NTNU – Trondheim
Norwegian University of
Science and Technology

Performance and Persistence in Norwegian Mutual Funds

Helge Hoff Hansen
Eirik Solli Haukaas
Kristoffer Gallefoss

Industrial Economics and Technology Management
Submission date: June 2012
Supervisor: Peter Molnar, IØT

Norwegian University of Science and Technology
Department of Industrial Economics and Technology Management

NTNU
Det skapende universitet

SAMARBEIDSKONTRAKT

1. Studenter i samarbeidsgruppen

Etternavn, fornavn	Fødselsdato
Gallefoss, Kristoffer	**02. feb 1987**
Etternavn, fornavn **Hansen, Helge Hoff**	Fødselsdato **17. mar 1987**
Etternavn, fornavn **Haukaas, Eirik Solli**	Fødselsdato **12. aug 1985**

2. Hovedveileder

Etternavn, fornavn	Institutt
Molnar, Peter	**Institutt for industriell økonomi og teknologiledelse**

3. Masteroppgave

Oppgavens (foreløpige) tittel
Mutual Fund Performance: empirical evidence from Norwegian mutual funds

4. Bedømmelse

Kandidatene skal ha *individuell* bedømmelse ☐
Kandidatene skal ha *felles* bedømmelse ☐

TRONDHEIM, 13.1.2012
Sted og dato

Kristoffer Gallefoss
Kristoffer Gallefoss

Molnar
Hovedveileder

Helge Hoff Hansen

Eirik Solli Haukaas
Eirik Solli Haukaas

Originalen oppbevares på instituttet.

NTNU
Det skapende universitet

MASTERKONTRAKT
- uttak av masteroppgave

1. Studentens personalia

Etternavn, fornavn **Gallefoss, Kristoffer**	Fødselsdato **02. feb 1987**
E-post **kristoffer.gallefoss@gmail.com**	Telefon **40874656**

2. Studieopplysninger

Fakultet **Fakultet for Samfunnsvitenskap og teknologiledelse**	
Institutt **Institutt for industriell økonomi og teknologiledelse**	
Studieprogram **Industriell økonomi og teknologiledelse**	Hovedprofil **Investering, finans og økonomistyring**

3. Masteroppgave

Oppstartsdato **16. jan 2012**	Innleveringsfrist **11. jun 2012**
Oppgavens (foreløpige) tittel **Mutual Fund Performance: Empirical Evidence from Norwegian Mutual Funds**	
Oppgavetekst/Problembeskrivelse This thesis will investigate the determinants for performance in Norwegian mutual funds, based on theories from historic literature and observed anomalies in the industry. We use a new data set to perform such analysis on topics as: -macroeconomic effects -fee structure -capital inflow/outflow -cross-sectional dispersion (performance)	
Hovedveileder ved institutt **Post doktor Peter Molnar**	Medveileder(e) ved institutt
Merknader **1 uke ekstra p.g.a påske.**	

4. Underskrift

Student: Jeg erklærer herved at jeg har satt meg inn i gjeldende bestemmelser for mastergradsstudiet og at jeg oppfyller kravene for adgang til å påbegynne oppgaven, herunder eventuelle praksiskrav.

Partene er gjort kjent med avtalens vilkår, samt kapitlene i studiehåndboken om generelle regler og aktuell studieplan for masterstudiet.

Trondheim, 23.01.12
Sted og dato

Student

Hovedveileder

Originalen lagres i NTNUs elektroniske arkiv. Kopi av avtalen sendes til instituttet og studenten.

NTNU
Det skapende universitet

MASTERKONTRAKT
- uttak av masteroppgave

1. Studentens personalia

Etternavn, fornavn **Hansen, Helge Hoff**	Fødselsdato **17. mar 1987**
E-post **helgehof@stud.ntnu.no**	Telefon **97701788**

2. Studieopplysninger

Fakultet **Fakultet for Samfunnsvitenskap og teknologiledelse**	
Institutt **Institutt for industriell økonomi og teknologiledelse**	
Studieprogram **Industriell økonomi og teknologiledelse**	Hovedprofil **Investering, finans og økonomistyring**

3. Masteroppgave

Oppstartsdato **16. jan 2012**	Innleveringsfrist **11. jun 2012**
Oppgavens (foreløpige) tittel **Mutual Fund Performance: Empirical Evidence from Norwegian Mutual Funds**	
Oppgavetekst/Problembeskrivelse This thesis will investigate the determinants for performance in Norwegian mutual funds, based on theories from historic literature and observed anomalies in the industry. We use a new data set to perform such analysis on topics as: -macroeconomic effects -fee structure -capital inflow/outflow -cross-sectional dispersion (performance)	
Hovedveileder ved institutt **Post doktor Peter Molnar**	Medveileder(e) ved institutt
Merknader **1 uke ekstra p.g.a påske.**	

4. Underskrift

Student: Jeg erklærer herved at jeg har satt meg inn i gjeldende bestemmelser for mastergradsstudiet og at jeg oppfyller kravene for adgang til å påbegynne oppgaven, herunder eventuelle praksiskrav.

Partene er gjort kjent med avtalens vilkår, samt kapitlene i studiehåndboken om generelle regler og aktuell studieplan for masterstudiet.

Trondheim 23.01.12
Sted og dato

Student | **Hovedveileder**

Originalen lagres i NTNUs elektroniske arkiv. Kopi av avtalen sendes til instituttet og studenten.

NTNU
Det skapende universitet

MASTERKONTRAKT

- uttak av masteroppgave

1. Studentens personalia

Etternavn, fornavn	Fødselsdato
Haukaas, Eirik Solli	**12. aug 1985**
E-post	Telefon
eirikhaukaas@gmail.com	**92484170**

2. Studieopplysninger

Fakultet	
Fakultet for Samfunnsvitenskap og teknologiledelse	
Institutt	
Institutt for industriell økonomi og teknologiledelse	
Studieprogram **Industriell økonomi og teknologiledelse**	Hovedprofil **Investering, finans og økonomistyring**

3. Masteroppgave

Oppstartsdato **16. jan 2012**	Innleveringsfrist **11. jun 2012**
Oppgavens (foreløpige) tittel **Mutual Fund Performance: empirical evidence from Norwegian mutual funds**	
Oppgavetekst/Problembeskrivelse This thesis will investigate the determinants for performance in Norwegian mutual funds, based on theories from historic literature and observed anomalies in the industry. We use a new data set to perform such analysis on topics as: -macroeconomic effects -fee structure -capital inflow/outflow -cross-sectional dispersion (performance)	
Hovedveileder ved institutt **Post doktor Peter Molnar**	Medveileder(e) ved institutt
Merknader **1 uke ekstra p.g.a påske.**	

4. Underskrift

Student: Jeg erklærer herved at jeg har satt meg inn i gjeldende bestemmelser for mastergradsstudiet og at jeg oppfyller kravene for adgang til å påbegynne oppgaven, herunder eventuelle praksiskrav.

Partene er gjort kjent med avtalens vilkår, samt kapitlene i studiehåndboken om generelle regler og aktuell studieplan for masterstudiet.

Trondheim, 13.1.2012
Sted og dato

[signature]
Student

[signature]
Hovedveileder

Originalen lagres i NTNUs elektroniske arkiv. Kopi av avtalen sendes til instituttet og studenten.

Preface

This master thesis concludes our Master of Science degree in Industrial Economics and Technology Management at the Norwegian University of Science and Technology (NTNU). The thesis is written as an academic research paper in the article format, thus intended for publishing. The references follow the guidelines of the American Psychological Association. The purpose of this paper is to evaluate the performance and persistence in the Norwegian mutual fund industry. The preceding master contracts contain the preliminary problem description and title, which has changed during the course of the semester, in agreement with our supervisor.

First and foremost, we thank our supervisor Peter Molnar for his guidance, feedback and conversations that we have greatly benefitted from. Additionally, we thank Børsprosjektet NHH, Verdipapirfondenes Forening and Bernt A. Ødegaard for generously providing us with the data.

Kristoffer Gallefoss | Helge Hoff Hansen | Eirik Solli Haukaas

Oppsummering

Ved hjelp av et datasett uten overlevelsesskjevhet undersøker vi prestasjonen og prestasjonens persistens til norske aksjefond i perioden 2000-2010. For å evaluere prestasjonen til fondene bruker vi en multiperiodisk versjon av Carhart sin 4-faktormodell for å estimere Jensens alfa. Vi vurderer den statistiske signifikansen av alfaestimatene ved å sammenligne dem med alfafordelinger generert fra bootstrap-simuleringer. Vi finner at investorer totalt sett realiserer en netto risikojustert avkastning som er dårligere enn den risikojusterte avkastningen til referanseindeksen med omtrent den gjennomsnittlige kostnaden til fondene. Videre finner vi at forvaltere av fondene som presterer best har god evne til å velge aksjer, mens de dårligste fondsforvaltere både har dårlig markedstiming og velger dårlige aksjer. Avhengig av ulike statistiske metoder finner vi ulike nivåer av persistens i fondsprestasjonene. Når vi reevaluerer prestasjonen på månedlig basis, antyder de empiriske bevisene kortsiktig persistens blant toppfondene. Videre finner vi at prestasjonen til bunnfondene vedvarer signifikant i opp til ett år.

Contents

Performance and Persistence in Norwegian Mutual Funds

Kristoffer Gallefoss, Helge Hoff Hansen, and Eirik Solli Haukaas*

June 6, 2012

Abstract

Using a dataset free of survivorship bias, we investigate the performance and performance persistence of Norwegian mutual funds in the period 2000-2010. To evaluate mutual fund performance we apply a multi-period version of the Carhart 4-factor model to obtain Jensen's alpha. We assess the statistical significance of the alpha point estimates by comparing them to alpha distributions generated from bootstrap simulations. We find that mutual fund investors in aggregate realize risk-adjusted net returns that underperform the benchmark by approximately the fund fees. Additionally, the evidence implies that the managers of the funds that lie in the right tail of the cross-section of mutual fund alpha estimates inhabit stock-picking skills, while the managers of the worst performing funds are both mistiming the market and picking bad stocks. Depending on various statistical methodologies we find different levels of performance persistence. When we reevaluate performance on a monthly basis, our evidence indicates short-term persistence among superior funds. Moreover, the performance of inferior funds strongly persists up to one year.

*Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway. We especially thank Peter Molnar for his guidance, feedback and conversations that we have greatly benefited from. Also, we thank Børsprosjektet NHH, Verdipapirfondenes Forening and Bernt A. Ødegaard for generously providing us with the data.

1 Introduction

The question whether actively managed funds are able to outperform their benchmarks, and if the managers of these funds inhabit superior skills, has been addressed in many papers. This attention is justified, considering the importance of the question in both academics' and investors' point of view. For academics, it is important since the existence – and persistence – of managerial skills would support a rejection of the semi-strong form of the efficient market hypothesis, as defined by Fama (1970). As for investors, it is in their interest to know whether investing in actively managed funds is worth the extra cost compared to investing in low-cost passive funds, and if so, which type of funds they should invest in.

Early research, starting with Jensen (1968), concludes that mutual funds are outperformed by a comparable passive market proxy when accounting for the funds' managing cost. Later academic studies, like Ippolito (1989), find that mutual funds are able to generate enough excess returns to offset the fees they charge. Sharpe (1991) argue that the average return before fees on actively managed financial portfolios must equal the return on passive portfolios in order for the market to clear, according to the equilibrium accounting theory. Indeed, most of the studies have shown little evidence that the aggregate mutual fund industry adds value to investors in terms of superior performance. Obviously, this does not mean that all the funds deliver average or below average performance. In most cases, some funds will perform well, whereas other funds will perform poorly. Therefore, two additional questions arise. The first address whether the differences in performance can be attributed to chance, or if it is due to managerial skill [as discussed in Cuthbertson et al. (2008)]. Second, if skilled managers exist, the question is whether performance persists – i.e. if mutual funds with an above (below) average performance in the last period also will have an above (below) average performance in the next period.

There are several studies on persistence in mutual fund performance. Hendricks et al. (1993) find that funds exhibit persistence over short time horizons, denoting the effect as hot hands. Carhart (1997) argues that the hot hands effect is mostly driven by the one-year momentum effect of Jegadeesh and Titman (1993), and finds no evidence of performance persistence after controlling for the momentum effect. A more recent paper by Berk and Green (2004) provide evidence that abnormal returns persist in mutual funds over short time horizons. In the long run, however, they find that the funds do not outperform passive benchmarks. They conclude that there exists managerial talent capable of generating abnormal returns, but this talent is offset in the long run by fund inflows that lead to diminishing returns to scale. Teo and Woo (2001) find that the performance of losing funds strongly persist for up to 6 years. Furthermore, Kosowski et al. (2006) applies a bootstrap analysis (rather than the standard t-test) to evaluate the significance of the fund persistence and find that performance does indeed strongly persist among the top performing funds.

Most academic studies on mutual funds are done on the US market due to the importance of the market and the data availability. There have been an increasing amount of studies of mutual funds in European countries since the 90s. Otten and Bams (2002) present a comprehensive study of European mutual funds (with a database of 506 mutual funds from five countries). In contrary to most US studies, their findings suggest that mutual funds are able to add value. In particular, four out of five countries exhibit significant out-performance at an aggregate level when management fees are added back. When it comes to persistence in performance, however, they find only weak evidence in all countries except the UK. A more recent study from the US, Bollen and Busse (2005), conclude that persistence in performance is a short-lived phenomenon that can only be measured using relatively short measurement windows. That is, investigating persistence using monthly data, like the before-mentioned studies, is unlikely to produce significant results. Therefore, whenever possible, more frequent data should be used.

Considering the fact that the Norwegian economy is one of the most developed in the world in terms of GDP per capita, Norwegian mutual funds are certainly a subject of interest. In spite of this, there are few academic studies on the Norwegian mutual fund market. The only research known to us is the unpublished work of Sørensen (2009) and Sandvik and Heitmann (2010), which investigate the performance and persistence in Norwegian mutual funds between 1982-2008 and 1993-2009, respectively. Both of them find that the aggregate mutual fund industry in Norway does not exhibit abnormal performance. Regarding persistence in performance, both papers conclude that there is no persistence for either superior or inferior performers. However, these studies use monthly data, in contrary to the argument presented by Bollen and Busse (2005). Therefore, finding evidence of persistence is less likely.

The purpose of this paper is to investigate the performance and persistence in Norwegian mutual funds. To do this we use a survivorship bias-free database that consist of 64 actively managed domestic funds with daily returns from January 2000 to December 2010.

We apply both unconditional and conditional versions of the Carhart (1997) 4-factor model to evaluate whether funds exhibit superior or inferior performance. As we find that the factor loadings vary significantly with time, we reevaluate the factor loadings on a yearly basis to better estimate the funds' performance. This results in a multi-period version of the 4-factor model. Adopting the method of Fama and French (2010) and Kosowski et al. (2006), we perform bootstrap analyses to determine the significance of the results. Furthermore, we investigate whether managerial skills, if present, are due to stock-picking or market timing ability. Managers with stock-picking ability are successfully predicting whether particular securities perform better than others, while those with market timing are predicting where the general market is going. An active fund manager may inhabit either one of these abilities, or a combination of the two.

Persistence in performance is assessed by various statistical methodologies. First, we evaluate the performance potential of investment strategies that seek to exploit short-term persistence, a method adopted from Hendricks et al. (1993). We use relatively short measurement periods from one month to two years to evaluate persistence. As argued by Bollen and Busse (2005), this identifies top and bottom performers more precisely. Second, we check for performance persistence by performing a series of nonparametric tests, both in a two-period and multi-period framework. Two-period tests compare the ranking of funds over two consecutive periods. Specifically, we check whether the winners (losers) of the last period are also the winners (losers) in the next period [similar to the method of Carhart (1997) and Karoui and Meier (2009)]. For the multi-period test, where more than two consecutive periods are considered, we use the Kolmogorov-Smirnov test, as described by Agarwal and Naik (2000) and Eling (2009).

Note that there is a fundamental difference between the persistence tests. The first method creates a portfolio that is long in the superior performers and short in the inferior performers from the last period. The aggregate risk-adjusted return from this strategy determines whether such a portfolio delivers abnormal returns relatively to the benchmark given by the 4-factor model. The nonparametric tests rank winners and losers relatively to all the funds in the sample. Thus, it measures performance persistence relative to other funds and it cannot conclude whether either winners or losers perform better than the market benchmark.

We find that the Norwegian mutual funds in aggregate deliver negative risk-adjusted returns. The size of this aggregate return is comparable to the average expense level charged by the funds, which is consistent with the equilibrium accounting theory of Sharpe (1991). The funds that lie in the right tail of the cross-section of alpha estimates are generating abnormal returns that, according to our bootstrap analyses, cannot be explained by luck alone. Our evidence implies that the managers of these funds have skills attributable to stock-picking ability. We find that the worst performing funds are realizing significant negative abnormal returns that seem to be explained by a combination of poor stock-picking ability and poor market timing ability. Furthermore, we find different level of performance persistence depending on the statistical methodologies we use. The nonparametric tests reveal that there is short-term persistence among superior funds for the entire sample as well as among individual funds. We conclude from both the parametric and nonparametric tests that the performance of the worst funds strongly persists up to one year. This coincides with the findings of Bollen and Busse (2005).

The remainder of this paper is organized as follows. Section 2 gives an overview of our data set and provides some descriptive statistics of the Norwegian mutual funds and our benchmarks. Section 3 consists of three parts. In the first two parts, we describe the performance model and the bootstrap method. In the third part, we present the results on mutual fund performance and discuss the presence of market timing and stock-picking ability amongst fund managers. Section 4 tests for persistence in mutual fund performance. Finally, section 5 concludes.

2 Data

2.1 Norwegian Mutual Funds

Our database contains daily net asset values (NAV) per share of 64 Norwegian actively managed open-end equity funds from January 2000 to December 2010. This is obtained from Børsprosjektet NHH, which is administered by the Norwegian School of Economics. The NAV include reinvestment of all distributions (e.g. dividends) and are net of expenses, but disregard load charges and exit fees. In aggregate, these 64 funds represented 74% of the total fund industry in Norway as of December 2010.[1] We restrict our sample to domestic equity funds with at least 36 months of data. That is, we exclude funds that invest in bonds, sectors and indices, as well as equity funds that invest more than 20% of their assets internationally. We calculate the fund returns as follows:

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

where $NAV_{i,t}$ is the net asset value of fund i at day t and $r_{i,t}$ is the return of fund i at day t.

As shown by for instance Brown et al. (1992), survivorship bias in a sample can severely affect the results. That is, if funds that are shut down or merged into another one within the sample period are excluded from the sample data, an overestimation of the average performance may occur. This is due to the fact that those funds that are shut down often are those with poor performance. Therefore, we include funds that have been terminated, and started, within the sample period, thus avoiding the survivorship issue in our dataset.

Additionally, we have monthly data on asset under management (AUM), inflow and outflow for each fund for the entire time period. This data is obtained from the Norwegian Fund and Asset Management Association.[2]

[1] Statistics Norway

[2] Verdipapirfondenes Forening

Table 1
Mutual Fund Database Summary Statistics

This table reports the number, total asset under management (AUM), returns and performance of Norwegian open-end equity funds in the period 2000-2010. Specifically, the first column of this table reports the number of funds in existence in each year. The second (third) column reports the number of funds that started (liquidated) during each year. The fourth column reports the total asset under management (in million NOK) in the 64 mutual funds at the end of each year. The fifth column reports the excess return calculated as the return of the equally weighted portfolio excess of the risk-free rate in percent per year. Column six reports the unconditional 4-factor model alpha in percent per year for the equally weighted portfolio of funds for each year.

	Number of funds			AUM	Excess Return of Equally Weighted Portfolio	4-Factor Alpha of Equally Weighted Portfolio
Year	End of year	Born	Liquidated	(End of Year, MNOK)		
2000	51	6	0	26,960	8.69%	1.95%
2001	54	3	0	22,936	3.57%	-5.02%
2002	60	6	0	15,615	-8.73%	-16.20%
2003	60	1	1	23,268	8.18%	-2.70%
2004	58	0	2	27,280	-0.49%	-2.34%
2005	56	0	2	33,508	-2.87%	0.89%
2006	53	3	6	44,137	-0.25%	0.16%
2007	52	0	1	46,443	3.37%	-2.19%
2008	52	0	0	22,146	2.68%	-14.35%
2009	52	0	0	48,362	12.77%	15.12%
2010	52	0	0	59,483	6.41%	4.22%

2.2 Benchmarks

We construct the daily excess market return by deducting the Oslo Stock Exchange All-Share Index (OSEAX) from the Norwegian 3-month Treasury bill index (ST1X), both obtained from Ecowin Reuters. The time-series of daily returns for the remaining factors of Carhart's 4-factor model, i.e. the size, book-to-market and momentum factor, are acquired from Professor Bernt Arne Ødegaard. These risk factors are constructed using stocks at the Oslo Stock Exchange (Næs et al., 2009). SMB is the returns of a portfolio that has a long position in small companies and a short position in large companies, while HML is the returns of a portfolio that is long in companies with high book-to-market value and short in companies with low book-to-market value. The momentum factor, identified by Jegadeesh and Titman (1993), is constructed by ranking stocks based on their prior 11-month return and sorting them into three portfolios. Then, the portfolios are held for one month, before they are rebalanced again according to the past return. The return difference between the top third portfolio return and the bottom third portfolio return is the PR1YR factor.

The OSE All-Share Index is plotted in Figure 1 together with the equally weighted fund portfolio. The aggregate return of the market index is 31 % higher than the aggregate return of the equally weighted portfolio. Additionally, Figure 1 displays the returns of the equally weighted portfolio of funds that is terminated within the sample period. We observe that the portfolio of dead funds have lower returns than the portfolio including all funds. This illustrates that survivorship bias would be present in the dataset if the defunct funds were to be excluded.

We use time-series of returns for oil price, 3- and 12-month Treasury bill (ST1X and ST5X, respectively), and 12-month Norwegian Inter Bank Offered Rate (NIBOR) as explanatory variables in the conditional model in the next section. These time-series are obtained from Ecowin Reuters.

Table 2
Performance Measurement Model Summary Statistics

MKT is the market return (given by the OSEAX) excess of the risk-free rate (given by the ST1X). SMB and HML are returns of the factor mimicking portfolios for size and book-to-market equity of Fama and French (1992). PR1YR is the factor-mimicking portfolio for one-year return momentum. Column one reports the yearly average return of the four portfolios. The second column reports the standard deviation of the return per year for each portfolio. Column three to six reports the correlations between the factor portfolios.

Factor Portfolio	Yearly Average Return	Standard Deviation	Correlations MKT	SMB	HML	PR1YR
MKT	4.86%	24.51%	1			
SMB	11.38%	21.43%	-0.70	1		
HML	8.32%	21.37%	-0.28	-0.08	1	
PR1YR	3.71%	19.16%	-0.06	0.07	0.09	1

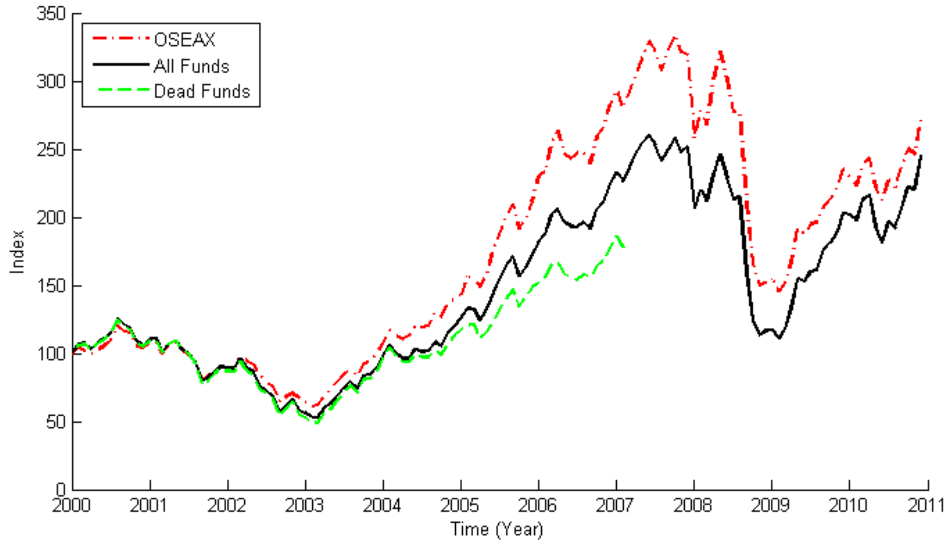

Figure 1. Return development of the OSE All-Share Index and the equally weighted portfolio of all funds and dead funds. The figure illustrates the returns of the OSEAX and the equally weighted portfolio of all 64 funds, in addition to the equally weighted portfolio of the funds that are terminated within the sample period.

3 Performance

In this section, we investigate whether Norwegian mutual funds exhibit performance excess of what could be expected by the given level of risk in the fund portfolios. To answer this, we first seek out the most appropriate model to evaluate performance. Further, we turn to simulations that use fund returns to infer the existence of skills, or lack thereof, among fund managers. Finally, we investigate whether these skills, if existent, are due to market timing abilities or stock-picking abilities.

3.1 Mutual Fund Performance Models

We employ both unconditional and conditional versions of the Carhart (1997) 4-factor model. This section briefly describes these models, and evaluates their performance estimates on the equally weighted portfolio of our dataset.

3.1.1 The Unconditional 4-Factor Model

First presented by Sharpe (1964), the capital asset pricing model (CAPM) is a basic model to explain variability in returns for a given asset. The intercept of the regression model, α, gives us Jensen's alpha, which is interpreted as a measure of performance relative to the market. The model is formulated as:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i \left(r_{m,t} - r_{f,t} \right) + \epsilon_{i,t} \tag{2}$$

where $r_{i,t}$ is return of fund i in period t, $r_{f,t}$ is the risk-free rate of return in period t, $r_{m,t}$ is the market return in period t and $\epsilon_{i,t}$ is the error term. The market return excess of the risk-free rate is the market risk premium and β_i is the exposure to the market risk for fund i. In particular, if α is significantly positive the fund manager is able to earn returns which are higher than one could expect from the given level of riskiness of the portfolio, according to the model. Opposite, a negative α indicates poor performance by the fund manager.

Prior to the 90s, the CAPM was the most common model in academic research to measure mutual fund performance. However, in addition to the market risk, there are other well-known additional risk factors in the stock market. These factors are often used by fund managers as stated investment strategies. To explain co-variation between fund returns and these risk factors, the model can be expanded to a multi-factor asset-pricing model. There exists an abundance of different proposed factor models. However, the 4-factor model of Carhart (1997) is the most common in financial literature. The model is specified as follows:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{iMKT} MKT_t + \beta_{iSMB} SMB_t + \beta_{iHML} HML_t \\ + \beta_{iPR1YR} PR1YR_t + \epsilon_{i,t} \quad (3)$$

where MKT is the market risk premium, SMB is the small-minus-big size factor, HML is the high-minus-low book-to-market factor, $PR1YR$ is the prior one-year momentum factor, and $\beta_{iMKT}, \beta_{iSMB}, \beta_{iHML}, \beta_{iPR1YR}$ is asset i's exposure to the MKT, SMB, HML, and PR1YR factor, respectively. The error term ϵ_i, has expectation of zero and represents the risk that is unrelated to any of these factors. Carhart (1997) finds that the explanatory power of the 4-factor model is significantly higher than the CAPM. Additionally, the 4-factor model explain asset pricing without the risk of over-explaining by adding factors that exhibit significance only as a result of data-mining, and thus risking the possibility of committing serious pricing errors (Ghysels, 1998).

To gain some insight whether the 4-factor model is appropriate for our use, we estimate the factor loadings for the equally weighted portfolio and each individual fund. The result is that the Norwegian mutual fund market in aggregate is not loading heavily on any of the three factors SMB, HML and PR1YR. This observation is natural considering that that the three factors are constructed as the return of holding a long and short position in stocks with certain characteristics (i.e., small-minus-big, high-minus-low, and prior winners-minus-prior losers). In order for the market to clear, investors must hold both positions. For instance, it is likely that some fund managers must hold non-momentum stocks, since all fund managers cannot hold the top-momentum portfolio. Thus, we observe that the aggregate mutual fund market (i.e. the equally weighted portfolio) exhibit near-zero coefficients. The regressions on each of the 64 funds reveal that the factor loadings for the SMB, HML and PR1YR indeed vary across funds (see Table 3). Therefore, the 4-factor model seems to be appropriate to evaluate mutual fund performance.

Table 3
Summary of Factor Loadings for Individual Funds

The table gives an overview of the factor loadings (betas) on individual funds. The first column reports the average load in each of the four benchmark portfolios. The second and third column reports the minimum and maximum load for the 64 funds, respectively. The fourth column reports the cross-sectional standard deviation of the factor loadings.

	Average Load	Minimum Load	Maximum Load	Standard Deviation
MKT	0.936	0.690	1.113	0.078
SMB	0.009	−0.107	0.222	0.075
HML	0.003	−0.069	0.116	0.031
PR1YR	−0.050	−0.135	0.036	0.036

3.1.2 The Conditional 4-Factor Model

When applying the unconditional 4-factor model to eleven years (2000-2010) of data on fund returns, one makes the assumption that funds have a constant exposure to the risk factors in the model. In other words, using this single-period model one assumes that the funds' investment styles remain the same throughout the whole sample period. When analyzing actively managed funds we could expect this assumption to be wrong. That is, general macroeconomic tendencies could influence a fund manager to change the exposure to the risk factors. For example, a manager might evaluate his portfolio holdings differently in a bull market than in a bear market. Likewise, if a fund manager expects a promising future for a specific small business, the exposure to the SMB factor would not remain unchanged if the manager buys a considerable portion of stocks from that firm.

By further questioning underlying assumptions of standard performance models, we analyze our sample using a conditional model. If manager actions are based on the state of the economy and are conditional on publicly available information, standard models may produce spurious results (Ferson and Schadt, 1996). The conditional 4-factor model can be specified:

$$r_{i,t} - r_{f,t} = \alpha_i + \sum_j \beta_{i,j} r_{j,t} + \beta_i' Z_{t-1} MKT + \epsilon_{i,t} \tag{4}$$

Here, Z_{t-1} is a vector of information available at time $t-1$, and β_i' is a vector of response coefficients of the conditional beta with respect to the instruments in Z_{t-1}. Thus, the time-varying beta $\beta_{i,t} = \beta_{iMKT} + \beta_i' Z_{t-1}$, is a linear relation between the average beta and the conditional instruments.

Specifically, conditional models seek to measure the variation of betas with time using some information vector that managers may use as a decision tool when adjusting fund portfolios. Previous papers [see for example Otten and Bams (2004), Bhuvaneswari (2011), and Barras et al. (2010)] introduce variables such as the T-bill rate, market dividend yield, interest term structure, lagged market returns and other macroeconomic variables. We test whether the aggregated mutual fund portfolio's market exposure depends on lagged market return, interest spread (12-month ST5X less the ST1X), oil price, and the 12-month NIBOR interest rate.

We do not find that the conditional model improves significantly in terms of R-squared. The market beta is very sensitive when combined with the oil price and NIBOR factors, which has implications for the interpretability of the model. The 12-month NIBOR has some explanatory power, but it does not contribute significantly to the goodness of fit of the model. Moreover, the economic significance of the 1-year interest factor is marginal, attributable to the relatively small factor return. A conditional model would be of much greater use if we know the variables on which the managers trade upon. Staff changes and other non-economic variables might affect portfolio decision-making, information that is not easily obtainable. Without further information on funds' trading decisions, we conclude that the presented conditional model does not add more insight to performance than the unconditional model presented in the preceding section. For details, see Appendix 1.

3.1.3 The Multi-Period Unconditional 4-Factor Model

Performing regressions on a year-to-year basis, instead of the entire sample, allows for betas to vary on a yearly basis independent of conditional information and should therefore provide more accurate results.[3] The multi-period model is expressed as follows:

$$r_{i,t} = \alpha_{i,k} + \sum_{j} \beta_{i,j,k}\, r_{j,t} + \epsilon_{i,t} \tag{5}$$

where $\alpha_{i,k}$ is alpha of fund i in year k and $\beta_{i,j,k}$ is the coefficient of risk factor j for fund i at year k for every year k $\in$ [2000 : 2011]. Further, $r_{j,t}$ is the return on the risk factor j at day t. The average multi-period alpha is calculated as:

$$\alpha_i = \frac{\sum_{i} \alpha_{i,k}}{k} \tag{6}$$

Because we regress on logarithmic returns, this average alpha will have the same interpretation as the alpha for the single-period model.

The multi-period unconditional model must not be confused with conditional models where factor betas vary with some variables. It is constructed as the single-period unconditional model, only that we are allowed to regress on shorter time intervals and still get reliable results, due to daily data. By construction, the multi-period model will have a closer fit with the observed fund returns than the single-period model, which increases explanatory power. We find that the funds' factor coefficients change quite substantially throughout the period 2000-2010. This is evident from Figure 2, which illustrates how each factor loading vary through time for a selection of funds. Thus, assuming constant beta coefficients over the whole sample period will most likely provide biased results. Considering that the trading dynamics of a fund might change over the course of 11 years, we conclude that it is appropriate to apply a multi-period version of the unconditional four-factor model when evaluating fund performance.

[3] We choose a yearly evaluation of fund performance since fund managers are usually measured on yearly performance by the public and their superiors (Chevalier and Ellison, 1997).

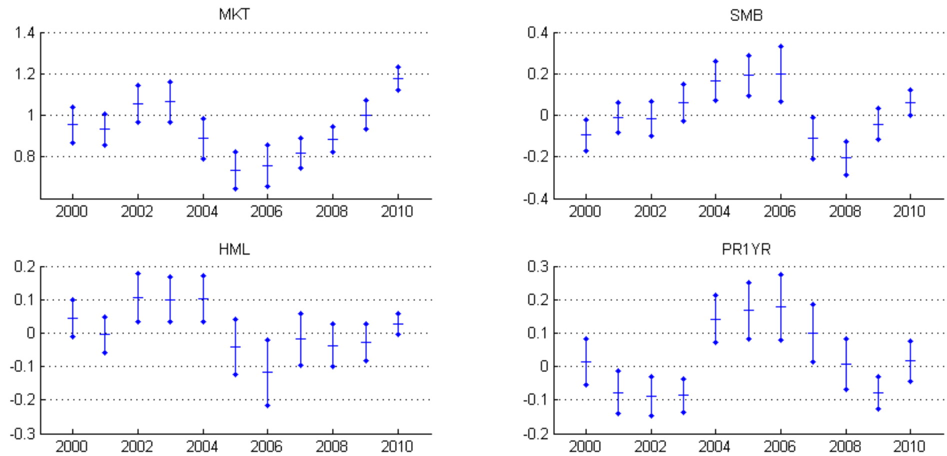

Figure 2. Development of the factor loadings through the sample period. The figure shows the four factor loadings for fund number five each year in the sample (see Appendix 10 for the list of funds). The fund is randomly chosen to illustrate how the factor loadings vary through time. All the funds show time-varying tendencies. The middle point in each vertical line is the point estimate of the factor loading at the given year, while the vertical lines provide the 95% confidence interval of the estimate. See Appendix 10 for an overview of the funds in our sample.

3.2 The Bootstrap Method

When analyzing mutual fund alphas using the OLS approach, one assumption is that the residual vector from the estimation has a normal distribution. There are several properties that would lead to a rejection of these assumptions. First, individual stocks within a portfolio often yield returns distributed with significantly different skewness and kurtosis compared to a normal distribution. The central limit theorem implies that an equally weighted portfolio of non-normal and independently distributed stocks will approach a normal distribution. However, returns are usually cross-correlated, and fund managers also do not always hold balanced portfolios. Thus, fund portfolios usually exhibit non-normal returns. Second, financial assets often exhibit heteroscedastic variance. Third, individual stocks exhibit varying levels of time-series autocorrelations in returns. Finally, funds may implement dynamic strategies that involve changing their levels of risk-taking when the risk of the overall market portfolio changes, or in response to their performance ranking relative to similar funds. Because each of the above-mentioned irregularities can contribute to non-normally distributed residuals, the normality assumption could lead to invalid alpha t-statistics in an OLS-estimation.

To provide more accurate levels of precision in our analyses of performance, we perform residual resampling simulations to better address the complex nature of financial data. As shown by e.g. Horowitz (2003), Fama and French (2010), Bickel and Freedman (1984), Kosowski et al. (2006) and Hall and Martin (1988), the bootstrap method gives a more accurate assessment of the significance levels of the alphas. That is, the bootstrap can significantly decrease the difference between true and nominal probabilities of correctly rejecting a given null hypothesis (i.e., that no superior fund managers exist). For example, by recognizing the presence of thick tails in individual fund returns, the bootstrap does not reject abnormal performance among mutual funds as often as the standard t-test.

Furthermore, consider a distribution of residuals obtained from a parametric regression on an individual fund's returns. This distribution may or may not be normally distributed. When modeling the mixed distribution of mutual fund residuals we cannot assume normality in the distribution. It must be treated as a complex, unique distribution based on two properties: heterogeneous risk-taking across funds, and individual non-normalities in individual distributions of residuals (Kosowski et al., 2006). A mixed distribution of two individual residual distributions with different standard deviations will not be normal, even if the two distributions are normal themselves. Consider a selection of funds with heterogeneous levels of risk such that, across funds, residual variances range uniformly between 0.5 and 1.5 (i.e., the average variance is 1). In this case, the tails of the cross-sectional distribution of residuals are now fatter than those of a normal distribution. The intuition here is clear: As we move further to the right in the right tail, the probability of these extreme outcomes does not fall very quickly, as high-risk funds more than compensate for the large drop in such extreme outcomes from low-risk funds. Given the intractability of parametrically modeling the mixed distribution of such mutual fund estimations across 64 funds, the bootstrap is a very attractive approach to use in analyzing a cross-section of mutual funds. For an illustration of the non-normal implication of the mixed distribution, see Appendix 2. The remainder of this sub-section will explain the implementation of the bootstrap method.

We seek to derive the combined distribution of alphas under the null hypothesis of zero alpha (Berk and Green, 2004). To obtain this, we save the residuals from the 4-factor regression in a vector $\epsilon_{i,t} = [\epsilon_{i,1}...\epsilon_{i,T}]$ for each individual fund. Then we generate 10,000 new random sets of returns based on the true data, but with randomly selected residuals from the vector $\epsilon_{i,t}$, where the alpha is equal to zero by construction:

$$r_{i,t}^{b} = \sum_{j} \beta_{i,j,k}\ r_{j,t} + rand(\epsilon_{i,t}) \tag{7}$$

where b is the bootstrap index, and $rand(\epsilon_{i,t})$ is a residual resampling. This ensures a random distribution of residuals. Finally, this pseudo-time series is regressed on the Carhart(1997) four factors, obtaining an alpha estimated from simulated data:

$$r^b_{i,t} = \widetilde{\alpha^b_i} + \sum_j \beta_{i,j,k}\ r_{j,t} + rand(\epsilon_{i,t}) \tag{8}$$

Repeating the above steps across all funds $i = 1, ..., N$, we arrive at a draw from the cross section of bootstrapped alphas. Repeating this for all bootstrap iterations, $b = 1, ..., 1,000$, we then build the distribution of these cross-sectional draws of alphas, $\{\alpha_i, i = 1, ..., N\}$, or their t-statistics, $\{t_i, i = 1, ..., N\}$, that result purely from sampling variation while imposing the null of a true alpha that is equal to zero. Percentiles from this distribution are used for statistical inference. For example, a bootstrapped distribution of the top decile portfolio is obtained by extracting the top decile alphas from each b-th iteration.

3.3 Results and Interpretation of Fund Performance

The previous sections suggested that the multi-period unconditional 4-factor model is an appropriate model to investigate fund performance, and that the bootstrap method is appropriate to evaluate the statistical significance of the performance[4]. Therefore, the results from the multi-period model will be the focus in what follows. However, for comparative purposes, we perform regressions on the equally weighted portfolio and each individual fund using both the single-period and multi-period unconditional 4-factor model and briefly comment on the differences.

3.3.1 Aggregate Mutual Fund Performance and Distribution of Alphas for Individual Funds

We find that the aggregate mutual fund market in Norway delivers a yearly alpha of -1.903% (-0.725%) on average using the multi-period (single-period) model. The yearly average management fee on the Norwegian mutual funds is approximately 1.69%, excluding performance fees.[5] Comparing these numbers to the alpha of -1.903%, it seems that the aggregated fund market is delivering approximately zero alpha before fees. These results coincide with the equilibrium accounting theory presented by Sharpe (1991). He states that markets can only clear when the expected pre-cost return to investors in actively managed funds equals the expected return in alternative investment opportunities.

[4]The possible non-normalities in the fund returns is not the only reason why bootstrapping is a more accurate method to determine statistical significance. As the coefficients in the multi-period model are averaged over eleven yearly coefficients, we have to apply a bootstrap approach to derive the corresponding p-value for the average alpha and betas.

[5]See Appendix 3 for a summary of fund fees in our sample.

Table 4
Regression Results on the Equally Weighted Portfolio

The table reports the results of the regression on the equally weighted portfolio for both the single-period and multi-period version of the 4-factor model. The first and fifth column presents the coefficients for the estimates of alpha and the loadings of the four factor-mimicking portfolios for the two models. Additionally, the table reports the R-squared and the Durbin-Watson statistic. The second and third column reports the t-statistic and the p-value given by the standard t-test. Column four and six reports the bootstrapped p-values for the two models.

	Single-Period Model				Multi-Period Model	
	Coefficients	t-statistic	p-value	Bootstrapped p-value	Coefficients	Bootstrapped p-value
Alpha	−0.725	−0.390	0.652	0.350	−1.903	0.150
MKT	0.930	121.550	0.000	0.000	0.909	0.000
SMB	0.004	0.485	0.314	0.314	0.023	0.009
HML	−0.001	−0.218	0.586	0.424	0.007	0.175
PR1YR	−0.044	−7.177	0.000	0.000	0.004	0.297
R^2	0.933				0.959	
DW	2.094				2.084	

Figure 3 illustrates the distribution of alphas, which is obtained by regressing on each of the 64 funds in our dataset. From the figure, the slight negative shift in alpha when regressing on the multi-period model compared to single-period model is evident. Our set does not have a sufficient cross-sectional size to conclude about the distribution of these alphas. However, the histograms indicate some clustering in the tails, i.e. there seems to be both superior and inferior performers compared to the equally-weighted portfolio. The question is whether managers of funds that exhibit positive (negative) alpha possess superior (inferior) skills, or if the alpha distribution is generated purely by chance. This will be addressed in the following subsection.

3.3.2 Separating Luck from Skill

From the distribution of individual alphas in the previous section, we find that there are some funds who are performing well (right tail) and some are performing poorly (left tail). This might be the case purely by chance or as a result of either superior or inferior managerial skills. To illustrate an example, consider a hypothetical database of 1,000 mutual funds. Performing regressions on each of the funds' return series with the 4-factor model, which does not explain return variation perfectly, one should expect that some funds portray alpha significantly different from zero purely by chance. Our objective is to separate luck from skill in performance using a bootstrap method similar to Fama and French (2010) and Kosowski et al. (2006). In the following, we apply the multi-period 4-factor model.

We find that the alpha of top and bottom performers cannot be attributed to chance alone. That is, the fund managers possess skills that affect the overall

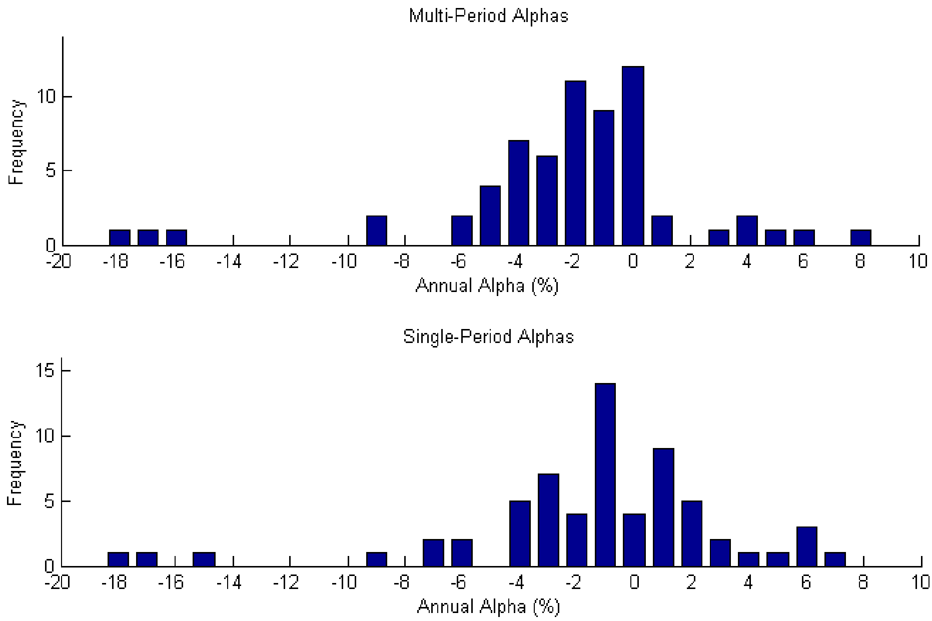

Figure 3. Alpha distribution for the multi-period and single-period model. The figure displays the alpha distribution of the 64 funds in our sample for both the multi-period and single-period version of the Carhart (1997) 4-factor model. Each fund alpha is estimated using daily data for as long as the fund was active within the sample period (2000-2010). The bars represent the number of fund alphas within the given bin, i.e. within $\pm 0.5\%$ of the alpha given by the x-axis.

performance of the funds. This is evident from Figure 4. We see that the bottom performing funds from the true data set lies in the left tail of the distribution of bottom funds obtained from the bootstrap simulation. This indicates that these funds may possess some negative attributes that affect their performance. Meanwhile, the opposite is the case for the top performing funds. While we cannot reject that the single top performing fund delivers positive alpha purely by chance, there exist clusters of top performing funds that deliver positive alpha due to managerial skills.

To confirm that there exist skills (and lack thereof) amongst fund managers, we report the probabilities that the estimated alphas of the top and bottom fund portfolios could occur purely by chance in Table 5. From the table we can conclude, at least on a 5% significance level, that the top and bottom portfolios of funds indeed are managed by skilled and unskilled managers, respectively. Ranking the funds according to the t-statistic yields approximately similar results.[6]

Table 5
Performance of Top and Bottom Funds

The table reports the performance, i.e. alpha, of the single top- and bottom fund, as well as the performance of the portfolios of the 5% and 10% top- and bottom funds. Column one and three reports in percent per year the alpha of the top and bottom funds when the funds are ranked according to the alpha and the alpha t-statistic, respectively. The second and fourth column reports the p-value given by the bootstrap simulations. All alphas in the table are significant at least on a 5% level, except for the single top fund.

	Ranked on Alpha		Ranked on T-Statistic	
	Alpha	p-value	Alpha	p-value
Top fund	8.17	0.334	5.74	0.581
Top 5%	6.44	0.011	6.44	0.012
Top 10%	4.46	0.034	4.46	0.031
Bottom fund	−17.76	0.011	−18.81	0.004
Bottom 5%	−16.53	0.000	−17.18	0.000
Bottom 10%	−12.45	0.000	−12.45	0.000

One must keep in mind that the fund return dataset is net of fees, so that one should expect a negative shift from zero in both the estimated alphas and the bootstrapped alpha distributions that is equal to the average management fee if one truly believes that the equilibrium accounting theory is valid. This could explain some of the relative stronger rejection of the null for the bottom funds. For example, if we add the average management fee of 1.69% back to the top fund, the corresponding p-value is reduced to 0.147.

3.3.3 Market Timing and Stock-Picking Abilities

Wermers (2000) argue that abnormal performance that cannot be explained by luck may be due to either stock-picking skill, market timing, expense levels, trading

[6] For added detail on this ranking method, see Appendix 4.

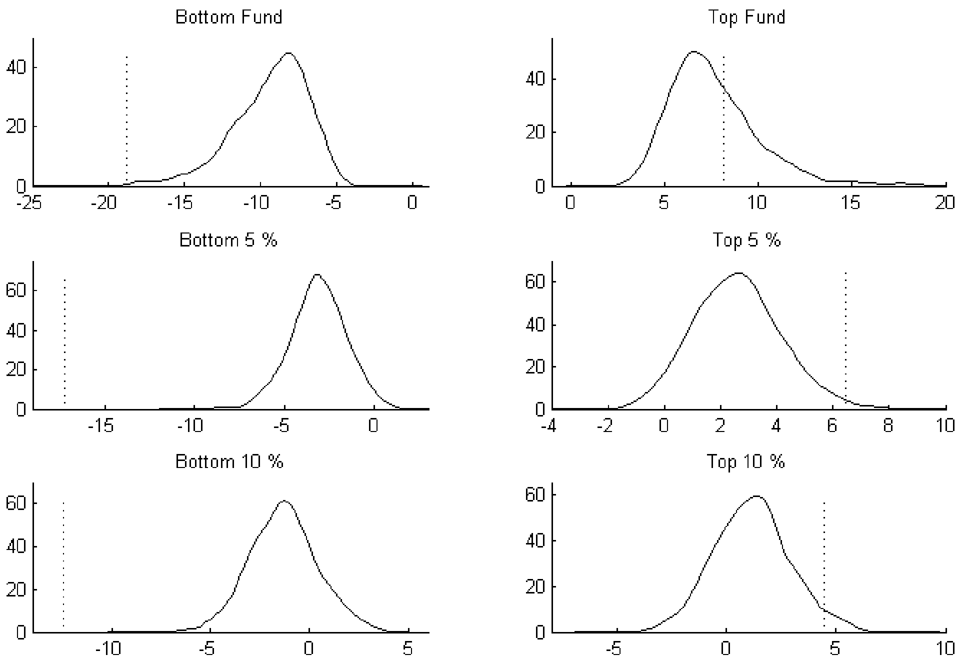

Figure 4. Estimated true alphas versus bootstrapped alpha distributions for individual funds. This figure plots kernel density estimates of the bootstrapped unconditional 4-factor model alpha distributions (solid line). To produce the distribution of the alpha estimates, we first create random returns given by the residuals in the in the regression on the actual return data. Then, we generate 10,000 alpha estimates from the OLS regression on those random returns. The x-axis shows the alpha performance measure per year and the y-axis show the kernel density estimate. The vertical dotted line shows the actual (estimated) fund alpha. The fund portfolios are ranked on alpha values, where the highest ranked fund has the highest estimated alpha over the sample period. The same ranking is performed on the bootstrapped alphas to create bootstrapped portfolios.

styles or a combination of the four. Trading styles is not a likely explanatory variable of performance in our case, since we are only including Norwegian equity funds in our dataset and control for the most common trading styles by applying the Carhart (1997) 4-factor model. Further, if fee levels vary significantly across funds, funds with low fees are more likely to form a winner portfolio than funds charging higher fees. We do not have access to the fee structure of the defunct funds (which are more often than not underperformers), thus we cannot investigate whether expense level is a determinant of performance. However, in their search for variations in managerial skills in European equity funds, Abinzano et al. (2010) find that their results remain unexplained by fee variations. In light of this, we investigate whether the performance that cannot be explained by luck alone is due to market timing abilities or stock-picking abilities.

It is possible that the fund managers, by adjusting their portfolio's market exposure, may generate superior performance by timing the market correctly. This would imply that the fund managers are capable of correctly assessing the direction of market (i.e., whether bull or bear). In case they do, they would be positioning their portfolios accordingly. If a mutual fund manager increases (decreases) a portfolio's exposure to the market in advance of positive (negative) excess market returns, the portfolio's risk adjusted returns will outperform its market return benchmark. Conversely, a manager could have superior knowledge on a selection of stocks by processing and analyzing information, and thereby obtain stock-picking abilities.

We evaluate the market timing ability using the Treynor and Mazuy (1966) (TM) approach:

$$r_{i,t} = \alpha_i + \sum_j \beta_{i,j}\ r_j + \gamma_i\ r_m^2 + \epsilon_{i,t} \tag{9}$$

and the Henriksson and Merton (1981) (HM) model:

$$r_{i,t} = \alpha_i + \sum_j \beta_{i,j}\ r_j + \gamma_i I_t r_m + \epsilon_{i,t} \tag{10}$$

In TM, market timing is measured by adding a squared market return factor. It captures the convex (concave) relationship between market- and fund return that would occur if a fund manager is timing (mistiming) the market. HM uses a more discrete approach, where I_t is equal to 1 (0) when the market return is positive (negative). Thus, γ_i measures the degree of market timing ability. In both TM and HM, α_i is interpreted as risk-adjusted return controlled for market timing ability. That is, a positive (negative) alpha can be interpreted as superior (inferior) stock-picking skills.

The TM and HM equations' objective is to describe the nonlinear relationship between the market return and the fund portfolio return. TM assume that the fund managers adjust their leverage towards the market proportionally to the predicted market return, equivalent to $\gamma_i \, r_m$. However, HM assume that fund managers operate with thresholds when attempting to time the market, implying discrete adjustments in risk compared to the TM model's continuous risk adjustment. Kon and Jen (1979) provide empirical results which suggest that fund managers adjust their portfolios with discrete changes in their risk levels, an indication that the HM model might be better specified.

HM's and TM's measures of market timing have been subject of critique in many recent papers. Ferson and Schadt (1996) find negative market timing coefficients for the unconditional TM and HM models, but not as negative results for the conditional versions of the models. They argue that the negative trend in market timing coefficients are most likely due to biases in the performance models. Bollen and Busse (2005) argue that funds exhibit short term market timing ability, but that these coefficients are perversely negative in a long term evaluation window, and relate this to the positive covariance between return and investor cash flow. Warther (1995) discusses the cash flow issue of mutual funds, where fund betas are pulled downwards by high temporary cash positions due to increased cash inflows to the mutual fund market in bull markets. Jagannathan and Korajczyk (1986) introduce two effects that could lead to over- or underestimation of market timing. One is the dynamic trading effect, which arises due to certain dynamic trading strategies implemented by mutual funds, such as volatility hedging strategies which would lead to a downward bias of the market timing coefficient. The other is the passive timing effect, which refers to the fact that the returns of a passive portfolio can also be nonlinearly related to market returns.

In light of the arguments above, we are therefore not explicitly interested in the individual fund market timing coefficients alone, but implicitly through the relative difference in these coefficients between the top- and bottom funds. We also allow for variation in the stock-picking and market timing measures by applying the multi-period model. An analysis by Kacperczyk et al. (2011) reveals that managers indeed change their skillset focus (stock-picking or market timing) depending on economic conditions. They find that market timing skills is more evident in recessions relative to booms. For example, one could imagine that it is hard for a portfolio manager to focus on the nuances of stock selection when the prospects of a recession keep rising. Simply put, the macroeconomic impact would overwhelm the micro. We find clear signs of a time-varying skillset,[7] which emphasizes the importance of a multi-period evaluation.

Table 6 reports the results from the two models. It is evident that the market

[7] See Appendix 5.

timing coefficient are negative for all the fund portfolios. We suspect that the market timing coefficients may suffer from some negative downward bias, since it is very unlikely that the fund market is consistently perversely mistiming the market (Ferson and Schadt, 1996). The relative difference between fund portfolios within the data set controls for some of the common bias that could exist. As shown in Table 6, the worst performing portfolios have significantly worse timing coefficients than the top ranked portfolios. The stock-picking indicators, α_{TM} and α_{HM}, portray similar decaying tendencies.

Table 6
Stock-Picking and Market Timing Ability in Decile Portfolios

The table reports the stock-picking ability and market timing ability in the decile portfolios using both the Traynor-Mazuy (TM) model and the Henriksson-Merton (HM) model. The first column reports the alpha of the portfolios using the multi-period 4-factor model. Column two and three report the stock-picking alpha and the market timing gamma, respectively, by applying the TM model. The fourth column reports the stock-picking alpha and the fifth column reports the market timing gamma, both given by the HM model. All alpha values are in percent per year. Numbers marked with three stars (***) are statistically significant on a 1% level, two stars (**) on a 5% level, and one star (*) on a 10% level.

		TM		HM	
Decile Portfolio	4-Factor Alpha	α	γ	α	γ
1 (Top)	4.46***	5.98***	-0.13	9.44***	-0.03*
2	0.31	4.12**	-0.34*	11.29***	-0.07**
3	-0.33	1.84	-0.15*	5.79***	-0.04*
4	-1.02	1.91	-0.27*	8.97***	-0.06**
5	-1.72	1.33	-0.51*	7.84***	-0.07**
6	-2.31	1.28	-0.41*	9.91***	-0.08**
7	-2.94*	0.70	-0.57*	8.95***	-0.09**
8	-3.72**	-1.25	-0.24	5.36***	-0.06*
9	-4.96***	-0.55	-0.88**	7.11***	-0.09**
10 (Bottom)	-12.45***	-6.21***	-1.57***	7.77***	-0.18***

Our evidence imply they there is stock-picking ability among top decile funds. Due to the uncertainty of the market timing measures, γ, we cannot draw conclusions on superior market timing ability among the top funds. Furthermore, we find that the bottom performers are exhibiting worse stock-picking skills than what could be expected by chance. As discussed, the TM and HM models could exhibit some biases in the estimation, and it is therefore difficult to conclude upon the presence of market mistiming among the bottom performers with confidence. Nevertheless, our results imply that the managers of the bottom funds are to some extent mistiming the market relative to the top funds.

4 Persistence

In the preceding section, we argue that, in aggregate, Norwegian mutual funds deliver negative alpha comparable to the fees they charge, and that there exists funds that exhibit superior and inferior performance on an individual level due to managerial skills. Now we investigate whether funds are exhibiting abnormal performance that persists over consecutive time periods. To evaluate this, we apply both parametric and nonparametric tests.

4.1 Performance Potential of Strategies That Exploit Short-Term Persistence

Adopting the method of Hendricks et al. (1993), we evaluate the performance of executable strategies that exploit persistence in performance during 2000-2010. This way we can assess – both economically and statistically – whether persistence is present in our sample of Norwegian mutual funds. In short, we evaluate the risk-adjusted return from a zero-investment strategy of buying past winners and short-selling past losers. An in-depth explanation of the method is described in the following paragraph.

First, we rank all the funds based on their performance in the past M months. Previous literature indicates that the evaluation period on which the funds are ranked could be crucial for the outcome of the test. If the evaluation period is too short, the outcome may be too noisy due to chance factors, thus managerial skill may be hard to recognize. If the evaluation periods are too long, recent abnormal performance may be unrecognizable. Therefore, we let the evaluation period, M, vary from 1 to 24 months. Furthermore, the funds' performance is measured in three ways, i.e. according to their alpha estimated from the 4-factor model, alpha t-statistic and return, in order to compare the results. Funds with the highest performance in the selection period go into the winner portfolio (top decile) and funds with the lowest performance in the selection period go into the loser portfolio (bottom decile). These portfolios are then held for N months, where N varies from 1 to 24 months. The daily return in both the winner and loser portfolio is the average return of the funds in each portfolio. The portfolios are rebalanced afterwards according to the highest and lowest performance based on the past M months. This process continues until the end of the sample. This way we obtain a time series of post-ranking average daily returns on both the winner and loser portfolio. In addition we create a time series for the spread portfolio, calculated as the difference in returns between the winner and the loser portfolio. Finally, abnormal return, i.e. alpha, is estimated for each (winner, loser and spread portfolio) of the portfolios using the multi-period unconditional 4-factor model.[8] Table 7 reports the resulting alpha. To evaluate the statistical significance of our results, we apply the bootstrap method as described in Section 3.2.

We find significant evidence of short-term persistence in performance when funds are ranked on alpha. Rebalancing the portfolios every month results in the most significant performance spread between the top and bottom decile portfolios. For instance, rebalancing the portfolios each month based on the preceding 3 months abnormal return yields a 9.09% spread in yearly risk-adjusted return on average. The performance spread decrease as the rebalancing frequency decrease, but remains statistically significant for up to a 12-month holding period. After that the spread become statistically indistinguishable from zero. This is consistent with the findings of Bollen and Busse (2005).

Kosowski et al. (2006) argue that he is able to control better for differences in risk taking across funds by ranking funds on their alpha t-statistic rather than alpha itself. Our evidence shows that ranking on alpha t-statistic also yield significant short-term persistence. Furthermore, we also find evidence of short-term persistence when ranking on return. This is in contrast to the findings of Bollen and Busse (2005) which report that the post-ranking performance spread disappears when winners and losers are selected on funds' return. Still, the abnormal return generated from ranking on return is lower than those generated from ranking on abnormal return. In our sample data, an investor looking to utilize short-term persistence in fund performance should rank on past alpha in order to maximize profits.

Figure 5 illustrates the post-ranking performance of the top and bottom deciles for the entire sample period. Clearly, rebalancing monthly rather than yearly results in a larger performance spread. This confirms the finding of Bollen and Busse (2005) that persistence in performance is a short-lived phenomenon. Furthermore, we observe that the top decile portfolio in aggregate delivers abnormal return that is slightly positive (negative) when rebalancing monthly (yearly). From Table 8, we see that the aggregate alpha is statistically indistinguishable from zero. In other words, we cannot conclude from this test that there is persistence among superior funds.

The bottom decile portfolio is generating strongly negative abnormal return for the entire sample period, indicating strong persistence amongst inferior funds. Looking at Table 8, we observe that this result is statistically significant. In fact, underperformance amongst inferior funds persists in our sample for two years. These findings are in line with Teo and Woo (2001). Since the returns are net of fees, some of the negative alpha can be attributed to fund fees. However, the

[8] We also apply the single-period 4-factor model to this test. The results are reported in Appendix 6.

Table 7

Performance Spread of Portfolios Formed on Lagged Performance

The table reports the aggregate performance spread of the top- and bottom decile portfolios that are formed on lagged (a) alpha, (b) alpha t-statistic, and (c) return. For example, for table (a), funds are ranked according to their past alpha, where the length of the evaluation period, in number of months, varies for each row in the table. Then, a portfolio that is long in the top 10% funds and short in the bottom 10% funds is formed. This portfolio is held for a varying number of months, given by the columns in the table. This process is continued through the entire sample, and the alpha obtained from this investment strategy is displayed in the table. Numbers marked with three stars (***) are statistically significant on a 1% level, two stars (**) on a 5% level, and one star (*) on a 10% level.

a) Rank on Alpha

	Holding Period				
Evaluation Period	1	3	6	12	24
1	5.69***	3.18**	2.07*	4.01***	1.50
3	9.09***	7.21***	3.67**	4.15***	1.56
6	6.73***	5.95***	3.52**	4.68***	-0.49
12	7.78***	6.59***	6.45***	4.28***	-2.21*
24	6.30***	5.22***	3.56**	1.93	2.84**

b) Rank on Alpha T-Statistic

	Holding Period				
Evaluation Period	1	3	6	12	24
1	4.81***	2.97**	4.21***	6.53***	1.36
3	5.72***	4.57***	-0.44	2.18*	-0.50
6	5.26***	5.03***	3.94***	4.61***	-0.32
12	7.57***	6.22***	6.50***	3.26**	-2.14*
24	5.43***	4.77***	3.78**	2.72**	3.02**

c) Rank on Return

	Holding Period				
Evaluation Period	1	3	6	12	24
1	3.93***	1.26	-0.49	2.15*	0.01
3	4.79***	3.93***	1.32	2.00	0.57
6	4.82***	4.88***	2.06*	3.97***	-1.30
12	6.62***	5.28***	4.40***	2.44*	-1.50
24	2.44*	3.32**	2.09*	1.06	2.33*

average annual fund fee is 1.77% and the maximum fee 2.00%, while the annual alpha for the loser portfolio is -7.07% (when rebalancing the portfolio monthly based on the performance the past 12 months). Thus, the fee charged by fund managers cannot fully explain the negative alpha delivered by the inferior funds. The negative alpha is therefore attributable to persistence in lack of skill, i.e. fund managers are consistently either mistiming the market or picking bad stocks.

From Figure 5 we observe tendencies of co-variation between the portfolios. This is especially evident in the period 2001-2002 and 2008-2010. This draws us to conclude that the abnormal return for the top and bottom decile portfolios are somewhat correlated through the sample period. Thus, there is some common factor that describes fund returns that is not captured in the 4-factor model. For instance, this common factor might be attributable to fund industry characteristics, e.g. legal restrictions on how the equity funds are allowed to invest, or it can relate to inflow and outflow to the funds. However, we do not investigate this issue further.

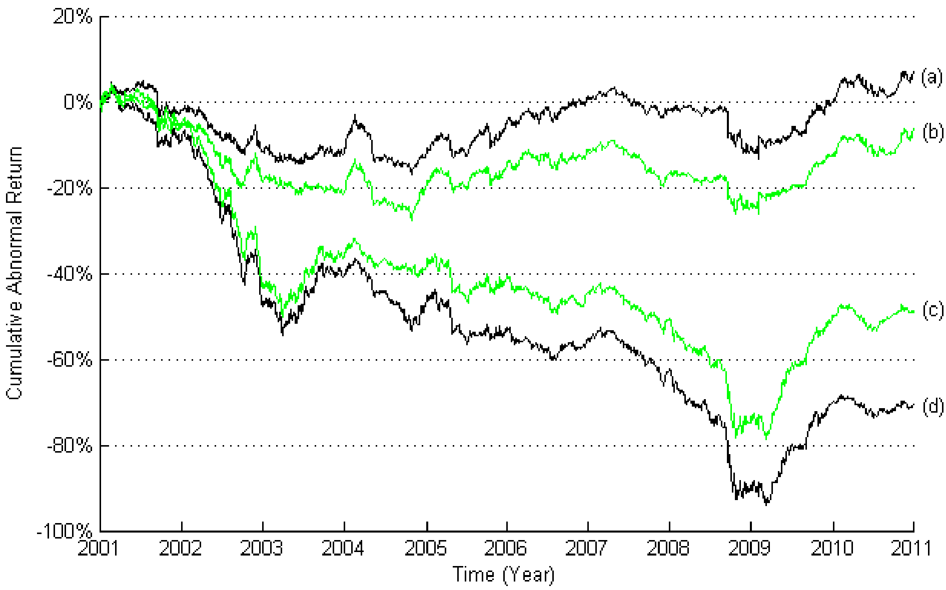

Figure 5. Cumulative abnormal returns for the top and bottom decile portfolios. The figure plots the cumulative abnormal return for (a) the top decile portfolio with monthly rebalancing, (b) the top decile portfolio with yearly rebalancing, (c) the bottom decile portfolio with yearly rebalancing and (d) the bottom decile portfolio with monthly rebalancing. For all cases, the funds are ranked on alpha estimate for the past 12 months. From the regression using the multi-period 4-factor model, we obtain an alpha estimates and a vector with daily residuals for each year. The abnormal return for each day is calculated as the sum of the alpha for the given year and the residual for that day.

Table 8

Performance of Top and Bottom Portfolios Formed on Lagged Alpha

The table reports the aggregate performance of (a) the top decile portfolio and (b) the bottom decile portfolios. The portfolios are formed on lagged alpha with a varying length of the evaluation periods (in number of months), given by the rows in the tables. To create table a), the funds are ranked according to past (1-24 months) alpha. A portfolio that consists of the top 10% funds are then created and held for a given number of months (varying from 1-24 months), before rebalancing the portfolio. This process is continued through the entire sample, and the alpha reported in the table is the aggregate alpha from applying this investment strategy through the entire sample period. Table b) is created the in the same manner, except that the portfolio consists of the bottom 10% funds. Numbers marked with three stars (***) are statistically significant on a 1% level, two stars (**) on a 5% level, and one star (*) on a 10% level.

a) Top Decile

	Holding Period				
Evaluation Period	1	3	6	12	24
1	0.53	-0.72	-0.97	2.21**	-0.47
3	1.34	1.30	-1.30	-0.49	-0.79
6	-0.07	0.26	-1.88**	-1.48*	-4.01***
12	0.70	0.08	0.29	-0.63	-4.76***
24	1.43*	1.02	-0.38	-1.14	-2.58**

b) Bottom Decile

	Holding Period				
Evaluation Period	1	3	6	12	24
1	-5.15***	-3.89***	-3.03***	-1.80*	-1.97**
3	-7.75***	-5.92***	-4.97***	-4.64***	-2.34**
6	-6.80***	-5.69***	-5.40***	-6.16***	-3.52***
12	-7.07***	-6.51***	-6.16***	-4.91***	-2.55**
24	-4.87***	-4.20***	-3.95***	-3.07***	-5.42***

In summary, we conclude that there exists persistence among losers in Norwegian mutual funds. We find, however, no abnormal return on superior funds. Still, we cannot conclude that there is no persistence in the performance among the superior funds. For example, some funds may be exhibiting superior persistence relative to the sample, even though they are not producing a positive alpha. Therefore, in the next sections we perform nonparametric tests which rely only on relative ranking of funds.

4.2 Nonparametric Two-Period Tests

In the previous section, we were not able to conclude about the persistence among superior funds relative to the benchmark given by the 4-factor model. However, there might be some funds that are persistently superior relative to the average fund industry. In this section, we investigate persistence in performance using two-period nonparametric tests.[9] By construction, these tests measure funds' persistence relative to the other funds in the sample. This allows us to draw conclusions on whether certain funds perform consistently better or worse than other funds. Specifically, we address whether funds are, over short-term horizons, consistently ranked in the top and bottom portfolios. Eling (2009) argues that nonparametric tests are more robust than parametric tests.

Following Carhart (1997) and Karoui and Meier (2009), we construct a contingency table of the decile rankings over two subsequent intervals. The funds are ranked based on their risk-adjusted performance and put into deciles over an initial and subsequent time window. Then, we count the number of times a fund ends up in one of the deciles over the second period conditional on its ranking over the first period. We create a contingency table for four different time windows. The contingency tables are displayed in Figure 6.

From the figure, it is apparent that winners are more likely to remain winners and losers are more likely to remain losers. Consistent with our results in the previous section, especially the bottom performing funds displays persistence in subsequent periods. Our most convincing result is that of 1-month time windows. There are two explanations for this. First, as we increase the time window our sample data diminishes and we may have too few observations to find statistical significant evidence of persistence. Second, a larger time window could lead to fewer observations of persistence in performance since persistence tends to be a short-term phenomenon (Bollen and Busse, 2005). Further, we observe from Figure 6 (a) that last month's winners are frequently becoming next month's losers and vice versa. This is consistent with gambling behavior by mutual funds (Carhart, 1997).

We also perform the cross-product ratio (CPR) test to check whether there exists persistence in our sample. Whereas Figure 6 presents the results visually, the CPR test quantifies the statistical significance of the persistence. Similar to Agarwal and Naik (2000), we denote those funds that are winners (losers) over two consecutive periods WW (LL). Those funds that are winners in the first period and losers in the second period are denoted WL, and LW for the opposite. The CPR is the cross product ratio of the funds which show persistence to the funds that do not:

$$CPR = \frac{(WW \cdot LL)}{(WL \cdot LW)} \tag{11}$$

The null hypothesis in the test is that each of the four cases, WW, LL, WL, and LW, is equally likely, i.e. there is no persistence in performance and the CPR is equal to one. The statistical significance of CPR can be tested using the standard

[9] Performance evaluation in nonparametric tests is addressed using the unconditional 4-factor model. Since the measurement windows are from one month to a year, the multi-period model and the single-period model are the same for the purpose of nonparametric tests.

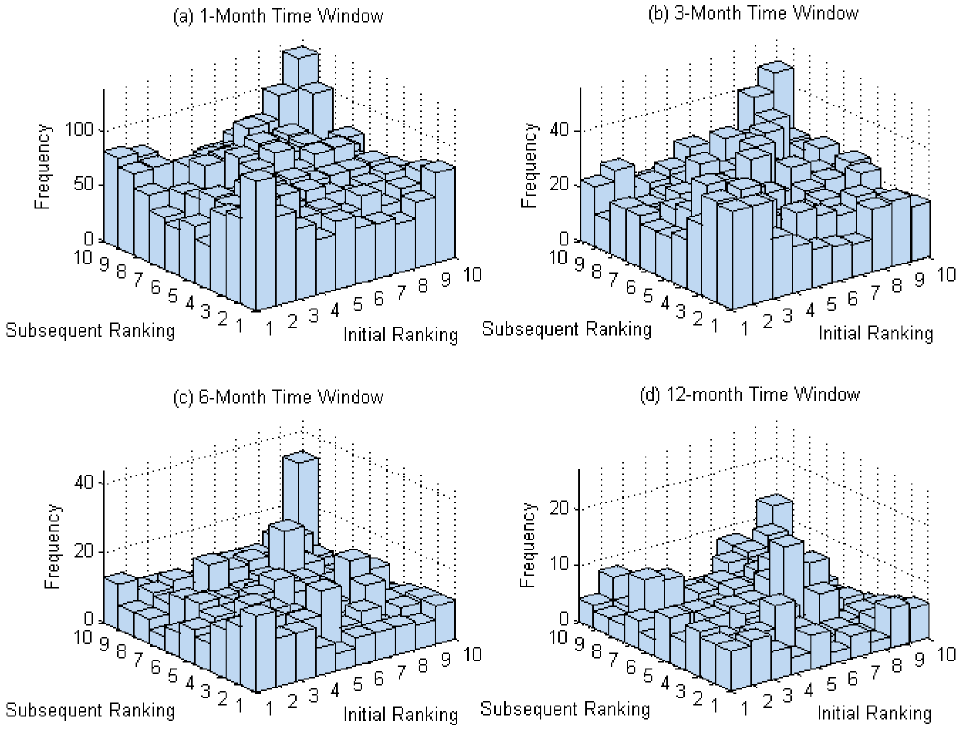

Figure 6. Contingency tables of initial and subsequent performance rankings. The figure illustrates to which degree the funds in our sample are consistently ranked in two consecutive periods. Each plot (a)-(d) has different length of the time periods. The top decile portfolio has rank 1 and the bottom decile portfolio has rank 10. It is evident from the figure, particular in a) with one-month time windows, that funds are most likely to achieve a subsequent ranking of 10 given an initial ranking of 10. The same can be observed for the top decile portfolio.

error $\sigma_{\ln CPR}$ of the natural logarithm of CPR. The resulting Z-statistic is the ratio of the natural logarithm of the CPR to the standard error of the natural logarithm:

$$Z = \frac{\ln CPR}{\sigma_{\ln CPR}} = \frac{\ln CPR}{\sqrt{\frac{1}{WW} + \frac{1}{WL} + \frac{1}{LW} + \frac{1}{LL}}} \tag{12}$$

Corresponding to the standard normal distribution, a value of Z greater than 1.96 indicates significant persistence at the 5% confidence level. We vary the number of portfolios from two to ten portfolios.

We find significant evidence of persistence in performance when we perform the CPR test on our sample of 64 funds. The statistical significance of the test weakens as the measurement horizon is extended (see Table 9), which implies short-term

Table 9
Persistence Results from the Cross-Product Ratio Test

The table reports the Z-statistic obtained from the cross-product ratio (CPR) test with a varying portfolio size (given by the rows in the table) and a varying time-period length (given by the columns). The observed Z-statistic is compared to the critical value given by the Z-test. That is, if the absolute value of the test statistic is greater than 1.96, then there is persistence in the sample for the given size of the portfolio and time-window at a 5% significant level. Numbers marked with three stars (***) are statistically significant on a 1% level, two stars (**) on a 5% level, and one star (*) on a 10% level.

	Length of Time Window			
Number of Portfolios	1	3	6	12
2	4.06***	4.35***	3.22***	1.80*
4	5.27***	4.64***	3.30***	2.34**
6	4.00***	5.07***	3.47***	1.77*
8	4.09***	4.18***	3.25***	2.09**
10	4.05***	3.48***	3.36***	1.67*

persistence. We find persistence in mutual fund performance at short horizons of up to six months significant at a 1% level.

Performing the CPR test on individual funds does not provide very significant results. On the most, 10 funds portray persistence on a 10% significance level when dividing into 4 portfolios. These weak results are due to few observations when testing on individual funds opposed to the joint sample test, which make the CPR test inappropriate for this purpose.

In summary, the results on the two-period nonparametric framework show short-term persistence in relative performance among the funds in our sample. To further investigate the presence of short-term persistence, we apply a multi-period framework in the following section.

4.3 Nonparametric Multi-Period Test (Kolmogorov-Smirnov Test)

Agarwal and Naik (2000) distinguish between two-period and multi-period nonparametric tests that can be used to examine performance persistence. In the first case, two consecutive time periods are compared to each other, while in the multi-period case, more than two consecutive time periods are considered. In this section, we complement our results from the two-period framework with a multi-period approach. That is, we divide into winners and losers within every evaluation period and count consecutive wins and losses for each fund throughout the sample period. Eling (2009) advocates the use of multi-period tests to obtain more robust results on performance persistence. As thoroughly discussed by Eling, both two-period and multi-period tests may produce spurious results by the nature of the tests. When dividing the sample into winners and losers, we have a very different treatment of very similar fund performance at the thresholds [see also Blake and Timmermann (2003)]. This is most evident when dividing into only two portfolios, but the same effects arise when ranking into higher number of portfolios. Yet, while two-period tests rely on a single threshold, multi-period tests depend on series of thresholds and will thus not discriminate funds with similar performance that severely. Since we are now tracking the history of individual fund rankings across the whole sample period, in contrast to only two consecutive periods, we are better able to separate true persistence from chance.

Agarwal and Naik (2000) and Eling (2009) both perform multi-period tests for performance persistence of hedge funds by using the Kolmogorov-Smirnov (KS) test. Their results coincide in finding a lower level of persistence in the multi-period framework than under the traditional two-period framework, but argue that the multi-period framework provide the most correct results, due to the robustness of the KS test.

4.3.1 Implementation of the KS Test

First, we rank funds on alpha and t-statistic[10] from a monthly unconditional 4-factor model based on daily data. Based on this fund ranking we assign every fund with the label W (winner), L (loser) or N (neutral) for each month. Funds will be assigned with the label N if they are not ranked in either the winner- or loser portfolio. In contrast to Agarwal and Naik (2000) and Eling (2009), we do not study our sample by only dividing into two portfolios. We also repeat the test for 3 to 10 number of portfolios. We investigate rankings based on both t-statistic alpha and alpha, and perform the test on our entire sample as a whole, as well as for individual funds.

Now we want to compare observations of consecutive wins in our (observed) data with the ones in a random sample. The random sample is obtained by assigning funds labels W, L or N randomly.[11] For the observed and the random sample we construct arrays (array O for observed, and R for random, respectively) counting numbers of consecutive wins. If a fund has a following history: WLWLNWWNLWWW, then the corresponding array for that fund is 1123. To obtain

[10] Agarwal and Naik (2000) and Eling (2009) perform similar multi-period tests on hedge fund performance, and apply the *appraisal ratio* as a performance ranking as well as alpha, to better control for volatility in fund returns. Appraisal ratio is defined as the alpha divided by the residual standard deviation resulting from a regression of the fund returns. The appraisal ratio accounts for the differences in the volatility of returns. When alphas have non-zero values, as is the case in our sample, the appraisal ratio is proportional to t-statistic and both measures will then lead to identical rankings (Lehmann and Modest, 1987). Therefore, for the purpose of the KS test, our results on t-statistic are then comparable to the studies using appraisal ratios.

[11] Note that the original unbalanced structure of the set is conserved in the random sample. That is, if fund i in our set is dead, or not born, in time t, then it will not be assigned with a fund label. This will ensure that the test does not suffer from biases due to set misspecifications.

array O for all the funds, we simply merge arrays for individual funds. The KS-test compares the empirical cumulative distribution functions (empirical CDF) of these two arrays, and the test statistic D is obtained as the maximum distance between the two empirical CDFs. For the given empirical CDFs, $F_{random,n}$ and $F_{observed,n'}$, the test statistic is:

$$D_{n,n'} = \sup_x \{F_{random,n} - F_{observed,n'}\} \tag{13}$$

Where $\sup_x$ is the supremum of the set of distances. Note that this test does not assume anything about the underlying distribution of the samples.

The test is formulated as follows:

> H_0: Consecutive wins (losses) in the observed data and in the random sample are generated from the same CDF.
>
> H_1: Cumulative distribution function for the consecutive wins (losses) in the observed data $F_{observed,n'}$ lies below[12] the cumulative distribution function for the consecutive wins (losses) of the randomly generated sample $F_{random,n}$.

Therefore, rejecting H_0 in favor of H_1 indicates presence of persistence (That is, a fund stays in the winner (loser) portfolio longer than can be explained by chance).

We apply the KS test both for the entire sample, and for individual funds. To perform the test on the entire sample, we construct a joint CDF containing the winner strings from all individual funds. When testing on individual funds, the winner-strings are tested separately for each fund. This will provide insight to whether there is significant persistence in the sample on aggregate, and if individual funds exhibit stand-alone significant persistence. As the above-mentioned method is constructed to find persistence among winners, we also perform the equivalent test for loser funds to detect persistence in inferior funds.

4.3.2 Results on the KS Test

Our results from the testing on the entire sample are summarized in Table 10. We find strong evidence of persistence among winners when we use monthly evaluation windows. However, when we increase the evaluation windows to 12 months, the persistence among winners disappears. This indicates that the persistence of superior funds is a short-lived phenomenon. Moreover, our evidence imply that the

[12] We use one-sided test because we expect persistence in our sample. This is due to the fact that a sample with persistence in performance should have less observations of a low number of consecutive wins and losses (e.g. W and WW) than a sample without persistence. Hence, the CDF of the persistent sample would lie underneath the CDF of the random sample.

performance of inferior funds persist up to one year. This result coincides with our findings from the parametric test in Section 4.1.

Table 10
Persistence Results from the KS Test on All Funds

The table presents the results from the Kolmogorov-Smirnov test performed on all funds together. We report, with varying portfolio size for rankings based on 1- and 12-month alpha, whether the entire fund sample exhibit superior and/or inferior persistence (See Appendix 7 for results on 3- and 6-month alpha). "Yes", presented with the corresponding p-value, indicates rejection of the null hypothesis of no persistence in performance. "Yes" marked with three stars (***) is statistically significant on a 1% level, two stars (**) on a 5% level, and one star (*) on a 10% level.

	1-Month Alpha		12-Month Alpha	
Number of portfolios	Persistency in wins?	Persistency in losses?	Persistency in wins?	Persistency in losses?
2	Yes**/0.036	Yes***/0.001	No/0.594	No/0.358
3	Yes***/0.002	Yes***/0.001	No/0.555	Yes**/0.036
4	Yes***/0.000	Yes*/0.062	No/0.397	Yes**/0.035
5	Yes***/0.001	Yes**/0.039	No/0.639	Yes**/0.020
6	Yes***/0.000	No/0.112	No/0.941	No/0.106
7	Yes***/0.000	Yes***/0.002	No/0.756	Yes**/0.016
8	Yes***/0.000	No/0.247	No/0.423	Yes**/0.023
9	Yes***/0.000	No/0.194	No/0.603	Yes*/0.057
10	Yes***/0.000	No/0.145	No/0.477	Yes**/0.011

Next, we investigate performance persistence of individual funds. Table 11 reports the results using 1-month evaluation periods. For comparative purposes, we present the persistence result on individual funds when ranking on both alpha and alpha t-statistic. Clearly, we find funds that consistently are ranked at the top and funds that consistently are ranked at the bottom.[13]

Results on individual funds seems to highlight a difference in the ranking criteria of funds. Ranking funds on alpha results in more funds exhibiting persistence than what is the case for the t-statistic ranking. This is most evident for the winner portfolios, where the number of funds showing significant persistence is reduced radically when ranking on t-statistics. Looking at the winner funds from the alpha-ranking that are excluded in the t-statistic ranking, one finds that the majority of these funds have an outspoken SMB strategy. Most of these funds, which are more exposed to the SMB factor than the average mutual fund, do not appear in the t-statistic results. T-statistics control for more volatility, and thus uncertainty, in the alpha measurements. There might be some common risk elements for small companies that are not accounted for in the SMB factor, or simply just the fact that small funds have more non-diversifiable firm-specific risk. This could lead

[13]Our analysis using 3-, 6-, and 12-month evaluation periods reveals that this also applies using longer evaluation periods (see Appendix 8 for an overview). Additionally, when using longer evaluation periods (e.g. 6-month) we observe that some funds are both appearing as consecutive losers and consecutive winners. This is consistent with the alternating tendencies observed in the contingency table in section 4.2. However, one should bear in mind that longer evaluation periods lead to fewer observations, which results in a less powerful KS-test.

Table 11
Persistence Results from the KS Test on Individual Funds

The table presents the results from the Kolmogorov-Smirnov test performed on all funds individually. We report, with varying portfolio size for rankings based on alpha and alpha t-statistic, whether individuals funds exhibit superior and/or inferior persistence based on monthly rankings (See Appendix 8 for results on 3-,6- and 12-month alpha and t-statistic). The appearance of a fund number indicates rejection of the null hypothesis of no persistence in performance for the given fund. Fund numbers in *italic* indicate SMB funds. These are funds that have an outspoken strategy of investing in small or medium sized companies, and may thus experience higher volatility in returns than the market average. All results are on a 5% significance level.

	1-Month Alpha		1-Month T-statistic	
Number of portfolios	Superior Funds	Inferior Funds	Superior Funds	Inferior Funds
2	54	2,6,17,22,27,60	35, 58	17, 40, 58
3	36,*53*,59	11,27,50	36, 59	-
4	10,*15*,32,*53*,59	7,24,37	-	4, 40
5	10,*15*,*31*,*53*,59	7.37	-	38
6	7,10,*15*,*31*,50,*53*	37	*53*	3
7	10,*15*,*31*,46,*53*	-	*53*	-
8	10,*15*,*31*,46,*53*	37	*53*	-
9	10,*15*,46,*53*	-	*53*	-
10	*15*,46,*53*	-	46	-

to higher variance in the residuals, and lower t-statistics relative to the alpha estimates, which would explain the absence of some SMB funds in the t-statistics ranking. We should not however conclude that the t-statistic ranking is superior to the alpha-ranking in our case. In general, there might be non-normal implications in fund return residuals that lead to biased t-statistic rankings. Additionally, if there exist common risk factors for the SMB-funds not accounted for in the 4-factor model, then these possible non-normalities would even be cross-correlated, and would lead to a biased t-statistic ranking for these particular funds. That is, the t-statistic is supposed to control for noise in alpha measurements by dividing with the standard deviation. However, if the estimate of the standard deviation is biased, this ranking method would only lead to more noise.

Teo and Woo (2001) argue that it might be unwise to compare funds which have different investment styles. Agarwal and Naik (2000) and Eling (2009) also control for different strategies among hedge funds. If fund managers are restricted in their investment decisions by the fund style, style persistence might be the reason for the performance persistence rather than superior managerial skill. If, as discussed above, the SMB factor does not fully capture the return commonalities of the SMB strategy, this could lead to bias in the fund ranking.

Finally, the results from our testing on the full sample and individual funds for different portfolio sizes give the following observation. When we divide the fund sample in two, the bottom-half performers are more persistently losers than the top-half performers are winners. For example, as shown in Table 11, this is evident from the six loser funds compared to the one winner (both from ranking on alpha). From the testing on smaller portfolios we find the opposite. This shows that while the funds in the bottom half alternate on being in the very bottom, the very best performers are more consistently on the top.

Based on the multi-period nonparametric tests, we conclude that there exists short-term persistence in our sample. Specifically, some funds exhibit persistence in superior and inferior performance relative to the other funds in our sample. In addition, inferior persistence seems to endure for longer time periods than superior persistence.

4.4 Robustness of the Results on Persistence

4.4.1 Effect of Dying Funds on Persistence

Persistence testing can potentially be influenced by characteristics in returns for dying funds. If terminated funds exhibit consistently inferior performance in the last months before fund termination, possible implications in the persistence tests arise. First, the inferior performance we find in our sample could to some extent be explained by the behavior of the dying funds the last months before extinction. Second, if the dying funds exhibit abnormal performance the last months before extinction, the losing string (e.g. LLLL) of the dead funds could possibly have been longer if they had not terminated, which in turn would mean that our results on inferior persistence is even stronger. However, we perform a probit analysis and find no significant relationship between performance and extinction. Size is a significant factor for fund survival, but this has no implications for our persistence analysis since alpha does not correlate with either size or fund extinction. See Appendix 9 for regression results.

4.4.2 Sample Size Implications for KS Test on Individual Funds

To understand the significance of our results on individual funds in the KS test, we test how many funds in our sample that are expected to exhibit significant persistence in performance by chance alone. Specifically, it might be the case that some of the funds would exhibit persistence only as a result of the sample size. The test is conducted by performing the KS test on random fund samples. We randomize our fund sample in 1,000 iterations, and then count the total number of funds rejecting the null hypothesis of no persistence for the random fund sample within each iteration. This is repeated for all portfolio sizes both for superior and inferior persistence. This way we obtain the theoretical distribution of funds rejecting the null by random for a given portfolio size, and find the corresponding probability that the observed number of funds in our sample of 64 funds rejecting the null is greater than what would be expected purely by chance.

Table 12
Significance of the KS Test Results on Individual Funds

The table shows the results from the Kolmogorov-Smirnov test on 1,000 random fund samples for different portfolio sizes with evaluation periods of one month. Column one and three show how many funds that exhibit persistence for winners and losers, respectively, in our sample. Column two and four show the probability that the number of funds exhibiting persistence in our sample would occur purely by chance. Numbers marked with three stars (***) are statistically significant on a 1% level, two stars (**) on a 5% level, and one star (*) on a 10% level.

	Superior Funds		Inferior Funds	
Number of portfolios	Number of funds	p-value	Number of funds	p-value
2	1	0.584	6	0.000***
3	3	0.022**	3	0.022**
4	5	0.001***	3	0.006***
5	5	0.000***	2	0.027**
6	6	0.000***	1	0.187
7	5	0.000***	-	1.000
8	5	0.000***	1	0.110
9	4	0.000***	-	1.000
10	3	0.000***	-	1.000

In Table 12 we address the significance of the winner and loser results when testing on the monthly alpha ranking. From the probabilities obtained on, we conclude that when dividing into 2 portfolios, the fund that appear as winner may actually be there by chance. As for the losers, we should not draw any conclusions upon the the results on portfolio sizes from 6 to 10 portfolios.

5 Conclusion

In this paper we investigate the performance and persistence of Norwegian mutual funds using a dataset of daily returns from 2000 to 2010 free of survivorship bias. To evaluate performance, we apply a multi-period version of the Carhart (1997) 4-factor asset-pricing model, as we find that the factor loadings vary significantly with time. Following Kosowski et al. (2006) and Fama and French (2010), we examine the statistical significance of the performance and persistence by means of a flexible bootstrap procedure that avoids issues of non-normality in the return residuals.

We find that Norwegian mutual fund investors in aggregate realize net returns that underperform the 4-factor benchmark by approximately the fees. When we study individual funds, we find that there exist fund managers that are able to generate higher returns than what is expected for the given level of risk in the stock portfolio. For instance, the top decile funds generate an abnormal return of 4.46% per year on average. We reject that the superior performance is due to luck, and conclude that the managers of these funds inhabit skills. Moreover, we find that these skills can most likely be explained by stock-picking abilities. Similarly, we find that there are funds that exhibit significant underperformance compared to the benchmarks. For the bottom decile funds the risk-adjusted negative return constitutes 12.45% per year on average. Our results imply that the managers of the worst performing funds both are picking bad stocks and to some extent are mistiming the market.

The search for persistence provides strong evidence of short-term persistence in mutual fund performance. Buying last year's top decile portfolio of funds and selling last year's bottom decile portfolio yields an annual abnormal return of 7.78% when the funds are ranked on past abnormal return and the portfolios are rebalanced every month. If an investor were to rebalance the portfolios every year, while still ranking on past 12 months abnormal return, the annual excess return decrease to 4.28% on average. Even when we form portfolios based on past excess return, i.e. we rank funds on past return that is not risk-adjusted, we find the same pattern of short-term persistence in the performance spread.

When we investigate the top and bottom decile portfolios separately, we find that abnormally bad performance of the worst funds persist strongly. This is consistent with Bollen and Busse (2005). We cannot, however, reject that the excess return of the superior performing funds are statistically different from zero. Thus, relative to the benchmark returns in the 4-factor model, we cannot conclude upon whether the top performers exhibit performance persistence or not.

Furthermore, we perform nonparametric tests in both a two-period and a multi-period framework. These tests allow us to infer whether funds are consistently ranked as winners or losers relative to each other. The result from the two-period tests imply short-term persistence for up to six months for the top and bottom performing funds when the funds are ranked on past abnormal return. To provide more robust evidence of short-term persistence, we complement our results from the two-period tests with the Kolmogorov-Smirnov test [as described by Agarwal and Naik (2000) and Eling (2009)], which evaluate performance persistence using a multi-period approach. Now, we find that the performance of top performing funds persist on a monthly basis. That is, the winners of the past month are likely to be ranked in the top in the next month. Further, the results from the KS test show that the worst performing funds exhibit persistence over longer time horizons. This coincides with our result from the investment strategies of buying past winners and selling past losers.

Successful active management of a stock portfolio requires that the market is not efficient in the semi-strong form. That is, if the market is perfectly efficient, it is impossible for fund managers to exploit mispricing and create abnormal returns. However, as argued by Grossman and Stiglitz (1980), the markets cannot always be fully efficient, i.e. there must occur periods of mispricing in the market. Our results indicate that fund managers indeed are able to exploit these periods of market mispricing, but still are not able deliver abnormal performance for its investors due to the fees they charge. In other words, a fund investor could theoretically profit from buying last month's winners and short-selling last month's losers. However, in practical terms, the economic impact of this finding is questionable, since frequently rebalancing the winner and loser portfolio would infer trading cost exceeding an investors profits. As a consequence, Norwegian investors seem to be better off by investing in low-cost passive funds that seek to replicate the market return. At the very least, investors should avoid those funds that have performed poorly in the past.

References

Abinzano, I., Muga, L., and Santamaria, R. (2010). Do managerial skills vary across fund managers? Results using european mutual funds. *Journal of Financial Services Research*, 38(1):41–67.

Agarwal, V. and Naik, N. Y. (2000). Multi-period performance persistence analysis of hedge funds. *Journal of Financial and Quantitative Analysis*, 35(03):327–342.

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

Berk, J. and Green, R. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, 112(6):1269–1295. doi: 10.1086/424739.

Bhuvaneswari, D. P. (2011). Market timing ability of indian mutual fund managers under conditional markets. Working Paper.

Bickel, P. J. and Freedman, D. A. (1984). Asymptotic normality and the bootstrap in stratified sampling. *The Annals of Statistics*, 12(2):470–482.

Blake, D. and Timmermann, A. (2003). Performance persistence in mutual funds: An independent assessment of the studies prepared by Charles River Associates for the investment management association. Working Paper.

Bollen, N. P. B. and Busse, J. A. (2005). Short-term persistence in mutual fund performance. *Review of Financial Studies*, 18(2):569–597.

Brown, S. J., Goetzmann, R. G. I., and Ross, S. A. (1992). Survivorship bias in performance studies. *Review of Financial Studies*, 5(4):553–80.

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

Chevalier, J. A. and Ellison, G. (1997). Risk taking by mutual funds as a response to incentives. *Journal of Political Economy*, 105(6):1167–1200.

Cuthbertson, K., Nitzsche, D., and O'Sullivan, N. (2008). UK mutual fund performance: Skill or luck? *Journal of Empirical Finance*, 15(4):613–634.

Eling, M. (2009). Does hedge fund performance persist? Overview and new empirical evidence. *European Financial Management*, 15(2):362–401.

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

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–65.

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

Ferson, W. E. and Schadt, R. W. (1996). Measuring fund strategy and performance in changing economic conditions. *Journal of Finance*, 51(2):425–61.

Ghysels, E. (1998). On stable factor structures in the pricing of risk: Do time-varying betas help or hurt? *The Journal of Finance*, 53(2):549–573.

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

Hall, P. and Martin, M. A. (1988). On bootstrap resampling and iteration. *Biometrika*, 75(4):661–671.

Hendricks, D., Patel, J., and Zeckhauser, R. J. (1993). Hot hands in mutual funds: Short-run persistence of relative performance, 1974-1988. *Journal of Finance*, 48(1):93–130.

Henriksson, R. D. and Merton, R. C. (1981). On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills. *The Journal of Business*, 54(4):pp. 513–533.

Horowitz, J. L. (2003). Bootstrap methods for markov processes. *Econometrica*, 71(4):1049–1082.

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

Jagannathan, R. and Korajczyk, R. A. (1986). Assessing the market timing performance of managed portfolios. *The Journal of Business*, 59(2):217–35.

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

Jensen, M. C. (1968). The performance of mutual funds in the period 19451964. *Journal of Finance*, 23(2):389–416.

Kacperczyk, M., Nieuwerburgh, S. V., and Veldkamp, L. (2011). Time-varying fund manager skill. Working Paper 17615, National Bureau of Economic Research.

Karoui, A. and Meier, I. (2009). Performance and characteristics of mutual fund starts. *European Journal of Finance*, 15(5-6):487–509.

Kosowski, R., Timmermann, A., Wermers, R., and White, H. (2006). Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis. *Journal of Finance*, 61(6):2551–2595.

Lehmann, B. N. and Modest, D. M. (1987). Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons. *Journal of Finance*, 42(2):233–65.

Næs, R., Skjeltorp, J. A., and Ødegaard, B. A. (2009). What factors affect the Oslo Stock Exchange? Working Paper.

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

Otten, R. and Bams, D. F. (2004). How to measure mutual fund performance: Economic versus statistical relevance. Open access publications from Maastricht University, Maastricht University.

Sandvik, S. and Heitmann, R. K. (2010). Norwegian mutual funds: Performance, capital flow and persistence. Master's thesis, NTNU.

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

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

Sørensen, L. Q. (2009). Mutual fund performance at the Oslo Stock Exchange. Working Paper.

Teo, M. and Woo, S.-J. (2001). Persistence in style-adjusted mutual fund returns. Working Paper.

Treynor, J. and Mazuy, K. (1966). Can mutual funds outguess the market? *Harvard Business Review*, 44(4):131–136.

Warther, V. A. (1995). Aggregate mutual fund flows and security returns. *Journal of Financial Economics*, 39(2-3):209–235.

Wermers, R. (2000). Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *Journal of Finance*, 55(4):1655–1703.

Appendix 1

Table 13

Regression Results on the Equally Weighted Portfolio Using the Conditional 4-Factor Model

The table reports a stepwise reduction of the conditional 4-factor model. The least significant information variable is removed in each step. Column one reports yearly alpha. Column two to five represent the standard market-, SMB-, HML, and momentum factors. In column six, the results for the lagged 12-month NIBOR rate are displayed. The lagged return on the oil price is shown in column seven, while the lagged- interest rate and market return is represented in column eight and nine, respectively. The lagged interest spread is calculated as the 12-month T-bill (ST5X) minus the 3-month T-bill (ST1X). In column ten, the R-squared checks for explanatory power, while Durbin-Watson (DW,) testing for autocorrelation in the residuals, is presented in column eleven. T-statistics are presented in the parenthesis. Numbers marked with three stars (***) are statistically significant on a 1% level.

	Annual Alpha	MKT	SMB	HML	PR1YR	12-Month NIBOR$_{t-1}$	Oil Price$_{t-1}$	Interest Spread$_{t-1}$	MKT$_{t-1}$	R-Squared	DW
Model 1	-0.052	0.879***	0.001	0.001	-0.041***	0.010***	-0.049	0.187	0.014	0.933	2.108
	(0.028)	(57.071)	(0.140)	(0.083)	(-6.566)	(3.760)	(-0.276)	(0.244)	(0.072)		
Model 2	-0.048	0.879***	0.001	0.001	-0.041***	0.010***	-0.043	0.179		0.933	2.096
	(0.026)	(57.085)	(0.143)	(0.088)	(-6.570)	(3.761)	(-0.271)	(0.235)			
Model 3	-0.060	0.879***	0.001	0.001	-0.041***	0.010***	-0.046			0.933	2.097
	(0.029)	(57.096)	(0.150)	(0.091)	(-6.576)	(3.764)	(-0.287)				
Model 4	0.024	0.878***	0.001	0.001	-0.041***	0.010***				0.933	2.096
	(0.013)	(57.216)	(0.155)	(0.087)	(-6.583)	(3.836)					

Appendix 2

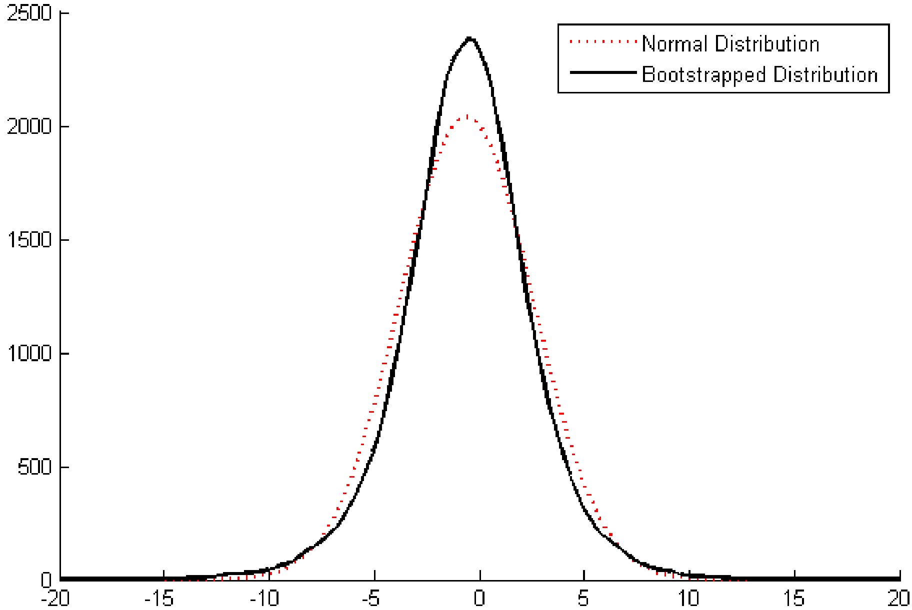

Figure 7. Normality check for the estimated distribution of alphas. The graph shows a simulation of the joint distribution of annual alphas for the 64 funds using the bootstrap method (residual resampling under a null hypothesis of zero alpha), compared to a theoretical normal distribution under the same null hypothesis. The distribution is obtained by resampling the residuals for each of the 64 funds in 10,000 iterations, and regressing on these simulated time-series with the multi-period 4-factor model. A kernel distribution of these alphas is depicted as the black solid line, while the normal distribution is illustrated with a dotted red line. The x-axis represents yearly alpha in percent.

Appendix 3

Table 14
Fund Fee Summary

The table reports the average, median, minimum and maximum fee for the funds in our sample that were alive as of December 2010. Specifically, column one reports the sales fee, i.e. the fee charged by funds when investors buy fund shares. The second column report the redemption fee, i.e. the fee charged when investors sell fund shares. The third column reports the annual management fee charged by funds.

	Sales Fee	Redemption Fee	Annual Management Fee
Mean	1.34%	0.36%	1.69%
Median	1.00%	0.30%	1.80%
Min	0.00%	0.00%	0.90%
Max	3.00%	1.50%	2.00%

Appendix 4

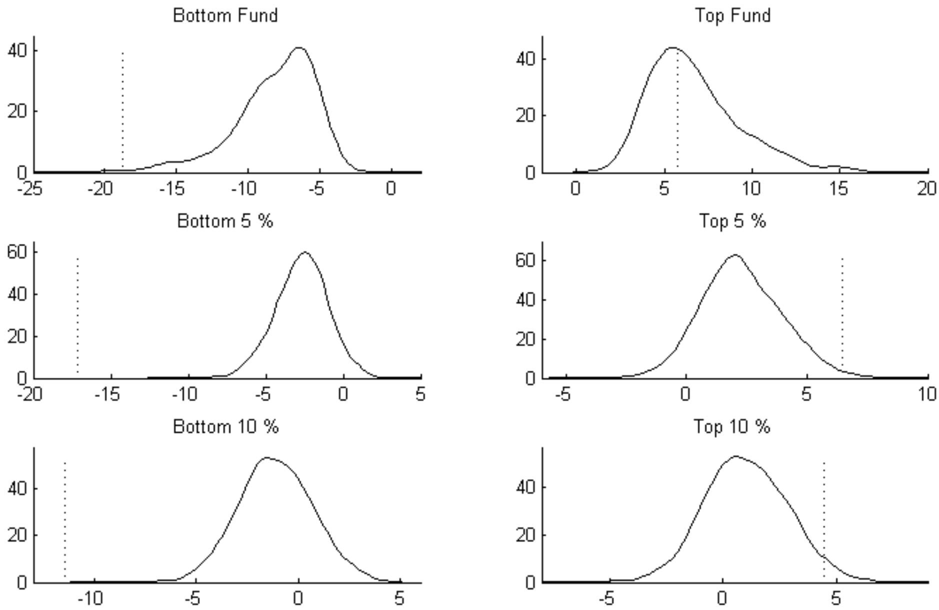

Figure 8. Estimated alphas (ranked on t-statistics) versus bootstrapped alpha distributions for individual funds. This figure plots kernel density estimates of the bootstrapped unconditional 4-factor model alpha distributions (solid line). To produce the distribution of the alpha estimates, we first create random returns given by the residuals in the in the regression on the actual return data. Then, we generate 10,000 alpha estimates from the OLS regression on those random returns. The x-axis shows the alpha performance measure per year and the y-axis show the kernel density estimate. The vertical dotted line shows the actual (estimated) fund alpha. The fund portfolios are ranked on alpha t-statistics, where the highest ranked fund has the highest t-statistic over the sample period. The same ranking is performed on the bootstrapped alphas to create bootstrapped portfolios.

Appendix 5

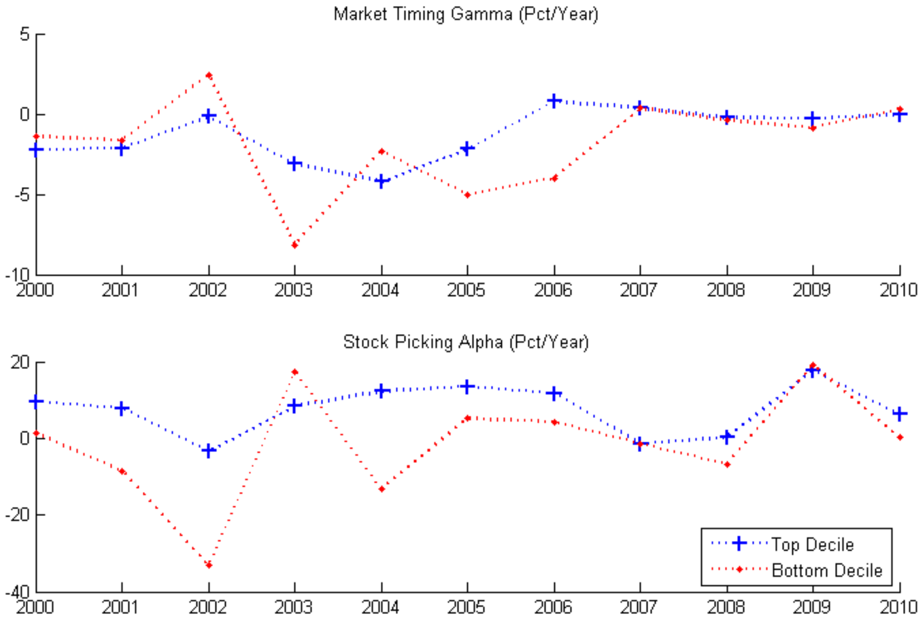

Figure 9. Illustration of time-varying market timing measures. The figure shows the estimated TM market timing coefficients and stock-picking alphas for the top- and bottom decile portfolios through 2000-2010. The fund ranking is based on the average alpha obtained from the multi-period 4-factor model applied on the entire sample period. The blue line depicts the average estimates for the top decile portfolio, while the red line shows the bottom decile portfolio. It is evident that market timing and stock-picking are negatively correlated, pointing out that fund managers shift their focus (micro or macro) with time. The top portfolio is on average performing better both with respect to market timing and stock-picking. In addition, the top portfolio has less volatility in the skill measurements than the bottom portfolio.

Appendix 6

Table 15
Performance Potential of Portfolios Formed on Lagged Alpha

We replicate the procedure from section X were we investigate the performance potential of exploiting short-term persistence. However, we now regress on the daily return series using the single-period version of the 4-factor model. The table reports the aggregate performance of (a) the top decile portfolio, (b) the bottom decile portfolios and (c) the spread portfolio. The portfolios are formed on lagged alpha with a varying length of the evaluation periods. To create table (a), the funds are ranked according to past (1-24 months) alpha. A portfolio that consist of the average return of the top 10% funds are then created and held for a given number of months, before rebalancing the portfolio. This process is continued through the entire sample, and the alpha reported in the table is the aggregate alpha from applying this investment strategy through the entire sample period. Table (b) is created the in the same manner, except that the portfolio consist of the average bottom 10% funds. The spread portfolio [Table (c)] is created as a long position in the top 10% funds and a short position in the bottom 10% funds.

a) Top Decile Portfolio

	Holding Period				
Evaluation Period	1	3	6	12	24
1	1.06	-0.30	-0.53	-1.50	-2.40**
3	1.77*	1.87*	-0.48	-0.99	-0.95
6	0.33	0.40	-1.79*	-2.00**	-4.10***
12	1.39	0.62	0.32	0.10	-2.55**
24	2.08**	1.61*	0.49	0.03	-1.05

b) Bottom Decile Portfolio

	Holding Period				
Evaluation Period	1	3	6	12	24
1	-3.56***	-2.19**	-1.79*	-1.90*	-2.28**
3	-5.69***	-3.90***	-3.22***	-4.11***	-2.14**
6	-4.33***	-3.09***	-3.30***	-4.34***	-1.84*
12	-4.65***	-4.14***	-3.82**	-3.10***	-1.56
24	-2.48**	-1.85*	-1.67*	-1.66*	-2.68**

c) Spread Portfolio

	Holding Period				
Evaluation Period	1	3	6	12	24
1	4.63***	1.89	1.26	0.41	-0.12
3	7.46***	5.76***	2.74**	3.13**	1.20
6	4.66***	3.49**	1.51	2.33*	-2.26*
12	6.04***	4.76***	4.14***	3.20**	-0.98
24	4.56***	3.46**	2.16*	1.69	1.63

Appendix 7

Table 16

Persistence Results from the Kolmogorov-Smirnov Test on All Funds

The table presents the results from the Kolmogorov-Smirnov test performed on all funds together. We report, with varying portfolio size for rankings based on 3- and 6-month alpha, whether the entire fund sample exhibit superior and/or inferior persistence. "Yes", presented with the corresponding p-value, indicates rejection of the null hypothesis of no persistence in performance. "Yes" marked with three stars (***) is statistically significant on a 1% level, two stars (**) on a 5% level, and one star (*) on a 10% level.

	3-Month Alpha		6-Month Alpha	
Number of portfolios	Persistency in wins?	Persistency in losses?	Persistency in wins?	Persistency in losses?
2	No/0.116	Yes*/0.053	Yes**/0.029	Yes**/0.042
3	Yes**/0.013	No/0.142	No/0.100	No/0.118
4	Yes**/0.022	No/0.211	Yes*/0.071	No/0.171
5	Yes*/0.058	Yes**/0.020	Yes*/0.084	Yes**/0.045
6	Yes*/0.095	Yes**/0.024	Yes*/0.094	Yes**/0.031
7	Yes**/0.016	Yes**/0.029	No/0.150	Yes**/0.033
8	Yes**/0.011	No/0.274	No/0.255	Yes*/0.060
9	Yes**/0.023	No/0.126	No/0.240	Yes**/0.011
10	No/0.112	No/0.112	No/0.391	Yes***/0.009

Appendix 8

Table 17
Persistence Results from the KS Test on Individual Funds

The table presents the results from the Kolmogorov-Smirnov test performed on funds individually. We report, with varying portfolio size for rankings based on alpha and alpha t-statistic, whether individuals funds exhibit superior and/or inferior persistence from 3-, 6-, and 12-month rankings. The appearance of a fund number indicates rejection of the null hypothesis of no persistence in performance for the given fund. Fund numbers in *italic* indicate SMB funds. All results are on a 5% significance level.

	3-month Alpha		3-month T-statistic	
Number of portfolios	Superior Funds	Inferior Funds	Superior Funds	Inferior Funds
2	61	29,37,41	16,61	37.41
3	16,54,59	31,37,41,45	54,59	37,41,42,58
4	7,24,38,54	37,41,45, *53*	38,54	*31*,37,41,42
5	38,54	37,41,45	54	*31*,34,37,41
6	38,54	37,41,*53*	54	41,53
7	38,54	37,41,*53*	54	41
8	38,54	37,41	54	-
9	*25*,38,54	37,41	54	-
10	*25*,38	37,41	54	-
	6-month Alpha		**6-month T-statistic**	
Number of portfolios	Superior Funds	Inferior Funds	Superior Funds	Inferior Funds
2	10, *15*,35,44,55	10,17,19	*15, 44*	4
3	6,10,*53*	10,41	-	10
4	10,*15*	10,*15*	-	-
5	*53*	10, *15,41*	*15*	*15*
6	10,*31,53*	10, *15*,37,41	-	*15*,37,41,42
7	10,*31,53*	*15*,37,41,42	-	37,41,42
8	*53*	37,41,42	52	37,41,42
9	-	37,41,42	-	37,41,42
10	-	37,41,42	-	37
	12-month Alpha		**12-month T-statistic**	
Number of portfolios	Superior Funds	Inferior Funds	Superior Funds	Inferior Funds
2	14,44,48,54	4,11,39	20,44,54	*39*
3	36	-	36.44	-
4	26,27,54	*53*	54	37
5	54	37,41,42,*53*	54	37.42
6	54	37,41,42	-	37,41,42
7	-	37,41,42	-	37,41,42
8	-	37,41,42	-	37,41,42
9	-	37,41,42	-	31,37,41,42
10	-	37,41,42	-	*31*,37,41

Appendix 9

Table 18
Probit Analysis of Fund Extinction

The table shows the results from the pooled probit regression on fund extinction. The coefficients are interpreted as determinants for fund extinction. The model is described as Pr(Y=1|X)=$\Phi(X_i'\beta_i)$ where Pr denotes probability and Φ is the Cumulative Distribution Function (CDF) of the standard normal distribution. The parameters β are estimated by maximum likelihood. The regression is based on yearly observations, e.g. the size of the fund, and the return of the fund for each year. If a fund dies during the sample period, it will be denoted as a dependent variable Y with value 0, and the corresponding independent factors will be evaluated for 252 days prior to extinction. If a fund survives, the dependent variable will be denoted as 1. Positive coefficients imply that the factor is positively correlated with fund survival. T-statistics are presented in the parenthesis. Numbers marked with three stars (***) are statistically significant on a 1% level.

	ln(Size)/Avg Size	Alpha_{t-1}	Raw Returns/Avg Returns	Raw Returns	ln(Net Flows)	Alpha
Coefficient	6.534***	-0.025	3.00E-04	1.00E-04	1.00E-04	-0.071
	(4.376)	(-0.729)	(2.893)	(1.131)	(1.160)	(-0.959)
Coefficient	7.869***	-0.033	2.00E-04	1.00E-04		
	(5.379)	(-0.951)	(1.930)	(1.388)		
Coefficient	9.405***	0.003	1.00E-04			
	(9.861)	(0.161)	(1.768)			

Appendix 10

Table 19
List of Mutual Funds

The list shows the fund names of all 64 funds in our sample, sorted by the date of formation. Column one and three shows the fund number assigned to the fund, and column 2 and 4 displays the corresponding fund name. Section 4.3.2 uses fund numbers to describe which individual funds that exhibit persistence.

Fund Number	Fund Name	Fund Number	Fund Name
1	Storebrand Norge	33	Romsdal Fellesbank Aksjefond
2	Avanse Norge (I)	34	Delphi Vekst
3	Nordea Vekst	35	Alfred Berg Norge +
4	Nordea Avkastning	36	Storebrand Verdi
5	Orkla Finans Investment Fund	37	Globus Norge
6	DnB NOR Norge (I)	38	Atlas Norge
7	Alfred Berg Gambak	39	Terra SMB
8	Alfred Berg Norge	40	Terra Norge
9	Avanse Norge (II)	41	Globus Norge II
10	ODIN Norge	42	Globus Aktiv
11	Storebrand Vekst	43	Nordea Kapital 2
12	Orkla Finans 30	44	KLP AksjeNorge
13	Danske Invest Norge II	45	Alfred Berg Humanfond
14	Danske Invest Norge I	46	ABIF Norge ++
15	Danske Invest Norge Vekst	47	Storebrand Norge I
16	Delphi Norge	48	Danske Invest Norske Aksjer Institusjon I
17	DnB NOR Norge Selektiv (III)	49	Nordea Kapital 3
18	PLUSS Markedsverdi (Fondsforvaltning)	50	Fondsfinans Aktiv II
19	Nordea Kapital	51	Storebrand Optima Norge A
20	Handelsbanken Norge	52	Holberg Norge
21	Carnegie Aksje Norge	53	DnB NOR SMB
22	Postbanken Norge	54	Pareto Aksje Norge
23	KLP Aksjeinvest	55	DnB NOR Norge Selektiv (IV)
24	Alfred Berg Aktiv	56	Romsdal Fellesbank Plussfond
25	Nordea Norge Verdi	57	Alfred Berg Norge Etisk
26	DnB NOR Norge (III)	58	Storebrand Norge A
27	DnB NOR Norge Selektiv (I)	59	Fondsfinans Spar
28	Storebrand Aksje Innland	60	DnB NOR Norge (IV)
29	NB Aksjefond	61	WarrenWicklund Norge
30	Pluss Aksje	62	Pareto Verdi
31	Nordea SMB	63	Landkreditt Norge
32	Alfred Berg Aktiv II	64	Danske Invest Norske Aksjer Institusjon II