. *UMD* is constructed in the same manner as *HML*, except using the previous 11 month return instead of *B/M* (Fama and French 2010). The formula for UMD is as follows:

$$UMD = \frac{1}{2}(SU + BU) - \frac{1}{2}(SD + BD) \tag{13}$$

In which U and D comprise the 30% top performers (up) and 30% poorest performers (down) respectively.

Several researchers such as Acharya and Pedersen (2005) and Sadka (2006) have suggested that deviations related to CAPM could stem from different levels of liquidity among traded companies. Motivated by the findings of Næs, Skjeltorp and Ødegaard (2009), who test this factor on the Norwegian stock market, we chose to include a liquidity factor (*LIQ*) instead of the value factor (HML) in some of our models (these results can generally be found in the appendix, except for the 'Injecting Alpha'-part, which is included in the main body of the text). According to the authors, a model containing a liquidity factor in combination with the market and a size-factor *provides a reasonable fit for the cross-section of Norwegian stock returns.* The liquidity factor is constructed by sorting a portfolio which is based on relative bid-ask spread, calculated as the closing bid-ask spread relative to the midpoint price.

The portfolio is a zero investment which is long in the least liquid companies and short in the most liquid companies.

Interest rates

Throughout our analysis we have used the interest rates provided from Bernt Arne Ødegaard's online resources (Ødegaard 2016). The interest rates are forward looking for borrowing in the following month. For the period subsequent to 1986 monthly NIBOR is used as the estimate for the risk free rate. Monthly NIBOR is not available prior to 1986, and for this period the overnight NIBOR is used as an approximation. Figure A1 shows the evolvement of the 1-month risk free rate starting in 1983 up until 2015.

Summary Statistics

The above described factors have all been constructed for the Norwegian equity market, following the methodology put forward by Fama and French (1998) and Carhart (1997). Panel A of Table 1 shows descriptive statistics for five explanatory factors for the Norwegian market from January 1983 through December 2015, as well as the risk free rate and an equal weighted portfolio consisting of all actively managed Norwegian mutual funds. The equal weighted portfolio exhibits the highest average monthly return of 1.26% ($t = 4.03$). For the independent variables, the size-factor *SMB* exhibits the highest average monthly return, 0.79% (t = 3.57) per month, and is the only factor statistically different from zero at a five per cent level. The average values of the monthly market premium ($R_m - R_f$) and the momentum portfolio *UMD* are also quite large, though not statistically significant 0.62% (t = 1.88) and 0.56% (t = 1.95) respectively. The liquidity factor has the lowest average return, 0.14% per month (t = 0.60). Panel B of Table 1 reports the correlation-matrix of the above mentioned variables. The greatest correlation is not surprisingly between the equal weighted portfolio of mutual funds and the market portfolio (0.97). The

Table 1: Descriptive statistics

This table provides selected descriptive statistics for the Norwegian 1-month risk free rate, an equal weighted portfolio of the funds in our dataset as well as all the factors considered and used throughout our analysis. All non-standard measurements are reported as percentages on a monthly basis. The average return is computed as the monthly arithmetic average. All relevant measures are reported as percentages. The market portfolio is a combination of the Total Index (Oslo Børs Totalindeks) up until 1995 and OSEFX (Oslo Børs Mutual Fund Index) which was initiated at the beginning of 1996. A thorough description of how the remaining factors are constructed is provided under 'Factor construction' *in* section 5.

	Risk free rate	Excess return equal weighted portfolio	R_m - R_f	*SMB*	*HML*	*UMD**	*LIQ*
Panel A: Summary Statistics							
Average Return	0.56	1.26	0.62	0.79	0.32	0.56	0.14
Standard Deviation	0.37	6.23	6.59	4.39	4.91	5.72	4.68
t -statistic	30.03	4.03	1.88	3.57	1.29	1.95	0.60
Max	2.07	17.39	16.51	22.22	18.46	25.48	16.42
Min	0.08	-25.49	-28.68	-17.08	-16.65	-24.27	-17.66
Skewness	0.66	-0.73	-1.00	0.47	-0.11	-0.19	0.13
Kurtosis	-0.35	1.85	2.89	3.38	1.26	1.93	0.90
Panel B: Cross-correlations							
Risk free rate	1.00						
Excess return equal weighted portfolio	-0.03	1.00					
R_m - R_f	-0.09	0.97	1.00				
SMB	0.02	-0.32	-0.42	1.00			
HML	0.11	0.04	0.07	-0.13	1.00		
*UMD**	-0.09	-0.09	-0.10	0.13	-0.07	1.00	
LIQ	0.12	-0.57	-0.60	0.58	0.03	-0.06	1.00
Overall time period: 1983M01 - 2015M12							

** We do not have data for UMD for December 2015. Statistics relating to UMD are over the time period 1983M01 - 2015M11* second greatest correlation, in absolute terms, is between the market premium and the liquidity factor (-0.6). The relatively large negative correlation is in accordance with expectations, as the *LIQ*-portfolio consists of a short position in the most liquid companies, which constitutes a great proportion of the market portfolio. Similar reasoning could be used to explain the large negative correlation (-0.42) between the size-factor (*SMB)* and the market as well. Moreover, *LIQ* and *SMB* exhibit a positive correlation of 0.58. This is in line with our anticipation, as large companies tend to be relatively more liquid while smaller companies tend to be less liquid.

6. Empirical Results

Equal weight portfolio regression results

To get an initial overview of the overall performance of the Norwegian mutual fund industry, we report the results from regressions of five different model specifications, with the excess return of an EW (equally weighted) mutual fund portfolio as the dependent variable. Results are shown in Table 2 below. The portfolio is constructed using all of the funds in our data sample, taking an EW average of the fund returns available at a given point in time. Ideally, we would report the same for a value weighted portfolio, but unfortunately limitations in our dataset, specifically the lack of data on assets under management, prevents this (current AUM is easily obtainable for funds that still exist, while historical data, especially for defunct funds, is hard to obtain). It is important to note that equally weighted returns of mutual funds can be misleading and should be handled with care. We nevertheless report results using EW, but use them only as interesting observations, while refraining from drawing conclusions from the results.

For the basic Jensen (1968) 1-factor model (1), the annualized alpha is actually positive and statistically significant (annualized alpha is equal to 1.5%, with t-statistic equal to 5.6), in contrast to what we hypothesise for net returns. However, when controlling for additional risk factors in the other models (2-5), the alpha becomes negative, albeit not significantly so. For example for model (4), which has the lowest t(α), the annualized alpha is -0.2% (not significantly different from 0), which is actually high given that returns are net of fees; for gross returns, we would expect average alpha close to 0 as above, but for net returns a number closer to the negative of the average annual fee, currently equal to ~1.4% (Strøm 2014), is plausible. But again, this could be a result of using EW returns rather than VW returns. With annualized standard error of 0.23%, the annualized model (4) alpha is significantly higher than -1.4%.

The coefficient on excess market return is significantly lower than 1 for all models. For example, for our the Fama-French 3-factor model (2) with β_M equal to 0.96, the equal weight portfolio increases/decreases with 0.96% when the market increases/decreases with 1%. The deviations from 1 are interesting, and might be related to the fact that Norwegian mutual funds cannot borrow money, but rather need to hold some very liquid assets such as T-bills. Additionally, as mentioned we use an equal-weight portfolio which does not accurately reflect the actual overall return of the sector. We also use OSEAX as the market benchmark before 1996, and the differences between OSEFX and OSEAX (see 'Market return' under section 5 for details) could be a reason; we find that regressing the EW portfolio on OSEFX only, over the post-1996 period in which it is available, yields a coefficient slightly higher than 1 (but not significant).

The adjusted R-squared is fairly high (almost 95%) for all models. It is slightly lower for model (1) than for the other specifications. We observe that going from the FF 3-factor model (2) to the FF 4-factor model (3), only increase explanatory power very slightly, as does going from model (2), (3) or (4) to the full 5-factor specification (5). Neither of the coefficients added (HML and UMD) in going from (4) to (5) are statistically significant (t = -0.9 and t = -0.53 respectively), and adding the liquidity factor used in models (4) and (5) to model (3) increases power slightly (with ~0.1%).

The coefficient on the size portfolio β_{SMB} is positive and statistically significant in all models it is included, while coefficients on the value portfolio β_{HML} and on the momentum portfolio β_{UMD} are never significant. The coefficient on the liquidity portfolio β_{LIQ} is negative and statistically significant when included in models (4) and (5).

Table 2: Regression results of various models specifications for equal-weight portfolio of actively managed Norwegian mutual funds

This table shows time series regression results for different models on net returns of an equal weight portfolio of the actively managed Norwegian mutual funds in our sample. The number of funds in our sample used to calculate the equal-weight mean return per period ranges from minimum two in 1983M01 to maximum 66 in 2002M10 with an average of 56 over the whole sample. Explanatory variables used are the market excess return (M), a size factor (SMB), a value/growth factor (HML), a momentum factor (UMD) and a liquidity factor (LIQ) (see 'Factor construction' under section 5 for descriptions of the factors). Regression results shown are the intercept and coefficient estimates with corresponding t-statistics, the regression R2 and adjusted R2. For the market slope, the t-statistics tests whether the coefficient β_M is different from 1, while the other t-statistics test whether coefficients are different from 0. We use the OLS estimator and standard errors corrected for heteroscedasticity and autocorrelation with the Newey and West (1986) procedure.

Model specification		α	β_M	β_{SMB}	β_{HML}	β_{UMD}	β_{LIQ}	R^2	Adj. R^2
(1) Jensen's Alpha / CAPM	Coefficient	0.001	0.92					93.9%	93.9%
	t(Coefficient)	1.62	-4.69						
(2) Original Fama-French 3-factor model	Coefficient	0.000	0.96	0.14	-0.02			94.8%	94.7%
	t(Coefficient)	-0.11	-2.31	7.22	-1.34				
(3) Fama-French 4-factor model (incl. momentum)[1]	Coefficient	0.000	0.96	0.14	-0.02	0.00		94.8%	94.7%
	t(Coefficient)	-0.17	-2.25	7.29	-1.29	0.07			
(4) Model specification based on Næs, Skjeltorp & Ødegaard (2009)	Coefficient	0.000	0.94	0.18			-0.07	94.9%	94.9%
	t(Coefficient)	-0.26	-2.77	6.43			-2.02		
(5) Fama-French extension (5-factor model incl. liquidity and momentum)[1]	Coefficient	0.000	0.94	0.18	-0.02	-0.01	-0.07	94.9%	94.9%
	t(Coefficient)	-0.15	-2.67	6.45	-0.90	-0.53	-2.00		

Overall time period: 1983M01 - 1983M12

1) Regressions including UMD are over the time period 1983M01 - 2015M11, as we do not have data for UMD for December 2015.

Individual fund regression results

Table 3 shows results from a subset of the benchmark regressions (step one of the bootstrap procedure in section 4), using the Fama-French 3-factor model (2). We are mostly interested in the tails of the distribution: the top and bottom ten funds, ranked from worst to best according to $t(\alpha)$, are included in the table. All of the ten worst performing funds have a negative constant (α), the nine worst statistically significant at the 1 %-level, with the worst being *Nordea SMB II* with a monthly α of -1.44% per month. This fund was closed down after only 69 months of operation, making it relatively short-lived compared to an average lifetime in our dataset of 153 months, supporting our findings of significant survivorship bias (see 'Survivorship bias in mutual funds' under section 5).

All of the best performing funds have delivered a positive α, the eight best statistically significant at the 5 %-level, and the four best even at the 1 % level. *Omega Investment Fund B* and *C* are the two best performing funds, probably with very similar holdings as all coefficients are close to equal. The two funds have provided an impressive alpha of approximately 1.2 % per month (14.7% per year) and with equal $\beta_M s$ of only 0.51 they are the two funds with the by far least exposure to market risk. We note however that the top four performing funds are new (only 20-34 months old) and that especially returns of the top two are not as well explained by the model as most other funds, both with R^2 = 44.8% (it will be interesting to follow the new, top performing funds in the future, to see whether they are able to sustain their track records over a longer period of time). Actually, there is an inverse u-pattern in R^2. It is relatively low in both extreme tails of the distribution (i.e. for the worst and best funds), whereas it seems to rise towards the center of the distribution. This could suggest that extreme performers, high or low, are more active and simultaneously less diversified than the average of mid-performers. Fully passive index funds, if included, should get a place at the middle of the distribution, with α and t(α) unaffected by skill and luck and close zero, while very active funds could end up anywhere depending on their individual levels of positive/negate skill and degrees of good/bad luck. Another proposition could be that the extreme performers are exposed to other risk factors which are not well captured in our chosen model (4), but this should not matter much for the presented bootstrap methodology. The funds with low R^2 in the benchmark regressions above, based on actual returns, will on average have relatively high absolute values of residuals. And, as residuals (as well as factor returns) are 'drawn' in the simulations, these funds will more often get low levels of R^2 and be more probable to get extreme levels of α and t(α) in the simulations as well.

Some interesting observations from Table 3b are the relative difference in risk factor exposure between the top and the bottom funds as groups. For example, it seems like the bottom funds are relatively more exposed to market risk and to small companies, compared to the best: the equal weighted average of β_M for the bottom ten is 1.05, while the same measure is 0.77 for the best ten, and the equal weighted average of β_{SMB} for the bottom ten is 0.29, while it is 0.07 for the ten best. For β_{HML} there is not a clear pattern, with equal weighted averages being -0.05 for the bottom funds and -0.01 for the top ten funds.

Another interesting observation is that all of the 20 worst performing funds have ended their operations, while all of the funds among the 20 best performing were all still operating at the end of 2015, further supporting our findings in the survivorship bias section.

The results are also in line with those of Sørensen (2009), who provides a similar table, also using the Fama-French model (2). The funds presented as the bottom four in his paper, are still the bottom four in our research (they were all shut down within his period of interest). And, among the top four from 2009, three are still among the top ten today (all four are alive at the end of both his and our periods).

In Table A4 in the appendix, the same table based on our second model specification (4), with market return, a size portfolio and a liquidity portfolio as factors, is provided. Overall the results are very similar, and there are no major discrepancies between these and the results based on the FF 3-factor model (2). The fund rankings are similar, levels of both α- and t(α)-values are of approximately the same magnitude, and we see the same inverse u-pattern in R^2.

Table 3: Results from regressions on the Fama-French 3-factor model (2) for individual actively managed Norwegian mutual funds

This table shows time series regression results for the Fama-French 3-factor model estimated on net returns of the individual actively managed Norwegian mutual funds in our sample. The explanatory variables used are the market excess return (M), a size factor (SMB) and a value/growth factor (HML) (see 'Factor construction' under section 5 for descriptions of the factors). Results are shown for the top and bottom ten funds, ranked by the t-statistic of alpha. In addition to regression results, the table shows the following information about each fund from left to right: rank of fund (by t-stat of alpha), ticker of fund, name of fund, the number of returns and the time span the fund is present in our sample. Regression results shown are the intercept and coefficient estimates with corresponding t-statistics and the regression R^2. For the market slope, the t-statistics test whether the coefficient β_M is different from 1, while the other t-statistics test whether intercepts/coefficients are different from 0. We use the OLS estimator and standard errors corrected for heteroscedasticity and autocorrelation with the Newey and West (1986) procedure.

Rank	Ticker	Name	# of obs	Alive from - Alive to	Coefficients				T-statistics				R^2
					α	β_M	β_{SMB}	β_{HML}	$t(\alpha)$	$t(\beta_M)$	$t(\beta_{SMB})$	$t(\beta_{HML})$	
101	KF-SMBII	Nordea SMB II	69	1997M07 - 2003M03	-0.0144	0.96	0.57	-0.12	-3.92	-0.56	4.63	-1.30	76.9%
100	SK-SMB	Skandia SMB Norge	96	1994M12 - 2002M11	-0.0112	1.03	0.46	-0.11	-3.35	0.27	6.20	-1.61	85.6%
99	SU-NORGE	Globus Norge II	94	1998M12 - 2006M11	-0.0105	1.26	0.32	-0.12	-3.15	3.31	3.80	-1.90	87.4%
98	GF-AKSJE	GJENSIDIGE AksjeSpar	151	1987M02 - 1999M08	-0.0035	0.94	0.05	0.03	-3.10	-1.88	1.99	0.76	93.1%
97	KF-SMB	Nordea SMB	212	1997M06 - 2015M01	-0.0060	1.04	0.55	-0.04	-3.02	0.89	8.64	-0.80	82.0%
96	GF-INVES	GJENSIDIGE Invest	103	1992M04 - 2000M10	-0.0048	0.98	0.19	0.09	-2.92	-0.63	3.38	3.10	93.2%
95	DI-RVKST	DnB Real-Vekst	156	1989M12 - 2002M11	-0.0036	0.93	0.11	0.01	-2.81	-4.46	4.30	0.40	95.8%
94	SU-GLNO	Globus Norge	103	1998M03 - 2006M11	-0.0090	1.19	0.34	-0.13	-2.78	2.40	4.71	-1.69	87.3%
93	DK-NORII	Avanse Norge (II)	286	1991M01 - 2014M10	-0.0017	0.95	0.04	-0.02	-2.55	-3.28	1.31	-1.24	97.5%
92	SU-AKTIV	Globus Aktiv	87	1998M12 - 2006M04	-0.0081	1.26	0.29	-0.12	-2.32	3.21	3.49	-1.90	87.5%
10	SP-VERDI	Storebrand Verdi	216	1998M01 - 2015M12	0.0018	0.91	-0.02	0.15	1.75	-3.37	-0.49	4.26	91.4%
9	FV-TRNDR	FORTE Tr?nder	32	2013M05 - 2015M12	0.0053	0.70	-0.01	-0.10	1.81	-2.09	-0.07	-1.11	53.8%
8	AC-NWECD	Arctic Norwegian Equities Class D	34	2013M03 - 2015M12	0.0054	0.74	0.00	-0.03	2.00	-2.33	0.00	-0.58	74.2%
7	CA-AKSJE	Carnegie Aksje Norge	245	1995M08 - 2015M12	0.0021	0.95	0.05	-0.13	2.15	-2.21	1.55	-3.88	94.0%
6	FF-NOAI2	Danske Fund Norge Aksj. Inst 2	109	2006M12 - 2015M12	0.0030	0.94	0.05	0.02	2.31	-1.69	1.02	0.77	97.1%
5	FF-NOIII	Danske Fund Norge Aksj. Inst 1	188	2000M05 - 2015M12	0.0021	0.93	0.01	0.01	2.39	-3.82	0.54	0.76	97.3%
4	IS-UTBYT	Landkreditt Utbytte	34	2013M03 - 2015M12	0.0059	0.66	0.23	0.13	2.62	-2.78	2.94	2.04	69.3%
3	AI-NORGI	Alfred Berg Norge Inst	20	2014M05 - 2015M12	0.0072	0.84	0.12	0.00	3.43	-1.84	1.49	-0.09	91.3%
2	OR-INVFB	Omega Investment Fund B	25	2013M12 - 2015M12	0.0119	0.51	0.13	-0.10	3.62	-5.07	1.17	-1.61	44.8%
1	OR-INVFC	Omega Investment Fund C	25	2013M12 - 2015M12	0.0123	0.51	0.13	-0.10	3.74	-5.08	1.16	-1.61	44.8%

Bootstrap results

The first column ('Actual') in Table 4 shows results from the benchmark regressions, based on the same numbers as in Table 3. The table also shows results from the bootstrap simulations for corresponding ranks/percentiles. Panel A is based on and reports values for α, whereas Panel B is for $t(\alpha)$. According to (Fama and French 2010) the t-statistic is a more accurate measure since it incorporates the measurement precision, and consequently our main focus in the analysis will be on Panel B. The results are for ranks and percentiles in ascending order, where percentiles are based on interpolations between the ranks closest to the given percentile. The cross sectional distribution from the estimated benchmark model can easily be compared with the average values of the corresponding rank or percentile of simulated α and $t(\alpha)$, based on the average of 10 000 bootstrap simulations. For example, the 5th worst and 5th best values of actual $t(\alpha)$ estimates are -3.02 and 2.39, whereas the average value of the corresponding ranks from simulations are -1.76 and 1.66 respectively.

The last column in the two panels provide the fraction of simulations yielding a lower result than the actual observations. For example, only 1.1% of the simulated values are lower than the actual value for the 5th worst fund. This indicates 'bad skill' or value destruction. Actually, most of the left side of the distribution (from worst rank to above the 40th percentile), the simulated $t(\alpha)$ estimates are greater than the actual values in more than 90% of the draws, and recurrently more than 95%. This leaves little evidence for misfortune as the main explanation for poor fund performance, i.e. we reject a null hypothesis stating that bad results are only due to bad luck. Also, all the way from the end of left tail (i.e. 1st percentile) and up until the 80th percentile, we observe that the simulated averages are higher than the corresponding actual observations from the benchmark regressions.

In the right tail, the results are more encouraging. From somewhere below the 90th percentile and up, every average simulation value is below the actual observation, and the proportion of simulated values below the actual frequently exceeds 95% among the best performing funds. This allows us to reject a null hypothesis that good performance is only due to luck, and thus acknowledge that there exists some skill among the best performers.

Table 4: Ranks and percentiles of α- and t(α)-estimates for actual and simulated mutual fund returns based on the Fama-French 3-factor model (2)

Panel A of this table shows estimated values of α at selected ranks and percentiles for actual fund returns of actively managed Norwegian mutual funds, while Panel B shows estimated values of t(α). The panels are produced separately and α-values of specific ranks do not necessarily correspond t(α)-values of the same rank. The simulated average is the average of α or t(α) at selected percentiles from the simulation. The % < Act columns show the percentage of simulations runs which produce lower values of α or t(α) at the given rank/percentile than those observed for actual fund returns. The explanatory variables used are the market excess return (M), a size factor (SMB) and a value/growth factor (HML) (see 'Factor construction' under section 5 for descriptions of the factors). We use the OLS estimator and standard errors corrected for heteroscedasticity and autocorrelation with the Newey and West (1986) procedure.

	Panel A: α				Panel B: t(α)		
Rank/ percentile	Actual	Simulated average	% < Act	Rank/ percentile	Actual	Simulated average	% < Act
Worst	-0.0144	-0.0101	16.0	Worst	-3.92	-2.83	9.4
2nd	-0.0112	-0.0061	3.5	2nd	-3.35	-2.32	5.7
3rd	-0.0109	-0.0048	0.5	3rd	-3.15	-2.06	3.4
4th	-0.0105	-0.0042	0.1	4th	-3.10	-1.89	1.6
5th	-0.0090	-0.0037	0.1	5th	-3.02	-1.76	1.1
10%	-0.0055	-0.0023	0.1	10%	-2.29	-1.31	1.6
20%	-0.0027	-0.0013	2.6	20%	-1.72	-0.86	2.2
30%	-0.0018	-0.0008	4.9	30%	-1.21	-0.54	4.8
40%	-0.0011	-0.0004	10.2	40%	-0.80	-0.28	9.3
50%	-0.0005	0.0000	17.5	50%	-0.38	-0.03	18.3
60%	0.0001	0.0003	32.3	60%	0.04	0.22	32.6
70%	0.0004	0.0007	32.8	70%	0.41	0.49	43.2
80%	0.0012	0.0012	51.6	80%	0.91	0.80	61.9
90%	0.0027	0.0022	75.5	90%	1.63	1.24	83.2
5th	0.0054	0.0035	92.9	5th	2.39	1.66	93.7
4th	0.0059	0.0039	91.7	4th	2.62	1.78	95.2
3rd	0.0072	0.0046	93.1	3rd	3.43	1.94	99.4
2nd	0.0119	0.0057	98.5	2nd	3.62	2.17	98.5
Best	0.0123	0.0089	82.2	Best	3.74	2.64	92.0

One observation worth mentioning is that the funds just below the best performing fund actually have a greater proportion of simulated values below the actual value.

For example, this proportion is 99.4% for the 3rd best fund, whereas the same fraction amounts to a slightly lower 92.0% for the best fund. Although this may appear as an inconsistency, it is in fact not. It can be explained by the relative performance among the funds, e.g. that the third best fund performs extremely well contingent on being the third best fund. The opposite is true when focusing on the left tail of the distribution. It is important to emphasize that the actual results of each individual fund in the is contingent on the results of all the other funds in the comparison with simulated numbers. Thus, one cannot draw conclusions for individual funds, but rather need to look at the general results of the tails as a whole.

Figure 1 shows the empirical cumulative distribution function for simulated and actual values of α and t(α). This can be perceived as a visualization of the results from the *Actual* and *Simulated average*-columns of Table 4. We see that the line based on simulated values lie everywhere to the right of the line based on actual values (i.e. are higher) up until around the 80th percentile, consistent with our observations in the second paragraph of this section, while the top 20% of the funds do have an actual alpha which is greater than the averages from simulations.

Figure 2 shows histograms for simulated values and actual value of $t(\alpha)$ at selected ranks and percentiles. For example, the bottom right subfigure shows the results for the best fund. The dotted line represents the actual $t(\alpha)$-value for the best fund, here amounting to 3.74 (again, numbers are the same as observed in table 3a), whereas the histogram shows the distribution of the best $t(\alpha)$ values collected from 10 000 simulations. From the histogram it is easy to see that the majority of simulated values place below the actual $t(\alpha)$-value. In the rightmost column in the bottom row of Table 4 it can be seen that exact fraction of simulated values which are less than the actual value amounts to 92.0%.

In Figure 3 we provide an alternative visualization of the bootstrap results, motivated by the methodology put forward by Cuthbertson, Nitzsche and O'Sullivan (2008). The main focus is still on the t-statistic, which is reported on the right hand side. The figure shows the Kernel density estimates of the frequency of funds one might expect

Figure 1: Empirical cumulative distribution function of actual and simulated alpha and t(alpha) using Fama-French 3-factor model (2)

This figure shows actual and simulated empirical cumulative distribution functions (ECDF) for alphas and their corresponding t-stats. The left panel shows the ECDFs for the actual and simulated alphas based on a three factor model with factors for market, size and value as explanatory variables. The right panel shows the same graphs for the t-statistics corresponding to the alphas of the same model. The actual and simulated alphas with corresponding t-statistics used are the same as described in Table 4 above.

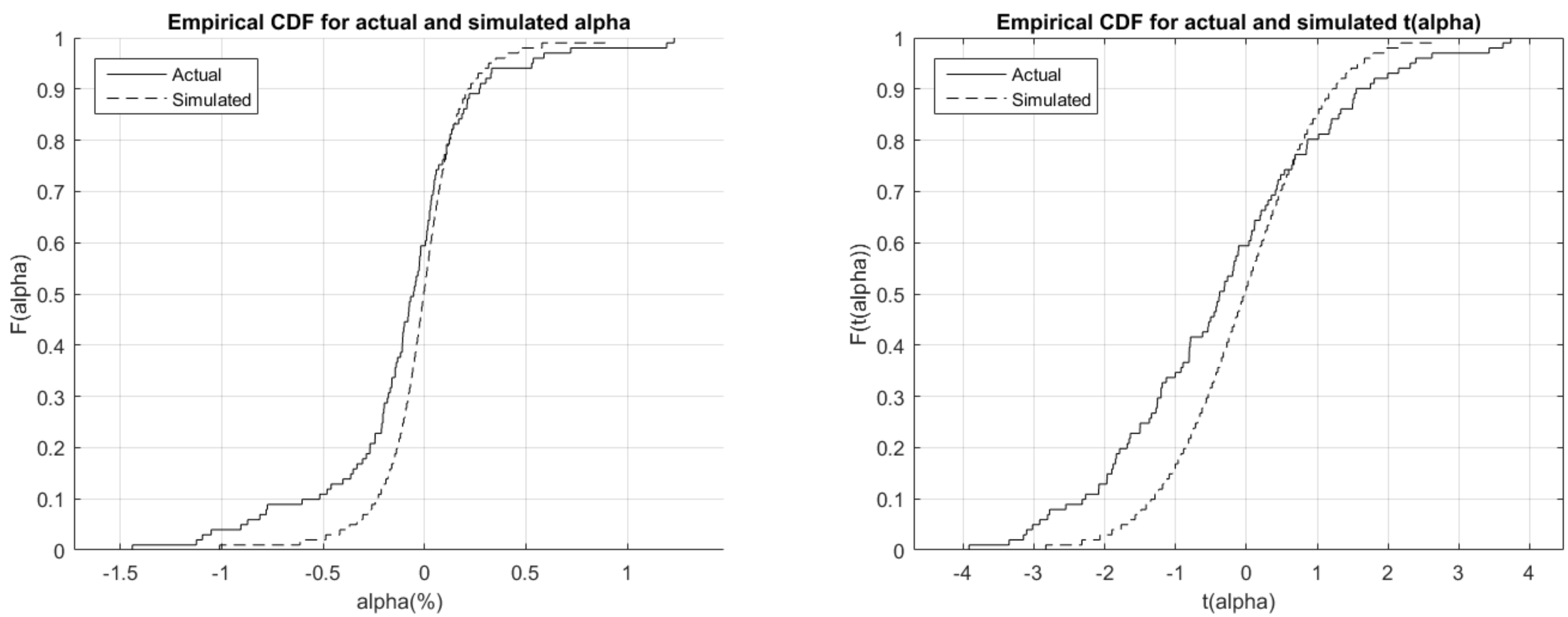


Figure 2: Histogram of different ranks/percentiles of the simulated t(alpha) using Fama-French 3-factor model (2)

This figure shows histograms for t-statistic of simulated alpha. Each panel displays the histogram of a specific rank/percentile from each of the bootstrap simulations, as specified in the titles. The dashed line in each panel displays the actual t-statistic of the corresponding rank/percentile. Actual alphas with corresponding t-statistics are estimated with the observed historical returns of each fund, while simulated alphas are the average of each rank/percentile of all the 10.000 basic bootstrapped simulations (described under 'The bootstrap procedure: 5 steps' under section 4). The top and bottom ranks of actual alpha and of the corresponding t-statistic and the average of the top and bottom simulated ranks can be found in Table above. We use the OLS estimator and standard errors, corrected for heteroscedasticity and autocorrelation with the Newey and West (1986) procedure, to estimate alpha and corresponding t-statistic.

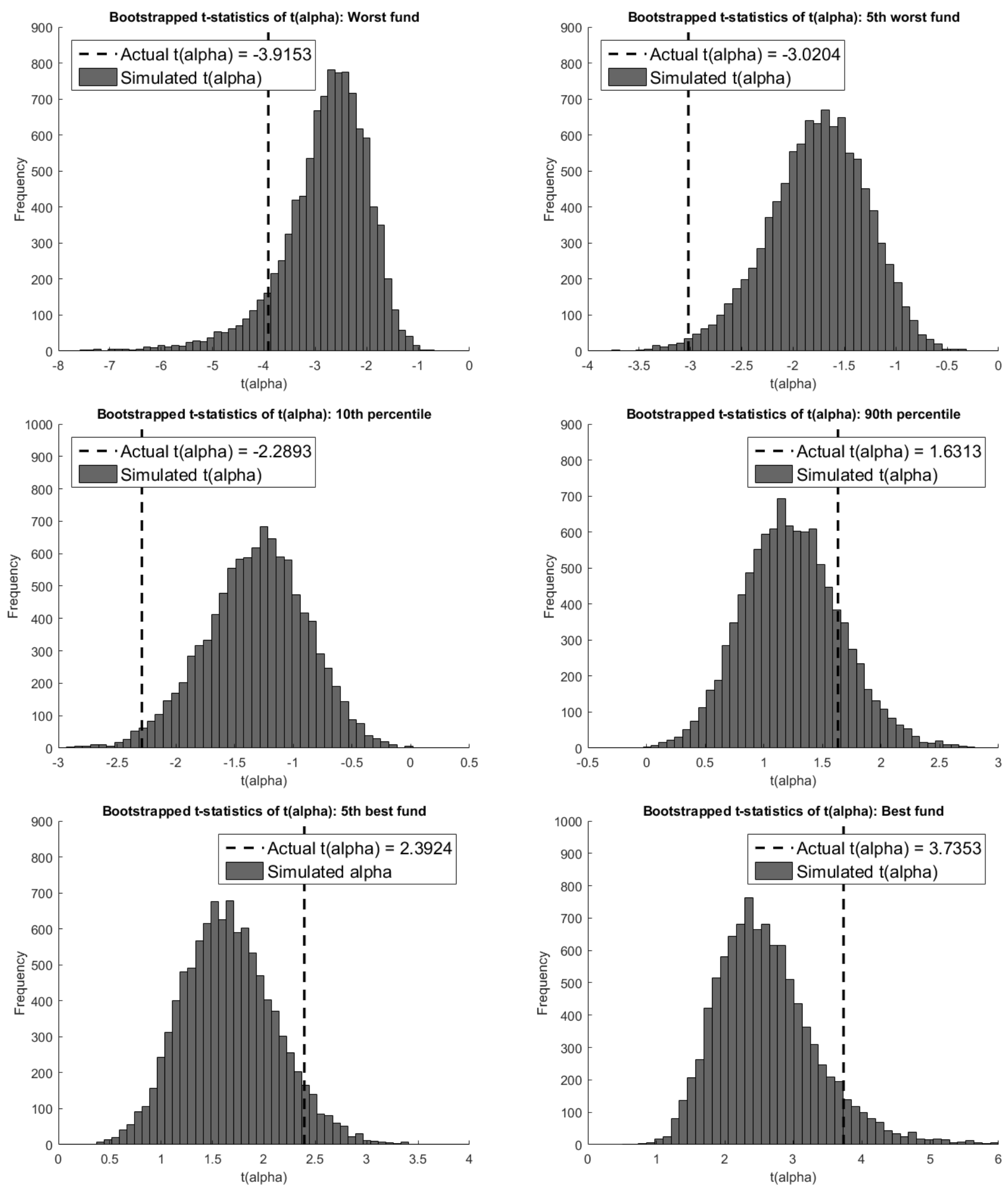
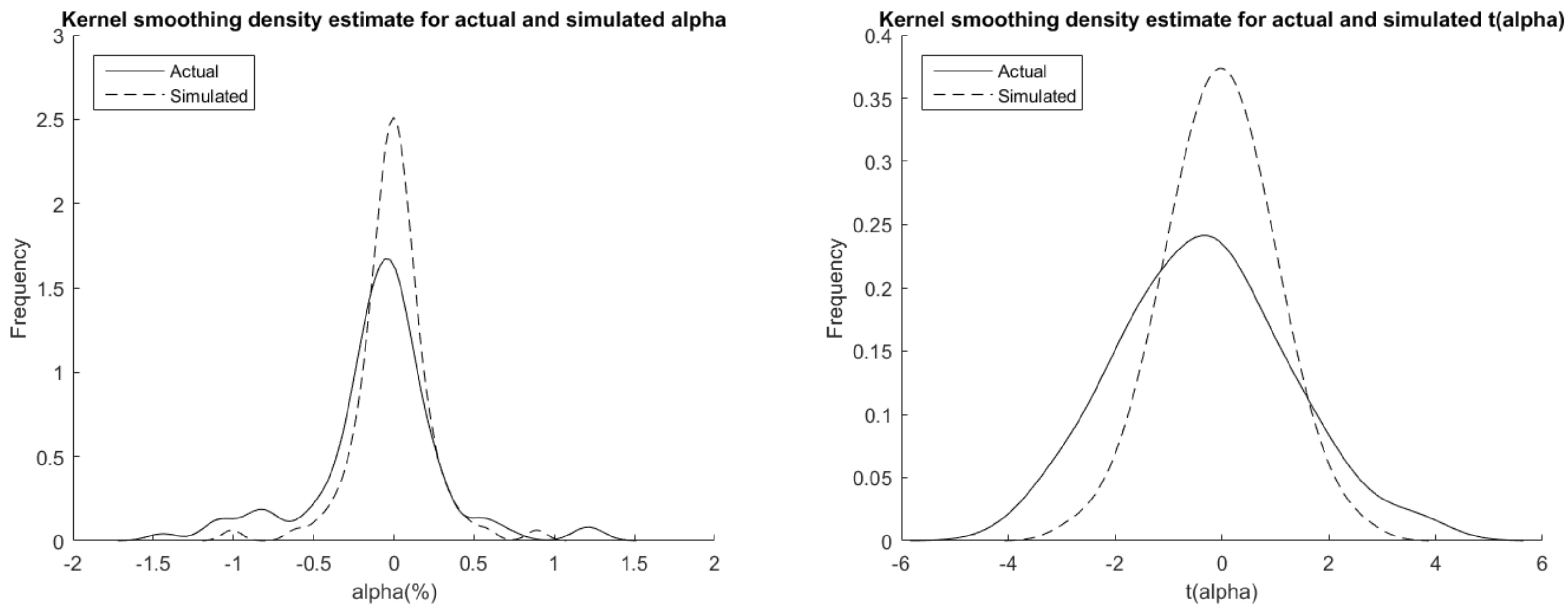


Figure 3: Kernel smoothing function estimate of actual and simulated alpha and t(alpha) using Fama-French 3-factor model (2)

This figure shows actual and simulated kernel smoothing density function (KSDF) estimates for alphas and their corresponding t-stats. The left panel shows the KSDFs for the actual and simulated alphas based on a three factor model with factors for market, size and value as explanatory variables. The right panel shows the same graphs for the t-statistics corresponding to the alphas of the same model. The actual and simulated alphas with corresponding t-statistics that are used to estimate the Kernel smoothing densities are the same as described in Table 4 above and we use the standard bandwidth optimal for normal densities.



to achieve a given level of $t(\alpha)$ as a result of luck alone, compared to the actually observed frequencies of achieved $t(\alpha)$. One of the most remarkable observations from Figure 3 is the apparent overpopulated left-tail of the actual $t(\alpha)$-distribution (solid line), compared to the luck-distribution (dashed line). This observation proposes that a large fraction of the funds cannot use poor luck as an explanation to bad performance. On the opposite side of the distribution there seems to be evidence of outperformance which is not due to luck. Additionally, the actual distributions are shifted slightly to the left compared to the simulated distributions, implying that the mean is lower than simulated mean. This could be explained by fees; we use net returns for the actual numbers, fees are captured in alpha, and in the simulations alpha, and fees with it, are left out. The simulated distribution is approximately symmetric around zero and looks relatively normal, while the actual distribution has negative mean and seems to be platykurtic.

Overall, the bootstrapping results give evidence for the existence of both inferior and superior fund management, i.e. that performance cannot be explain by good/bad luck alone.

Bootstrap results with injected alpha

In the methodology used for estimating the distribution true α (explained at the end of 'Extension of the bootstrap procedure: Injecting alpha' under section 4), a normality assumption is used. This will not be completely accurate, and even when allowing for different σ for each tail, one cannot expect single values of σ to capture the tails of $t(\alpha)$ estimates for actual returns completely (Fama and French 2010). Our simulations with injection of alpha, summarized in Table 5 below, suggest that σ in the area ~2.0% to ~2.5% captures most of the extreme left tail (depending on which point in the tail is used), with 2.25% as the best estimate. While σ in the area ~1.75% to ~2.25%, with 2% as the best estimate, captures the extreme right tail of actual $t(\alpha)$ for net fund returns. As a perspective, taking the average of the standard errors of the individual alpha estimates from the benchmark regressions (subset shown in table 3) yields an annualized number of 2.3%, very close to the estimates from the simulations with injected alpha.

The estimates above imply some ability in generating returns above fees charged among fund managers. E.g. with an estimate of σ equal to 2% for the right tail, ~16% of funds have true annual α above 2%, and ~0.6% have true annual alpha above 5%. Similarly, for the left tail, with an estimate of σ equal to 2.25% for the left tail, ~16% of funds have true annual α below -2.25%, and ~1.3% have true annual alpha below -5%. Alternatively, we could say that the top 5% of managers generate true alpha of around 3.3% or more, i.e. they have the skill to earn at least 3.3% more than the fees they charge, and the top 1% of managers generate true alpha of 4.7% or more. And for the left tail, we could say that the bottom 5% of managers generate true alpha of -3.7% or less, i.e. they represent a value destruction of at least 3.7% including the fees they charge, and the bottom 1% of managers generate true alpha of -5.2% or less.

Our estimates for σ are somewhat higher than the average numbers found by Fama and French (2010) for the US market with gross returns. It is important to remember that datasets are not remotely comparable (due to different regions, time periods, number of funds and net vs. gross returns). All else equal, the direction of the difference is opposite of what we would expect given the assumption that there exist a positive relationship between the skill of managers and the fees they are able to charge. Values of net returns α should be less dispersed than those based on gross returns, as high levels of gross return α will effectively be reduced more by the fee than low levels of gross return α. We thus expect that the difference between our results and those by Fama and French (2010) would be even larger with gross returns.

However, the mean should go down as well when going from gross to net returns. When we impose the mean to remain constant (and equal to zero), the implied right tail σ may be additionally biased down (right tail values of net return α are closer to the imposed mean of α (zero) than to the actual and unknown mean of α). For the left tail σ the effect of holding the mean constant is opposite, it increases the left tail σ-estimate; while the overall effect may be ambiguous depending on the magnitudes of the change in mean and σ when going from gross to net returns.

Table 5: Ranks and percentiles of α- and t(α)-estimates for actual and simulated mutual fund returns with injected α

This table first shows estimated values of α at selected ranks and percentiles for actual net fund returns of actively managed Norwegian mutual funds. Panel A shows the average of t(α) and Panel B the percentage of simulations runs which produce lower values t(α) at the same ranks/percentiles from the simulation, given 15 different levels of σ (average standard deviation of injected alpha). The columns 'Actual t-stat' and 'No inj. α' are the same as in Table 4. A randomly drawn alpha is injected into returns for each simulation and for each fund, with mean equal to zero (See 'Extension of the bootstrap procedure: Injecting alpha' under section 4 for a detailed explanation of how alpha is injected). σ varies according to the table below (from 0% to 3.5%, in steps of 0.25%), while each individual funds injected alpha is scaled by multiplying with the following ratio:

$$\left(\frac{\textit{the standard error of actual residuals of the fund}}{\textit{average of the standard error of actual residuals of all funds}}\right)$$

The explanatory variables used to estimate actual and simulated t(α) are the market excess return (M), a size factor (SMB) and a liquidity factor (LIQ) (see 'Factor construction' under section 5 for descriptions of the factors). We use the OLS estimator and standard errors corrected for heteroscedasticity and autocorrelation with the Newey and West (1986) procedure.

Rank/	Actual	Average of the annual standard deviation of injected α														
percentile	t-stat	No inj. **α**	0.25%	0.50%	0.75%	1.00%	1.25%	1.50%	1.75%	2.00%	2.25%	2.50%	2.75%	3.00%	3.25%	3.50%
Panel A: Average of simulated t(α)																
Worst	-3.90	-2.83	-2.85	-2.92	-3.03	-3.18	-3.37	-3.60	-3.87	-4.16	-4.46	-4.79	-5.13	-5.49	-5.85	-6.21
2nd	-3.51	-2.35	-2.37	-2.43	-2.52	-2.65	-2.82	-3.01	-3.22	-3.45	-3.71	-3.97	-4.24	-4.52	-4.81	-5.11
3rd	-3.15	-2.07	-2.09	-2.15	-2.24	-2.36	-2.51	-2.68	-2.87	-3.07	-3.29	-3.52	-3.76	-4.01	-4.26	-4.51
4th	-3.03	-1.90	-1.91	-1.97	-2.05	-2.16	-2.30	-2.45	-2.63	-2.82	-3.01	-3.22	-3.43	-3.66	-3.88	-4.11
5th	-3.00	-1.77	-1.78	-1.83	-1.91	-2.01	-2.14	-2.28	-2.44	-2.61	-2.80	-2.99	-3.18	-3.39	-3.59	-3.80
10%	-2.55	-1.31	-1.33	-1.36	-1.42	-1.50	-1.59	-1.70	-1.81	-1.93	-2.06	-2.19	-2.33	-2.47	-2.61	-2.75
20%	-1.77	-0.86	-0.87	-0.89	-0.93	-0.98	-1.04	-1.11	-1.18	-1.25	-1.33	-1.41	-1.49	-1.58	-1.66	-1.75
30%	-1.30	-0.54	-0.54	-0.56	-0.58	-0.61	-0.65	-0.69	-0.73	-0.78	-0.82	-0.87	-0.92	-0.97	-1.02	-1.07
40%	-0.86	-0.27	-0.27	-0.28	-0.29	-0.31	-0.32	-0.34	-0.36	-0.38	-0.40	-0.43	-0.45	-0.47	-0.50	-0.52
50%	-0.57	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
60%	-0.11	0.23	0.23	0.24	0.25	0.26	0.28	0.30	0.32	0.34	0.36	0.39	0.41	0.43	0.46	0.48
70%	0.35	0.49	0.50	0.51	0.54	0.57	0.61	0.64	0.69	0.73	0.78	0.82	0.87	0.92	0.97	1.03
80%	0.80	0.81	0.82	0.84	0.88	0.93	0.99	1.05	1.12	1.20	1.27	1.35	1.43	1.52	1.60	1.69
90%	1.67	1.25	1.26	1.30	1.36	1.44	1.53	1.64	1.75	1.87	2.00	2.13	2.27	2.41	2.55	2.70
5th	2.53	1.67	1.69	1.74	1.82	1.93	2.05	2.20	2.36	2.53	2.72	2.91	3.11	3.31	3.52	3.73
4th	2.54	1.80	1.82	1.87	1.95	2.07	2.20	2.36	2.54	2.73	2.93	3.14	3.36	3.58	3.80	4.04
3rd	3.40	1.96	1.98	2.03	2.12	2.25	2.40	2.57	2.77	2.98	3.20	3.43	3.67	3.92	4.18	4.44
2nd	3.43	2.19	2.21	2.27	2.37	2.51	2.68	2.87	3.09	3.33	3.59	3.85	4.13	4.42	4.71	5.00
Best	3.54	2.67	2.69	2.76	2.87	3.02	3.22	3.45	3.72	4.01	4.32	4.65	5.00	5.36	5.72	6.09
Panel B: % of simulated t(α) < Actual t(α)																
Worst	-3.90	9.5	9.8	10.7	12.8	16.2	**22.6**	32.2	43.2	55.5	67.4	**77.2**	**85.5**	91.1	94.6	96.9
2nd	-3.51	5.0	5.0	5.5	6.7	9.3	13.7	**21.1**	31.2	43.7	56.6	68.5	**79.0**	86.7	92.1	95.6
3rd	-3.15	3.7	3.9	4.5	6.0	8.6	13.3	**20.6**	30.5	42.7	55.7	67.9	**78.8**	86.3	91.8	95.4
4th	-3.03	2.0	2.1	2.6	3.7	5.6	8.9	**14.7**	**23.1**	34.1	46.7	59.4	71.2	**80.5**	87.6	92.4
5th	-3.00	1.1	1.1	1.4	1.9	3.0	5.2	8.8	**15.3**	**23.9**	34.7	46.9	59.5	70.6	**79.8**	87.0
10%	-2.55	0.4	0.4	0.5	0.7	1.3	2.1	3.4	6.1	10.3	**16.0**	**23.6**	32.5	42.6	53.0	63.1
20%	-1.77	1.4	1.5	1.7	2.2	3.0	4.1	5.7	8.0	11.3	15.5	20.9	26.6	32.9	39.9	47.1
30%	-1.30	2.9	3.1	3.2	3.6	4.3	5.0	6.3	7.8	9.8	12.1	15.0	18.1	21.7	25.2	29.6
40%	-0.86	6.8	6.6	6.9	7.3	7.6	8.3	9.0	10.1	11.2	12.6	13.8	15.7	17.3	19.0	20.7
50%	-0.57	7.5	7.4	7.3	7.2	7.0	7.1	7.3	7.4	7.5	7.5	7.7	7.9	8.2	8.6	8.9
60%	-0.11	18.9	18.6	18.3	17.5	16.2	15.1	14.3	13.1	12.2	11.1	10.3	9.6	9.0	8.3	7.9
70%	0.35	35.6	35.7	34.3	31.8	29.2	26.1	23.1	19.6	16.6	14.1	11.8	9.9	8.2	6.8	5.5
80%	0.80	50.1	49.2	46.6	42.7	37.7	32.5	27.2	22.3	17.2	12.9	9.8	6.8	4.7	3.4	2.4
90%	1.67	84.7	83.9	**81.8**	**77.7**	71.5	63.6	54.4	44.3	34.7	**25.7**	**17.8**	11.8	7.3	4.5	2.5
5th	2.53	96.3	95.9	94.9	92.6	89.2	**83.3**	**75.1**	64.2	51.6	38.5	**26.9**	**17.9**	11.1	6.5	3.6
4th	2.54	93.4	93.0	91.5	88.4	**83.3**	**75.5**	65.1	52.3	39.0	**26.6**	**17.4**	10.6	6.0	3.1	1.5
3rd	3.40	99.2	99.2	99.2	98.6	97.8	96.1	92.7	**86.0**	**76.2**	64.0	50.4	36.7	**25.2**	**16.7**	10.7
2nd	3.43	97.6	97.3	96.9	95.5	93.6	89.4	**82.0**	71.1	58.3	44.5	31.5	**20.5**	12.6	7.4	4.3
Best	3.54	88.2	87.9	86.5	**83.9**	**78.7**	70.2	59.1	46.4	33.8	**22.6**	14.2	8.3	4.6	2.5	1.2

As explained in the last paragraph of 'Extension of the bootstrap procedure: Injecting alpha' under section 4, we can also use the simulation results for drawing inference on unlikely/rejectable levels of σ. Here, we accept an approximate ~20% change of setting a lower bound for σ that is too high, and similarly an approximate ~20% chance of setting upper bound for that is too low, which are same thresholds as used by Fama and French (2010). The appropriate interval for σ is different from the left to the right tail, and varies within the tails depending on which points in the tails are used. The distance from the estimated σ, or the size of the interval, is much more consistent than its position; we find that using a band of plus and minus 0.75% fits fairly well. In the left tail, this implies a band ranging from ~1.25% to ~2.75% for the lower estimates of σ and from ~1.75% to ~3.25% for the higher estimates of σ. In the right tail, it implies a band ranging from ~1.0% to ~2.5% for the lower estimates of σ and from ~1.5% to ~3% for the higher estimates of σ. For the best estimate of 2.25% in the left tail, the band ranges from 1.5% to 3%, suggesting that values of σ below or above this are not very likely. For the right tail, with a best estimate of 2%, the band ranges from 1.25% to 2.75%.

Persistence in performance

In the following section we will discuss the result from tests of persistence in returns of the Norwegian mutual fund market.

In order to disclose any performance persistence among Norwegian mutual funds, we chose to make use of the methodology put forward by Brown et al. (1992), Goetzmann and Ibbotson (1994) and Brown and Goetzmann (1995) using a Cross Product Ratio *(CPR)* test. The results could be found in Table A4 and Table A5 in the appendix. The test is used to analyze the persistency relative to a benchmark among mutual funds, with a nonparametric methodology based on contingency tables. In our tests we use two benchmarks. For the first one, the median return of funds that exist throughout the year in which the test is conducted is calculated. This is used as a benchmark which allows us to test whether there is a tendency that some funds persistently outperform or underperform relative to its peers. Another possible benchmark is simply the market return. This allows us to test for a tendency of funds persistently out-performing or losing relative to the market.

Each year mutual funds are sorted in one of four categories. Two categories comprise persistent performing funds, namely winner-winner and loser-loser, which comprise funds that persistently outperform or underperform the benchmark respectively in two subsequent years. The remaining categories comprise the non-persistent performers, namely winner-loser and loser-winner, which contain funds that change their relative ranking between two consecutive years. Each year we compute a CPR-value, which is defined as:

$$CPR = \frac{N_{WW} \cdot N_{LL}}{N_{WL} \cdot N_{LW}} \tag{14}$$

In which N_{WW} is the numbers of funds in the WW-group and so on. This is the odds ratio with the product of the two persistent performing groups divided by the product of the two groups which does not repeat prior year's relative ranking. In a large sample consisting of independent observations, we would expect an unconditional probability 25% for ending up in one of the groups. This will result in a CPR-value equal to one. Alternatively, if the data does show a tendency of persistence we would expect a greater product in the nominator, and hence a CPR greater than one. Conversely, a tendency of reversal would manifest as a CPR-value between zero and one. Table A4 and Table A5 reports the results from Z-tests with H_0 of no persistence, corresponding to a CPR-value equal to 1. In Table A4 the median return is used as benchmark, i.e. test if funds persistently outperform or underperform relative to peers. The test fails each time at least one of the categories amount to zero, and the result is reported as *NA.* The Z-statistic is significant at a 5% level in only 6 of a total of 32 years. A remarkably lower fraction than the results of e.g. Brown and Goetzmann (1995). Furthermore, half of these test results are significantly negative, proposing a reversal. Even though 3 out of 32 is slightly higher than what we would expect in such a test if returns were purely random, we do not feel safe to conclude that the Norwegian mutual fund market exhibit persistence in performance relative to peers.

Table A5 shows the results from the test using the market return as the relevant benchmark. The tests for persistence report a significant result in only 2 of the 32 years, which are both statistically positive. This is close to what we would expect only by chance, and consequently we do neither find evidence of persistent performance of fund managers relative to the market index.

These results do not contradict our findings in the bootstrapping analysis. Even though some funds exhibited a significantly positive alpha term, several of these are funds with relatively short return histories. The average lifetime of the three best performing funds is only just below two years, and their presence coincides with a significant and positive Z-statistic.

7. Conclusion

As expected, for the Norwegian mutual fund managers as a whole, we do not find any evidence that managers in general possess sufficient skills necessary to cover the costs imposed on investors. Some active fund managers have been able to outperform the market, but in accordance with the equilibrium accounting theory, they do so at the expense of other active managers.

In our bootstrapping analysis we find evidence of significant inferior skills among the worst performing funds and of significant superior skills among the best performing funds: In the left tail of the distribution, the results are discouraging, and we conclude that there is some existence of significant 'bad skill' (after fees are deducted) or value destruction. In the right tail however, we do find evidence of superior performance among the top performers. They have delivered returns better than what would be expected purely by luck. In line with the Efficient Market Hypothesis, we do not find evidence in the data suggesting any general persistence in performance. This implies a difficulty for investors in making any reliable ex ante decisions of which fund to choose.

Our results deviates from those of Sørensen (2009) mostly with respect to the best performing funds. His study finds week evidence of some superior performance among the very best funds, while the best funds perform very well in our data. We believe this difference is mainly due to the fact that the majority of the best performing funds in our dataset initiated after 2008. However, we have some empirical complications with our best funds, due to the fact that six of the ten best funds have existed for less than three years. With fewer observations superior performance becomes more likely, and this performance may not be due to actual skills.

With the assumption that true alpha is approximately symmetric around 0, we suggest that the level of the annual standard deviation of true α is likely to be around 1.75%-2.50% and relatively unlikely to be less than 1.00% or more than 3.25%.

Since the data we have used is based on net returns, it difficult to draw any conclusion whether fund managers as a group are actually skilled or unskilled. It is a possibility that there exist several highly skilled managers, who deliver significantly positive alphas measured on gross return, but not on net returns. The weakness with net returns is that high positive skill may be concealed by high fees, and hence superior skill will not be detected in our bootstrapping methodology. Overall however, this is not a major problem, as the main question posed instead becomes whether fund managers, as a group and individually, possess skills sufficient to defend their actual levels of fees. And, for an investor, this is actually more interesting than skill vs. luck per se. We do however encounter some problems using net returns when trying to estimate the σ of true α, due to the assumption that the mean is equal to 0, which would be more reasonable for gross returns.

Future research proposals

In our opinion, future research on this particular topic would benefit the most from concentrating on improving the existing dataset. One important improvement would be to include gross returns. This could potentially provide powerful contributions such as a more solid conclusion of the existence of actual fund manager skill (instead of skill to cover costs), and also reveal a possible relationship between skills and fees (i.e. if managers with higher skills are able to charge higher fees). Moreover, extending the dataset by including each fund’s assets under management would allow for the construction of value weighted returns, as well sorting the funds by size. A further improvement of the data would be to map each individual fund's holdings and portfolio weights. This could, among other, be used to compute Tracking Error and Active Share in order to make a better assessment of whether a fund is passive or active.

It would also be interesting to duplicate the results using a selection of different market benchmarks, e.g. using OSEAX over the whole period or OSEBX instead of OSEFX, in order to document the impact on performance measurement. Additionally, other model specifications might be investigated (for example by including the FF 3-factor model in the alpha injection), and one could look at how the results vary between different subsets of the time period.

In terms of extending the methodology, we believe that a study which implements conditional factor models, which allows for time-varying coefficients, also would generate valuable insight. Using unconditional models, as in this thesis, implicitly assumes the factor loadings to be constant for a fund throughout the whole lifetime, which is not realistic. It would also be interesting to extend the research to include all active Norwegian mutual funds, not just the ones with a Norwegian mandate.

Bibliography

Acharya, Viral V and Lasse Heje Pedersen. 2005. "Asset pricing with liquidity risk." *Journal of financial Economics* 77 (2): 375-410.

Banz, Rolf W. 1981. "The relationship between return and market value of common stocks." *Journal of financial economics* 9 (1): 3-18.

Bhandari, Laxmi Chand. 1988. "Debt/equity ratio and expected common stock returns: Empirical evidence." *Journal of finance*: 507-528.

Black, Fischer. 1972. "Capital market equilibrium with restricted borrowing." *The Journal of Business* 45 (3): 444-455.

Brown, Stephen J, William Goetzmann, Roger G Ibbotson and Stephen A Ross. 1992. "Survivorship bias in performance studies." *Review of Financial Studies* 5 (4): 553-580.

Brown, Stephen J and William N Goetzmann. 1995. "Performance persistence." *The Journal of finance* 50 (2): 679-698.

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

Cuthbertson, Keith, Dirk Nitzsche and Niall O'Sullivan. 2008. "UK mutual fund performance: Skill or luck?" *Journal of Empirical Finance* 15 (4): 613-634. doi: 10.1016/j.jempfin.2007.09.005.

Dahlquist, Magnus, Stefan Engström and Paul Söderlind. 2000. "Performance and characteristics of Swedish mutual funds." *Journal of Financial and quantitative Analysis* 35 (03): 409-423.

Dimson, Elroy and Massoud Mussavian. 1998. "A brief history of market efficiency." *European Financial Management* 4 (1): 91.

Fama, Eugene F and Kenneth R French. 1998. "Value versus growth: The international evidence." *The journal of finance* 53 (6): 1975-1999.

Fama, Eugene F. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *The Journal of Finance* 25 (2): 383-417. doi: 10.2307/2325486.

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

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

Gjerde, Øystein and Frode Sættem. 1991. "Performance evaluation of Norwegian mutual funds." *Scandinavian Journal of Management* 7 (4): 297-307.

Goetzmann, William N and Roger G Ibbotson. 1994. "Do winners repeat?" *The Journal of Portfolio Management* 20 (2): 9-18.

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

Jensen, Michael C. 1968. "The Performance of mutual funds in the period 1945-1964." *Journal of Finance* 23 (2): 389-416.

———. 1978. "Some anomalous evidence regarding market efficiency." *Journal of financial economics* 6 (2): 95-101.

Koller, Tim, Marc Goedhart and David Wessels. 2010. *Measuring and Managing the Value of Companies*.

Kosowski, Robert, Allan Timmermann, Russ Wermers and Hal White. 2006. "Can Mutual Fund "Stars" Really Pick Stocks? New Evidence from a Bootstrap Analysis." *Journal of Finance* 61 (6): 2551-2595.

Lintner, John. 1965. "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets." *The review of economics and statistics*: 13-37.

Markowitz, Harry M. 1959. *Portfolio Selection: Efticient Diversification of Investments*.

Næs, Randi, Johannes Skjeltorp and Bernt Arne Ødegaard. 2009. "What factors affect the Oslo Stock Exchange." *Norges Bank (Central Bank of Norway), Working Paper*.

Newey, Whitney K and Kenneth D West. 1986. A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix. National Bureau of Economic Research Cambridge, Mass., USA.

Newey, Whitney K. and Kenneth D. West. 1987. "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix." *Econometrica* 55: 703.

Ødegaard, Bernt Arne. 2016. "Asset pricing data at OSE." http://finance.bi.no/~bernt/financial_data/ose_asset_pricing_data/index.html.

OSE. 2016. "Oslo Børs Benchmark Index." 2016. http://www.oslobors.no/ob_eng/Oslo-Boers/Products-and-services/Market-data/Equity-indices.

Rosenberg, Barr, Kenneth Reid and Ronald Lanstein. 1985. "Persuasive evidence of market inefficiency." *The Journal of Portfolio Management* 11 (3): 9-16.

Sadka, Ronnie. 2006. "Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk." *Journal of Financial Economics* 80 (2): 309-349.

SEC. *"Mutual Funds. A Guide for Investors"*. U.S. Securities and Exchange Commission. https://www.sec.gov/investor/pubs/sec-guide-to-mutual-funds.pdf.

Sharpe, William F. 1964. "Capital asset prices: A theory of market equilibrium under conditions of risk." *The journal of finance* 19 (3): 425-442.

Sharpe, William F. 1991. "The arithmetic of active management." *Financial Analysts Journal* 47 (1): 7-9.

Sørensen, Lars Qvigstad. 2009. "Mutual fund performance at the Oslo Stock Exchange." *Available at SSRN 1488745*.

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

Strøm, Karl Oscar. 2014. "Hvordan velge riktig fond, og tjene mer penger?". Nordnet Accessed 08.01.2016 2016. http://www.nordnetbloggen.no/hvordan-velge-riktig-fond-og-tjene-mer-penger/15/05/2014/.

Summers, Lawrence H. 1986. "Does the Stock Market Rationally Reflect Fundamental Values." *Journal of Finance* 41 (3): 591-601.

VFF. 2015. "Tenk helhet, spre risiko." 2016. http://www.altomfond.no/Fondshandboken/Hva_MA_du_vite+/Tenk_helhet_og_spre_risiko/.

Appendix

Figure A1: Norwegian 1-month risk free rate

This figure shows the evolvement of the 1-month Norwegian risk free rate with daily observations. Data are retrieved from the Norwegian Central Bank.

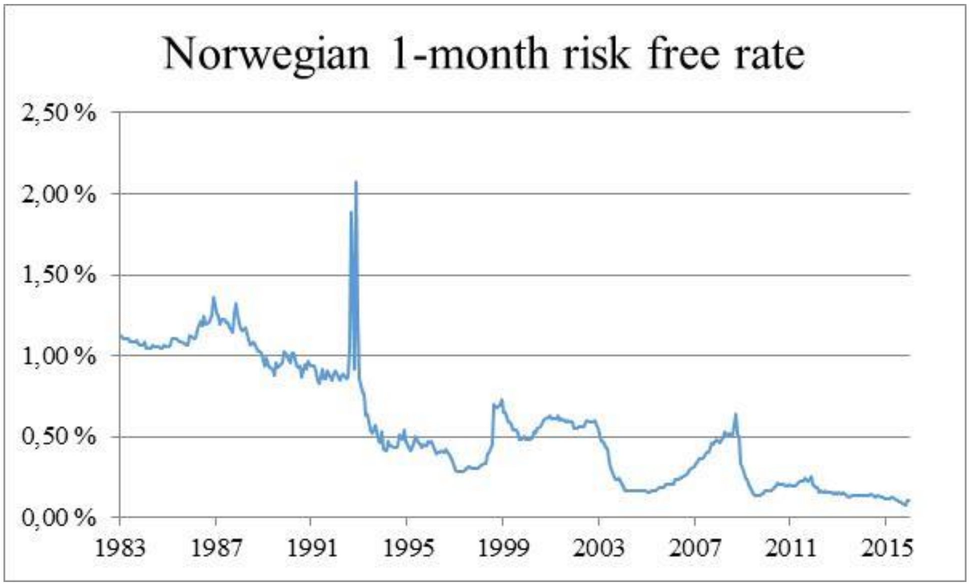


Figure A2: Number of actively managed Norwegian mutual funds

This illustration shows the involvement of the number of funds included in our dataset.

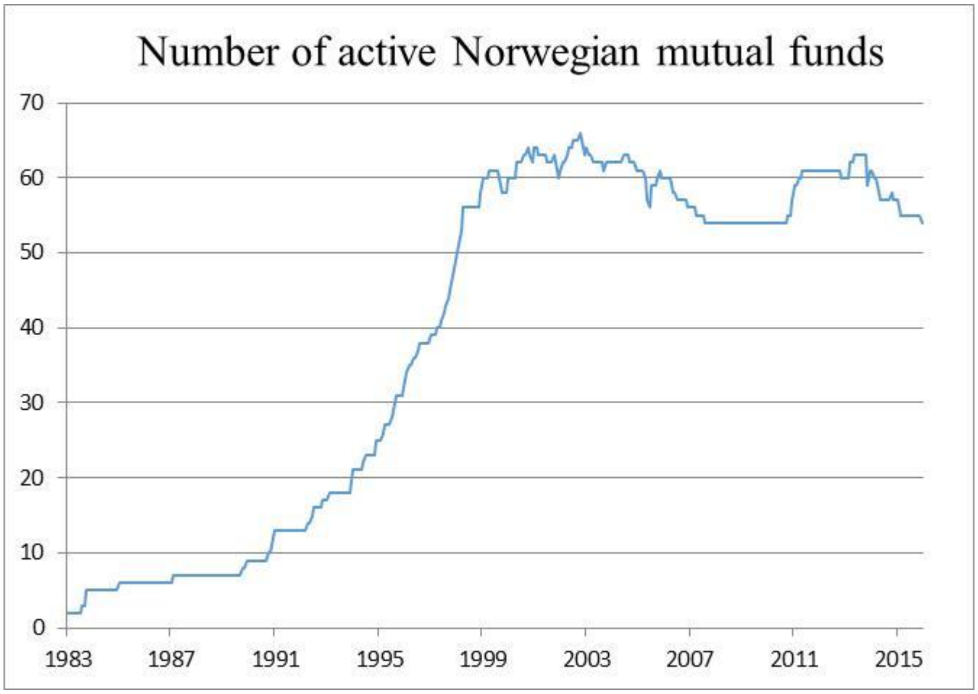


Figure A3: Return on an equal weighted portfolio consisting of funds active at the end of 2015 and an equal weighted portfolio comprising all funds

The blue line shows the evolvement of a1 NOK investment in an equal weighted portfolio consisting of the all funds in our data set. The dark line is a similar investment in a portfolio in which all defunct funds are excluded.

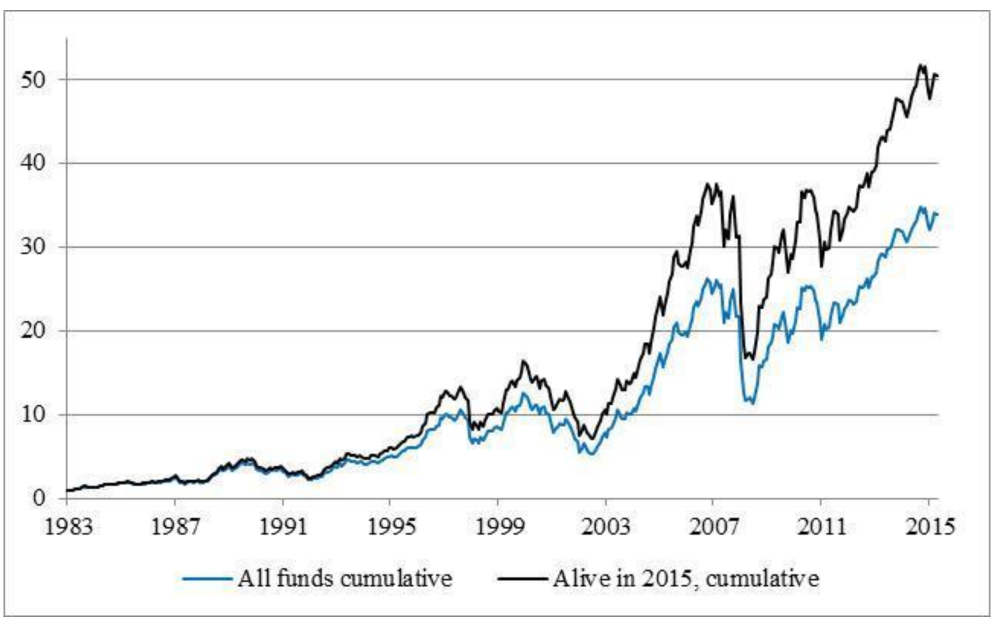


Table A1: Descriptive statistics on Norwegian mutual fund industry

This table shows selected statistics for Norwegian mutual funds which invests solely in Norwegian assets. All data, except the rightmost column, concerns the complete Norwegian mutual fund industry, hence passively managed funds are included. The third column shows the proportion of Norwegian funds which solely invests in Norway as a percentage of the total Norwegian mutual fund industry which includes Norwegian funds investing in foreign assets.

	Number of Norwegian Mutual Funds	Pct of total equity fund market	Assets under management	Total net inflow	Average fund size	Number of funds included *(active)*
1995	68	91.9	14 929	518	220	31
1996	60	86.1	26 512	5930	442	38
1997	76	80.1	45 892	10621	604	48
1998	81	67.3	32 636	357	403	58
1999	65	46.1	37 220	453	573	61
2000	76	38.3	34 915	-1686	459	65
2001	73	37.0	27 280	-813	374	66
2002	79	37.1	17 007	-843	215	67
2003	72	35.8	25 278	-116	351	65
2004	72	31.8	30 336	-3745	421	63
2005	76	26.2	38 283	-4652	504	63
2006	80	24.5	50 767	1267	635	62
2007	71	23.1	52 941	-3110	746	56
2008	71	19.7	25 482	-70	359	54
2009	70	24.8	57 540	10730	822	54
2010	73	26.6	77 576	4362	1 063	57
2011	73	24.6	60 800	-1305	833	61
2012	72	24.5	68 060	-725	945	61
2013	75	22.4	81 548	-945	1 087	65
2014	78	20.9	85 055	-1919	1 090	63
2015	78	19.9	86 746	-3928	1 112	57

Source: Norwegian Fund and Asset Management Association (VFF)
All relevant numbers in NOK millions

Table A2: t-test for difference in means

This table shows descriptive statistics and test results for the analysis considering survivorship bias, testing for a difference in means. *Defunct by 201512* and *Extant by 201512* is portfolios comprising funds that were "dead/alive" at the end of our dataset, respectively. Hence, the two portfolios combined makes up the *Whole sample*-portfolio. The t-test is carried out in three different starting points. The first test covers the whole timespan considered. The second test starts at the beginning of 1996 which is consistent with the beginning of OSEFX. The last test starts subsequent to the Global Financial Crisis. Still, findings of a survivorship bias are consistent across the three tests.

	# Obs	Mean (Pct/Month)	Std. Dev.	t-value
Excess return, time period 198301 - 201511				
Defunct by 201512	5 682	-0.14	2.32	-4.53
Extant by 201512	9 672	0.09	2.23	4.11
Whole sample	15 354	0.01	2.26	0.40
Difference		-0.23		-6.14
Excess return, time period 199601 - 201511				
Defunct by 201512	4 686	-0.17	2.32	-4.86
Extant by 201512	8 905	0.08	2.10	3.61
Whole sample	13 591	-0.004	2.18	-0.23
Difference		-0.25		-6.23
Excess return, time period 200901 - 201511				
Defunct by 201512	868	-0.22	1.81	-3.55
Extant by 201512	3 934	-0.01	1.60	-0.54
Whole sample	4 802	-0.05	1.64	-2.15
Difference		-0.20		-3.33

Table A3: Average yearly total and excess return for all, extant, and defunct funds

This table shows the equal weighted average return calculated for each year for three samples. The first sample comprises the entire dataset. The two following are subsamples extracted from the total sample. The sample in the middle of the table contains all the funds which were still operating at the end of 2015, whereas the last sample consists of funds which had closed down prior to this date. December 2015 are excluded from all samples since none of the defunct funds were operating at this point in time.

	Whole sample		Active 201512		Defunct 201512	
	Total return	Excess return	Total return	Excess return	Total return	Excess return
1983	5.15	0.51	1.74	-0.70	5.85	0.77
1984	3.26	1.25	3.30	1.30	3.22	1.22
1985	1.94	-0.49	2.16	-0.27	1.72	-0.71
1986	-0.27	0.41	-0.06	0.63	-0.49	0.19
1987	0.29	0.73	1.14	1.55	-0.36	0.10
1988	2.13	-0.69	1.41	-1.42	2.68	-0.14
1989	4.15	0.34	4.65	0.74	3.81	0.06
1990	-1.45	-0.26	-1.36	0.08	-1.49	-0.45
1991	-0.65	-0.08	-0.92	-0.35	-0.48	0.10
1992	-1.09	-0.42	-0.66	-0.14	-1.34	-0.59
1993	4.87	0.46	5.03	0.63	4.77	0.36
1994	0.48	-0.29	0.48	-0.28	0.49	-0.30
1995	1.43	0.43	1.78	0.73	1.08	0.11
1996	2.97	0.52	3.12	0.67	2.79	0.32
1997	2.37	-0.08	2.83	0.24	1.87	-0.43
1998	-2.80	-0.66	-2.90	-0.83	-2.72	-0.53
1999	3.80	0.28	4.05	0.51	3.62	0.12
2000	0.42	0.11	0.63	0.30	0.25	-0.04
2001	-1.18	0.00	-0.99	0.14	-1.36	-0.13
2002	-3.27	-0.29	-3.02	-0.07	-3.54	-0.53
2003	3.78	-0.08	3.75	-0.15	3.81	0.02
2004	2.74	-0.16	2.87	-0.03	2.56	-0.34
2005	3.31	0.41	3.47	0.53	3.03	0.20
2006	2.27	-0.23	2.42	-0.07	1.95	-0.55
2007	1.02	0.12	1.09	0.20	0.81	-0.12
2008	-5.46	0.51	-5.40	0.57	-5.67	0.31
2009	4.86	0.15	4.89	0.17	4.79	0.07
2010	1.88	-0.09	1.90	-0.09	1.83	-0.10
2011	-1.65	-0.06	-1.56	0.03	-1.93	-0.35
2012	1.33	-0.35	1.34	-0.32	1.26	-0.42
2013	1.82	-0.03	1.84	-0.01	1.76	-0.12
2014	0.66	0.18	0.72	0.24	0.03	-0.50
2015	0.74	-0.14	0.75	-0.13	0.53	-0.67

Table A4: Results from regression using model (4) for indidual actively managed Norwegian mutual funds

This table shows time series regression results for a three factor model estimated on net returns of the individual actively managed Norwegian mutual funds in our sample. The explanatory variables used are the market excess return (M), a size factor (SMB) and a liquidity factor (LIQ) (see 'Factor construction' under section 5 for descriptions of the factors). Results are shown for the top and bottom ten funds, ranked by the t-statistic of alpha. In addition to regression results, the table shows the following information about each fund from left to right: rank of fund (by t-stat of alpha), ticker of fund, name of fund, the number of returns and the time span the fund is present in our sample. Regression results shown are the intercept and coefficient estimates with corresponding t-statistics and the regression R^2. For the market slope, the t-statistics test whether the coefficient β_M is different from 1, while the other t-statistics test whether intercepts/coefficients are different from 0. We use the OLS estimator and standard errors corrected for heteroscedasticity and autocorrelation with the Newey and West (1986) procedure.

Rank	Ticker	Name	# of obs	Alive from	-	Alive to	Coefficients				T-statistics				R^2
							α	β_M	β_{SMB}	β_{LIQ}	$t(\alpha)$	$t(\beta_M)$	$t(\beta_{SMB})$	$t(\beta_{LIQ})$	
101	KF-SMBII	Nordea SMB II	69	1997M07	-	2003M03	-0.0148	0.95	0.66	-0.11	-3.90	-0.56	4.54	-0.85	76.9%
100	SK-SMB	Skandia SMB Norge	96	1994M12	-	2002M11	-0.0119	0.95	0.60	-0.25	-3.51	-0.72	6.49	-2.21	85.6%
99	SU-NORGE	Globus Norge II	94	1998M12	-	2006M11	-0.0108	1.12	0.42	-0.33	-3.15	1.81	6.47	-4.13	87.4%
98	KF-SMB	Nordea SMB	212	1997M06	-	2015M01	-0.0062	1.03	0.56	-0.03	-3.03	0.57	8.13	-0.47	82.0%
97	SU-GLNO	Globus Norge	103	1998M03	-	2006M11	-0.0098	1.09	0.45	-0.27	-3.00	1.00	6.19	-2.53	87.3%
96	GF-AKSJE	GJENSIDIGE AksjeSpar	151	1987M02	-	1999M08	-0.0033	0.95	0.06	-0.02	-2.97	-1.78	1.52	-0.52	93.1%
95	DI-RVKST	DnB Real-Vekst	156	1989M12	-	2002M11	-0.0036	0.94	0.10	0.02	-2.85	-3.49	3.91	0.67	95.8%
94	DK-NORII	Avanse Norge (II)	286	1991M01	-	2014M10	-0.0018	0.94	0.05	-0.04	-2.69	-3.25	2.26	-1.73	97.5%
93	EO-NORDN	Nordnet Superfondet Norge	15	2014M10	-	2015M12	-0.0043	1.08	-0.04	0.01	-2.60	1.37	-0.51	0.17	97.7%
92	GF-INVES	GJENSIDIGE Invest	103	1992M04	-	2000M10	-0.0044	1.00	0.13	0.06	-2.59	0.12	2.55	1.36	93.2%
10	KF-NOEQM	Nordea Norw Eq Mark Fund	123	2005M10	-	2015M12	0.0016	0.92	-0.16	0.05	1.71	-2.67	-5.28	1.63	97.5%
9	FV-TRNDR	FORTE Tr?nder	32	2013M05	-	2015M12	0.0055	0.69	0.01	-0.02	1.76	-2.48	0.03	-0.19	53.8%
8	SP-VERDI	Storebrand Verdi	216	1998M01	-	2015M12	0.0022	0.94	-0.07	0.13	1.81	-1.99	-1.54	2.54	91.4%
7	AC-NWECD	Arctic Norwegian Equities Class D	34	2013M03	-	2015M12	0.0054	0.74	0.00	0.00	2.02	-2.28	0.02	0.04	74.2%
6	FF-NOAI2	Danske Fund Norge Aksj. Inst 2	109	2006M12	-	2015M12	0.0022	0.88	0.10	-0.16	2.09	-2.78	2.57	-2.65	97.1%
5	IS-UTBYT	Landkreditt Utbytte	34	2013M03	-	2015M12	0.0062	0.79	0.05	0.33	2.53	-2.45	0.69	4.22	69.3%
4	FF-NOIII	Danske Fund Norge Aksj. Inst 1	188	2000M05	-	2015M12	0.0020	0.89	0.03	-0.09	2.54	-3.83	1.04	-2.08	97.3%
3	AI-NORGI	Alfred Berg Norge Inst	20	2014M05	-	2015M12	0.0073	0.85	0.11	0.02	3.40	-1.94	1.33	0.29	91.3%
2	OR-INVFB	Omega Investment Fund B	25	2013M12	-	2015M12	0.0118	0.47	0.14	-0.06	3.43	-4.01	1.12	-0.42	44.8%
1	OR-INVFC	Omega Investment Fund C	25	2013M12	-	2015M12	0.0121	0.47	0.14	-0.06	3.54	-4.02	1.12	-0.42	44.8%

Table A5: Ranks and percentiles of α- and t(α)-estimates for actual and simulated mutual fund returns based on regression (4)

Panel A of this table shows estimated values of α at selected ranks and percentiles for actual fund returns of actively managed Norwegian mutual funds, while Panel B shows estimated values of t(α). The panels are produced separately and α-values of specific ranks do not necessarily correspond t(α)-values of the same rank. The simulated average is the average of α or t(α) at selected percentiles from the simulation. The % < Act columns show the percentage of simulations runs which produce lower values of α or t(α) at the given rank/percentile than those observed for actual fund returns. The explanatory variables used to estimate actual and simulated α and t(α) are the market excess return (M), a size factor (SMB) and a liquidity factor (LIQ) (see 'Factor construction' under section 5 for descriptions of the factors). We use the OLS estimator and standard errors corrected for heteroscedasticity and autocorrelation with the Newey and West (1986) procedure.

	Panel A: α				Panel B: t(α)		
Rank/ percentile	Actual	Simulated average	% < Act	Rank/ percentile	Actual	Simulated average	% < Act
Worst	-0.0148	-0.0088	8.9	Worst	-3.90	-2.83	9.5
2nd	-0.0120	-0.0059	1.6	2nd	-3.51	-2.35	5.0
3rd	-0.0119	-0.0048	0.1	3rd	-3.15	-2.07	3.7
4th	-0.0118	-0.0041	0.0	4th	-3.03	-1.90	2.0
5th	-0.0108	-0.0037	0.0	5th	-3.00	-1.77	1.1
10%	-0.0061	-0.0023	0.0	10%	-2.55	-1.31	0.4
20%	-0.0032	-0.0013	0.3	20%	-1.77	-0.86	1.4
30%	-0.0021	-0.0008	1.5	30%	-1.30	-0.54	2.9
40%	-0.0014	-0.0004	2.9	40%	-0.86	-0.27	6.8
50%	-0.0009	0.0000	4.7	50%	-0.57	-0.02	7.5
60%	-0.0002	0.0003	15.0	60%	-0.11	0.23	18.9
70%	0.0003	0.0007	25.0	70%	0.35	0.49	35.6
80%	0.0012	0.0013	49.5	80%	0.80	0.81	50.1
90%	0.0026	0.0022	71.6	90%	1.67	1.25	84.7
5th	0.0055	0.0036	92.7	5th	2.53	1.67	96.3
4th	0.0062	0.0041	92.6	4th	2.54	1.80	93.4
3rd	0.0073	0.0048	91.8	3rd	3.40	1.96	99.2
2nd	0.0118	0.0060	97.9	2nd	3.43	2.19	97.6
Best	0.0121	0.0092	80.3	Best	3.54	2.67	88.2

Table A6: Performance persistence using median return as benchmark

This table shows the results from a performance persistence test based on the methodology of Brown and Goetzmann (1995). The *Total*-column reports the number of funds present in each specific year. *New*- and *Gone*-columns reports number of funds which entered or exited the market respectively. Funds in the *winner-winner* (*WW*) group have a yearly return greater than or equal to the median return among the funds present the specific year, and the subsequent year. Consequently, a fund in the *loser-winner* (*LW*) group in year 2005 had a return less than the median return in 2005 and greater or equal to the median in 2006, and so on. The *Cross product ratio (CPR)* is computed as the product of persistent funds *(WW * LL)* divided by the product of non-persistent funds *(LW * WL)*. The *Z-stat* is computed as: $\frac{\ln(CPR_t)}{\sigma_{\ln(CPR)_t}}$ in which $\sigma_{\ln(CPR)} = \sqrt{\frac{1}{N_{WW}} + \frac{1}{N_{LL}} + \frac{1}{N_{WL}} + \frac{1}{N_{LW}}}$ and the Z-stat reports the result in a test with H_0 of no persistence (i.e. *CPR* = 1). Significant test results are reported in **bold.**

	TOTAL	NEW	GONE	WINNER - WINNER	LOSER - LOSER	WINNER - LOSER	LOSER - WINNER	Cross product ratio	Z-stat
1983	5	3	0	0	0	1	1	0.00	*NA*
1984	5	0	0	1	1	2	1	0.50	-0.37
1985	6	0	0	3	3	0	0	*NA*	*NA*
1986	6	0	0	2	2	1	1	4.00	0.80
1987	7	1	0	1	0	2	3	0.00	*NA*
1988	7	0	0	2	1	2	2	0.50	-0.44
1989	9	2	0	3	2	1	1	6.00	1.06
1990	12	3	0	2	0	3	4	0.00	*NA*
1991	13	0	0	4	3	3	3	1.33	0.26
1992	17	4	0	3	3	4	3	0.75	-0.26
1993	18	1	0	5	5	4	3	2.08	0.74
1994	25	4	0	6	7	5	3	2.80	1.12
1995	31	6	0	8	8	5	4	3.20	1.39
1996	38	5	0	12	11	5	5	5.28	**2.20**
1997	48	9	0	6	7	14	12	0.25	**-2.03**
1998	58	8	0	7	9	18	16	0.22	**-2.49**
1999	61	1	3	18	18	11	10	2.95	**1.97**
2000	65	5	3	18	17	11	11	2.53	1.70
2001	66	2	6	18	16	11	13	2.01	1.31
2002	67	6	4	9	8	20	20	0.18	**-2.96**
2003	65	1	3	13	14	18	16	0.63	-0.89
2004	63	1	2	15	17	15	13	1.31	0.52
2005	63	2	3	13	16	16	13	1.00	0.00
2006	62	2	6	15	17	12	10	2.13	1.36
2007	56	0	2	15	15	12	12	1.56	0.81
2008	54	0	0	11	11	16	16	0.47	-1.35
2009	54	0	0	17	17	10	10	2.89	1.88
2010	57	3	0	11	12	16	15	0.55	-1.09
2011	61	2	0	14	13	16	16	0.71	-0.65
2012	61	0	1	16	18	14	12	1.71	1.03
2013	65	5	4	12	17	16	11	1.16	0.27
2014	63	2	6	19	20	9	7	6.03	**3.01**

Median return as benchmark

Table A7: Performance persistence using market return as benchmark

This table uses the same methodology as described for the table above, except that the relevant benchmark is switched to market return, instead of median fund return.

	TOTAL	NEW	GONE	WINNER - WINNER	LOSER - LOSER	WINNER - LOSER	LOSER - WINNER	Cross product ratio	Z-stat
1983	5	3	0	2	0	0	0	*NA*	*NA*
1984	5	0	0	0	0	5	0	*NA*	*NA*
1985	6	0	0	0	2	0	4	*NA*	*NA*
1986	6	0	0	4	1	0	1	*NA*	*NA*
1987	7	1	0	1	0	4	1	*NA*	*NA*
1988	7	0	0	2	1	0	4	*NA*	*NA*
1989	9	2	0	1	1	5	0	*NA*	*NA*
1990	12	3	0	1	3	0	5	*NA*	*NA*
1991	13	0	0	0	6	7	0	*NA*	*NA*
1992	17	4	0	0	6	0	7	*NA*	*NA*
1993	18	1	0	2	5	8	2	0.63	-0.41
1994	25	4	0	3	11	1	6	5.50	1.35
1995	31	6	0	12	4	1	8	6.00	1.48
1996	38	5	0	13	6	14	0	*NA*	*NA*
1997	48	9	0	2	14	13	10	0.22	-1.77
1998	58	8	0	7	20	10	13	1.08	0.12
1999	61	1	3	14	20	12	11	2.12	1.38
2000	65	5	3	17	12	9	19	1.19	0.32
2001	66	2	6	18	17	19	4	4.03	**2.16**
2002	67	6	4	4	28	16	9	0.78	-0.37
2003	65	1	3	3	33	10	15	0.66	-0.57
2004	63	1	2	15	10	4	31	1.21	0.28
2005	63	2	3	18	10	29	1	6.21	1.67
2006	62	2	6	16	13	4	21	2.48	1.37
2007	56	0	2	37	0	2	15	0.00	*NA*
2008	54	0	0	40	0	12	2	0.00	*NA*
2009	54	0	0	21	8	21	4	2.00	*NA*
2010	57	3	0	8	13	17	16	0.38	-1.69
2011	61	2	0	3	24	25	7	0.41	-1.19
2012	61	0	1	2	30	8	20	0.38	-1.17
2013	65	5	4	14	19	8	15	2.22	1.42
2014	63	2	6	22	16	12	5	5.87	**2.83**

Market return used as benchmark

MatLab Scripts / Code

The main MATLAB-code we have produced and used in our thesis is included below. Some of the code may rely on the 'Econometrics'- and the 'Statistics and Machine Learning'-toolboxes of MATLAB being activated. Given that the file 'z_data_for_import.mat' is stored in the current working directory, the code should work without modifications, but scripts need to be executed in the same order as given below. We thank Matteo Ottaviani, Education Account Manager at MathWorks, for providing us with MATLAB licences.

1. *Constructing matrix of dependent variables (fund returns)*

```
%% Script that loads data, extracts fund returns and constructs matrix ...'Y_all' of
excess fund ruturns.
% Raw fund returns are constructed in excel prior to importing to matlab.
% The imported returns are already limited to the specific period we are
% looking at, 396 months beginning with Jan 1983 and ending with Dec 2015.

load('z_data_for_import.mat'); % Load the file containing tables with data.
% |||| Contact 'mathias.krafft@outlook.com' for access to the file. |||

date_begin = 198301; % First date, format: YYYYDD
date_end = 201512; % Last date, format: YYYYDD (max=201511 if UMD is a X-factor).

% Create date-series, extracted from fund returns table:
dates = fund_r.date(fund_r.date >= date_begin & fund_r.date <= date_end);

% Convert tables to array, limited to the relevant dates:
fund_returns = table2array...
    (fund_r(fund_r.date >= date_begin & fund_r.date <= date_end,2:end));
rf_returns = risk_free_r.Rf1m...
    (risk_free_r.date >= date_begin & risk_free_r.date <= date_end);

ii = size(fund_returns,2); % Number of funds in data matrix
n = size(dates,1); % Total number of periods in data matrix

Y_all = fund_returns - repmat(rf_returns,1,ii); % Calculate excess returns

n_i = sum(~isnan(Y_all),1); % Number of periods per fund

%% Option to convert 'Y_all' to series of equally weighted fund returns (remove '%'
below):
% Calculate EW returns:  |     Redefine 'ii'   |   Redefine 'n_i'    %
% Y_all = nanmean(Y_all,2) ;  ii = size(Y_all,2) ;  n_i = sum(~isnan(Y_all),1);

%% Option to include first 20 observations only of all funds (remove % from for-loop):
% for i_i = 1:ii
%     a = 0;
%     for nn = 1:n
%         if isnan(Y_all(nn,i_i)) == 0
%             a = a + 1;
%             if a > 20
%                 Y_all(nn,i_i) = NaN;
%             end
%         end
%     end
% end

%% Option to include last 20 observations only of ALL funds:
% for i_i = 1:ii
%     a = 0;
%     for nn = 1:n
%         if isnan(Y_all(nn,i_i)) == 0
%             a = a + 1;
%             if a < n_i(i_i)-20+1
%                 Y_all(nn,i_i) = NaN;
%             end
%         end
%     end
% end
```

```
%% Option to include last 20 observations only of SURVIVING funds:
% for i_i = 1:ii
%     a = 0;
%     if isnan(Y_all(n,i_i)) == 1
%         Y_all(:,i_i) = NaN;
%     end
%     for nn = 1:n
%         if isnan(Y_all(nn,i_i)) == 0
%             a = a + 1;
%             if a < n_i(i_i)-20+1
%                 Y_all(nn,i_i) = NaN;
%             end
%         end
%     end
% end

clear i_i nn a
```

2. *Constructing matrix of explanatory variables (factor returns)*

```
%% Script that extracts factor returns from tables and construct explanatory variable
matrix 'X-mat'.
% Extract market and risk free returns:
mrkt_returns = market_r.OSE_AX_FX...
    (market_r.date >= date_begin & market_r.date <= date_end);
mrkt_prem = mrkt_returns-rf_returns; % Construct excess returns

% Extract other risk factors (add/remove '%' to disclude/include factors);
SMB = more_factors_r.SMB(more_factors_r.date >= date_begin & more_factors_r.date <=
date_end);
% HML = more_factors_r.HML(more_factors_r.date >= date_begin & more_factors_r.date <=
date_end);
% UMD = more_factors_r.UMD(more_factors_r.date >= date_begin & more_factors_r.date <=
date_end);
LIQ = more_factors_r.LIQ(more_factors_r.date >= date_begin & more_factors_r.date <=
date_end);

X_mat = [mrkt_prem SMB LIQ]; % Construct 'X-mat'. Include variables from above.

k = size(X_mat,2); % Calculate number of factors:

clear SMB HML UMD LIQ % Clear "excess" variables
clear date_begin date_end dates % Clear "excess" variables
clear fund_r market_r more_factors_r risk_free_r % Clear "excess" variables
clear fund_returns mrkt_returns mrkt_prem ref_returns % Clear "excess" variables
```

3. *Performing initial regressions on actual fund returns (HAC)*

```
%% Script that constructs HAC benchmark regression results for all funds.
tic; % Start timing
% Populate target output vectors to be filled with loop:
orig_SE_coefs = NaN(k+1,ii); orig_coefs = NaN(k+1,ii);

n_i = sum(~isnan(Y_all),1); % Redifine number of periods per fund

% Calculate the lag selection parameter for the standard HAC Newey-West
% (1994) plug-in procedure:
maxLag = floor(4*(n_i/100).^(2/9));

% Loop through all funds and estimate the standard Newey-West OLS
% coefficient covariance using 'hac(__)' by setting the bandwidth to
% 'maxLag+1'. Exctract the OLS coefficient estimates and their standard
% errors.
for jj = 1:ii
    if (n_i(jj) ~= 0)
        [~ , orig_SE_coefs(:,jj), orig_coefs(:,jj)] = ...
            hac(X_mat,Y_all(:,jj),'display','off'),'bandwidth',maxLag(jj)+1);
    end
end

% Calculate t-statistics (with the null that coefficient = 0).
orig_t_stats = orig_coefs./orig_SE_coefs;

% Calculate standard error of residuals of all funds:
orig_resids = Y_all - [ones(size(X_mat,1),1), X_mat]*orig_coefs;
orig_SSR = nansum((orig_resids.^2),1);
orig_SE_resid = sqrt(orig_SSR ./ (n_i - ones(1,ii)*(k+1))); clear orig_SSR

toc; % End timing
clear n_i jj MaxLag % Clear temporary / redundant variables

% Sort alphas and provide index of sorted alpha:
[sort_a, sort_index_a] = sort(orig_coefs(1,:),'ascend');
% Sort t-stats of alphas and provide index of sorted t-stats:
[sort_t, sort_index_t] = sort(orig_t_stats(1,:),'ascend');
```

4. *Constructing simulation indices*

```
%% Script that creates the random number indices used for all our simulations
rng(0); % Set RNG for reproducibility
s = 10000; % Set number of simulations

% 'n' is calculated in the script 'a1_..' and is equal to the total number
% of periods/months that we use in our calculations.

% Construct a matrix of 's' vectors of 'n' random integers ranging from
% minimum '1' to to maximum 'n':
sim_indices = randi(n,n,s);
```

5. Constructing series of simulated returns

```
%% Construct simulated series based on "sim_indices"
tic; % Begin timer
% This script used the simulated index numbers to:
%   1) Pick corresponding numbers from factors and residuals, and
%   2) Construct series of fund returns (potentially including injected alpha)
%   3) Series are "alpha free" if 'std_alpha' below is set to '0'.
%
% The constructed returns will be the basis for new regressions to
% calculate simulated alphas.

% From before:  ii = total number of funds
%               k = total number of factors
%               n = total number of time periods
%               s = total number of simulations, s = 1 here refers sim #1

% Check if the value for annual "average" standard deviation is already
% defined. If it is, dont't touch it. If it isn't, define a chosen
% value (usually '0') below. We do this to avoid overriding the std of alpha
% in the loop running through different values of std of alpha.
if exist('annual_std_alpha') == 0;
    annual_std_alpha = 0; % Set a desired value of std of alpha
end

std_alpha = annual_std_alpha / sqrt(12); % convert std to monthly
clear annual_std_alpha; % clear no longer needed variable

% Construct matrix of betas from coefficient matrix (excluding alphas):
orig_betas = orig_coefs(2:(k+1),:);

% Populate matrices for simulted residuals and factor- and fund returns:
constructed_X_mat = NaN(n,k,s);
constructed_resids = NaN(n,ii,s);
constructed_Y_all = NaN(n,ii,s);

temp_avg_orig_std_resid = mean(orig_SE_resid,2); % Find std of orig reg residuals
temp_std_resid_ratio = orig_SE_resid/temp_avg_orig_std_resid; % Ratio residual std to
avg resid std for each fund
rng(0); % reset random number generator
% construct series of alphas for injection (constant over time, scaled per
% fund, independent per simulation). These numbers become zero when we set
% desired injection of average annual alpha to zero above.
temp_alpha =
std_alpha*repmat(randn([1,ii,s]).*repmat(temp_std_resid_ratio,1,1,s),n,1,1);
clear      temp_avg_orig_std_resid      temp_std_resid_ratio

% Construct matrices of all simulated factor and fund returns:
for ss = 1:s
constructed_X_mat(:,:,ss) = X_mat(sim_indices(:,ss),:);
constructed_resids(:,:,ss) = orig_resids(sim_indices(:,ss),:);
constructed_Y_all(:,:,ss) = temp_alpha(:,:,ss) + constructed_X_mat(:,:,ss)*orig_betas
+ constructed_resids(:,:,ss);
end

clear ss; % Clear loop counter:
toc; % End timer
```

6. Performing regressions on simulated series

(N.B. this script may take long time, mainly due to the total number of total regressions (1.01 million) and because the hac-formula is slow)

```
% Script that does bootstrap regression results for all funds.
tic; % Begin timer

% Set minumum number of observations(n) required in simulation for the
% regression to be valid:
sim_cutoff = 15;

% Populate target output vectors to be filled in with loop:
sim_SE_resid = NaN(k+1,ii,s); sim_coefs = NaN(k+1,ii,s);

% Calculate number of observations per fund per simulation for future
% reference:
n_i_s = sum(~isnan(constructed_Y_all),1);

% Calculate the lag selection parameter for the standard Newey-West HAC
% estimate (Andrews and Monohan, 1992), one number per fund per simulation:
maxLag_s = floor(4*(n_i_s/100).^(2/9));

% Loop through each simulation, and for each simulation loop through each
% fund and use OLS-regression from sheet 'c1_..' on constructed
% returns/factors based simulated time indices. Fill in results in matrices
% defined above.

% Loop through each simulation run:
for ss = 1:s
    % Loop through each fund:
    for jj = 1:ii
        if n_i_s(1,jj,ss) >= sim_cutoff
            [~, sim_SE_resid(:,jj,ss), sim_coefs(:,jj,ss)] =
hac(constructed_X_mat(:,:,ss),constructed_Y_all(:,jj,ss),'bandwidth',maxLag_s(1,jj,ss)
+1,'display','off');
        end
    end
end

% Calculate t-statistics (with the null that coefficient = 0):
sim_t_stats = sim_coefs./sim_SE_resid;

clear n_i_s sim_cutoff jj ss; % Clear temporary / redundant variables:

toc; % End timer
```

7. *Estimating/constructing the bootstrapping results/tables*

```
% This script calculates the averages of alphas/t-stat in different
% ranks/percentiles of the simulation runs. Additionally, it estimates the
% percent of simulated alphas/t-stats of each rank/percentile, that are
% lower than the 'actual' (historically observed) alpha/t-stat at the same
% rank/percentile.

% NOTE: With 101 funds to calculate percentiles from, for the low and high
% percentiles, we rather use the value of funds with rank #1-5 and 97-101.
% For the 'middle' percentiles, we use matlab's 'prctile(___)'-function
% which interpolates linearly between observations.

% Construct matrix of relevant percentages (10 through 90);
percentages = [10,20,30,40,50,60,70,80,90]';

% Sort original alphas and t-values in order to extract top/bottom ranked
% values:
temp_sorted_orig_a = sort(orig_coefs(1,:))';
temp_sorted_orig_t = sort(orig_t_stats(1,:))';

% Construct tables of relevant ranks and percentiles of alpha & t-values:
percentiles_orig_a = ...
    [temp_sorted_orig_a(1:5) ; prctile(orig_coefs(1,:),percentages) ;
temp_sorted_orig_a((end-4):end)];
percentiles_orig_t = ...
    [temp_sorted_orig_t(1:5) ; prctile(orig_t_stats(1,:),percentages) ;
temp_sorted_orig_t((end-4):end)];
clear temp_sorted_orig_a temp_sorted_t_alpha

% Find averages of simulated ranks/percentiles for alphas and t-values:
% NOTE: For each simulation run, we find the percentiles corresponding to the
% percentages above and the top/bottom 5 ranked values. For each
% rank/percentile, we then take the average of the alphas/t-stats over
% all simulation runs with ('prctile'-formula treats 'NaN' values as
% missing and removes them, so this is not a problem). Some simulation runs
% may have less than 101 valid regressions (due to some funds with short
% original series and our required number of observations for a regression
% to be valid).

% Descending sort of simulated alphas per simulation (thorny due to NaN's
% which are considered to be 'largest' in matlab's sort function; we change
% the NaN's to nagative infinite to make the smallest instead of largest):
temp_sort_desc_sim_a = permute(sim_coefs(1,:,:),[2 3 1]);
temp_sort_desc_sim_a(isnan(temp_sort_desc_sim_a)) = -Inf;
temp_sort_desc_sim_a = sort(temp_sort_desc_sim_a,1,'descend');
temp_sort_desc_sim_a(isinf(temp_sort_desc_sim_a)) = NaN;
% Ascending sort of simulated alphas per simulation:
temp_sort_asc_sim_a = sort(permute(sim_coefs(1,:,:),[2 3 1]),1,'ascend');

% Descending sort of simulated t-stats per simulation (thorny due to NaN's
% which are considered to be 'largest' in matlab's sort function; we change
% the NaN's to nagative infinite to make the smallest instead of largest):
temp_sort_desc_sim_t = permute(sim_t_stats(1,:,:),[2 3 1]);
temp_sort_desc_sim_t(isnan(temp_sort_desc_sim_t)) = -Inf;
temp_sort_desc_sim_t = sort(temp_sort_desc_sim_t,1,'descend');
temp_sort_desc_sim_t(isinf(temp_sort_desc_sim_t)) = NaN;
% Ascending sort of simulated t-stats per simulation:
temp_sort_asc_sim_t = sort(permute(sim_t_stats(1,:,:),[2 3 1]),1,'ascend');

% Construct matrix containing the top/bottom 5 ranks and 9 different
% percentiles of alpha each simulation (result is a 21 x 10000 matrix):
temp_percentiles_sim_a = ...
        [temp_sort_asc_sim_a(1:5,:) ; ...
            prctile(permute(sim_coefs(1,:,:),[2 3 1]),percentages,1) ; ...
                temp_sort_desc_sim_a(5:-1:1,:)];
% Construct matrix containing the top/bottom 5 ranks and 9 different
% percentiles of t-stat in each simulation (result is a 21 x 10000 matrix):
temp_percentiles_sim_t = ...
    [temp_sort_asc_sim_t(1:5,:) ; ...
        prctile(permute(sim_t_stats(1,:,:),[2 3 1]),percentages,1) ; ...
            temp_sort_desc_sim_t(5:-1:1,:)];

% Take the means across each rank/percentile:
```

```
mean_percentiles_sim_a = mean(temp_percentiles_sim_a,2);
mean_percentiles_sim_t = mean(temp_percentiles_sim_t,2);

% Calcualte the percentage of simulated alphas in each rank/percentile that
% are smaller than actual alpha (and same for t-values):
sim_smaller_a = ...
    sum(temp_percentiles_sim_a < repmat(percentiles_orig_a,1,s),2) /s*100;
sim_smaller_t = ...
    sum(temp_percentiles_sim_t < repmat(percentiles_orig_t,1,s),2) /s*100;

% Combine the above in one table (one for alpha and one for t-stat):
table_a = array2table([[(101:-1:97)';percentages/100;(5:-1:1)'], ...
    percentiles_orig_a, mean_percentiles_sim_a, sim_smaller_a],...
    'VariableNames',{'Percentile_Rank','Act','Sim','Perc_sim_lower'});
table_t = array2table([[(101:-1:97)';percentages/100;(5:-1:1)'], ...
    percentiles_orig_t, mean_percentiles_sim_t, sim_smaller_t], ...
    'VariableNames',{'Percentile_Rank','Act','Sim','Perc_sim_lower'});

% Clear 'excess' variables no longer needed:
clear percentages
clear temp_sort_desc_sim_a temp_sort_asc_sim_a
clear temp_sort_desc_sim_t temp_sort_asc_sim_t
clear mean_percentiles_sim_a mean_percentiles_sim_t
clear sim_smaller_a sim_smaller_t
```

8. Generating CDF plots

```
% Generate CDF plot for alphas:
figure('name','Empirical CDF for actual and simulated alpha');
hold on
plot_orig = cdfplot(orig_coefs(1,:)*100);
plot_sim = cdfplot(mean(prctile(permute(sim_coefs(1,:,:),[2 3
1])*100,(100*0.5/ii:100/ii:100-50/ii),1),2));
hold off
title('Empirical CDF for actual and simulated alpha')
axis([min(orig_coefs(1,:))*100*1.2 max(orig_coefs(1,:))*100*1.2 0 1])
set(plot_orig,'LineStyle','-','Color','Black')
set(plot_sim,'LineStyle','--','Color','Black')
legend('Actual','Simulated','Location','NW')
xlabel('alpha(%)') % x-axis label
ylabel('F(alpha)') % y-axis label
clear plot_orig plot_sim;

% Generate CDF plot for alpha t-stats:
figure('name','Empirical CDF for actual and simulated t(alpha)');
plot_orig = cdfplot(orig_t_stats(1,:));
hold on
plot_sim = cdfplot(mean(prctile(permute(sim_t_stats(1,:,:),[2 3
1]),(100*0.5/ii:100/ii:100-50/ii),1),2));
hold off
title('Empirical CDF for actual and simulated t(alpha)')
axis([min(orig_t_stats(1,:))*1.2 max(orig_t_stats(1,:))*1.2 0 1])
set(plot_orig,'LineStyle','-','Color','Black')
set(plot_sim,'LineStyle','--','Color','Black')
legend('Actual','Simulated','Location','NW')
xlabel('t(alpha)') % x-axis label
ylabel('F(t(alpha))') % y-axis label
clear plot_orig plot_sim;
```

9. *Generating Kernel smoothing density estimate plots*

```
% Generate Kernel smoothing density estimate plot for alphas:
figure('name','Kernel smoothing density estimate for actual and simulated alpha');
title('Kernel smoothing density estimate for actual and simulated alpha')
hold on
[ksd_orig_a, x1] = ksdensity(orig_coefs(1,:)*100);
[ksd_sim_a, x2] = ksdensity(mean(prctile(permute(sim_coefs(1,:,:),[2 3
1]),(100*0.5/ii:100/ii:100-50/ii),1),2)*100);
plot(x1, ksd_orig_a,'Black-',x2, ksd_sim_a,'--Black');
hold off
legend('Actual','Simulated','Location','NW')
xlabel('alpha(%)') % x-axis label
ylabel('Frequency') % y-axis label
clear ksd_orig_a x1 ksd_sim_a x2;

% Generate Kernel smoothing density estimate plot for t-stats of alpha:
figure('name','Kernel smoothing density estimate for actual and simulated t(alpha)');
title('Kernel smoothing density estimate for actual and simulated t(alpha)')
hold on
[ksd_orig_t, x1] = ksdensity(orig_t_stats(1,:));
[ksd_sim_t, x2] = ksdensity(mean(prctile(permute(sim_t_stats(1,:,:),[2 3
1]),(100*0.5/ii:100/ii:100-50/ii),1),2));
plot(x1, ksd_orig_t,'Black-',x2, ksd_sim_t,'--Black');
hold off
legend('Actual','Simulated','Location','NW')
xlabel('t(alpha)') % x-axis label
ylabel('Frequency') % y-axis label
clear ksd_orig_t x1 ksd_sim_t x2;
```

10. *Generating histograms*

(Includes a subset of the code only, i.e. only the code for generating a histogram for the *best* fund)

```
%% Generate histogram of simulated t(alpha)'s for various ranks/percentiles:
% Includes vertical line representing the actual performance of fund at
% equivalent rank/percentile:

% Different index INPUTS used to make histograms from
% 'temp_percentiles_sim_t' and vertical lines from 'percentiles_orig_t'
% represent different ranks/percentiles to extract, as specified below:
% input: 1:5 - represents ranks: 101(worst):-1:97
% input: 6:14 - represents percentiles 10:10:90
% input: 15:19 - represents ranks 5:-1:1(best))

%% BEST FUND
figure('name','Bootstrapped t-statistics of t(alpha): Best fund');
title('Bootstrapped t-statistics of t(alpha): Best fund','FontSize',16)
hold on
% plot_orig = cdfplot(orig_coefs(1,:));
temp_input = 19; % SET INPUT value chosen from table at top
line_hight = 900; % ADJUST this to just above max of highest bar
line([percentiles_orig_t(temp_input) percentiles_orig_t(temp_input)],[0
line_hight],'LineWidth',2,'Color','Black','LineStyle','--');
plot_sim = histogram(temp_percentiles_sim_t(temp_input,:),...
    'BinLimits',[0.5,7]); % Adjust max/min values here (start with large numbers to
see range)
hold off;
plot_sim.FaceColor = 'k'; plot_sim.EdgeColor = 'k';
xlabel('t(alpha)','FontSize',14) % x-axis label
ylabel('Frequency','FontSize',14) % y-axis label
legend({['Actual t(alpha) = ' num2str(percentiles_orig_t(temp_input))],'Simulated
t(alpha)'},...
    'Location','NE','FontSize',14); % legend font size.
clear plot_sim temp_input
```