Norwegian School of Economics
Bergen, Autumn, 2015

Performance, Persistence and Business Cycle Asymmetries in Norwegian Mutual Fund Returns

Do mutual funds perform when it matters the most?

Lars Petter Utseth and Kristoffer Bergquist Sandvik

Supervisor: Torfinn Harding

Master Thesis, MSc, Finance

NORWEGIAN SCHOOL OF ECONOMICS

Preface

This master thesis concludes our Master of Science (M.Sc.) in Economics and Business Administration at the Norwegian School of Economics (NHH). The purpose of this thesis is to investigate the performance, persistence, and business cycle asymmetries in Norwegian mutual funds during the period from 1983 to 2014. The thesis is written in the same article format as the many inspiring academic research papers we have encountered in completing our research. We would like to thank our supervisor Torfinn Harding for his guidance and feedback which is much appreciated. Additionally, we thank the Oslo Stock Exchange Information Services, the Norwegian Fund and Asset Management Association and Bernt A. Ødegaard for their help in providing us with the data necessary for the research conducted in this study.

Contents

Performance, Persistence and Business Cycle Asymmetries in Norwegian Mutual Fund Returns

Lars Petter Utseth and Kristoffer Bergquist Sandvik*

Monday 21st December, 2015

Abstract

This paper investigates the performance, persistence, and business cycle asymmetries in active Norwegian mutual funds using a dataset free of survivorship bias between 1983 and 2014. Fund performance is evaluated using both unconditional and conditional versions of Carhart's (1997) four-factor model. To determine the statistical significance of our result, we adopt a cross-sectional bootstrap methodology. We find that actively managed Norwegian mutual funds on aggregate produce returns that underperform the four-factor benchmark net of costs. When we examine individual funds, our bootstrap simulations provide no evidence of skilled fund managers in the right tail of the cross-sectional performance distribution, but several inferior performing fund managers in the left tail. Tests for persistence in performance provide no evidence of risk-adjusted performance persistence among previous winners, but short-term persistence among previous losers. Additionally, we perform a series of non-parametric two-period tests that allow us to infer whether some funds perform consistently better or worse compared to other funds in the sample. These tests reveal evidence of short-term performance persistence among both recent winners and losers. Moreover, we use two different methodologies to explicitly link fund performance to recessionary and non-recessionary states in the Norwegian business cycle. We find weak evidence of asymmetric performance of actively managed Norwegian mutual funds.

*Norwegian School of Economics (NHH), NO-5045 Bergen, Norway. We would like to thank our supervisor Torfinn Harding for his guidance and feedback which is much appreciated. Additionally, we gratefully acknlowedge the Oslo Stock Exchange Information Services, the Norwegian Fund and Asset Management Association and Bernt A. Ødegaard for their help in providing us with the data necessary for the completeness of our research. Main programmes in this study use MATLAB®. Codes are available from the authors upon request.

1 Introduction

There are two key issues on mutual fund performance that have been subject for academic debates over the years. The first issue concerns whether active mutual funds are able to add value by generating risk-adjusted returns net of costs. According to the efficient market hypothesis (EMH) brought forward by Fama (1970), any attempts to outperform the market is essentially a game of *chance* rather than *skill*, as current prices should reflect all available information. Still, active fund managers try to add value by attempting to "beat" the market by exploiting temporary mispricing. Grossman and Stiglitz's (1980) equilibrium model states the markets cannot be fully efficient all the time; thus, there is reason to believe that skilled fund managers are able to exploit periods in time where mispricing in the market occur. However, most previous studies document significant underperformance of actively managed mutual funds and argue that active fund managers do little besides collect fees (See e.g. Jensen, 1968; Grinblatt and Titman, 1989b; Elton et al., 1995; Ferson and Schadt, 1996; Carhart, 1997; Edelen, 1999). So why do investors buy actively managed mutual funds when empirical evidence suggest no superior managerial skill? The reasons remain a puzzle.

The second issue concerns whether it is possible to identify abnormal performance ex-ante, and for how long it persists. Persistence in performance is interesting from both an academic and practical point of view. From an academic point of view, evidence of persistence in performance would support a rejection of the semi-strong form of the EHM. The practical implication is that evidence of persistence could allow investors to earn risk-adjusted returns by exploiting past performance. There have been some discrepancies regarding the presence of persistence in mutual fund returns in previous literature. The majority of recent studies suggests that identifying funds with superior future performance is a difficult task, unless portfolio rebalancing is frequent and the performance is evaluated over short time horizons (See e.g. Hendricks et al., 1993; Brown and Goetzmann, 1995; Carhart, 1997; Bollen and Busse, 2005).

Motivated by the discussion of asymmetries in light of changing economic circumstances on mutual fund performance, Kosowski (2011) explicitly investigates the performance of US mutual funds in recession and non-recession periods. Kosowski (2011) argue that mutual fund investors may be willing to trade off some overall performance for superior performance in bad states of the economy when the marginal utility of wealth is high. Thus, he aims to provide an answer to the puzzle why investors keep investing in actively managed mutual funds despite the documented underperformance. Previous literature on asset pricing suggests that investors are more willing to pay premiums for assets whose returns are negatively correlated with consumption. When we have economic contractions in the business cycle (i.e. recessions), consumption tends to be particularly low (See e.g. Breeden, 1979; Rubinstein, 1976; Grossman and Shiller, 1981).

These implications give rise to a third issue on mutual fund performance, concerning whether active fund managers are able to add value for investors during recessionary states in the economy when consumption tends to be low, and the marginal utility of wealth is high. Up to this date, research on this issue has been quite sparse.

Given the practical importance to the average investor, Norwegian mutual funds have received little consideration, which makes Norwegian mutual funds truly a subject of interest. To our knowledge, there exist only a handful studies that have conducted comprehensive research on Norwegian mutual funds. The paper closest to ours is Sørensen (2009a) who examine all Norwegian equity mutual funds from 1982 to 2008.[1] Sørensen (2009a) find no significant evidence of superior performance at the aggregate level. His bootstrap simulations document virtually no evidence of superior performance at the individual fund level but provide evidence of inferior performing funds. Furthermore, Sørensen (2009a) find no evidence of performance persistence amongst either winner or loser funds. Although his results are in line with the theoretical concepts in finance theory, Sørensen (2009a) do not shed light on the third issue that is, whether mutual funds perform well in bad states of the economy when it matter the most for investors.

The purpose of this paper is threefold. The paper investigates the performance and persistence in actively managed Norwegian mutual fund returns. Additionally, the paper aims to answer a hitherto unanswered question regarding how active Norwegian fund managers perform in state of recessions when it matters the most to investors. Specifically, we aim to answer the following questions to ensure a thorough evaluation: 1) Do managers of active Norwegian funds generate risk-adjusted returns (i.e. alpha) net of costs, and if so, is the performance attributable to *skill* or *luck*? and 2) Does performance persist among extreme winners and extreme loser funds? and 3) Do actively managed Norwegian mutual funds deliver alpha in the state of recessions when performance matters the most to investors?

To address these issues, we use a dataset free of survivorship bias comprising 98 actively managed domestic equity mutual funds with monthly net returns from January 1983 to December 2014. We apply both unconditional and conditional versions of the four-factor model of Carhart (1997) to examine the existence of superior and inferior fund managers. To ensure proper statistical inference of our results, we adopt a bootstrap methodology similar to Kosowski et al. (2006), Cuthbertson et al. (2008) and Fama and French (2010). In addition to account for complex distributional properties, the bootstrap allows us to separate skill from luck in individual mutual fund performance. To investigate the existence of persistence in performance we adapt some of the most prominent statistical tests proposed in the literature. Specifically, we employ a recursive portfolio formation approach to examine the existence of risk-adjusted performance persistence. Additionally,

[1] Sørensen (2009a) wrote his paper as a part of his doctoral dissertation at The Norwegian School of Economics. He applies a bootstrap methodology similar to ours in his study.

we perform a series of non-parametric two-period tests to assess whether there are funds in our sample that consistently perform better compared to other funds in our sample.

To answer the question on how fund managers perform in recessions, we apply two different methodologies. First, we explicitly examine aggregate fund performance in different states of the business cycle using a binary classification of recessions and non-recession periods in the Norwegian economy. Specifically, we construct separate sub-samples of recession and non-recession periods in the Norwegian economy based on Aastveit et al. (2014) classification of recession dates in Norway. Second, following Kosowski (2011), we apply a novel conditional performance measurement methodology based on a Markov regime-switching model where we let the data determine the indicator of the recession and non-recessionary state. The main advantage of this model is that it allows for the involvement of multiple equations in a system that characterizes time-series behaviors in different states, and is permitting switching between these equations. This enables the model to capture more complex dynamic patterns. The switching mechanism between the equations (or states) is controlled by an unobservable state (latent-state) variable that is assumed to follow a first-order Markov chain.

We find that managers of active Norwegian mutual funds, on aggregate, do not have sufficient skill to generate risk-adjusted returns to cover the costs they are imposing on investors. When we study individual funds, our bootstrap simulations suggest no evidence of superior fund mangers. On the other hand, we find significantly negative risk-adjusted performance in the left tail of the performance distribution, which cannot be explained by random chance alone. Thus, our results indicate that there exist a large number of inferior performing fund managers in the universe of Norwegian mutual funds. Moreover, our recursive portfolio formation test reveals no evidence of dependable performance persistence when adjusting for risk among top performing funds. This result implies that investors cannot exploit past performance to earn positive risk-adjusted returns, a result that coincides with the semi-strong form of the EHM. On the contrary, we find that performance amongst loser funds strongly persists for short time horizons before it largely disappears, a result in line with the major consensus in previous literature (See e.g. Berk and Green, 2004; Bollen and Busse, 2005; Huij and Verbeek, 2007). Non-parametric two-period tests reveal short-term persistence amongst extreme winners and extreme losers relative to other funds in our sample.

Furthermore, our tests for asymmetric performance reveals that actively managed Norwegian mutual funds, on aggregate, show some indications to perform better in recession periods compared to non-recession periods. Specifically, from our Markov regime-switching model, we find that the difference in alpha between recession and non-recession periods is 1.89% per year. Differences in alpha estimates between recessions and non-recession periods are robust to the binary classification of recession dates based on Aastveit et al. (2014). Although statistical tests show evidence of asymmetries in the returns of actively managed Norwegian funds, we cannot reject the hypothesis that the alpha in recession and non-recession periods is independently statistically significantly different from zero.

Our paper makes two main contributions to the existing literature. First, it provides the most comprehensive performance analysis on Norwegian mutual funds up to date, covering almost the whole Norwegian mutual funds market's period of existence. Second, to our knowledge, it is the first paper to employ a regime-switching methodology to calculate risk-adjusted performance measures during recessionary and non-recessionary states in the Norwegian business cycle. Thus, our study provides answers to the question whether Norwegian funds are able to add value for investors when it matters the most.

The remainder of this paper is structured as follows. Section 2 provides a literature review that covers important academic papers on topics similar to ours. Section 3 presents our dataset and considers various data properties and selection criteria. Section 4 presents empirical results on the performance of actively managed Norwegian mutual funds, whereas Section 5 tests for persistence in the performance. Section 6 present empirical evidence on Norwegian mutual fund performance in recession and non-recession periods, and provides in-depth explanations of our implementation of a regime-switching framework used to capture asymmetries in mutual fund returns. Section 7 provides concluding remarks.

1.1 The Structure of Mutual Funds

A mutual fund is a collective investment vehicle that pools money from many investors to purchase securities. It has separate legal entity and is owned by its unitholders, whereas an investment company with concession manages the money in the fund. Fund management companies are paid a fee for this service, which is usually a percentage of funds under management, but it may also be linked to performance. The fund's Net Asset Value (NAV) is the price you have to pay to take part in this investor community. The investment manager then adds your money into the same pot as the other investors, and the sum of all these investments is called Assets Under Management (AUM). Based on investment goals set by the fund management, the fund constructs a portfolio consisting of stocks, bonds, short-term money-market instruments, other securities or assets, or some combination.

The Norwegian Fund and Asset Management Association (VFF[2]) classifies the division of funds into four main types, and a variety of sub-groups. The point of division is to make it easier to compare returns, risks and costs between comparable funds. The four main categories are stock or equity funds, bond or fixed income funds, money market funds and balanced funds. Equity funds invest most of the unitholder's capital in the stock market, which represent an ownership share (or equity) in the companies. Equity funds

[2]From here on referred to as VFF (Verdipapirfondenes Forening).

are divided into different sub-groups, each depending on what kind of investment universe the particular fund invests in. Stock or equity funds may invest primarily in Norwegian securities (Domestic or Norwegian equity funds), in Nordic securities (Nordic funds), in both Norwegian and foreign securities (Global funds), in foreign securities (International funds) or in assets in the European equity market (EU and EFTA countries; European funds), among others. These funds may also differ with respect to the share distribution method used. In addition, the funds may focus only on specific industries or sectors.

Bond funds invest in long-term fixed income securities. Since it is a fixed income fund, it has less volatility than equity funds and balanced funds. The major difference between bond and money market funds is that bond funds have greater price risk, which emerges as a result of changes in interest rates. Money market funds invest primarily in short-term fixed income securities, i.e. securities that have a maturity of less than a year. These funds are subject to strict requirements regarding liquidity and credit quality. This means that the funds are only permitted to invest in securities that have been considered to be of good quality by an analysis bureau. Balanced funds are funds that invest in a combination of both equities and fixed income securities. For example, a balanced fund may invest 50% of its total assets in equities and 50% in fixed income securities. This allocation can vary across the many different balanced funds, and over time. Because of the smaller proportion of stocks, the volatility is less, and it has lower fluctuations in the value.

Norwegian equity mutual funds are funds whose investment mandate are to normally have 80-100% exposure to domestic equities, and are regulated by "Verdipapirfondloven" (LOV)[3]. Norwegian equity mutual funds are open-end, meaning that the shares in the funds can be issued and redeemed at any given point of time. §6-6 in LOV states that a mutual fund cannot allocate more than 5% of the assets to a single security. However, up to 10% is allowed if the total sum of the allocations does not exceed 40% of the fund's total assets. Under certain regulations given by the Ministry of Finance, mutual funds are allowed to use derivatives (§6-1), but shorting stocks or engage in the futures and option markets is not permitted. The practical implication of this means that Norwegian mutual funds must have a spread of at least 16 single securities in their portfolios. Moreover, the four largest individual investments cannot exceed 40%, whereas the remaining 60% must consist of minimum twelve single securities (since maximum allocation is 5%). Thus, the potential ability to generate positive abnormal returns is rather limited due to the reduced hedging opportunities. Moreover, Norwegian equity mutual funds are open-end, meaning that the shares in the funds can be issued and redeemed at any given point of time.

[3]Norwegian mutual funds are also regulated by the European Union's Undertakings for Collective Investment in Transferable Securities Directive (UCITS). The directive (adopted in 1985) does not directly regulate mutual funds in the European Union, but is implemented in "Verdipapirfondloven" (LOV).

More importantly, Norwegian equity mutual funds can be passively or actively managed. In active management, the fund manager pursues his own strategy and invests in companies that he believes will provide the best returns in order to *beat* a given benchmark index. The different strategies involve e.g. future predictions about the market and other fundamental analyses in the quest to beat the index. The costs of investing in an active fund are therefore quite sizable because of the fees imposed for this service. In passive management, the capital is invested to *track* a given benchmark index. Thus, the cost of investing in passively managed funds is relatively low compared to that of actively managed funds.

1.1.1 The Norwegian Mutual Fund Industry

Worldwide, there has been a remarkable increase in the mutual fund market. The Norwegian mutual fund industry is still in an early phase in comparison to other more established markets, but has grown quite rapidly throughout the years of existence. From 1982 to 2014 the total market value of Norwegian equity mutual funds increased from NOK 290 million to NOK 85 billion.[4] Table I below reports some interesting features about Norwegian equity mutual funds for the period 1994-2014.[5]

As can be observed from Column 1 in the table, the average number of customers per fund each year is steadily decreasing from 1998 and throughout 2014. This coincides with the last two columns of the table, which shows the development in Norwegian equity funds as a percentage of the total equity fund market, and as percentage of the total fund market. In 1994, 92% of the total equity fund market consisted of Norwegian equity funds, whereas 37% was attributable to the total mutual fund market. At the end of 2014, the same numbers decreased to 20.9% and 10.2%, respectively. It may seem like investors have gradually turned their investments towards global equity funds, and sought the diversification benefits that funds with wider investment mandates provides. Figure 1 at the end of this section illustrates the development in asset allocations of Norwegian mutual funds from 1994 to 2014 and puts this observation into perspective. From the figure it becomes evident that the percentage of assets invested in Norwegian equity mutual funds has decreased considerably. This decline has mainly been at the expense of international equity funds, which possessed almost 40% of the total equity fund market in 2014 and only 3% in 1994.

[4]Prior to 1982, there was only a single fund in existence at the Oslo Stock Exchange. Gjerde and Sættem (1991) report a total market value of NOK 290 million at the end of 1982. VFF reports a total market value of NOK 85 billion at the end of 2014.

[5]VFF did not report any data prior to 1994. We would like to thank Ida Aamodth-Hansen at VFF for generously providing us with characteristics data on Norwegian mutual funds for the period between 1994 and 2014.

Table I
Characteristics of the Norwegian Mutual Fund Market

This table reports characteristics for Norwegian equity mutual funds registered in Norway between 1994 and 2014. The data is collected from The Norwegian Fund and Asset Management Association (VFF). Column 1 shows the average number of customers per fund each year, whereas Column 2 refers to average assets under management. Column 3 reports average net inflows. The last two columns refer to assets under management in percent of the total Norwegian equity fund market, and in percent of the total Norwegian fund market. AUMs, inflows and outflows are reported in million NOK.

Year	Average customers	Average AUM	Average net inflow	% of total equity fund market	% of total fund market
2014	4,138	1090	-25	20.9	10.2
2013	4,634	1087	-13	22.4	12.3
2012	5,745	945	-10	24.5	12.2
2011	6,017	833	-18	24.6	12.5
2010	6,281	1063	60	26.6	15.6
2009	6,874	822	-4	24.8	13.9
2008	6,571	359	-1	19.7	8.7
2007	6,726	746	-44	23.1	12.9
2006	6,175	635	16	24.5	14.8
2005	6,854	504	-61	26.2	14.0
2004	8,342	421	-52	31.8	16.8
2003	9,281	351	-1	35.9	17.3
2002	9,024	215	-11	37.1	15.8
2001	11,302	374	-11	37.0	20.7
2000	11,537	459	-23	38.3	24.6
1999	14,255	573	7	46.1	30.8
1998	15,878	403	4	67.3	38.4
1997	14,858	604	140	80.1	47.8
1996	13,354	422	99	86.1	41.4
1995	9,689	227	7	91.9	34.1
1994	10,987	235	8	92.0	37.0

From Column 2 in Table I, it can be seen that the average AUM grew significantly from NOK 215 million in 2002 to NOK 746 million by the end of 2007. In 2008 the average AUM decreased drastically to NOK 359 million, largely attributable to a sharp drop in equity prices as a result of the Global Financial Crisis (GFC). The average AUM quickly recovered to NOK 822 million at the end of 2009. Since 2009, the average AUM has increased quite steadily with a minor drop in 2011. With NOK 1090 million in average AUM by the end of year 2014, the compounded annual growth rate during the twenty-year period has been 8%.

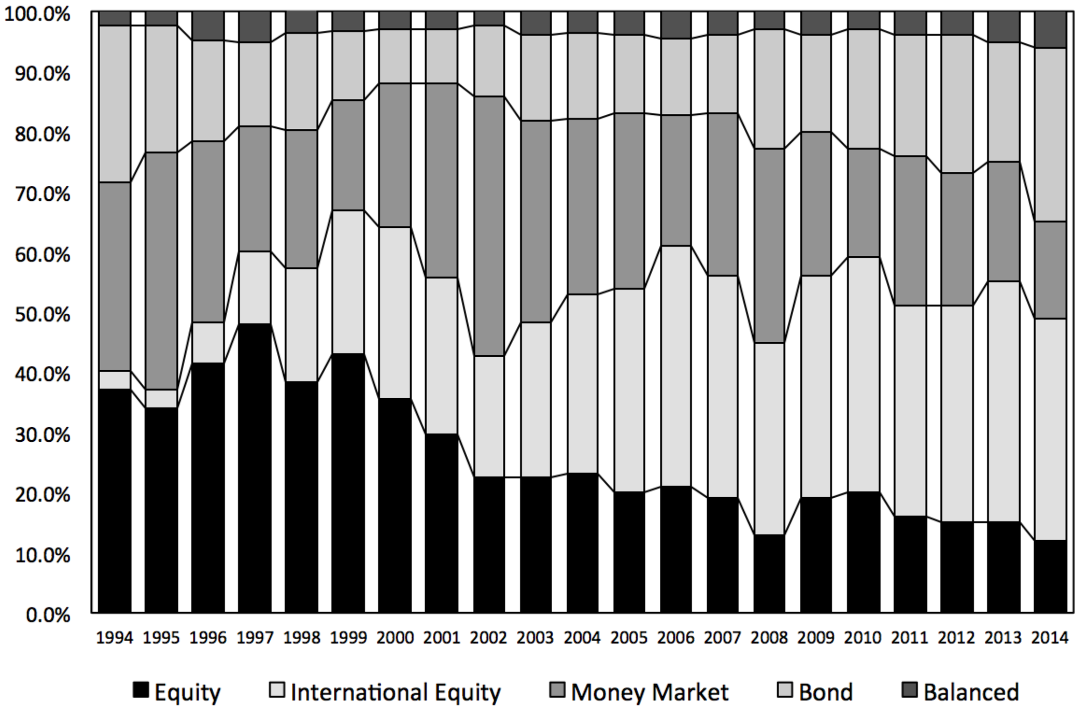

Figure 1. Asset allocation of Norwegian mutual funds through time, 1994-2014. The figure provides average asset allocations between five categories of Norwegian mutual funds, namely equity, international equity, money market, and balanced funds. The x-axis presents the respective years, whereas the y-axis presents the market share for each of the five categories. The data from 1994 to 2014 is obtained from The Norwegian Fund and Asset Management Association (VFF).

2 Literature Review

In this section, we review previous studies on the performance, persistence, and time-variability in mutual fund returns. First and foremost, we aim to establish expectations to our findings by assessing the most important previous literature on subjects similar to ours. The following sections are structured in the same manner as the remainder of this thesis. That is, we first survey the most important literature on mutual fund performance. Second, we survey the relevant literature on performance persistence in mutual funds. Finally, we examine the literature on time-variability in mutual fund performance.

2.1 Mutual Fund Performance

Mutual fund performance is a widely researched topic within finance. First out was Jensen (1968), who developed a single-factor model based on the earliest version of the capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965) and Mossin (1966). In the single-factor model of Jensen (1968), the intercept (alpha) represents the abnormal performance of fund managers. The benchmark used to compute this measure is assumed to be mean-variance efficient from the perspective of an uninformed investor. The perception is that an actively managed fund is expected to generate a positive alpha, whereas a passive fund is expected to generate an alpha of zero. By using data of 115 US mutual funds in the period 1945 - 1964, Jensen (1968) was the first to find solid evidence on the performance of actively managed mutual funds. He concluded that US mutual funds were on average not able to outperform a passive market proxy when accounting for management fees. In an updated study, Ippolito (1989) find results that contradict with Jensen (1968). Using a sample of US funds spanning over 20 years, Ippolito (1989) concludes that U.S mutual funds are able to outperform the passive benchmark net of expenses. More specifically, Ippolito (1989) find that 12 funds have significantly positive alphas net of expenses, and that actively managed funds on average outperform the S&P500 by 88 basis points. In light of Jensen's (1968) results, several research papers debate the use of appropriate benchmarks when evaluating mutual fund performance (See e.g. Roll, 1978; Lehmann and Modest, 1987; Grinblatt and Titman, 1989a; Connor and Korajczyk, 1991; Sharpe, 1992; Elton et al., 1993; Pástor and Stambaugh, 2002b).

Roll (1978) criticizes the use of CAPM market proxies as performance benchmarks since the model assumes that all investors have common beliefs and information, hence that any measured abnormal performance can only occur when the market is inefficient. Lehmann and Modest (1987) provide results on whether the choice of benchmarks affects Jensen's alpha. In particular, their empirical research shows how sensitive the choice of arbitrage pricing theory benchmarks concerns Jensen's (1968) measure. In particular, Elton et al. (1993) argue that Ippolito's (1989) positive alpha emerge as a result of inappropriate benchmarks.

They find that the funds included in Ippolito's sample invest heavily in small stocks that are not included in the S&P500 benchmark used in the study. These stocks outperform the S&P500 considerably during the sample period. When correcting for this, Elton et al. (1995) concludes that the positive alpha found by Ippolito (1989) becomes negative. Malkiel (1995) examine all diversified mutual funds between the period 1971 and 1999 each year and find that mutual funds underperform both net and gross of expenses. The conclusion of Malkiel (1995), however, is also sensitive to the choice of benchmark. Overall, this led to the rise of extended multifactor models that controls for various anomalies in the stock market. For instance, Fama and French (1993, 1996) establish a three-factor model by extending the single-factor model of Jensen (1968) adding size (SMB) and value (HML) factors in addition to the single market factor. Carhart (1997) extends the three-factor model of Fama and French (1993) further by including the one-year return momentum factor of Jegadeesh and Titman (1993).

In his study on mutual funds from 1985 to 1994, Gruber (1996) was among the first to implement a multi-index model for mutual fund performance evaluation. His multi-index model consists of four variables, namely excess market return, the difference in return between a small cap and large cap portfolio, the difference in return between high growth and a value portfolio, and excess return on a bond index. The model suggests that mutual funds underperform an appropriately weighted average of the indices by about 65 basis points per year. More interestingly, Gruber (1996) argue that mutual fund managers are able to generate abnormal performance (i.e. positive risk-adjusted returns) gross of expenses by looking at the average expense ratio. He finds an average expense ratio of 1.33%, suggesting that mutual fund managers have superior stock-picking abilities. These selection skills come at a great cost, however, which is too high for the average investor.

Daniel et al. (1997) conducts a comprehensive evaluation of 2500 US equity mutual funds from the period 1975-1994, and investigate whether fund managers have sufficient stock-picking abilities to earn back some of the costs they generate. Specifically, they construct characteristic measures based on the market capitalization, book-to-market, and prior one-year return portfolio benchmarks, and decompose the funds' excess returns into *Characteristic Selectivity* and *Characteristic Timing* measures. Daniel et al. (1997) show that mutual funds, in contrary to most previous research, exhibit some stock-picking skills, in particular, aggressive-growth funds. The average abnormal performance of 0.8% per year in the paper, however, is close to the average management fee, which indicates net neutral performance. Furthermore, they find no evidence of characteristic timing ability.

Wermers (2000) also examine the performance of US equity mutual funds between 1975 and 1994, but decompose the performance based on net returns and stock holdings. He finds a difference of 2.3% between the return on stock holdings and net returns for the average mutual fund. Specifically, the stock-holdings approach indicates that mutual funds outperform the market by 1.3%, almost enough to cover their costs.[6] Moskowitz (2000) discuss that the abnormal returns based on the characteristic selectivity measure in Wermers's (2000) paper might be due to the use of an inappropriate benchmark, and argue that portfolio-based benchmarks only consist of small, illiquid and risky firms. Thus, overstating the stock-picking abilities of fund managers.

Edelen (1999) examine 166 US mutual funds, and documents a significantly negative average alpha of -1.63% per year based on a single-factor market model using the CRSP value-weighted index. The negative alpha is close to the expenses of 1.72%, which indicates *prima facie* that fund managers do little besides collect fees. When controlling for the effects of flow-related liquidity trading costs, he finds no evidence of superior performance (i.e. alpha) or bad market timing.[7] Edelen (1999) argue that underperformance is not due to fund managers's inability to generate alpha, but results from the costs of providing investors with liquidity service.

The evidence regarding mutual fund performance reviewed so far is based on studies conducted by the use of unconditional performance measures. Ferson and Schadt (1996) argue that unconditional performance measures are inappropriate as they fail to account for the fact that fund managers change their portfolios over time, based on observable information variables. Ferson and Schadt (1996) encourage the use of conditional performance measures for two reasons; the first argument being that traditional measures are unable to handle the dynamic behavior of returns, the second being the possibility that trading behavior of managers results in more complex and interesting dynamics than those of the underlying assets traded. Ferson and Schadt (1996) modify Jensen's alpha and the market timing measures of Treynor and Mazuy (1966) and Merton (1981) to incorporate conditioning information, and by doing so allows for time-varying risk exposures (i.e. betas). By examining monthly data of 67 mutual funds over the period 1968-1990, Ferson and Schadt (1996) find that the conditioning information is both statistically and economically significant. At the aggregate level, their results show that the funds' unconditional alphas are negative more often than positive, which is similar to the evidence that Jensen (1968) and Elton et al. (1993) interpret as an indicator of *poor* average performance. Using conditional models that allow for time-varying risk exposures, they find that the distribution of mutual fund alphas has a mean value of zero, and that the distribution of mutual fund alphas is consistent with the neutral performance for the group.

[6] In the difference of 2.3%, 0.7% is attributable to lower average returns of non-stock holdings, 1.6% to expense rations and the transaction costs of the funds.

[7] Edelen (1999) reports a negative alpha of -0.20% when controlling for the effects of flow-related liquidity trading costs, which is statistically insignificant.

Otten and Bams (2004) uses a survivorship-bias free dataset of US mutual funds provided by the Center for Research in Security Prices (CRSP), and shows in a comprehensive comparison of factor models that conditional models are a significantly better choice than their unconditional counterparts.[8] The four-factor model of Carhart (1997) stands out being the statistically strongest of the models tested. Moreover, they show that the aggregate US mutual fund industry delivers an insignificant alpha of -0.42% and 0.04% net of expenses, measured against the unconditional and conditional Carhart (1997) four-factor model, respectively.

Undoubtedly, the research discussed so far provides little evidence that the aggregate mutual fund industry has created value for its investors. Nonetheless, this does not imply that every fund underperforms their given benchmarks. In general, some fund managers will underperform and other funds will outperform from time to time in accordance to Grossman and Stiglitz's (1980) equilibrium model, i.e. that the markets cannot be fully efficient all the time and that temporary mispricing in the market must occur. This implication raises the question whether differences in fund performance is attributable to managerial *skill* or simply due to *luck*.

Kosowski et al. (2006) and Fama and French (2010) addresses the question whether superior performance in individual funds are attributed to skill or are simply due to luck by employing an innovative bootstrap approach that account for non-normality in fund returns.[9] Kosowski et al. (2006) examine US mutual funds between 1975 and 2002, and finds that approximately 10% of the funds have significant stock picking ability to cover their costs. Fama and French (2010) use an alternative bootstrapping technique to evaluate the performance of US mutual funds, and in contrast to Kosowski et al. (2006), they find no evidence of performance among the top funds.[10] Fama and French (2010) agrees with Kosowski et al. (2006) regarding the worst funds, which they both argue is due to poor skill, and not due to bad luck.

Barras et al. (2010) also argue that standard tests designed to identify mutual funds with non-zero alphas are problematic. That is, the standard tests does not adequately account for the presence of lucky funds. By applying new measures built on the False Discovery Rate (FDR), Barras et al. (2010) quantifies the impact of luck and find that about one-fifth of the funds in their sample truly yields negative alphas.[11] More specifically, this technique separates funds into unskilled, zero-alpha and skilled funds by controlling for false discoveries.

[8]Otten and Bams (2004) use a Likelihood ratio test to determine whether the differences in explanatory power between the models are statistically significant.

[9]In-depth explanations of the bootstrapping approach will be discussed in Section 4.2.

[10]Fama and French's (2010) adjusted approach implicitly assume no autocorrelation in individual mutual fund returns.

[11]In their study, they also find that a small proportion of funds yield positive performance prior to 1996, concentrated in the extreme right tail of the alpha distribution.

2.1.1 Non-US Studies

Blake and Timmermann (1998) conduct a comprehensive study on 2300 UK mutual funds during the period 1972-1995, and find that the average UK equity fund underperforms by around 1.8% on a risk-adjusted basis. However, they find short-lived outperformance of 0.8% during the first year of the funds' existence. Cuthbertson et al. (2008) uses a bootstrap methodology similar to Kosowski et al. (2006) on a dataset comprising UK equity mutual funds from 1975 to 2002, and find evidence of stock picking abilities among a relatively small number of the top performing funds.[12]

Otten and Bams (2002) conduct a comprehensive study of 506 mutual funds in five different European countries, and compare results from both unconditional and conditional factor models.[13] When the unconditional four-factor model of Carhart (1997) is used on net returns of the European countries, only UK mutual funds seem to exhibit a significantly positive alpha. The rest of the countries have positive alphas (although not significant) except for Germany, which exhibit a negative alpha of -1.20%. The conditional model also indicates a significantly positive alpha on Dutch mutual funds, while the results on the other countries remain unaltered. In contrary to most US studies, Otten and Bams's (2002) study show that before costs, all of the countries except Germany exhibit significantly positive alphas. In a more recent study, Ferreira et al. (2012) examine equity mutual funds in 27 countries, in which the five countries in Otten and Bams's (2002) study are included, and show that mutual funds underperform the market overall. Their findings suggest that the adverse scale effects in the US are related to liquidity constraints faced by funds that, by virtue of their style, have to invest in small and domestic stocks. In addition, they find that funds located in countries with liquid stock markets and strong legal institutions display higher performance.

Previous research regarding the performance of mutual funds in Scandinavian countries, however, is quite sparse. The most recent include Dahlquist et al. (2000), Korkeamaki and Smythe (2004), Sørensen (2009a), Christensen (2005, 2013), and Gallefoss et al. (2015). Dahlquist et al. (2000) investigate performance and characteristics of Swedish mutual funds, and documents neutral performance for special equity funds, bond, and money market funds, using both unconditional and unconditional performance measures. In contrary, the regular equity funds in their sample seem to have generated abnormal returns, thus indicating outperformance. Korkeamaki and Smythe (2004) examine the Finnish mutual fund market from 1993 to 2000, and show that Finnish mutual funds, in general, exhibit neutral performance. The equity funds in their sample seem to have provided negative performance. The unpublished work of Christensen (2005) documents no significant performance among 47 Danish mutual funds during the period 1996 to 2003. The funds are split between equity funds and fixed income funds, and are amongst the

[12]Specifically, Cuthbertson et al. (2008) provides evidence of skill in 3-8% of UK mutual funds.

[13]The five European countries include France, Italy, Germany, UK and the Netherlands.

funds in Europe with the lowest expenses. Still, they have delivered neutral to slightly negative performance. Christensen (2013) also investigates the performance of 71 Danish mutual funds between 2000 and 2010 individually and on aggregate, separating them into different categories by using equally weighted portfolios. His findings show that 80% of individual Danish mutual funds exhibit negative alpha estimates, of which 42% prove to be significant. Only 7% of the funds in the sample yielded significantly positive alphas.

Sørensen (2009a) uses a dataset free of survivorship bias comprising all available Norwegian equity mutual funds between 1982 and 2008. By adjusting for risk in the Norwegian market, he finds no significant evidence of superior performance at the aggregate level. Moreover, his bootstrap method shows weak signs of skill in the right tail of the cross-sectional distribution of individual fund alphas, and several inferior performing funds in the left tail that are not attributable to bad luck. Gallefoss et al. (2015) examine actively managed Norwegian mutual funds during the period 2000-2010 using daily data, and restrict their sample to funds with minimum 36 months of observations. Gallefoss et al. (2015) find that actively managed mutual funds on aggregate underperform the benchmark by approximately their fund fees. Furthermore, they find that funds in the right tail (i.e. top performing funds) of the performance distribution inhabit genuine stock-picking skills, and that the performance of the worst funds is not a result of bad luck. Ferreira et al. (2012) include Denmark, Norway, and Sweden in their study, and also provide evidence of underperformance for these countries. Thus, confirming the findings of Dahlquist et al. (2000), Christensen (2005, 2013), Sørensen (2009a) and Gallefoss et al. (2015).

Most of the previously reviewed literature documents underperformance of mutual fund managers, which is not improved upon in studies regarding the ability of mutual funds to time the market. Most of the studies regarding market-timing is conducted in the US market (See e.g. Treynor and Mazuy, 1966; Kon and Jen, 1979; Kon, 1983; Henriksson, 1984; Chang and Lewellen, 1984; Connor and Korajczyk, 1991; Ferson and Schadt, 1996). The majority of these studies document perverse negative market-timing abilities among mutual fund managers. Treynor and Mazuy (1966) invented a market-timing model by adding a quadratic square function of the market factor in Jensen's (1968) model. In their study, he investigates 57 mutual funds and provides evidence on timing ability of only one of these funds. In a later study, Veit and Cheney (1982) examine whether mutual funds, in general, appear to change their characteristic lines in bull and bear markets. They conclude that of the funds, which in fact, changed their characteristic lines, only three succeeded in timing the market. Connor and Korajczyk (1991) and Hendricks et al. (1993) extends the Henriksson and Merton (1981) model, and also conclude on the absence of market-timing abilities in US mutual funds. Goetzmann et al. (2000) adjusts Henriksson and Merton's (1981) method further, and supplements the literature by providing additional evidence of negative timing abilities in US mutual funds.

2.2 Performance Persistence

The previously reviewed literature provides little evidence of superior performance among mutual funds at the aggregate level. However, it is still possible that some fund managers are able to outperform their benchmark from time to time, and that this performance might persist over subsequent periods. Thus, persistence in mutual fund returns is of principal importance from both an academic and practical point of view. From an academic point of view, persistence is important as the efficient market hypothesis is tested. If past performance cannot be an indicator of future performance, the practical importance to investors is that they might be better off by engaging in passive management. The literature on performance persistence aims to test this hypothesis.

Sharpe (1966) initiated the research on persistence in mutual fund performance by studying rank correlations on the basis of his performance measure; the *Sharpe Ratio*. By ranking funds according to their Sharpe ratio over two consecutive periods, Sharpe (1966) find significant positive correlations, which indicates that past performance might be an indicator of future performance. Grinblatt and Titman (1992) investigate 279 US funds during the period 1975-1984 using multiple portfolio benchmarks with evaluation periods of five years, and provide evidence of persistence in consecutive five-year periods. Following up their previous work, Grinblatt and Titman (1993) examine CRSP listed quarterly holdings of mutual fund portfolios in the period 1974-1984, and find evidence of persistence among the funds in their sample. Specifically, they find that top performing funds in the first half of the sample also performs well in the last part of the sample, thus suggesting that superior performance to a certain extent is predictable ex-ante. The strongest evidence of persistence was centered amongst funds in the category aggressive growth.

Goetzmann and Ibbotson (1994) show that past returns and past risk-adjusted returns predict future performance for the period 1976 to 1988. Brown and Goetzmann (1995) follow up this study by examining the same 1976 to 1988 period. Their results suggest that relative abnormal performance of US mutual funds seem to persist, but mostly due to funds that repeatedly lag the passive benchmark (the S&P500). Specifically, they suggest two possible reasons for performance persistence, the first being that persistence seems to be correlated across managers due to trading strategies that are not captured by style categories or risk-adjustment techniques. The second suggestion is that the market is unable to fully discipline the worst performing funds, and that their presence in the sample contributes to a pattern of relative persistence.

The earliest studies that provide evidence of performance persistence, however, might be prone to survivorship bias as Malkiel (1995) suggest, thus the evidence is less valid. Malkiel (1995) utilizes a unique data set comprising all existing US equity mutual funds from 1971 to 1991 to account for the influence of survivorship bias. He finds evidence on performance persistence for seven out of nine periods in the 1970's.[14] Hendricks et al. (1993) investigate persistence in US mutual funds between 1974 and 1988 by regressing current performance on quarterly lags. He finds evidence of persistence for up to four quarters, denoting the effect as "Hot hands". Using a dataset free of survivorship bias, Carhart (1997) argues that the "Hot hands" phenomenon found by Hendricks et al. (1993) is mostly driven by the one-year momentum effect of Jegadeesh and Titman (1993).[15] His results suggest that fund managers possess little stock selection skill since superior funds generate their returns simply by holding stocks that recently have had abnormal returns. When controlling for the momentum effect, Carhart (1997) finds no evidence of persistence among the top performing funds. However, Kosowski et al. (2006) applies a bootstrap approach to assess significance on the same data sample as Carhart (1997) and find that performance seems to persist among the top performing funds. In his study, Wermers (1997) support the findings of Carhart (1997), and argue that active use of momentum strategies is the reason for short-term persistence. He concludes that top performing funds during one year also are the top performers the following year, and that this pattern corresponds exactly to the pattern found in the momentum effect in stock returns.

Moreover, several studies provide evidence on significantly positive alphas (gross of costs) when following a hypothetical momentum strategy that implies buying prior winners and selling prior loser funds (See e.g. Hendricks et al., 1993; Carhart, 1997; Kosowski et al., 2006). Additionally, Busse et al. (2010) provides evidence on weak performance persistence for institutional funds. Other studies have, more specifically, found stronger evidence of persistence amongst early-phase funds, small-cap growth funds and funds with no load (See e.g. Gruber, 1996; Blake and Timmermann, 1998; Bollen and Busse, 2005; Huij and Verbeek, 2007).

Most of the previously reviewed studies are focused on long-run performance persistence.[16] More recent studies, however, provide evidence that performance seems to persist in the short run (See e.g. Berk and Green, 2004; Bollen and Busse, 2005; Huij and Verbeek, 2007). Berk and Green (2004) find that abnormal performance persists over shorter evaluation periods. Over longer time periods, they find no persistence among the top performing funds in their sample. Bollen and Busse (2005) use daily frequency data on mutual fund returns to allow for short evaluation periods. Specifically, they establish quarterly rankings based on the funds' abnormal returns, and then measure the performance over subsequent quarters. Bollen and Busse (2005) show that performance persistence exists

[14] In the 1980's, however, Malkiel (1995) find only three periods with statistically significant evidence of return reversals.

[15] By including a momentum factor in his four-factor model, Carhart (1997) finds that persistence largely disappears. Among the lowest performers, however, persistence arises from persistently high expenses.

[16] These studies differ in respect to the methodologies used, but the non-accessibility of short frequency data is common, thus making it hard to investigate short-run performance persistence.

among top funds when using short evaluation periods, but that it seems to disappear when longer evaluation periods are used. Huij and Verbeek (2007) investigate short-run performance persistence between 1984 and 2003 by using monthly frequency data. They employ a Bayes approach to cope with short ranking periods, and find that performance is persistent even beyond load fees when the funds are sorted into decile portfolios based on their 12-month past performance.[17] Overall, empirical evidence shows that post-ranking returns largely disappear when longer evaluation periods are used. Hence, superior performance persistence is considered to be, if any, a short-lived phenomenon.

2.2.1 Non-US Studies

Blake and Timmermann (1998) investigate performance persistence by examining 2300 UK mutual funds during 1972-1995. By using a similar recursive portfolio approach as of Hendricks et al. (1993), they find evidence of persistence in portfolios composed by prior winners. Prior losers, on the other hand, produced significantly negative alphas. These findings are in line with Otten and Bams (2002), who finds a spread between the two portfolios comprising prior winners and prior losers of 6.08% per year for UK funds.

In a more recent study, Vidal-Garcia (2013) investigates performance persistence of actively managed mutual fund returns for six European countries over the 1988-2010 period. He applies several conventional tests for persistence and applies the same bootstrap approach following Kosowski et al. (2006) to test for significance. His results indicate significant evidence of performance persistence among European mutual funds, and that these results are robust under the non-normality of the funds' return distribution. In addition, he finds that the performance spread between prior winners and prior losers is largest among UK mutual funds, thus confirming the findings of Otten and Bams (2002).

The studies of Dahlquist et al. (2000), Christensen (2005) and Sørensen (2009a) do not provide general evidence of persistence in the Scandinavian sector. More specifically, following the same approach as Carhart (1997), Sørensen (2009a) analyze persistence among Norwegian mutual funds during the period 1985-2008. By sorting funds into quintile portfolios based on lagged one-year returns, he find no evidence of persistence amongst top and bottom performing funds. Gallefoss et al. (2015) use daily data, and are thus able to allow for shorter ranking periods. They find evidence on performance persistence in the performance spread, and confirm the findings of Vidal-Garcia (2013). In addition, their results indicate that abnormally bad performance of the worst performing funds strongly persists, which is in line with the findings of Bollen and Busse (2005).

[17] Specifically, the top decile funds in Huij and Verbeek's (2007) sample that earns significantly abnormal returns are mainly early-phase small cap/growth funds.

2.3 Time-Variability

The evidence regarding the performance and persistence in mutual fund returns reviewed so far is based on studies conducted by the use of unconditional models that assumes constant betas, and conditional models that only account for potential time-variation in betas. Previous studies also show that time-variation in alphas can lead to biased OLS alpha estimates (See e.g. Grinblatt and Titman, 1989b; Glosten and Jagannathan, 1994; Christopherson et al., 1998).

Krueger and Callaway (1995) investigate persistence in the performance of 41 aggressive growth, 229 growth funds, and 35 equity income mutual funds by examining two consecutive three-year periods. By using the Sharpe (1966), Treynor (1965), and Jensen (1968) performance measures, they show that fund performance indeed varies by the period. Specifically, aggressive growth funds prove to be the riskiest of the categories.

Christopherson et al. (1998) propose an extension of Ferson and Schadt's (1996) conditional model that allows for both betas and alphas to be conditioned on public information. Specifically, they document that time-varying alpha measures are superior in predicting future performance as compared to unconditional alphas or raw returns, even though none of which allows for ex-ante detection of real investment skill. Avramov and Wermers (2006) exploit this further by incorporating public information in a Bayesian setting, and argue that actively managed funds add more value than documented in previous literature. Specifically, they analyze the performance of different portfolio strategies by incorporating predictability in managerial skill, fund's risk-loadings, and benchmark returns. Overall, Avramov and Wermers (2006) provide evidence on cross-sectional fund differences by showing that funds are superior within industry stock pickers.

However, even though the abovementioned studies confirm that it is important to account for time-variation in mutual fund alphas and betas, they do not explicitly examine the risk-adjusted performance of mutual funds in economic downturns and upturns. Most of the previous research assumes the functional relationship between excess returns and predetermined factors to be constant rather than vary through different states in the economy. Moskowitz (2000) argue that mutual funds may, in fact, add value by performing well during economic downturns. By computing performance measures over two subsamples by using the NBER classifications of recessionary and non-recessionary periods, he shows that active mutual funds generate an additional 6% per year during recessions. His results shows that funds earn an additional 1% per year during recessions also when adjusting fund returns for size, book-to-market equity, and momentum premium. This indicates that active managers deliver returns when investors need them the most, and that examining their unconditional performance may understate their abilities. However, Moskowitz (2000) is unconvinced of these results given the paucity of recessionary periods over the 20-year sample and suggest that his findings might be pure chance.

Kosowski et al. (2006), Ang and Chen (2007) and Glode et al. (2011) document that both the size and value factors in their conditional models turn out to be insignificant, and argue that benchmarks with multiple factors might be a better way to account for the time-variability in the market factor. As a suggestion to this problem, Mamaysky et al. (2008) make use of a Kalman filter to track alpha and beta dynamics, and allows the coefficients to depend on an unobservable variable that itself follows an AR(1) process. This improves upon the alpha and beta estimation from conventional OLS models that solely relies on macroeconomic variables to explain the variation in coefficients over time. In contrary to Ferson and Schadt (1996) and Christopherson et al. (1998), the model allows for ex-ante detection of real investment skill. However, despite the improved in- and out-of-sample properties of the Kalman filter, the alphas in the model are measured unconditionally.

A model that overcomes some of the problems inherent with the abovementioned approaches was introduced as early as in 1989. Hamilton (1989) developed a Markov regime-switching model for dealing with asymmetric business cycles and structural breaks in time-series data. The major advantage of this model is that it allows for a continuous state probability, where a first-order Markov process governs the transition between the states. By using maximum likelihood estimation, the transitions can be obtained recursively along with other parameters in the model.

The more recent study by Kosowski (2011) shows that traditional unconditional performance measures in fact understate the value added by active mutual fund managers during recessionary states in the economy, when the marginal utility of wealth is high. He conducts a comprehensive analysis on business cycle asymmetries in mutual fund performance by investigating US domestic equity funds in recessions and expansions from 1962 to 2005. Kosowski (2011) identify recessionary and non-recessionary periods using two methods; the NBER classification of business cycle dates and a two-state Markov regime-switching model. His research shows that the negative mutual fund underperformance documented in literature is attributable to expansion periods when funds have negative risk-adjusted alpha, and not during recessions when the risk-adjusted alpha is positive. However, by using the NBER recession dates as state indicators, a limitation is that it only becomes available ex post.

Kacperczyk et al. (2010) construct a model on fund manager's attention allocation and portfolio choice over the business cycle. They show that the portfolio dispersion is higher when skilled fund managers engage in market timing, and that these results are true both among skilled managers and between skilled and unskilled managers. Interestingly, their research indicates that recessionary states in the economy are the times when skilled managers outperform the most, which is consistent with Kosowski (2011). In addition, they find that mutual fund portfolios exhibit more cross-sectional dispersion and generate higher abnormal performance in recession periods.

Qiu et al. (2011) investigate business cycles and mutual fund timing performance of US mutual funds by examining daily data from the period 1998 to 2009. They incorporate a regime-switching framework into Treynor and Mazuy's (1966) model, and allow for switching between two regimes (e.g. up and down markets) that is governed by a first-order Markov process with time-varying transition probabilities. By stratifying the funds into nine categories based on their stated investment objective, they show that the regime-switching model captures the asymmetric timing performance, whereas the single-regime model does not.[18] Further, they find that fund managers have significant perverse timing abilities in expansions periods, but not in recessionary states in the economy.

3 Data

This section presents the data used in the empirical analysis of this study. Details regarding the data and its providers will be reviewed throughout the following sections.

3.1 Norwegian Mutual Funds

Our mutual fund data set comprises 98 Norwegian actively managed open-end domestic equity funds. These funds' investment mandate is to invest primarily in Norwegian equities (i.e. minimum 80% must be invested in domestic equities). We restrict our sample only to consider Norwegian equity mutual funds to be consistent regarding the choice of benchmarks used in this study. By excluding funds with different risk exposures, we only require one specific benchmark spanning the investment opportunity set. This will allow for more accurate benchmark returns when computing risk-adjusted performance. The data set consists of all available active Norwegian equity mutual funds at the Oslo Stock Exchange between 1983 and 2014, both surviving and non-surviving.[19] The choice of period is simple as only a few funds existed before 1983, and it covers almost the entire lifetime of the Norwegian mutual fund market. We omit funds that are passively managed as they only pursue neutral investment strategies.

To compute the funds' returns, we have obtained historical data on both daily and monthly Net Asset Value (NAV) for each fund from The Oslo Stock Exchange Information Services.[20] The NAV is computed by summing the current value of all stocks contained in the funds' portfolio, deducting expenses such as management fees and other ongoing

[18] The categories include; all funds, retail funds, retail aggressive, retail growth, retail income and growth, institutional, institutional aggressive, institutional growth, and institutional growth and income funds.

[19] Pareto Investment Fund B and Pareto Investment Fund C have been omitted from the data sample; these funds are practically the same fund as Pareto Investment Fund A. The differences are fee structures with respect to the amount invested in the fund.

[20] We thank Truls Henrik Hollen at The Oslo Stock Exchange Information Services (Oslo Børs Informasjon) for generously providing us with the data.

trading costs, then divide this figure by the funds' total outstanding shares. The NAV is net of expenses such as management costs and fees, but disregard load charges associated with purchases and redemptions. Since most of the funds report NAV on different trading dates during the earliest years of the sample, we have constructed monthly NAVs by using the last day of reported NAV in each month for all funds. We assume that intra-month dividend payments are reinvested in the fund. Consequently, the one-month simple return between t and $t-1$ is defined as follows:

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

where $NAV_{i,t}$ is the net asset value of fund i at day t, and $r_{i,t}$ is the simple return of fund i at day t. In total, this yields 14.937 observations of monthly returns, which amount to approximately 13 years of return history for each fund, on average. Table A.I in Appendix A displays the exact number of funds available at the end of each year, and how many funds that were born and liquidated throughout the entire sample period. Additionally, the table shows returns for an equally weighted portfolio of all funds in our sample compared to the Oslo Stock Exchange All Share index.[21]

3.2 Interest Rate

We construct excess returns by deducting a proxy of the risk-free interest rate. Treasury bills are widely used for this purpose in the literature, but Norwegian T-bills have proven to be far less liquid than Treasury bills in larger markets. Hence, T-bills might be an unsuitable proxy for the Norwegian market. Ødegaard (2015) argue that the Norwegian Interbank Offered Rate (NIBOR) is the most appropriate for this purpose. Following Ødegaard (2015), we construct a short-term (monthly) risk-free rate from the one-month NIBOR rate, which reflects the pricing of loans in the interbank market. The period before 1986 is however slightly messy regarding interest data. For the period between 1983 and 1986 we therefore use the overnight NIBOR rate as an approximation for the risk-free rate.[22] The one-month risk-free rate at time t is estimated as follows: [23]

$$rf_t = (1 + NIBOR)^{1/12} - 1$$

[21]The equally weighted portfolio is constructed by first calculating returns for each fund in period t. Then, these returns are concatenated into a return-vector and divided by the total number of observations in that period. This type of weighting gives the same importance to each fund in a portfolio, regardless of size. Thus, all of the funds are considered evenly.

[22]The data is obtained from The Norwegian Central Bank (Norges Bank). For more details and explanations about Norwegian interest rate data, see Eitrheim et al. (2004).

[23]Figure B.1 in Appendix B plots the monthly risk-free interest rate. The large spike between 1992 and 1993 as can be seen in the figure is attributable to the currency crisis during that period.

3.3 The Market Proxy

Due to non-observability of the true market portfolio (See e.g. Roll, 1977, 1980), we need to choose an appropriate benchmark as a reasonable approximation of the market. In practice, a market-wide index of stocks is usually applied as a representation of the true market portfolio. In Norway, the Oslo Børs Mutual Fund Index (OSEFX) serve as the benchmark for most of the Norwegian equity mutual funds registered at the OSE. This index would be a natural choice of benchmark, since it is designed to meet specific regulations and diversification requirements in compliance with the directives for fund investments given by UCITS. The OSEFX, however, cannot serve as a market proxy for the whole period between 1983 and 2014 as it originated in December 1995. We therefore disregard the OSEFX as a benchmark in this study, as we require only one specific benchmark spanning the sample period.

In this study, we apply the Oslo Stock Exchange All Share index (OSEAX) as the market proxy for use in the performance models.[24] The OSEAX contains all listed shares on the OSE, and is adjusted for dividends. This index is a widely used benchmark, and provides returns for the whole sample period between 1983 and 2014. Table II below displays average returns for the OSEAX, the equally weighted portfolio of all funds and the equally weighed portfolio of dead funds for various time periods in the sample period. Panel A shows that, on aggregate, the average mutual fund has slightly outperformed the market by 0.12% per year. Panel B shows the total returns for the first half of the sample between 1983 and 1998. The average fund performance was particularly strong during that period, outperforming the market by 0.73% per year. This is no surprise considering that no funds died during that part of the sample. Conversely, the market outperformed the equally weighted fund portfolio during the second half of the sample, as displayed by Panel C. During this period, the OSEAX outperformed the average mutual fund by approximately 0.50% per year. It should be noted that surviving funds outperformed non-surviving funds during both periods. This demonstrates the importance of including liquidated funds in the data sample, as failure to do so could impose the issue of survivorship bias in mutual fund returns.[25] Figure 2 at the end of the section plots the cumulative returns for the OSEAX and equally weighted portfolios comprising all funds and non-surviving funds only and illustrate these results graphically.

[24]The data is obtained as total return indices from Reuters Datastream for the period 1983 to 2014.

[25]The issue of Survivorship Bias in mutual fund returns is discussed in further detail in Section 3.5.

Table II
Descriptive Statistics of Benchmarks and Fund Returns

The table shows descriptive statistics on returns from the OSE All Share index (OSEAX), and returns on equally weighted portfolios comprising both all funds and non-surviving funds. Columns 1 - 6 display average return, standard deviation, maximum and minimum return, skewness and kurtosis. Panel A shows the whole sample period from 1983 to 2014, whereas Panels B and C show the first (1983 to 1998) and second half (1999 to 2014) of the sample. Average returns and standard deviations are annualized, whereas the remaining statistics are reported monthly.

	Average return	Standard deviation	Max	Min	Skewness	Kurtosis
Panel A: 1983:01 - 2014:12						
OSEAX	13.87	21.71	17.45	-27.43	-0.83	5.22
EW (All)	13.99	21.79	17.39	-25.61	-0.77	4.86
EW (Dead)	13.23	21.74	17.39	-25.54	-0.82	5.02
Panel B: 1983:01 - 1998:12						
OSEAX	15.32	22.71	17.45	-27.43	-0.83	5.24
EW (All)	16.05	21.62	17.39	-23.96	-0.69	4.78
EW (Dead)	15.73	21.16	17.39	-23.32	-0.73	4.90
Panel C: 1999:01 - 2014:12						
OSEAX	12.43	20.71	15.04	-23.93	-0.85	5.09
EW (All)	11.93	22.00	15.68	-25.61	-0.84	4.91
EW (Dead)	10.74	22.35	15.62	-25.54	-0.89	5.08

3.4 Risk Factors

To estimate the multifactor models employed in this study, we need return-series on the SMB (Small-Minus-Big), HML (High-Minus-Low), and PR1YR (Momentum) risk factors of Fama and French (1993) and Jegadeesh and Titman (1993). The Fama and French (1993) factors are constructed using value-weighted portfolios formed on size and value (i.e. book-to-market ratios). The SMB factor is the average return on portfolios with a long position in small capitalization companies minus the average return of portfolios consisting of large capitalization companies. HML is the average return on value portfolios (high book-to-market ratio) minus the average return on growth portfolios (low book-to-market ratio). The momentum factor of Jegadeesh and Titman (1993) is constructed by holding long positions in portfolios consisting of stocks with the highest one-year lagged returns, minus portfolios consisting of stocks with the lowest one-year lagged returns, i.e. prior one-year (PR1YR).[26]

In Norway, Professor Bernt Arne Ødegaard has constructed similar factors for the Norwegian equity market by using stocks at the Oslo Stock Exchange.[27] We have obtained monthly return-series on these factors from his website.[28] Table III below reports descriptive statistics on these factors for the entire sample period and the two sub-periods.

[26]See Kenneth French's website and Jegadeesh and Titman (1993) for more detailed explanations on how these factors are constructed.

[27]Details regarding the construction of Norwegian risk factors and their ability to explain differences in cross-sectional returns is discussed in Ødegaard (2015) and Næs et al. (2009).

[28]We thank Bernt Arne Ødegaard for the opportunity to use his asset pricing data at the OSE.

Over the entire sample period, it becomes evident from Panel A that the SMB and PR1YR factors generated the highest average returns of 9.76% and 8.69%, respectively. The size and value factors exhibits the highest returns in the first half of the sample, with average returns of 11.64% and 8.86%, respectively. Conversely, both SMB and HML show considerable lower performance during the second half of the sample with average returns of 7.88% and -0.05%, respectively. The market premium and the PR1YR factor, in contrary, show considerably higher returns in the second part of the sample. Panel B displays standard deviations, and we observe the volatility is highest for all factors during the first half of the sample. In the whole sample period, the market displays the highest volatility with a standard deviation of 22.33%, whereas SMB displays the lowest volatility with a standard deviation of 15.31%. Panel C reports cross-correlations between the factors over the entire sample period. We see that SMB and PR1YR are negatively correlated with the market proxy with correlation coefficients of -0.417 and -0.123, respectively. The HML factor, on the other hand, shows a slightly positive correlation of 0.083 with the market. Furthermore, HML is negatively correlated with the SMB factor and PR1YR is negatively correlated with the HML factor. The correlation coefficients are -0.137 and -0.060, respectively.

Table III
Descriptive Statistics of Factor Returns

The table provides summary statistics in various sample periods for the Norwegian factors used in the unconditional four-factor model of Carhart (1997). MKT in Column 1 is the excess return on the market proxy (i.e. the OSEAX minus the risk-free rate). Columns 2 and 3 show SMB and HML, which are the size and value factors of Fama and French (1993). Column 4 shows the one-year momentum factor of Jegadeesh and Titman (1993), PR1YR. Panel A shows average returns, whereas Panels B and C display standard deviations and cross-correlations between the factors. Average returns and standard deviations are annualized. Returns and standard deviations are reported in percent.

	MKT	SMB	HML	PR1YR
Panel A: Average returns				
1983:01 - 2014:12	6.60	9.76	4.40	8.69
1983:01 - 1998:12	5.74	11.64	8.86	7.37
1999:01 - 2014:12	7.46	7.88	-0.05	10.01
Panel B: Standard deviations				
1983:01 - 2014:12	22.33	15.31	17.17	17.03
1983:01 - 1998:12	22.74	16.97	18.00	17.56
1999:01 - 2014:12	21.98	13.47	16.25	16.51
Panel C: Correlation matrix				
MKT	1.000			
SMB	-0.417	1.000		
HML	0.083	-0.137	1.000	
PR1YR	-0.123	0.145	-0.060	1.000

3.5 Potential Biases in Mutual Fund Returns

Previous studies on mutual fund performance have indicated that characteristics and sample selection regarding the data set could produce biased results. It is important to include both surviving and non-surviving funds in mutual fund performance evaluations, as failure to do so may impose *Survivorship Bias* (See e.g. Brown et al., 1992; Malkiel, 1995; Elton et al., 1996b). Survivorship bias is a property of sample selection, and arises when liquidated funds (e.g. due to bad performance) are removed from the data sample. Existing funds in a data set will typically consist of a mixture of different strategies regarding management style and risk exposure. By eliminating non-surviving funds, strategies that have been proven to be unsuccessful ex-post are excluded from the analysis. This imposes complications since strategies that have yielded high returns tend to survive. Thus, the average fund returns will be biased upwards, which makes the estimate of aggregate mutual fund performance unrealistically high.

Furthermore, a look-ahead bias might arise if one requires a fund to survive a minimum period. Carpenter and Lynch (1999) discover that the look-ahead bias also might occur when year-end returns are missing or excluded in the data sample due to requirements regarding fund size. Elton et al. (1996b) propose treatments for these biases, specifically a "follow the money" strategy that implies tracing the fund after its disappearance.

To gain the most accurate understanding of fund performance, our mutual fund data set includes both surviving funds (54) and non-surviving funds (44). Note that non-surviving funds may have been merged with other funds as a result of acquisitions in the fund industry, or simply been closed down due to bad performance. Following Elton et al. (1996b), we assume that if a fund is merged with another fund, the money is invested in the acquiring funds according to merger terms. Additionally, we impose no specific requirements regarding fund size or number of observations.[29] Figure 2 below displays the cumulative returns development of equally weighted portfolios of all funds and non-surviving funds. As can be observed from Panel A in the figure, the portfolio comprising only non-surviving funds has considerably lower returns than the equally weighted portfolio comprising all funds. We also emphasized this matter in Table II, where the equally weighted portfolio of liquidated funds underperformed all funds in the full sample, as well in both sub-periods. This illustrates the importance of including both surviving and non-surviving funds in our data sample, as failure to do so would clearly impose survivorship bias in the mutual fund data set.[30]

[29]In our sample, GAMBAK Oppkjøp is the smallest fund with 17 monthly observations. Nordea Vekst is the largest with a total of 377 monthly observations.

[30]If non-surviving funds were not to be included in our sample, the potential survivorship bias will be around 0.428 percentage points p.a. See Appendix C for further details and specifications.

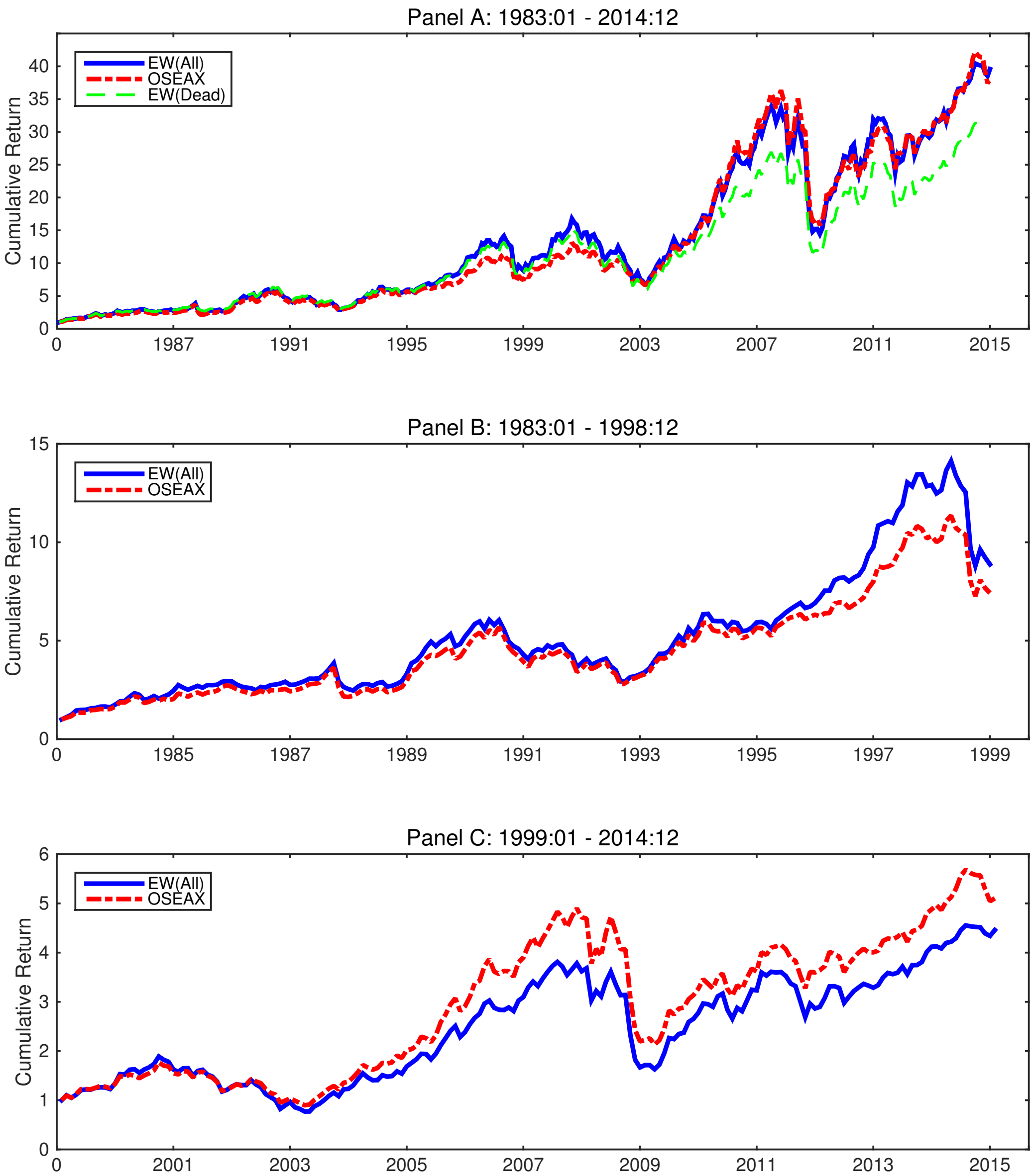

Figure 2. Cumulative returns on equally weighted portfolios and the OSEAX. This figure plots cumulative returns of the OSE All Share index, the equally weighted portfolio that contains all available actively managed funds at the Oslo Stock Exchange between 1983 and 2014, and an equally weighted portfolio comprising only funds that have died during that period. Panel A depicts the entire sample period between 1983 and 2014, including the equally weighted portfolio of non-surviving funds. Panels B and C refer to the first and second half of the sample, 1983-1998 and 1999-2014, respectively.

4 The Performance of Norwegian Mutual Funds

In this section, we investigate whether Norwegian mutual funds generate risk-adjusted returns (i.e. alpha), and if so, whether the abnormal performance is attributable to skilled or lucky fund managers. First, we discuss the theoretical foundation for the performance study, and rationalize our choice of performance model. Second, we investigate the aggregate performance of the funds in our sample. Finally, we disentangle skill from luck in the performance distribution, and investigate whether Norwegian mutual fund managers exhibit superior stock-picking abilities or if the performance is attributable to random chance as economic theory suggest.

4.1 Model Selection

Traditional methods for performance evaluations usually entail employing factor models for returns, and interpret the intercepts from time-series regressions. These factor models serve as performance benchmarks, which we need to specify to evaluate performance, and thereby the stock-picking skills of fund managers. The following subsections briefly discuss our choice of multifactor performance models.

4.1.1 The Unconditional Four-Factor Model

The Capital Asset Pricing Model (CAPM) as first presented by Sharpe (1964), Lintner (1965) and Mossin (1966) aims to describe the relationship between risk and expected returns for a given asset. The single-factor model of Jensen (1968) rests upon the CAPM equation, and is the foundation for all risk-based performance measures. Jensen's alpha, α_i, of fund i is given by the intercept of the model as presented below and is the measure of performance relative to the market at time t:

$$r_{i,t} - r_{f,t} = \underbrace{\alpha_i}_{\text{selection skills}} + \underbrace{\beta_i(r_{m,t} - r_{f,t})}_{\text{risk premium}} + \underbrace{\varepsilon_{i,t}}_{\text{idiosyncratic risk}} \tag{1}$$

where $r_{i,t}$ is the return on fund i in period t, and $r_{f,t}$ is the risk-free rate at time t. The single factor, the market risk premium is given by the market return in excess of the risk-free rate, $r_{m,t} - r_{f,t}$, and β_i is the market risk exposure for fund i. The β_i estimate is the fund portfolio's exposure to non-diversifiable risk (i.e. systematic risk) that the market factor proxy for. The error term, ε_t, has an expectation of zero, and represents the idiosyncratic risk unexplained by the model. If α is positive and significant, the fund manager is able to earn returns that are higher than expected given the portfolio's level of risk as implied by the CAPM. Conversely, a negative α indicates poor performance by the fund manager.

By using the single-factor model for mutual fund performance evaluation, one implicitly assumes empirical validity of the CAPM. This means that the model characterize the true data generating process of the excess returns produced by fund managers. The single-factor model, however, only takes into account one specific risk factor, the market factor. This means that by using the single-factor model for mutual fund performance evaluation one implicitly assumes that a single market proxy can approximate the fund manager's investment behavior. Several studies have questioned the adequacy of the single-factor benchmark for performance evaluation and provided evidence that the single-factor model of Jensen (1968) is not appropriate, as it do not capture cross-sectional differences in average stock returns (See e.g. Elton et al., 1993; Fama and French, 1993; Carhart, 1997). Research on the behavior of expected stock returns lead to the development of multifactor asset pricing models that accounts for several non-diversifiable risk factors in expected stock returns.

In their famous paper, Fama and French (1993) augment Jensen's model to explain the cross-sectional pattern of average returns in the US stock market. They show that the market risk is not the only relevant risk factor in the cross-section of asset returns, and propose a three-factor model to better be able to describe the behavior of expected stock returns. Their model can be run as a regression that includes two additional risk factors, the size (the higher average return of small cap stocks relative to large cap stocks) and value (the higher average return of value stocks relative to growth stocks) premiums. The three-factor model of Fama and French (1993) can be estimated as follows:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{1i} \cdot MKT_t + \beta_{2i} \cdot SMB_t + \beta_{3i} \cdot HML_t + \varepsilon_{i,t} \tag{2}$$

where SMB_t (Small-Minus-Big) and HML_t (High-Minus-Low) are the size and book-to-market risk factors of Fama and French (1993) at time t. The beta coefficients, β_{1i}, β_{2i}, and β_{3i} is the corresponding exposure to the MKT, SMB and HML factors for fund i, respectively. That is, the coefficients relate to the exposure to the sources of systematic risk each of the factors behaves as a proxy for. Fama and French (1996) argue that two additional premiums proxy for non-market systematic risk factors. As in the single-factor model, the intercept α_i is the measure of abnormal performance relative to the fund portfolio's exposure to the risk factors. Motivated by the three-factor model's inability to explain cross-sectional variations in momentum-sorted portfolio returns, Carhart (1997) augments the three-factor model by including the one-year momentum factor (PR1YR) of Jegadeesh and Titman (1993). The PR1YR factor considers the higher expected return of stocks that have performed well during the prior year relative to poor performing stocks (i.e. contrarian stocks). The four-factor model of Carhart (1997) is specified as follows:

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

We use Carhart's (1997) model as our primary performance model due to the following reasons. First and foremost, the four-factor model is able to control for four common risk factors in stock returns that cannot be removed by diversification. Consider a fund manager that increases his exposure to one (or more) of these factors. The increased returns he gains will be offset by increased systematic risk that cannot be removed by diversification. From finance theory, it follows that non-diversifiable risk is the only relevant risk in a well-diversified portfolio. Thus, the increase in expected returns that the fund manager gains by taking on this added risk would only be a reward for the increased systematic risk relevant for the portfolio, and not a reflection of selection skills. Furthermore, the model is advantageous as it can be interpreted as a performance attribution model, were the four risk coefficients can be seen as factor-mimicking portfolios that represent mean returns attributable to four elementary trading strategies. That is, high-versus-low beta stocks, large-versus-small capitalization stocks, value-versus-growth stocks, and one-year return momentum versus contrarian stocks. In this way, the model is used to estimate the added value of active fund managers by measuring the fund's return that cannot be explained by the exposure to systematic risk factors. When evaluating mutual fund performance it is important that the benchmarks used in the performance models include all risk factors relevant for the various investment strategies of fund managers. Thus, Carhart's (1997) model is preferable as it allows us to estimate *alpha* in a more correct sense, i.e. by controlling for the added systematic risk fund managers face by following four elementary investment strategies.

Moreover, Carhart (1997) finds that the four-factor model substantially improves on the average pricing errors of the single-factor CAPM and three-factor model. To our knowledge, there has been conducted little on how the risk factors in the four-factor model behave concerning Norwegian equity mutual funds. Thus, it is imperative to examine whether the model is suitable to our data. To gain some insight, we estimate the four-factor model of Carhart (1997) for each of the 98 funds in our sample and investigate fund loadings on the four factors. The results are presented in Table IV below. The average load on the market proxy is close to 1, which is not surprising considering that the funds in our sample invest primarily in domestic equities. Furthermore, it is evident from the table that the average loading on the SMB, HML and PR1YR factors are quite small. This is, however, the result of a mixture of individual funds that exhibit high and low factor loadings. Thus, even though active Norwegian funds on average are not loading heavily on these factors, the loadings fluctuate heavily across the individual funds in our sample. Table IV also reports the proportion of funds whose factor loading is statistically significant at various levels. For example, the PR1YR factor is significant at the 5% level for approximately 54% of the funds in our sample. Based on these results and the discussion above, we conclude that the four-factor model of Carhart (1997) seems suitable to use for performance evaluations in the Norwegian mutual fund market.

Table IV
Individual Factor Loadings

This table provides a summary of the individual factor loadings (i.e. beta coefficients) obtained by running time-series regressions on all 98 funds in our sample. Column 1 reports the average load, whereas Columns 2 and 3 report the maximum and minimum load. Columns 4 and 5 report the fraction of funds significant at the 5% and 10% significance level, respectively. Column 6 provides the cross-sectional standard deviation of the coefficients. Significance levels are reported in percent. The sample period is 1983 to 2014.

	$\beta_{AVERAGE}$	β_{MAX}	β_{MIN}	Significant *5% level*	Significant *10% level*	σ^2
MKT	0.995	1.235	0.547	100.00	100.00	0.100
SMB	0.097	0.582	-0.194	38.78	42.86	0.149
HML	-0.086	0.193	-0.430	43.88	59.18	0.117
PR1YR	-0.044	0.404	-0.423	54.08	60.20	0.139

4.1.2 The Conditional Four-Factor Model

By using the unconditional four-factor model one implicitly assume the funds to have constant exposure to the four risk factors over time. In a real life scenario, fund managers trade upon varying market conditions, and have changing information about future expectations of risk and return and thus, reconstitute their fund portfolios accordingly. For example, general macroeconomic cycles may influence a fund manager's inclination to bear risk. Consequently, estimating average alphas for the entire sample period based on a fixed beta estimate might produce spurious results (Ferson and Schadt, 1996). Conditional models allow for the possibility that a fund's risk exposure might vary over time depending on lagged public information variables. The conditional version of Carhart's (1997) four-factor model that controls for time-varying market exposure can be specified as follows:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{1i}MKT_t + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}PR1YR_t + \sum_{j=1}^{K} \mathrm{B}_{i,j[Z_{j,t-1}MKT_t]} + \varepsilon_{it} \quad (4)$$

where $Z_{j,t-1}$ is a vector of predetermined information variables that represents public information available at time $t-1$, and $\mathrm{B}_{i,j}$ is a vector of response coefficients of the conditional beta with respect to the information in $Z_{j,t-1}$. Hence, the time-varying market beta is a linear relation between the average (unconditional) beta and the conditional instruments: $\beta_{i,j} = \sum_{j=1}^{K} \mathrm{B}_{i,j}[Z_{j,t-1}MKT_t] + \beta_{1i}$. Note that in this model, we have assumed that the information variables only affect the market exposure. We could augment the model further by adding interaction terms for the other three risk factors. For example, we can add $\sum_{j=1}^{K} \mathrm{B}_{i,j}[Z_{j,t-1}HML_t]$ if we think that the value premium varies over time.

Specifically, conditional models aim to measure the time-variation in risk exposures by using a vector of predetermined information variables that fund managers can use as a tool in their investment strategy. Several information variables have been suggested for this purpose in previous literature. Among others, some of these include variables such as the T-bill yield, market dividend yield, lagged market return, and a quality spread between government and corporate bonds (See e.g. Ferson and Schadt, 1996; Otten and Bams, 2004). In this study, we consider the following three variables: (1) lagged growth on industrial production, (2) lagged market return, and (3) lagged yield on the oil price.[31] These variables have also been proven to be useful for predicting stock returns (See e.g. Pesaran and Timmermann, 1995; Sørensen, 2009b).[32]

In order to keep the conditional model parsimonious and to ease the interpretability of the model, only a single conditional variable is considered in this study. To gain insight on which of the three information variables that best suits our data, we first examine whether the market exposure of the aggregate portfolio of all funds in our sample depends on the lagged growth on industrial production, lagged yield on the oil price and the lagged market return. Results are reported in Appendix D. In general, we find that the conditional variables do not improve the model in terms of goodness of fit. Moreover, none of the conditional variables are individually statistically significant, and they have marginal economic significance at the aggregate level. The aggregate portfolio's market exposure is most sensitive to the lagged yield on the oil price. This variable is also bordering on statistical significance. To gain further insight on the relevance of the variables, we also estimate conditional four-factor models that controls for time-varying MKT loadings for each individual fund. Table V below shows that there is considerable variation in individual loadings on the three conditional variables. The lagged yield on the oil price appears to be the variable that is significant for most funds at the individual level.

Table V
Individual Loadings on the Conditional Variables

The table reports individual loadings on the three information variables considered in our study; the lagged yield on the oil price, lagged industrial production growth, and lagged market return, respectively. Column 1 reports the average load, whereas Columns 2 and 3 report the maximum and minimum load. Columns 4 and 5 report the fraction of variables significant at the 5% and 10% significance level, respectively. Column 6 provides the cross-sectional standard deviation of the coefficients. Significance levels are reported in percent. The sample period is 1983 to 2014.

Information variables	$B_{AVERAGE}$	B_{MAX}	B_{MIN}	Significant *5% level*	Significant *10% level*	σ^2
Oil Price $_{t-1}$	-0.024	0.507	-0.265	19.39	27.56	0.276
Market Return$_{t-1}$	-0.010	0.274	-0.211	11.23	18.37	0.112
Industrial Production$_{t-1}$	-0.004	0.075	-0.131	10.21	17.35	0.082

[31] Return-series on industrial production for Norway from 1983 to 2014 is obtained from OECD Statistics. Data on the Brent Crude oil price is collected from Macrobond.

[32] Because data on the most commonly used information variables was difficult to obtain due to our long sample period, we advocate other information variables that have been proven to be useful for predicting stock returns.

Overall, our results suggest that the unconditional four-factor model explains returns on actively managed Norwegian mutual funds reasonably well. Conditional variables add little to the explanatory power of the model. Thus, the main focus in this study will be on results from the unconditional four-factor model. However, since the lagged yield on the oil price is statistically and economically significant for a sizeable fraction of the funds in our sample, we also perform our tests using a conditional model where we include the lagged yield on the oil price as an information variable and briefly comment on the results.

4.2 The Bootstrap Methodology

We employ a cross-sectional bootstrap methodology to evaluate the performance of Norwegian open-end, domestic equity mutual funds. There are several reasons for why the bootstrap is necessary for proper inference in this context. These include non-normalities in individual fund returns as well as non-normalities in the cross-section of mutual fund alphas. The bootstrap has the advantage that it provides a non-parametric approach to statistical inference about performance.

Conventional OLS inference relies on the assumption of normally distributed residuals. There are several properties that would lead to a rejection of the normality assumption when analyzing mutual fund alphas, making standard parametric test statistics invalid. First, individual stocks within the typical mutual fund portfolio tend to yield returns with significantly different skewness and kurtosis compared to a normal distribution. Even though the central limit theorem implies that a large equally weighted portfolio of non-normal stocks will approach normality, fund managers usually hold large positions in some stocks and returns tend to be cross-correlated. Second, individual stocks often exhibit various levels of serial correlation in returns and tend to have heteroscedastic variance. 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 (Kosowski et al., 2006). As a result, normality is a poor approximation for the typical mutual fund. Indeed, tests reveal that normality is rejected for 62% of the funds in our sample using the unconditional four-factor model of Carhart (1997). By exploring the issue further, we observe that funds in the extreme region of the performance distribution (i.e. exceptionally good or bad funds) exhibit even greater skewness and kurtosis than compared to funds closer to the center of the performance distribution.[33] The bootstrap can significantly improve the validity of inferences about performance, as it does not rely on any distributional assumptions, a feature particularly important in the extreme regions of the performance distribution.

[33] We test for normality in the residuals using the Jarque-Bera test at the 5% significance level. For details on the test results and information on skewness and kurtosis for individual funds we refer to Appendix E.

Bickel and Freedman (1984), Hall and Martin (1988), Horowitz (2003), Kosowski et al. (2006), and Fama and French (2010) all argue that the bootstrap provide more accurate evaluation of the significance of alpha estimates. Horowitz (2003) conducted Monte Carlo experiments that demonstrated that the bootstrap could significantly reduce the difference between the true and nominal probability of correctly rejecting a given null hypothesis. For example, by recognizing the presence of thick tails in individual fund returns, the bootstrap does not reject abnormal performance as often as the standard parametric t-test.

Furthermore, when modelling the cross-sectional distribution of mutual fund residuals, one cannot assume normality in this distribution. The cross-sectional distribution of mutual fund residuals will be a complicated mixture of individual fund distributions characterized by the following two features: higher moments in individual fund residuals and heterogeneous risk taking across funds. Even tough individual distributions of residuals are normally distributed; it does not necessarily imply that the cross-sectional distribution of these residuals is normal. To exemplify, consider a selection of funds with normally distributed residuals, with different levels of risk so that the residual variances vary uniformly between 0.5 and 1.5 (i.e. the mean variance is 1). Given these assumptions, the cross-sectional distribution of these residuals will have fatter tails relative to a normal distribution. To illustrate, as we move further into the tails of the distribution, the probability of extreme outcomes does not drop as fast as it would have if it had been normally distributed, since high-risk funds overcompensate for the large drop in extreme outcomes from low-risk funds (Kosowski et al., 2006). This example shows that the cross-section of mutual fund residuals can exhibit non-normal behavior regardless of the distribution of individual funds, as long as risk levels are heterogeneous across funds. Given the complex nature of modeling the joint distribution across all 98 funds in our sample, the bootstrap emerges as a very attractive method to analyze a cross-section of mutual funds. In the next subsection, we will describe the implementation of the bootstrap.

4.2.1 Implementation

We follow the bootstrap method developed by Kosowski et al. (2006). The method involves residual-only resampling under the null of no outperformance. The implementation of the bootstrap procedure will be explained in the remainder of this section.

First, for fund $i = \{1, 2, \cdots, N\}$ we estimate the unconditional four-factor model of Carhart (1997) using OLS (Equation 3), and save the coefficients (i.e. alphas and betas), alpha t-statistic, and the time-series of the estimated residuals, $\hat{\epsilon}_{i,t}, t = \{T_{i0}, \cdots, T_{i1}\}$. Where T_{i0} and T_{i1} are the first and last monthly returns available for fund i, respectively. Next, we draw a random sample (with replacement) from fund i's residuals $\hat{\epsilon}_{i,t}$ to generate a time-series of resampled residuals, $\tilde{\epsilon}_{i,t}^{(b)}$, with the same length as the initial residual vector. Where b represent an index of the bootstrap number (i.e. $b = 1$ for bootstrap resample number one and so on). The time-series of the *resampled* residuals are then used in combination with the factor returns and the estimated beta coefficients to construct a *pseudo* time-series of monthly excess returns for fund i, where the fund's performance (i.e. alpha) is set to zero by construction ($\hat{\alpha}_i = 0$):

$$\tilde{r}_{i,j}^{(b)} = 0 + \hat{\beta}_{1i} \cdot MKT_t + \hat{\beta}_{2i} \cdot SMB_t + \hat{\beta}_{3i} \cdot HML_t + \hat{\beta}_{4i} \cdot PR1YR_t + \tilde{\epsilon}_{i,t}^{(b)} \tag{5}$$

The Carhart (1997) four-factor model is then estimated using the *pseudo* excess return vector, $\tilde{r}_{i,j}^{(b)}$. If an abnormally high number of positive (negative) residuals are drawn in a given bootstrap sample, b, a positive (negative) alpha may result. This alpha emerges as a result entirely due to sampling variation around a true alpha of zero and is entirely due to *luck*. For bootstrap simulation 1, the above process is repeated across all funds to arrive at the first draw from the cross-section of bootstrapped alphas. The bootstrapped alphas, $\tilde{\alpha}_i^{(1)}\{i = 1, 2, \cdots, N\}$, are then ordered from the highest to the lowest $\tilde{\alpha}_{max}^{(1)}$ to $\tilde{\alpha}_{min}^{(1)}$. The process is then repeated for $b = 10.000$ bootstrap simulations to generate a cross-sectional distribution of alphas, $\hat{\alpha}_i^{(b)}\{b = 1, 2, \cdots, 10.000; i = 1, 2, \cdots, N\}$, resulting purely from sampling variation around a true alpha of zero (by construction). Percentiles (points) from this cross-sectional distribution of alphas are used to construct separate *luck distributions* for each of the ex-post ranked funds. As an example, the distribution of alphas for the ex-post top fund ($\hat{\alpha}_{max}$) is constructed of the maximum alpha across all bootstraps ($f(\tilde{\alpha}_{max})$).[34] It is important to note that the alphas in this distribution can be associated with a different fund for each bootstrap, depending on the outcome of the draw from each fund's residuals. This enables us to use valuable information of *luck* represented by all the funds in our sample. To infer the existence of inferior or superior managerial skill we compare any ex-post ranked fund with its appropriate *luck distribution.* For example, if we want to examine whether the performance of the ex-post top fund is attributable to *skill* or *luck*, we compare the estimated ex-post alpha for this fund ($\hat{\alpha}_{max}$) with its *luck distribution* ($f(\tilde{\alpha}_{max})$).[35] If the bootstrap simulations generate far fewer extreme positive alphas than compared to the ex-post estimated alpha, we conclude that the observed alpha are not due to sampling variation around a true alpha of zero (i.e. *luck*), and that real manageral skill exist.[36]

In this study we use the t-statistic of the estimated alpha as our performance statistic rather than alpha as it has some advantageous statistical properties (i.e., it is a pivotal statistic) when constructing bootstrapped cross-sectional distributions.[37] Although the

[34]To further specify, $\hat{\alpha}_{max}$ is the estimated ex-post alpha for the top fund while $f(\tilde{\alpha}_{max})$ is the appropriate *luck distribution* for this fund.

[35]This process can be repeated for any other point in the performance distribution - all the way down to the bottom ranked fund.

[36]For the top fund, this requires that only 5% of $f(\tilde{\alpha}_{max})$ is greater than $\hat{\alpha}_{max}$ (5% confidence level).

[37]The process involves the same steps as described above. The only difference is that we sort on t-statistic of alpha instead of alpha.

estimated alpha quantifies the economic size of abnormal performance, it tends to suffer from a lack of precision. Short-lived funds (i.e. funds with a small number of observations) or funds that engage in high levels of risk-taking are likely to have more imprecise estimated alphas and these alphas may be spurious outliers in the cross-sectional distribution of alphas. The t-statistic normalizes the estimated alpha by its standard error and inevitably account for differences in risk-taking across funds and different lifespans of funds. Thus, the bootstrapped distribution of the t-statistic is less dispersed than that of the alphas.

Furthermore, a small number of observations in the estimation are likely to increase the sampling variability in the bootstrap results and consequently widen the tails of the bootstrapped distribution. This could bias the results towards the conclusion that fund performance is not outside what is expected by mere chance. A way to improve the precision of the performance estimates is to impose a minimum requirement on the number of observations for a fund to be included in the analysis. A drawback of this approach, however, is that it imposes a certain survivorship bias by restricting the sample to only include funds with a given minimum number of observations. To address this issue, we assess the sensitivity of the bootstrap results for alternative restrictions on the minimum number of observations required for a fund to be included in the analysis.

4.3 Empirical Results of Fund Performance

We initiate our analysis on performance by examining the aggregate performance of all funds in our sample. In the previous sections, we concluded that the unconditional four-factor model of Carhart (1997) is appropriate to use for mutual fund performance evaluations in the Norwegian market. However, for completeness, we also evaluate aggregate fund performance using Jensen's (1968) single-factor CAPM, the three-factor model of Fama and French (1993) and the conditional version of Carhart's model. We then turn to bootstrap simulations to infer the existence of superior and inferior fund managers.

4.3.1 Aggregate Mutual Fund Performance

Table VI below reports estimates from regressions on the equally weighted portfolio of all funds in our sample using various performance models. The intercepts from the regressions inform us whether funds, on average, produce returns different from those implied by exposure to common risk factors. Panel A shows that the funds, on average, have generated a negative yearly alpha of -0.43% measured against the four-factor model of Carhart (1997).[38] The three-factor model of Fama and French (1993) yields a negative

[38]Measured against the conditional four-factor model, the alpha estimate increases to -0.19%. See Appendix D for details.

alpha of -0.57% per year. On the other hand, the estimated alpha from the single-factor CAPM is positive of 0.39% per year. However, the estimated alphas are not statistically significant in any of the models. The drop in risk-adjusted performance from the single-factor CAPM to the multifactor models is mainly attributable to the portfolio's exposure to the size (SMB) factor. The SMB factor is the equity benchmark with the highest average return in our sample period. Further, Panel A shows that the funds, on average, have a beta exposure close to one against the market (MKT) for all of the factor models. This result is not surprising considering that our sample consists of funds that invest primarily in Norwegian equities. By examining the factor loadings, we get an indication of the potential investment strategy the average fund in our sample is following. The positive sign on the coefficient in front of the SMB factor implies that the average fund tends to favor small capitalization stocks. The coefficient in front of the value (HML) factor is negative and significant. Thus, on average, active Norwegian mutual funds show some exposure to growth stocks. Moreover, the equally weighted portfolio has negative but insignificant exposure to the momentum (PR1YR) factor.

Panels B and C show that the aggregate performance is best during the first part of the sample, which is consistent with our previous findings.[39] It is important to note that the fund portfolio changes its exposure to the risk factors quite substantially from the first part to the second part of the sample. In particular, it becomes evident from the table that active Norwegian funds, on average, load more heavily on the SMB factor during the first part of the sample period than compared to the last part. Thus, the average fund tends to favor small capitalization stocks to a greater extent during the first part of the sample. Additionally, Table VI shows that the fund portfolio has positive but insignificant exposure to the HML and PRIYR factors during the first part of the sample period (Panel B) while the exposure towards these factors turns statistically significant and negative in the second part of the sample (Panel C). The negative and significant loading on the PR1YR factor in the second part of the sample indicates that the average fund follows a contrarian investment strategy during that period.

One might question why the best performance is assigned to the first part of our sample period. The rationale for this is somewhat unclear, but one explanation might be that no funds died during the first part of the sample, and that the competition among funds accelerated throughout the second part. Moreover, it is not inconceivable that the most successful fund managers over time have turned to more lucrative areas such as international equity funds, hedge funds or private equity funds. As discussed in Section 1.1.1, the decline in the market share of Norwegian equity mutual funds relative to the total equity market was indeed quite substantial at the beginning of our sample.

[39] Table II in Section 3.3 shows that the equally-weighted portfolio of all actively managed Norwegian mutual funds performs best in the first half of the sample period before adjusting for risk.

Table VI
Aggregate Fund Performance for Different Sample Periods

The table shows mutual fund alphas for an equally weighted portolio comprising all mutual fund returns, loadings on the four risk factors of Carhart (1997) for the Norwegian market, and adjusted R^2. The numbers are obtained by conducting time-series regressions by using the single-factor CAPM of Jensen (1968), the Fama-French (1993) three-factor model and Carhart's (1997) unconditional four-factor model. Panel A shows the results for the whole sample period between 1983 and 2014, whereas Panel B (C) shows results from the first (second) half of the sample. The t-statistics are reported in parentheses. Numbers assigned stars, ***, **, and * indicates significance at the 1%, 5%, and 10% levels, respectively. Alphas are reported in percent per year.

Model	α	β_{MKT}	β_{SMB}	β_{HML}	β_{PR1YR}	R^2_{Adj}
Panel A: 1983:01 - 2014:12						
CAPM	0.391	0.965***				0.923
	(0.33)	(67.60)				
Fama-French	-0.569	0.997***	0.098***	-0.049***		0.928
	(-0.56)	(66.07)	(4.51)	(-2.80)		
Carhart	-0.428	0.995***	0.101***	-0.050***	-0.029*	0.929
	(-0.33)	(66.01)	(4.66)	(-2.86)	(-1.69)	
Panel B: 1983:01 - 1998:12						
CAPM	1.339	0.910***				0.911
	(0.95)	(44.21)				
Fama-French	-0.463	0.937***	0.124***	0.022		0.919
	(-0.36)	(42.81)	(4.36)	(0.84)		
Carhart	-0.421	0.936***	0.119***	0.027	0.041	0.919
	(-0.55)	(42.88)	(4.17)	(0.99)	(1.58)	
Panel C: 1999:01 - 2014:12						
CAPM	-0.840	1.030***				0.941
	(-0.56)	(54.93)				
Fama-French	-1.398	1.038***	0.068**	-0.091***		0.946
	(-1.05)	(49.60)	(2.22)	(-3.84)		
Carhart	-0.762	1.026***	0.077**	-0.092***	-0.088***	0.950
	(-0.37)	(50.43)	(2.31)	(-4.08)	(-3.99)	

To examine the time-series dimension of the aggregate mutual fund performance we report estimates from rolling window and extending window regressions in Figure 3 below. This allows us to track how the fund portfolio's alpha estimates have evolved throughout the entire sample period. Panels A1 and A2 show rolling window and extending window estimates of the annualized alphas using Jensen's (1968) single-factor CAPM.[40] In Panels B1 and B2, the single-factor CAPM is replaced by Carhart's (1997) four-factor model. From the rolling window estimation in Panel A1, we see that the alphas computed against the CAPM spikes above zero several times for short intervals during the sample period. The CAPM alphas are particularly high at one period in the middle of the sample. At that period, the estimated alphas are also significant at the 5% level for a short while.

[40] In the rolling window estimation, the window length is set to 36 months. In the extending window estimation, the estimation window extends from 36 months initially to 384 months.

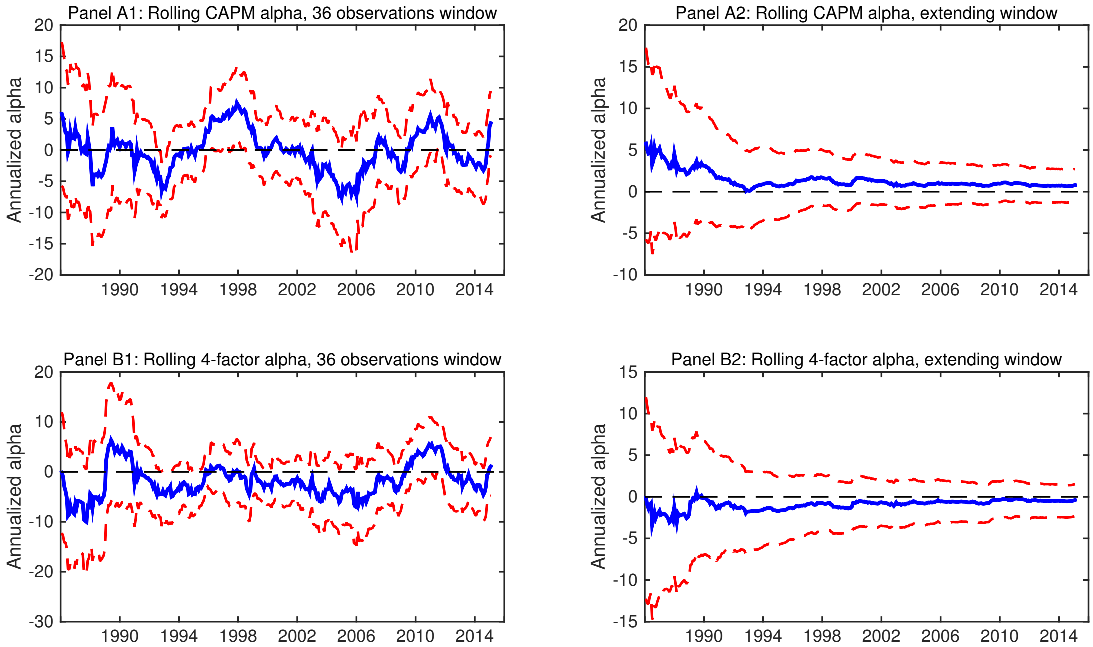

Figure 3. Rolling and extending window estimations of the equally weighted portfolio. This figure plots annualized alpha estimates in percent of the equally weighted portfolio. The left panels report rolling window estimates of alpha, where the window length is set to 36 months. The right panel reports extending window estimates, where 36 months is used to estimate the first regression. The top panels show alpha estimates versus Jensen's (1968) single factor CAPM, whereas the bottom panels show alpha estimates versus Carhart's (1997) four-factor model. The solid line shows the alpha estimate, whereas the dotted line shows the standard error bands. The sample period is 1983 to 2014.

The four-factor alphas (Panel B1) are more stable compared to the CAPM alphas and are never significant. Furthermore, from the extending window estimations (i.e. Panels A2 and B2), we observe that the CAPM alpha is particularly high at the beginning of the sample period and declines almost monotonically trough time. Conversely, the four-factor alpha is particularly low at the beginning of the sample period and increases somewhat through time. Both alpha estimates in the extending window estimations are, however, not statistically significantly different from zero at any point in time.

One explanation for the relatively high performance spread between the CAPM and the four-factor model during the earliest part of the sample period may be attributed to the fund portfolio's exposure to the SMB, HML and PR1YR factors. Although the CAPM explains most of the variations in returns, it does not directly control for exposure towards the SMB, HML and PR1YR factors. From Table VI, we observe that the aggregate fund portfolio exhibit relatively high exposure towards the SMB factor in the first part of the sample. Indeed, Table III in Section 3.4 shows that the SMB factor performed particularly well during that period. This may explain a good proportion of the difference between the CAPM alpha and the four-factor alpha, especially during the early part of the sample period. These implications indicate that the CAPM seems like a poor choice to evaluate "true" abnormal performance of mutual funds in the Norwegian market, as it fails to control for exposure towards the SMB, HML and PR1YR factors.

Based on the results, we conclude that active Norwegian mutual fund managers do not, on average, have sufficient skill to generate returns to cover the costs they impose on their investors. This outcome may emerge as a result of inferior and superior fund managers that balance each other out. That is, our sample might consist of both funds that perform well and funds that perform poorly. In the next section, we turn to individual fund performance and adapt bootstrap simulations that enable us to distinguish between skill and luck in individual mutual fund performance.

4.3.2 Individual Funds - Separating Skill from Luck

The tables in Appendix F provides results of regressions on each fund in our sample using the unconditional four-factor model of Carhart (1997). The tables show that three funds exhibit significantly positive alphas, whereas five funds exhibit significantly negative alphas. Given the large number of funds in our sample, it is reasonable to expect that some funds will perform well and some to perform poorly due to mere chance. The question then arises, whether the performance of actively managed funds, in fact, is credible evidence of genuine stock-picking skills, or if it is simply attributable to random chance. Moreover, are the bad performing funds only a result of bad luck, or is the performance a result of inferior managerial skill? In this section, our objective is to separate skill from luck in individual mutual fund performance using a bootstrap approach similar to Kosowski et al. (2006) and Cuthbertson et al. (2008). As we will see, when compared to the standard parametric t-test, the bootstrap can provide significantly different results regarding the significance of mutual fund performance in the tails of the performance distribution.

Table VII below shows the results of the bootstrap simulations where funds are ranked according to their actual (*ex-post*) t-statistic of alpha. Row 1 in each panel reports the actual t-statistic of alpha for funds at different percentile points in the performance distribution, and Row 2 shows the funds associated alpha estimate. Bootstrapped p-values of the t-statistics of alphas in Row 1 are reported in Row 3.[41] Panel A reports results from the unconditional four-factor model of Carhart (1997). The top-ranked fund has an actual t-statistic of alpha of 3.03 and a corresponding yearly alpha of 5.03%.[42] However, the bootstrapped p-value of the t-statistic of alpha is 0.175, indicating that out of the

[41]The bootstrapped p-values are calculated as follows. Top funds: proportion of bootstrapped alpha t-statistics > actual t-statistic at each percentile (point). Bottom funds: proportion of bootstrapped alpha t-statistics < actual alpha t-statistic at each percentile (point).

[42]One has to bear in mind that the fund returns is net of management fees. Adding back management fees will push the alpha estimate and the corresponding t-statistic upward. We are, however, unable to validate fund performance gross of fees since we were unable to obtain data on fee structures dating back to 1983.

10 000 bootstrap simulations under the null of no abnormal performance, the bootstrap generates 17.5% t-statistics of alpha that are greater than 3.03. Thus, we cannot reject the null that the alpha t-statistic of the top fund may be a result of luck alone. Moving further to the center of the performance distribution (i.e. the 2nd, 3rd, top 5% and 10% funds), we see that the t-statistics of alpha remain insignificant under the bootstrap. In the left tail of the distribution, the bottom ranked fund has an actual t-statistic of alpha of -3.59 and a corresponding yearly alpha of -14.25%. From Row 3, we see that the bootstrapped p-value for the bottom fund is 0.036. This indicates that only 3.6% the bootstrapped t-statistics of alpha is below -3.59. Hence, there is a very small probability that the bottom-ranked fund could have generated a t-statistic of alpha of -3.59 or higher (i.e. more negative) by pure chance. Thus, we reject the hypothesis that this fund does not underperform its benchmarks and conclude that this fund has *poor skill.* We also find significant negative abnormal performance that is not attributable to *bad luck* but is due to *poor skill* when we move further to the center of the performance distribution from the left tail (i.e. the 2nd, 3rd, and bottom 5% and 10% funds).

Table VII

Bootstrap Results of Norwegian Mutual Fund Performance

The table provides results from the cross-sectional bootstrap of Norwegian mutual fund performance for the whole sample period from 1983 to 2014. Panel A (B) shows statistics from the unconditional (conditional) four-factor model of Carhart (1997). Columns 1 to 6 report statistics for various points and percentiles of the performance distribution, ranging from the worst (bottom) fund to the best (top) fund. Row 1 in both panels reports the actual (estimated) t-statistics of alpha, whereas Row 2 reports to the associated alpha (annualized) for these t-statistics. The third row in each panel reports the bootstrapped p-values of the t-statistics of alpha based on 10,000 bootstrap resamples. Funds are ranked on their unconditional and conditional t-statistic of alpha in Panel A and B, respectively.

Panel A: Unconditional Four-Factor Model										
$r_{i,t} - r_{f,t} = \alpha_i + \beta_{1i}MKT_t + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}PR1YR_t + \varepsilon_{it}$										
	Bottom Fund	2^{nd}	3^{rd}	Bottom 5%	Bottom 10%	Top 10%	Top 5%	3^{rd}	2^{rd}	Top Fund
t-alpha	-3.59	-3.20	-3.19	-2.61	-1.93	1.33	1.80	2.37	2.70	3.03
Alpha	-14.25	-7.84	-19.32	-6.20	-6.71	2.05	4.90	3.98	6.58	5.03
p-tstat	0.036	0.005	0.005	0.000	0.000	0.415	0.261	0.091	0.075	0.175
Panel B: Conditional Four-Factor Model										
$r_{i,t} - r_{f,t} = \alpha_i + \beta_{1i}MKT_t + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}PR1YR_t + \mathrm{B}_{i,j[Z_{j,t-1}MKT_t]} + \varepsilon_{it}$										
	Bottom Fund	2^{nd}	3^{rd}	Bottom 5%	Bottom 10%	Top 10%	Top 5%	3^{rd}	2^{nd}	Top Fund
t-alpha	-3.25	-3.15	-3.13	-2.36	-1.65	1.39	1.83	2.40	2.74	3.22
Alpha	-13.56	-7.91	-20.27	-3.44	-9.99	2.39	3.50	3.77	4.52	5.60
p-tstat	0.091	0.008	0.000	0.001	0.016	0.280	0.220	0.074	0.064	0.104

Panel B of Table VII reports findings from the conditional four-factor model. It is evident that including oil price as a conditional variable shifts the performance distribution slightly to the right. This finding indicates that active Norwegian mutual funds seem to time the overall market according to the level of the oil price. With the exclusion of the bottom fund, inferences from the bootstrap are broadly consistent with those from the unconditional model. There is no evidence of skill in the right tail of the performance distribution, but with a bootstrapped p-value of 0.09, we can no longer conclude that the bottom ranked fund has delivered truly inferior performance. When moving further towards the center of the performance distribution from the left tail, the funds display *poor skill*.

As a robustness check of our bootstrap results, we conduct a sensitivity analysis that addresses the potential issue of wide tails in the bootstrapped distribution of alphas t-statistics caused by the inclusion of funds with a small number of observations. In the sensitivity analysis, we restrict our sample to include funds with at least 24 and 36 monthly return observations. When imposing a restriction of at least 36 months of return observations, we find that the bootstrapped p-values for funds in both tails of the performance distribution drops somewhat but remains insignificant. For details see Appendix G.

Figure 4 below shows distributions of alpha t-statistics for funds at various percentile points in the cross-section using the unconditional four-factor model.[43] For example, Panel A1 depicts the bootstrapped distribution of alpha t-statistics for the top-ranked fund across all bootstrap simulations. The mode of the distribution is about 2.2, but the distribution is heavily skewed to the right and include alpha t-statistics varying from around one to, in rare cases, somewhat above 6. Furthermore, Panel A1 shows that the bootstrap generates too many alpha t-statistics that are greater than the actual (estimated) alpha t-statistic of 3.03 for the top-ranked fund to conclude that the actual (estimated) t-statistic of alpha for this fund is due to genuine stock picking ability. Indeed, Panel A1 illustrates a case where the standard parametric t-test rejects the null of no outperformance, while the bootstrap does not. Conversely, Panel B3 illustrates a situation where the bootstrap rejects the null while the standard t-test does not. These apparent contradictions are due to the highly non-normal distribution of idiosyncratic risk across our top and bottom performing funds. Both these cases show that standard t-tests might give misleading inferences for funds in the extreme tails, and highlight the importance of a bootstrap when examining the statistical significance in the tails of the performance distribution, which can have complex distributional properties.

[43] A similar figure for the conditional four-factor model is reported in Appendix H.

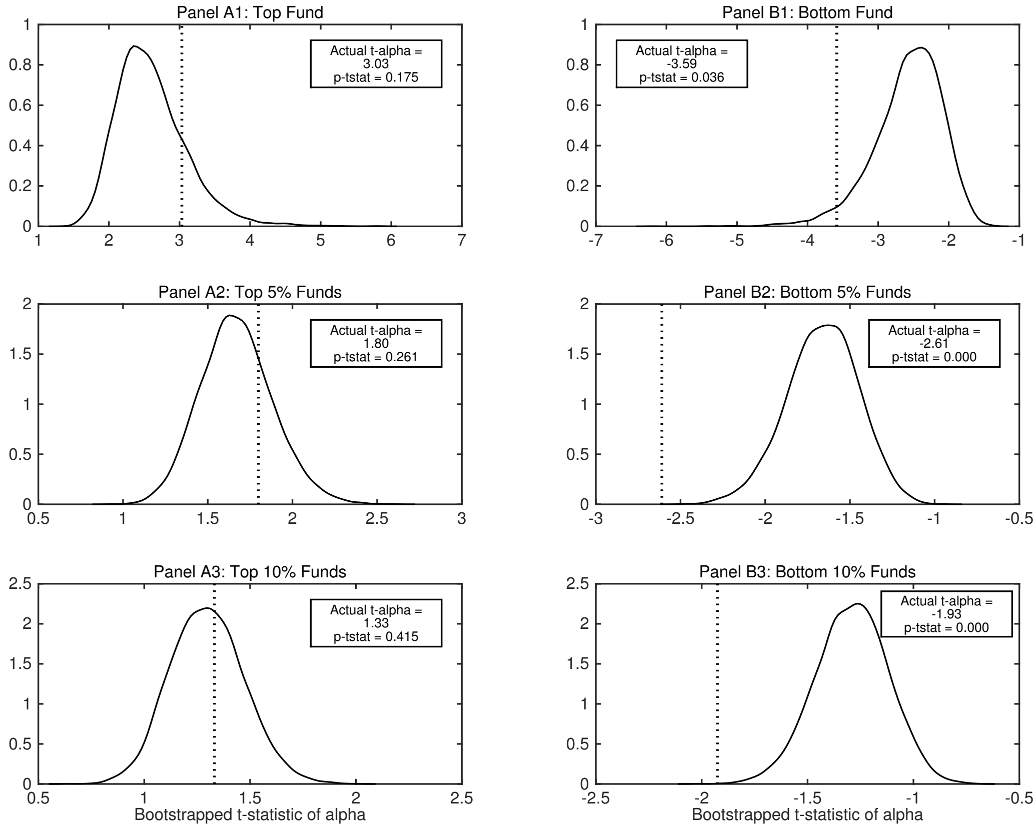

Figure 4. Actual (estimated) t-statistics of alpha vs. bootstrapped t-statistics of alpha distributions for individual funds. This figure plots kernel density estimates of the bootstrapped unconditional four-factor t-statistic of alpha distribution (solid line) for various percentile points in the cross-section. The x-axis shows the t-statistic of alpha (performance measure) and the y-axis the kernel density estimate. The vertical dotted line shows the actual (estimated) fund t-statistic of alpha. Panel A1-A3 (B1-B3) show marginal funds in the right (left) tail of the performance distribution. For example, "Top 5%" in Panel A2, refers to the marginal alpha t-statistic at the top 5 percentile of the performance distribution. The bootstrapped distributions of t-statistics of alpha (under $H_0 : \hat{t}_{\hat{\alpha}} = 0$) is based on 10,000 bootstrap resamples. Funds are ranked on their unconditional four-factor t-statistic of alpha, where the highest ranked fund has the highest t-statistic over the sample period.

To summarize, our bootstrap results suggest no evidence of superior fund management net of fees. Even for the top fund one would expect t-statistics of alpha at least as high as the observed ex-post t-statistic of alpha too many times by pure chance to conclude that this particular fund has genuine stock-picking skill (i.e. the funds performance may be explained by chance alone). In the left tail of the performance distribution, we find several funds that exhibit significant and negative t-statistics of alpha that cannot be explained by *bad luck*. Thus, our results provide evidence on the existence of a large number of *poorly performing* funds in our sample. This is bad news for investors, as it appears to be very difficult to earn risk-adjusted returns from investing in actively managed Norwegian mutual funds.

Although our analysis suggests little evidence of abnormal performance compared to benchmark returns from Carhart's (1997) unconditional four-factor model, we find funds that performed well and were able to generate returns before any risk-adjustments. As can be observed from the factor loadings on each individual fund in Appendix F, it seems as if the actual returns generated from these funds can be attributed to beta risk. In other words, our analysis gives no credit for taking on added systematic risk in terms of betas that drive the fund portfolios. If it is possible to buy this beta exposure at a lower cost than the fees imposed by mutual funds is a subject for debate, but it is unlikely that the average investor is able to form the right combinations and be able to benefit in a similar way. Thus, a practical implication to investors is that the negative alphas often inferred by multifactor models does not always imply *value destruction*, but can merely represent *transaction fees*. For sophisticated investors, mutual funds seems to be a preferable option to gain exposure to certain risk factors (See e.g. Pástor and Stambaugh, 2002a).

5 Persistence in Norwegian Mutual Fund Returns

In the previous section, we found no evidence of superior fund performance among actively managed Norwegian mutual funds. However, investors might still be able to identify fund managers who occasionally outperform the market and earn risk-adjusted returns by exploiting performance persistence. Persistence is also important from an academic point of view, because evidence of performance persistence would support a rejection of the semi-strong form of the Efficient Market Hypothesis (EMH). In efficient capital markets, evidence of persistence will not occur because all investors are perfectly informed. Previous studies have, however, found evidence of persistence in U.S. mutual funds.[44] To evaluate whether there exist persistence in the performance of actively managed Norwegian mutual funds during the period from 1983 to 2014, we apply some of the most prominent methods proposed in the academic literature.

5.1 Recursive Portfolio Formation Test

We begin examining performance persistence by using the recursive portfolio formation approach (See e.g. Hendricks et al., 1993; Grinblatt and Titman, 1993; Carhart, 1997). Carpenter and Lynch (1999) argue that the recursive portfolio approach is the most powerful test to detect performance persistence. In short, this approach involves forming portfolios of funds based on the fund's performance over a ranking period and evaluating how these portfolios perform over a future holding period. The following section provides an in-depth explanation of our implementation of the method.

[44]See e.g. Hendricks et al. (1993), Carhart (1997) and Pástor and Stambaugh (2002a) for previous evidence on long-term persistence. For previous research and evidence on short-term persistene, see e.g. Bollen and Busse (2005) and Huij and Verbeek (2007).

The first step in the recursive portfolio approach is to rank all funds based on their past M months performance. Previous research indicates that persistence is hard to recognize if the evaluation period on which the funds are ranked is too long (See e.g. Bollen and Busse, 2005). Thus, we establish both short and long ranking periods. In our baseline test we follow Carhart (1997) and rank funds based on their past raw returns, and let M vary from 1 to 24 months. Carhart (1997) argue that ranking funds based on their past raw returns has the advantage of allowing for short ranking periods, and has the capability to avoid potential estimation errors inherent with ranking on risk-adjusted returns. On the other hand, raw returns might not be an adequate measure of real investment skill since they are affected by different investment styles such as growth or value, the portfolio's risk level, and managerial skill as well as luck (Lückoff, 2011). If managerial skill exists, raw returns might be a noisy measure. To get a better representation of true investment skill, we also perform the recursive portfolio tests were we use the four-factor alpha and four-factor t-statistic of alpha as ranking measures. The drawback of using risk-adjusted return as ranking measures, however, is that we need a sufficient number of observations to estimate the four-factor model properly. This puts a limitation on how short we can keep the evaluation period on which the funds are ranked. For this reason, we let M vary from 24 to 36 months when funds are ranked on past four-factor alpha and four-factor t-statistic of alpha. The top performing funds in the ranking period are placed in the winner portfolio (Top quintile), and the worst performing funds are placed in the loser portfolio (Bottom quintile).[45] These portfolios are then held over N months before they are rebalanced according to the new historical M month performance. We let N vary from 1 to 24 months for all ranking measures. The monthly return for each quintile portfolio is the cross-sectional average return of all funds in the specific portfolio. This procedure is repeated until the end of the sample to construct a concatenated post-ranking time series for each quintile portfolio. Funds that disappear during the evaluation period due to mergers or liquidations are included in their respective equally-weighted quintile portfolio until their last monthly return observation, and then the portfolio weights are readjusted accordingly using the remaining alive funds to avoid any potential look-ahead bias. Finally, we estimate the four-factor alpha of the quintile portfolios, where the statistical significance is evaluated using the bootstrap method described in Section 4.2.

Table VIII below reports the results when funds are ranked on the basis of their past raw returns. In general, we evaluate persistence by looking at the top (Panel A) and bottom (Panel B) quintile portfolios. In addition, we construct a hypothetical self-financing portfolio (Panel C) that is long in prior winners and short prior losers (i.e. top quintile minus the bottom quintile). By looking at Panel A in the table, we find

[45] We order the portfolios by means of quintiles in contrast to Carhart (1997) who use deciles, since the average number of funds in our sample is considerably smaller. In our sample, the top (bottom) quintile contains the twenty percent best (worst) performing funds, et cetera.

no significant evidence of persistence amongst winner funds. This implies that investors cannot look to past performance to generate positive risk-adjusted returns in the near term, i.e. we find no presence of "Hot hands" in the Norwegian mutual fund industry. This result is in line with previous research conducted with the use of monthly data (See e.g. Sørensen, 2009a; Carhart, 1997). Furthermore, we observe that when the holding period increases, winner funds generate more frequent negative risk-adjusted returns. Specifically, it can be seen from Panel A that with a 12-month ranking period and a holding period of 12 months, the top quintile portfolio yields a negative yearly alpha that is not significantly different from zero. This finding is consistent with Sørensen (2009a).

Table VIII

Performance of Top, Bottom and Spread Portfolios Formed on Lagged Returns

This table reports annualized monthly alphas based on the unconditional four-factor model of Carhart (1997) for the top quintile portfolio (Quintile 1) in Panel A, for the bottom portfolio (Quintile 5) in Panel B and for a spread portfolio long in quintile 1 and short in quintile 5 in Panel C using different lengths of ranking and holding periods. Explicitly, the respective portfolios are ranked according to past (3-36 months) raw returns, and held for different evaluation periods (1-24 months). The portfolios are then rebalanced accordingly, and the process repeated throughout the entire sample. Panel A (B) reports the top (bottom) quintile containing the top (bottom) 20% funds, whereas Panel C reports the spread between these quintiles. Columns 2 - 6 refer to holding periods of 1, 3, 6, 12, and 24, respectively. Rows 1 - 5 refer to rankings based on prior 3, 6, 12, 24, and 36 month raw returns. Significance is validated by employing the bootstrap approach as described in Section 4.2. Numbers assigned stars, ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

	Holding Period				
Ranking Period	1	3	6	12	24
Panel A: Top Quintile					
3	0.46	0.64	0.13	-1.12	-1.07
6	1.06	0.29	-0.18	-0.91	-0.91
12	0.46	0.38	0.11	-1.15	-1.24
24	-0.72	-1.42	-1.25	-0.95	-1.03
36	-0.16	-0.61	-0.14	0.06	-0.64
Panel B: Bottom Quintile					
3	-2.29**	-2.24**	-1.79*	0.45	0.51
6	-2.42**	-1.39	-1.34	-0.39	-0.14
12	-3.17***	-3.08***	-2.54**	-0.84	0.02
24	-2.19**	-1.95**	-1.69*	-1.44	-1.24
36	-1.94**	-2.07**	-1.74*	-1.50	-1.41
Panel C: Spread Portfolio					
3	2.75**	2.88**	1.92*	-1.57*	-1.58*
6	3.48***	1.68*	1.15	-0.52	-0.77
12	3.63***	3.46***	2.65**	-0.31	-1.26
24	1.47*	0.53	0.43	0.48	0.21
36	1.78*	1.46*	1.60*	1.56*	0.77

Panel B reports risk-adjusted return for loser funds (i.e. funds belonging in the bottom quintile). We find significant evidence of short-term persistence among losers. Specifically, when the funds are held for one month, we find significantly negative alphas for all ranking periods. For example, monthly rebalancing based on past 12 months raw returns yields a negative yearly alpha of -3.17% on average, which is significant at the 1% level. The magnitude of the estimated alpha drops monotonically with increasing holding periods but remain significant for holding periods up to six months. This is consistent with the "Icy hands" phenomenon of Hendricks et al. (1993), which implies that poorly performing funds in an initial period continue to be inferior performers in the near term.

Moreover, as can be observed from Panel C in the table, we find evidence of significant risk-adjusted return on the hypothetical self-financing spread portfolio.[46] Specifically, using one-month holding periods based on the funds' prior 12 month performance yields a winner-minus-loser spread of 3.63% on average, which is statistically significant at the 1% level. The alpha decreases with longer holding periods but remains statistically significantly different from zero for holding periods up to six months. For longer holding periods, the positive risk-adjusted returns turn negative and statistically indistinguishable from zero

As previously mentioned, raw returns might not be an adequate measure of real investment skill. Ranking funds on their risk-adjusted returns (i.e. alpha or t-statistic of alpha) not only controls for risk but in the case of multifactor models also controls for differences in style exposures. In addition, ranking funds based on risk-adjusted returns should make it easier to distinguish between skilled and unskilled but lucky fund managers (See e.g. Elton et al., 1996a; Bollen and Busse, 2005). All things considered, we employ the same recursive portfolio test by using both risk-adjusted returns from the four-factor model (i.e. alpha) and the corresponding alpha t-statistic as ranking measures.

Tables IX and X below report results from the recursive portfolio test when funds are ranked into quintiles based on their past alpha and t-statistic of alpha, respectively. The tables suggest no evidence of performance persistence amongst past-winner funds when funds are ranked on risk-adjusted returns, which is in line with our previous findings. Furthermore, it can be seen from Table IX (i.e. when funds are ranked on past alpha) that past losers stay losers for holding periods up to six months for both ranking periods. By looking at Panel A, we find a negative alpha of -3.24% per year with monthly rebalancing. The magnitude of the alpha estimate decreases monotonically with the length of the holding period but stay statistically significantly negative for holding periods up to six months. This result holds for both ranking periods. When funds are ranked on past t-statistics of alpha (i.e. Table X), past-losers stays losers for holding periods up to three months. For longer holding periods, the estimated alpha becomes statistically indistinguishable from zero. This result also holds for both ranking periods. The magnitude of the alpha estimates when funds are ranked on past t-statistic of alpha is analogous to the alpha estimates when funds are ranked on past alpha. Moreover, by comparing these results with the results in Table VIII above, we find that the results are mostly in line, but

[46]The portfolio is only *hypothetical* because short-selling shares in Norwegian mutual funds is not possible. It is *self-financing* because it represents a trading strategy of going long in prior winners and shorting prior losers. Results from this spread portfolio are still interesting, however, as it exposes interesting dissimilarities between extreme funds (i.e. the top and bottom performers).

the magnitude of the alpha estimates increases somewhat when ranking is based on alpha or t-statistic of alpha. Overall, these results strengthen the evidence of no dependable performance persistence among winner funds, and further supporting the presence of "Icy hands" in the Norwegian mutual fund industry.

Furthermore, Tables IX and X show that hypothetical spread portfolio yields significantly positive alphas for holding periods up to three months when funds are ranked based on their past alpha or t-statistic of alpha. This result holds for both ranking periods and is in contrast to our findings when funds are ranked on the basis of their past raw returns (i.e. Table VIII). From Table VIII, we do not find significant alphas for any of the holding periods when the evaluation period on which the funds are ranked is 24 or 36 months. As we mentioned earlier, raw returns might be a noisy measure, and does not reflect true investment skill. Thus, we consider rankings based on past alpha and t-statistic of alpha to be more reliable.

Table IX
Performance of Quintile Portfolios Formed on Lagged Alpha

This table reports annualized alphas and bootstrapped p-values from the unconditional four-factor model of Carhart (1997) for each individual quintile portfolio (1 to 5), and for the spread portfolio portfolio long in quintile 1 and short in quintile 5 using different lengths of ranking and holding periods. In Panel A (B), the respective portfolios are ranked according to past 36 (24) months alpha, and evaluated using different holding periods (1 - 24 months). The process is repeated throughout the entire sample period. Columns 1, 3, 5, 7, and 9 refer to alpha estimates with respect to 1, 3, 6, 12, and 24 months holding periods, respectively. Columns 2, 4, 6, 8 and 10 report the bootstrapped p-values of the alpha estimates.

Quintile Portfolio	1 Month Holding period α	p-value	3 Month Holding period α	p-value	6 Month Holding period α	p-value	12 Month Holding period α	p-value	24 Month Holding period α	p-value
Panel A: 36 Month Ranking Period										
Top 20%	-0.39	0.371	-0.19	0.433	-0.15	0.448	0.03	0.492	0.05	0.487
2nd	0.80	0.236	-0.14	0.460	-0.94	0.211	-0.98	0.200	-1.46	0.112
3rd	0.51	0.304	0.63	0.272	0.48	0.332	-0.01	0.505	0.45	0.317
4th	-1.82	0.083	-2.30	0.051	-1.95	0.082	-1.30	0.171	-0.73	0.307
Bottom 20%	-3.24	0.002	-2.33	0.020	-1.93	0.047	-1.30	0.172	-1.05	0.234
Spread	2.85	0.016	2.13	0.025	1.78	0.066	1.33	0.160	1.10	0.215
Panel B: 24 Month Ranking Period										
Top 20%	0.16	0.441	-0.61	0.278	-0.42	0.364	-0.35	0.392	-0.43	0.349
2nd	-0.27	0.397	0.20	0.418	0.33	0.385	-1.18	0.127	0.23	0.409
3rd	-0.19	0.421	0.41	0.362	-0.51	0.309	-0.27	0.394	-0.82	0.219
4th	-1.34	0.168	-1.37	0.167	-1.75	0.103	-0.63	0.279	-1.52	0.134
Bottom 20%	-2.24	0.022	-2.18	0.029	-2.02	0.041	-1.55	0.124	-0.80	0.269
Spread	2.40	0.019	2.07	0.034	1.33	0.137	1.20	0.178	0.37	0.357

Taken together, with the exception of the hypothetical spread portfolio, we get rather consistent results by applying different ranking measures. Based on results from the recursive portfolio formation tests, we conclude that there exists no persistence in performance among recent top performers, even when portfolio rebalancing is frequent. This implies that investors cannot identify superior performance ex-post and use this information to earn risk-adjusted returns. In contrary, we find that performance among the worst performing funds strongly persist for short time-periods, indicating the presence of "Icy hands" in the universe of Norwegian mutual funds. Furthermore, our results suggest that it is possible to create a trading strategy that entails buying past winners and shorting past losers based on ex-post performance information, which yields positive risk-adjusted returns. However, for the average investor, shorting mutual funds seems like an impossible task. In addition, this strategy requires frequent portfolio rebalancing, which in a realistic scenario will impose substantial transaction costs. Thus, the economic implication of this finding is questionable. Overall, with the exception of winner funds, our results are consistent with Bollen and Busse (2005) who conclude that post-ranking abnormal returns largely disappears when funds are evaluated over longer periods, suggesting that persistence is a short-lived phenomenon. In the next section, we conduct non-parametric two-period tests to assess whether there are funds in our sample that perform consistently better or worse relative to other funds in our sample.

Table X

Performance of Quintile Portfolios Formed on Lagged t-stat of Alpha

This table reports annualized alphas and bootstrapped p-values from the unconditional four-factor model of Carhart (1997) for each individual quintile portfolio (1 to 5), and for the spread portfolio portfolio long in quintile 1 and short in quintile 5 using different lengths of ranking and holding periods. In Panel A (B), the respective portfolios are ranked according to past 36 (24) months alpha t-statistic, and evaluated using different holding periods (1 - 24 months). The process is repeated throughout the entire sample period. Columns 1, 3, 5, 7, and 9 refer to alpha estimates with respect to 1, 3, 6, 12, and 24 months holding periods, respectively. Columns 2, 4, 6, 8 and 10 report the bootstrapped p-values of the alpha estimates.

	1 Month		3 Month		6 Month		12 Month		24 Month	
Quintile	Holding period		Holding period		Holding period		Holding period		Holding period	
Portfolio	α	p-value	α	p-value	α	p-value	α	p-value	α	p-value
Panel A: 36 Month Ranking Period										
Top 20%	0.07	0.476	0.04	0.483	0.06	0.482	0.00	0.498	-0.26	0.404
2nd	0.23	0.416	-0.33	0.390	-0.18	0.436	-0.45	0.337	-0.44	0.355
3rd	0.20	0.435	0.32	0.393	-0.31	0.393	-0.35	0.382	-0.07	0.480
4th	-0.99	0.203	-1.82	0.061	-2.19	0.041	-0.92	0.221	-1.17	0.176
Bottom 20%	-3.52	0.001	-2.77	0.011	-1.81	0.066	-1.72	0.068	-1.69	0.072
Spread	3.59	0.004	2.80	0.009	1.87	0.059	2.05	0.052	1.43	0.124
Panel B: 24 Month Ranking Period										
Top 20%	0.20	0.431	-0.31	0.382	-0.30	0.389	-0.09	0.464	0.03	0.479
2nd	-0.27	0.413	0.38	0.357	0.48	0.361	0.32	0.377	-0.96	0.191
3rd	-0.16	0.436	-0.71	0.281	-0.81	0.272	-1.16	0.158	-1.47	0.110
4th	-1.62	0.095	-1.52	0.101	-0.76	0.294	-1.09	0.216	-2.32	0.024
Bottom 20%	-2.27	0.031	-2.12	0.037	-1.28	0.153	-0.38	0.368	0.67	0.285
Spread	2.07	0.039	1.91	0.049	0.98	0.189	0.30	0.378	-0.63	0.264

5.2 Non-Parametric Two-Period Tests

Despite that we did not find any positive risk-adjusted performance persistence among top performing funds, we cannot conclude that there is no persistence among these funds. Even though these funds delivered negative risk-adjusted returns, they might exhibit superior performance relative to the sample. In this section, we turn to non-parametric two-period tests that allow us to investigate persistence relative to the average mutual fund industry (i.e. whether some funds perform consistently better or worse compared to other funds in the sample).

5.2.1 Consistency in Ranking

Following Carhart (1997), we investigate the consistency in fund ranking by constructing contingency tables of initial and subsequent fund rankings. Specifically, we rank funds based on their past raw return and place them in quintiles over an initial and subsequent time window. Then, we count the number of times a fund ends up in one of the quintiles in the subsequent period, conditional on its ranking over the initial period. Funds that disappear during the subsequent period are placed in a separate category for dead funds. Consistency in rankings is then evaluated by looking at the conditional probability of achieving a subsequent ranking of quintile j (or dying) given its initial ranking of i. We construct contingency tables for four different time windows of 3, 6, 12, and 24 months, respectively. The contingency tables are illustrated in Figure 5.

From the figure, it becomes apparent that losers are more likely to remain losers. This can be seen from the relatively large contingent probabilities that the worst performing (bottom quintile) funds will remain in the bottom quintile. This is in line with our previous findings, thus further supporting the premise of persistence among loser funds and the presence of “Icy hands” among active Norwegian mutual funds. In addition, it can be observed from Figure 5 that the probability of disappearing from the database (i.e. to be placed in the bar for dead funds) is highest for funds initially placed in the bottom quintile, and it decreases almost monotonically with initial ranking. This is best illustrated in Panels A and B of the figure (i.e. time windows of 24 and 12 months).

Further, from Panel D (i.e. time windows of three months), it is evident that prior winners are more likely to become next period’s winners. Interestingly, it can be seen from the figure that the contingent probability for winners to remain winners decreases monotonically with the time window. Thus, the figure provides evidence of short-term persistence among winner funds before adjusting for risk. Furthermore, it becomes evident from the figure that last periods winners frequently become next period’s losers and vice versa. This is consistent with the typical gambling behavior among mutual funds (Carhart, 1997).

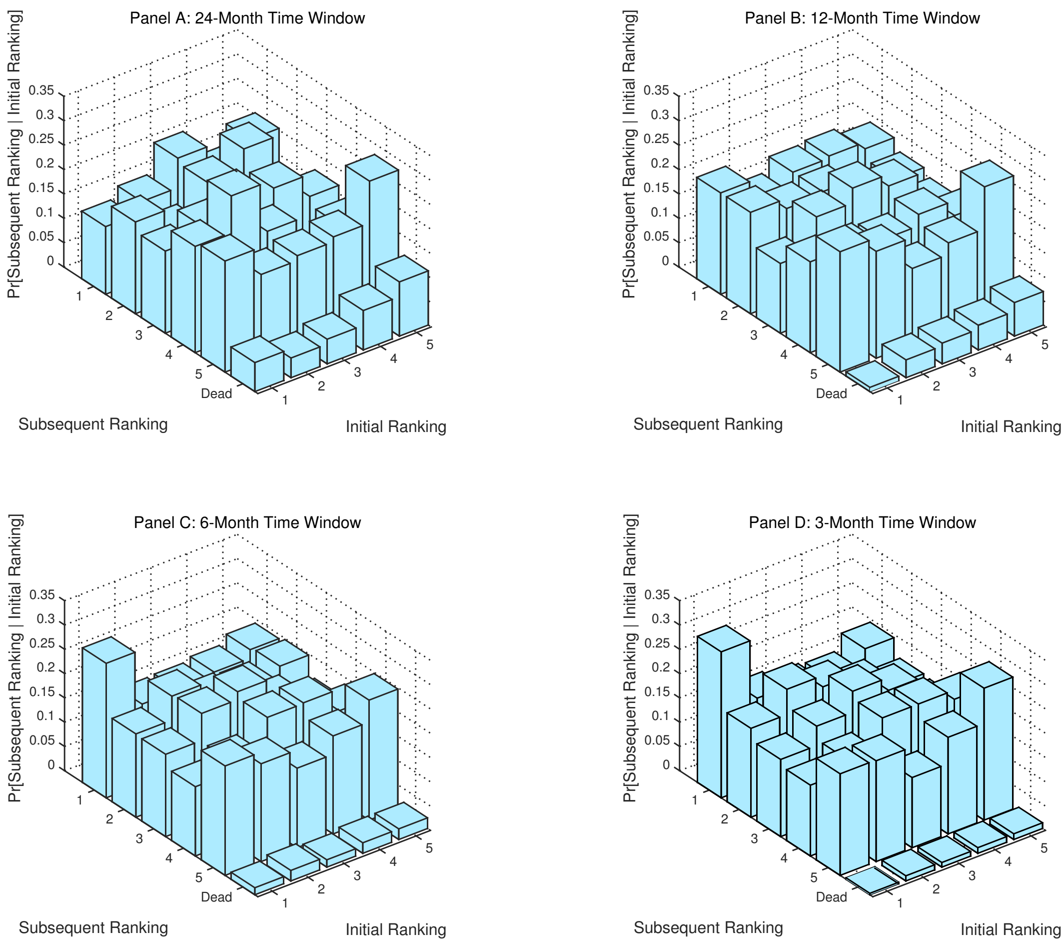

Figure 5. Contingency tables over two subsequent intervals. The figure depicts consistency in rankings over two consecutive periods; initial and subsequent rankings. Panel A - D has different length of time windows, that is, 24, 12, 6 and 3 months respectively. The top (worst) performing funds are placed in Quintile 1 (5). Funds that have died during the sample period are placed in a separate bar for dead funds. The bars in the figure represent the conditional probability of achieving a subsequent rank j (or dying) given its initial ranking i. The sample period is 1987 to 2014.

5.2.2 The Cross-Product Ratio

The contingency tables in Figure 5 provide a visual representation of persistence in our sample, but do not inform us about the statistical significance of the results. In this section, we perform a cross-product ratio (CRP) test to evaluate whether there exist persistence in our sample. In contrary to the first non-parametric test, the CRP test quantifies the statistical significance of persistence. Following Brown and Goetzmann (1995), we categorize funds as winners (W) if the funds' return for a given period is greater than the median in that same period. Conversely, a loser fund (L) is labeled if the funds' performance is lower than the median. Funds that are categorized as winners (losers) over an initial period and subsequent period are denoted WW (LL). If a fund is categorized as a winner (loser) in the first period and as a loser (winner) in the second period, the fund is denoted WL (LW). The cross-product ratio captures the odds ratio of the funds that show persistence in performance relative to other funds in the sample that do not. That is, the number of repeated performers to the number of those that do not repeat. The cross-product ratio (CPR) is defined as follows:

$$\text{CPR} = \frac{WW \cdot LL}{WL \cdot LW} \tag{6}$$

The null hypothesis in the cross-product ratio test is that there is no persistence, i.e. that WW, LL, WL, and LW has equal probability. That is, the performance in the initial period is unrelated to the performance in the subsequent period, which corresponds to a CPR of one. The statistical significance of the CPR can be evaluated using the standard error of the natural logarithm of the CPR (See e.g. Christensen, 1990). The resulting Z-statistic is the ratio of the natural logarithm of the CPR to the standard error of the natural logarithm:

$$Z = \frac{\ln CPR}{\sigma_{\ln CPR}} = \frac{\ln CPR}{\sqrt{\frac{1}{WW} + \frac{1}{WL} + \frac{1}{LW} + \frac{1}{LL}}} \tag{7}$$

We also use the Chi-square statistic to evaluate the statistical significance of our results as a robustness check. Carpenter and Lynch (1999) argue that the Chi-square test based on the number of winner and loser funds is well specified and powerful, and that it specifically proves to be more robust to the presence of survivorship bias. We compute the Chi-square statistic similar to Agarwal and Naik (2000):

$$\chi^2 = \frac{(WW - D1)^2}{D1} + \frac{(WL - D2)^2}{D2} + \frac{(LW - D3)^2}{D3} + \frac{(LL - D4)^2}{D4}, \tag{8}$$

where

$$D1 = (WW + WL) \cdot (WW + LW)/N, \quad D2 = (WW + WL) \cdot (WL + LL)/N$$
$$D3 = (LW + LL) \cdot (WW + LW)/N, \quad D4 = (LW + LL) \cdot (WL + LL)/N,$$

and N denotes the total number of observations. Table XI below reports the CPR and the corresponding z-statistic in addition to outputs from the Chi-square test for various time periods. From the table, we find statistically significant evidence of performance persistence in line with our previous findings for time windows up to six months. It should be be noted that the CPR is also above one for time windows of 12 months, but this result is not statistically distinguishable from zero. Furthermore, the CPR below one when using time windows of 24 months point toward reversed performance persistence, i.e. prior winners are expected to lose in the following period. Similar to the results from the contingency tables in the previous section, we find that the statistical significance of performance persistence decreases as the time windows increases. Overall, this sensitivity to the return measurement interval strengthens the evidence of short-term persistence among Norwegian mutual funds.

Table XI
Persistence Results from the Cross-Product Ratio Test

The table reports the cross-product ratio (CPR), the corresponding Z-statistic and Chi-square statistic. Row 1 reports the CPR, whereas Row 2 (3) reports the Z-statistic (Chi-square statistic). Columns 1-4 refer to different return measurement periods with time windows of 3, 6, 12, and 24 months, respectively. Numbers assigned stars, ***, **, and * indicates significance at the 1%, 5%, and 10% levels, respectively. The sample period is 1987 to 2014.

	3 Month Time Window	6 Month Time Window	12 Month Time Window	24 Month Time Window
CPR	1.36	1.31	1.09	0.80
Z-statistic	5.33***	3.20***	0.68	-1.21
χ^2-statistic	28.51***	10.23***	0.47	1.46

To conclude our analysis on persistence, we find no evidence of dependable performance persistence when adjusting for risk among top-performing active mutual funds in the Norwegian market. Our results imply that investors cannot exploit past performance to earn positive risk-adjusted returns, a result which coincide with the semi-strong form of the EHM. On the other hand, we find significant short-term persistence among loser funds. Interestingly, this is the most consistent conclusion that can be found in all major studies on performance persistence, and indicate that following a contrarian investment strategy of buying last year's losers to gain positive risk-adjusted returns is a long shot. The results are robust for various ranking and holding periods, and regardless of whether the funds are ranked on the basis of past raw return, four-factor alpha or four-factor t-statistic of alpha. In the non-parametric two-period tests, we find further support for the premise of persistence among loser funds. However, these tests also reveal persistence among prior winners in subsequent periods for short time windows. This effect diminishes as the time window increases. The fact that we only find persistence for short time horizons is consistent with Bollen and Busse (2005) who argue that persistence is, if any, a short-lived phenomenon.

6 Asymmetries in Norwegian Mutual Fund Performance

So far, based on the analyses in previous sections, we have found no empirical evidence of abnormal performance in Norwegian equity mutual fund returns. Kosowski (2011) argue that the premium investors are willing to pay for an asset is driven by the covariance of the assets payoff with marginal utility. Moreover, he suggests that investors more willingly pay premiums for assets that are negatively correlated with consumption. When we have economic contractions in the business cycle (i.e. recessionary states in the economy), consumption tends to be particularly low (See e.g. Breeden, 1979; Rubinstein, 1976; Grossman and Shiller, 1981). The practical implications to investors based on Kosowski's (2011) arguments, is that if a given asset performs poorly in recessions when the investor is poor, the asset is less attractive than assets that perform well in non-recessions when consumption is high (and the investor feels wealthy).

In general, there is mostly good news about firms during economic upturns. In line with the semi-strong form of market efficiency, most news about companies is revealed accordingly during non-recession periods, and information about companies is relatively symmetrically distributed in the market. As a consequence, the variance of information signals tends to be low throughout non-recessionary states in the economy (Kosowski, 2011). During economic downturns when there is more bad news, however, corporate managers tend to withhold information. This leads to asymmetric information and thus higher variance of information signals in the market. This hypothesis is supported by Kothari et al. (2009), who suggest that corporate managers withhold bad news, but immediately leak and reveal positive news. In other words, there should be reason for fund managers to possibly be better informed in recessions, because of the higher variance of information signals. Thus, under these assumptions, there is more potential to possibly outperform passive benchmarks during recessionary states. The question then arises, whether mutual funds in fact are able to perform particularly well during such recessionary states in the economy when the marginal utility of wealth is high, and performance really matters.

In the proceeding sections, we aim to answer a hitherto unanswered question in the Norwegian fund sector. That is, if Norwegian equity mutual funds perform well in recessionary states in the economy when performance matters the most to investors. To answer this hypothesis, we explicitly examine aggregate fund performance in different states of the business cycle using a binary classification of recession and non-recessionary dates in the Norwegian economy. Additionally, we apply a novel conditional performance measurement methodology based on a Markov Regime-Switching Model following Kosowski (2011), where we let the data determine the indicator of the states in the business cycle.

The remainder of this section proceeds as follows. First, we explicitly investigate aggregate fund performance using sub-periods based on a binary classification of Norwegian recession and non-recession periods. Then, we discuss advantages of the Markov regime-switching Model and describe the theoretical foundation of regime-switching performance measures. Finally, we integrate the regime-switching framework in our analysis on asymmetries in Norwegian mutual fund performance.

6.1 Norwegian Business Cycle Reference Dates

Following Burns and Mitchell (1946), we define business cycles as fluctuations in aggregate economic activity. This is the classic way to define business cycles, at which peaks and troughs describe developments in the level of economic activity across different sectors in the economy. Contractions (i.e. recessions) start at the peak of a business cycle, and end at the through. The periods from through to peak are considered to be non-recessions, i.e. periods in time with up-and-coming growth. Alternatively, one could refer to business cycles as “growth cycles”, where the economic fluctuations in the cycle is defined relative to trend growth. That is, “high” or “low” growth characterize the business cycle to be above (non-recession) or below (recession) the estimated trend. The classical business cycle approach is advantageous since there is no need to model unobserved trend growth.

The most widely used reference point in business cycle literature is the National Bureau of Economic Research (NBER). Based on available data (i.e. not normally revised), The Business Cycle Dating Committee of NBER decides when a turning point occurs, and provides specific dates (monthly) on when US business cycle recessions starts and ends. Since these decisions are made after much deliberation on the basis of available data, the announcements of the turning points are often made as late as one year after the actual turning point took place.

Most of the previous US literature that analyses asymmetries in fund returns use NBER dates. These dates, however, are not representable for the Norwegian economy (See e.g. Christoffersen, 2000). There exist several studies on Norwegian business cycles, but these papers conduct analyses primarily on growth cycles (See e.g. Eika and Lindquist, 1997; Bjørnland, 2000; Bjørnland et al., 2008). Since there is no official dating of classic business cycle dates in Norway, we turn to the only two papers to our knowledge that date Norwegian business cycle turning points by using classic programmed approaches similar to NBER.

Several research papers have been able to almost identically replicate the NBER recession dates by using programmed approaches. Specifically, these papers use the Bry-Boschan algorithm for programmed determination of turning points (See e.g. Bry and Boschan, 1971; Artis et al., 1995). Christoffersen (2000) use the same approach following Bry and Boschan (1971), and define turning points to four Nordic countries; Denmark, Finland, Norway, and Sweden. The study of Christoffersen (2000), however, use unrevised data on industrial production for each country. In Norges Bank Occasional Papers, Aastveit et al. (2014) conducts research that considers *revised* data. The employ a similar programmed approach as Bry and Boschan (1971) to date and forecast business cycle reference dates in Norway on the basis of Norwegian mainland GDP. It is important to note that GDP growth rates typically fluctuate around a higher level and are more persistent during non-recession periods. They stay at a relatively low lever, however, and are less persistent during recessions. For such data, it would not be reasonable to expect a single, linear model to capture these distinct behaviors, thus a programmed approach is more preferable. We consider the study of Aastveit et al. (2014) to be the most consistent and similar to the dates published by NBER, and apply these dates in our empirical analysis in the following sections as basis for comparison. Areas shaded in grey in Figure 6 below present dates at which recessions occur in our sample period from 1983 to 2014.

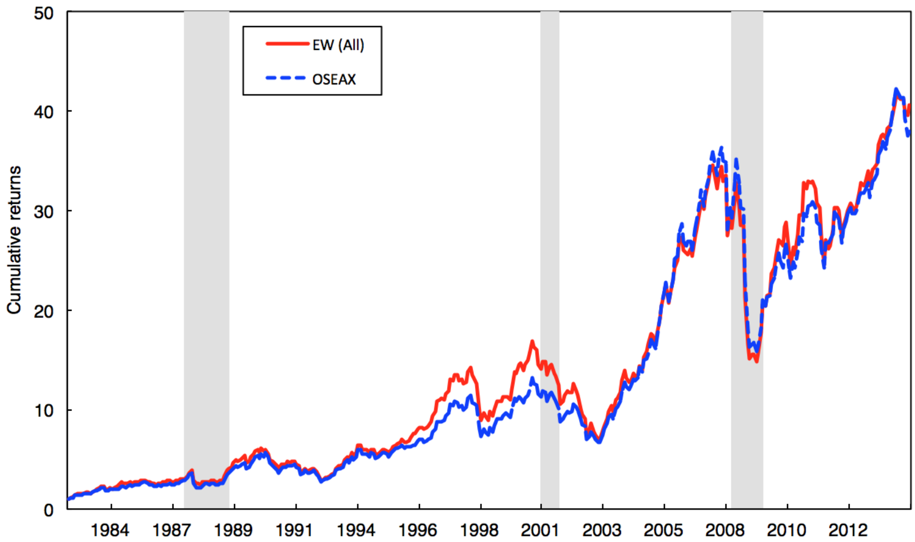

Figure 6. Norwegian Business Cycle Reference Dates. The figure plots cumulative returns for the equally weighted portfolio of all funds and the OSE All Share index for recessionary and non-recessionary periods in Norway. The Norwegian business cycle reference dates are obtained from Aastveit et al. (2014), and are represented by the areas in gray. The solid line plots mean returns for the equally weighted portfolio (i.e. returns of all funds that existed at some point over the sample period), whereas the dotted line plots mean returns for the OSEAX.

6.2 Empirical Results of Fund Performance in Recession and Non-Recession Sub-Periods

Based on the Norwegian business cycle reference dates defined in the previous section, we construct two separate sub-samples for recession and non-recession periods. Before we explicitly evaluate risk-adjusted fund performance in recession and non-recessions, we provide some summary statistics for the two sub-samples.

6.2.1 Summary Statistics in Recession and Non-Recession Periods

Table XII below reports summary statistics for the equally weighted portfolio excess return and equity benchmark returns for the whole period as well as recession and non-recession sub-periods. Panel A reports means and standard deviations, whereas Panel B reports skewness and kurtosis. From Panel A, we see that the mean return on the fund portfolio and the market proxy (MKT) are statistically significantly higher in non-recession periods than in recession periods. Furthermore, Panel A shows that the size (SMB) and momentum (PR1YR) factor perform better in non-recessions, whereas the book-to-market (HML) factor-mimicking portfolio benchmark performs best in recessions.

Table XII
Descriptive Statistics of Benchmarks and Mutual Fund Returns in Recessions and Non-Recession Periods

The table provides summary statistics for benchmark returns (i.e. the four risk factors from Carhart's (1997) model), and excess returns for the equally weighted portfolio comprising all funds in our sample. Panel A displays first and second moments, i.e. mean and standard deviation. Panel B reports higher order moments, i.e. skewness and kurtosis. In all Panels, values are reported for both recessions, non-recessions, and the full sample, respectively. Dates on recessions and non-recessions are obtained from Aastveit et al. (2014). Means and standard deviations are annualized and reported in percent, whereas the remaining statistics are reported monthly. '***', '***' and '*' represent 1%, 5% and 10% significance of the difference between recessionary and non-recessionary means. The sample period is 1983 to 2014.

Panel A: First and Second Moments						
	Mean			Standard Deviation		
	Full Sample	Recessions	Non-Recessions	Full Sample	Recessions	Non-Recessions
EW(All)	7.06	-11.38**	8.97	21.90	33.94	20.24
MKT	6.94	-11.57**	8.86	21.80	35.59	19.83
SMB	9.81	1.65	10.66	15.32	18.91	14.91
HML	4.41	16.83*	3.12	17.20	17.61	17.13
PR1YR	9.03	5.77	9.37	16.94	19.97	16.63
Panel B: Higher Order Moments						
	Skewness			Kurtosis		
	Full Sample	Recessions	Non-Recessions	Full Sample	Recessions	Non-Recessions
EW(All)	-0.81	-0.94	-0.56	4.92	3.75	4.00
MKT	-0.92	-1.04	-0.58	5.31	3.71	4.06
SMB	0.47	0.57	0.48	6.31	2.83	7.06
HML	-0.13	-0.44	-0.11	4.17	2.84	4.36
PR1YR	-0.40	-1.05	-0.29	4.32	4.90	4.10

Figure 7 below plots the annualized mean excess return of the equally weighted fund portfolio and the equity benchmarks over the recessions and non-recession periods in our sample. Panel A of the figure shows that the fund portfolio reach a through in each recession period, where the latest recession (the Global Financial Crisis). From Panel B and C, we observe that the market proxy and the momentum factor follow the same pattern as the fund portfolio. Also the size-mimicking portfolio follows a similar cycle, except for performing particularly well in the 2001 recession. Conversely, the mean return of the book-to-market mimicking portfolio peaks in all recessions period. This finding is in line with Lakonishok et al. (1994) who find that value strategies outperform growth strategies in bad states of the economy. Overall, these results are consistent with Kosowski (2011) who find similar patterns on mean returns of the equity benchmarks in Carhart's (1997) four-factor model in NBER recession and non-recession sub-periods. Based on this findings, we will in the next section, assess how active Norwegian mutual funds change their exposure to the equity benchmarks between recession and non-recession periods.

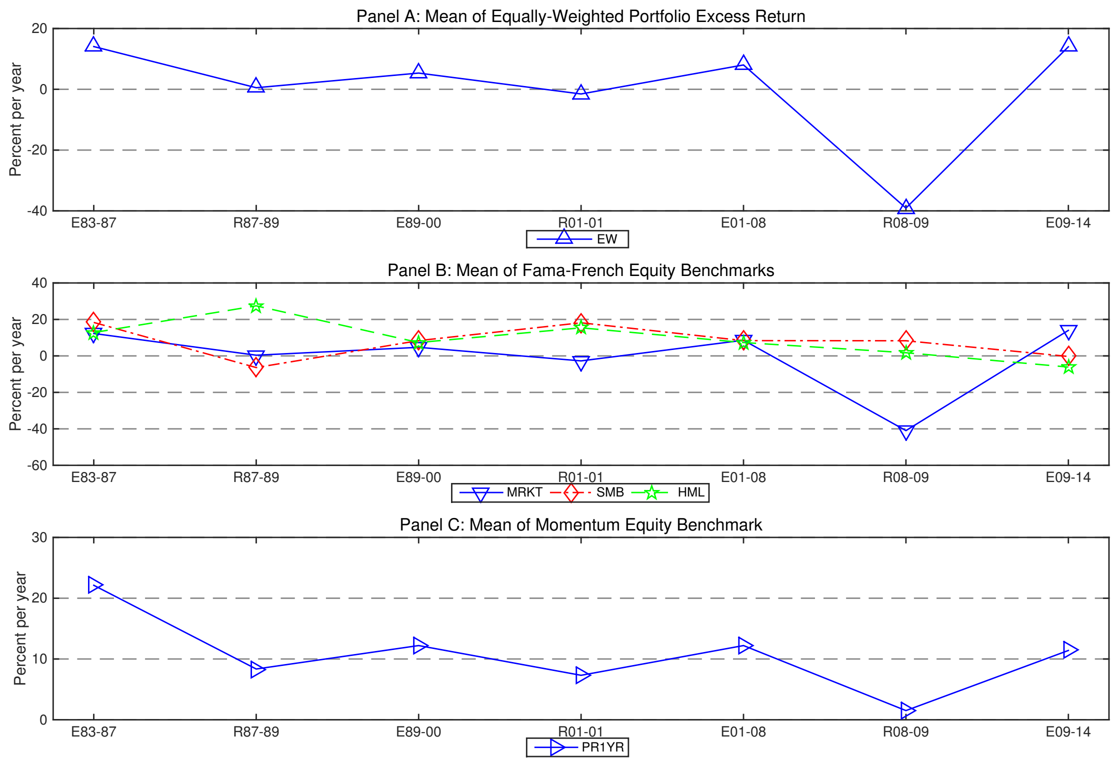

Figure 7. Mean excess returns for Norwegian mutual funds and equity benchmarks during recessions and non-recession periods. The figure plots annualized mean excess return for the equally weighted portfolio of all funds and equity benchmarks for recessionary and non-recessionary periods during 1983 - 2014. Panel A reports mean excess return for the equally weighted portfolio. Panel B shows the three risk factors in Fama-French's (1993) model, MKT, SMB, and HML. Panel C reports the one-year return momentum factor of Jegadeesh and Titman (1993). The x-axis displays recessions (R) and non-recession (E) periods, whereas the y-axis refer to the excess returns in percent per year.

In addition, it is evident from Panel A in Table XII that the standard deviations of the fund portfolio and the market proxy are substantially higher in recession periods. The volatility of the other equity benchmarks is fairly stable across recession and non-recession periods, but is slightly higher in recessions for all equity benchmarks. Figure 8 below plots the annualized standard deviation over recessions and non-recession periods and confirm these findings. From Panel A and B, it is apparent that both mutual fund and market return volatility peaks in every recession in our sample. The volatility of the other benchmarks is fairly stable around 20% per year in both recession and non-recessions. Furthermore, Panel B of Table XII reports skewness and kurtosis, and we observe only minor differences between recessions and non-recession periods. On average, all of the return series except from the SMB are negatively skewed, and they exhibit relatively high kurtosis. These results apply to the full sample, as well as for recessionary and non-recessionary periods.

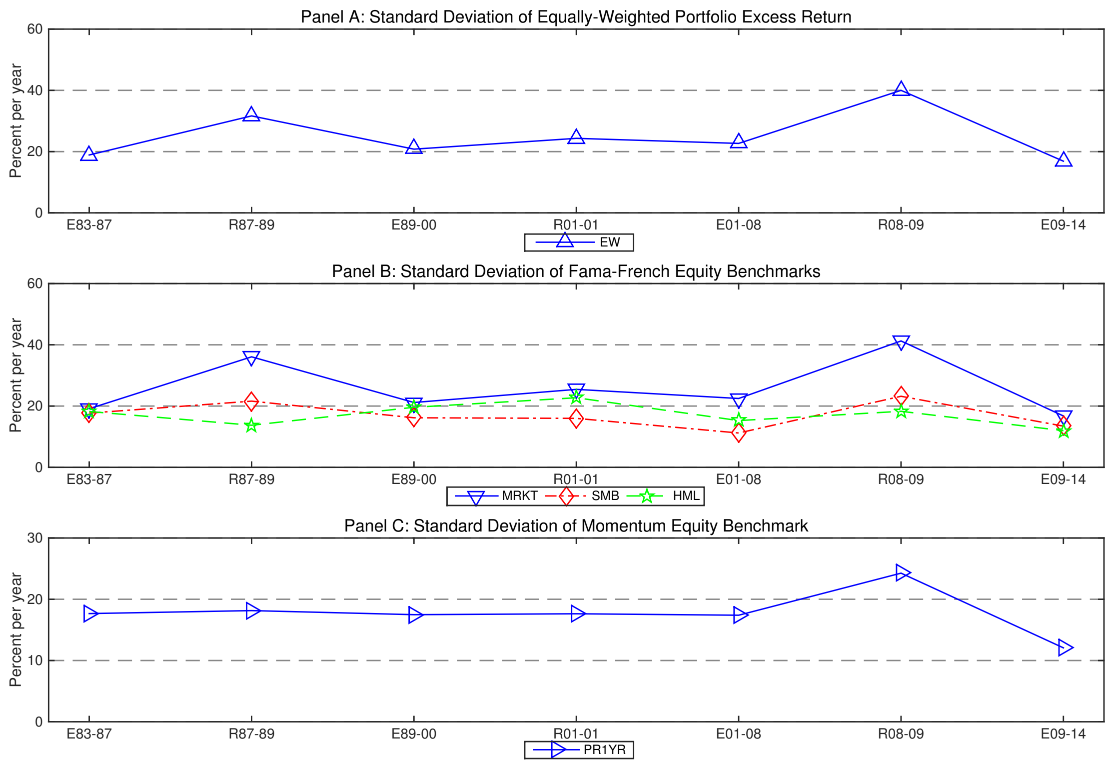

Figure 8. Standard deviation of excess returns for Norwegian mutual funds and equity benchmarks during recessions and non-recession periods. The figure plots annualized standard deviations of excess returns for the equally weighted portfolio of all funds and equity benchmarks for recessionary and non-recessionary periods during 1983 - 2014. Panel A reports standard deviation of excess return for the equally weighted portfolio. Panel B shows the three risk factors in Fama-French's (1993) model, MKT, SMB, and HML. Panel C reports the one-year return momentum factor of Jegadeesh and Titman (1993). The x-axis displays recessions (R) and non-recessions (E) periods, whereas the y-axis refer to the standard deviations in percent per year.

6.2.2 Risk-Adjusted Performance in Recession and Non-Recession Periods

In this section, we investigate the performance of Norwegian funds in Norwegian recession and non-recession sub-periods. Risk-adjusted Carhart (1997) four-factor alphas and beta loadings for the equally weighted fund portfolio in recession and non-recession sub-periods as well as the full sample period is reported in Table XIII below. The table shows that Norwegian mutual funds, on average, perform better in recessions compared to non-recession periods. Actively managed Norwegian mutual funds, on average, deliver a positive yearly alpha of 0.26% in recessions and a negative yearly alpha of -0.87% in non-recession periods. However, none of the estimated alphas are statistically significant. The adjusted R-squared is high in recession periods (95%), implying that the model explains most of the variation in returns. Hence, making it unlikely that a missing factor is causing the higher alpha in recession periods.

Table XIII
Risk-Adjusted Performance in Recession and Non-Recession Periods

This table reports risk-adjusted mutual fund performance based on the unconditional Carhart (1997) four-factor model for the full sample period (1983 to 2014), and recession and non-recession sub-periods. Dates on the recessionary and non-recessionary sub-periods are obtained from Aastveit et al. (2014). The alpha (α) is based on excess returns of the equally weighted portfolio comprising all funds in our sample. The table also reports regression coefficients (β's) for the factors in Carhart's (1997) model, MKT, SMB, HML and PR1YR, respectively. Columns 1, 3, and 5 report the regression coefficients, whilst Columns 2, 4, and 6, report their respective t-statistics (numbers in parentheses). The two bottom rows refer to the annualized residual SD and adjusted R^2. Numbers assigned stars, ***, **, and * indicates significance at the 1%, 5%, and 10% levels, respectively.

	Full Sample		Recessions		Non-Recessions	
α	-0.428%	(-0.30)	0.262%	(0.06)	-0.870%	(-0.79)
MKT	0.995***	(65.99)	0.882***	(16.31)	1.011***	(62.60)
SMB	0.101***	(4.67)	-0.056	(-0.54)	0.111***	(5.11)
HML	-0.050***	(-2.85)	-0.023	(-0.32)	-0.050***	(-2.98)
PR1YR	-0.029*	(-1.62)	-0.163**	(-2.44)	-0.011	(-0.59)
$\sigma^{\varepsilon}_{annual.}$	5.82		7.15		5.54	
$R^2_{Adj.}$	0.93		0.95		0.92	

An examination of the beta loadings in recession and non-recession periods reveals that actively managed Norwegian mutual funds, on average, reduce their exposure to the market-, size- and momentum factors in recession periods, while they increase their exposure (i.e. less negative) to the value factor in recession periods. As we observed from Table XII, the market-, size-, and momentum benchmarks perform best in non-recession periods, whereas the book-to-market mimicking portfolio performs better in recession periods. Thus, the changes in beta loadings on the equity benchmarks between recession and non-recession periods can be interpreted as evidence of positive market timing regarding all benchmark factors. However, portfolio betas can change as a result of changes in the betas of the underlying assets comprising the portfolio. Betas tend to change during economic downturns. Thus, our evidence of positive market timing must be interpreted with caution. Furthermore, fund managers appear to stay away from high-risk small firms by reducing their exposure to the size factor in high volatility states. This result is consistent with a *flight to quality* tendency typically caused by uncertainty in the financial markets.[47]

Table XIII also reports the standard deviation of the residuals in recession and non-recession periods. We see that the residual standard deviation is higher in recession periods than in non-recession periods. This is important evidence that would be used to identify the recessionary and non-recessionary states in the Markov regime-switching models in proceeding sections. As we will show later, one of the characteristics of the recession state in the Markov regime-switching models is that it is always the state with the highest volatility.

6.3 The Markov Regime-Switching Model

The Markov Regime-Switching model was first introduced by Hamilton (1989) and was developed as a tool for dealing with infrequent, but recurrent and endogenous shifts between regimes (or states) in time series data. One of the advantages of the model is that it allows for the involvement of multiple equations that characterize time-series behaviors in different regimes, and is permitting switching between these equations. This enables the model to capture more complex dynamic patterns. The switching mechanism between the equations (or regimes) is controlled by an unobservable state (latent-state) variable that is assumed to follow a first-order Markov chain.

Regime-switching models have several advantages over alternative techniques to capture structural breaks in data because the parameters in the model are not conditioned on pre-defined state indicators.[48] Instead, the model allows for a continuous time interval with state transition probabilities that contains information on the direction of variations on mutual fund returns (Kosowski, 2011). Binary classifications also fail to capture periods of economic growth slowdowns in the economy (Chauvet and Potter, 2000), which can be properly modelled in a Markov regime-switching model. Additionally, the binary classifications might have its own measurement issues, and only become available "after the fact" (Kosowski, 2011). Even more complex models, for example Kalman filter models, are not able to estimate mutual fund performance in recessions and non-recession periods directly, without being combined with binary indicators.

[47] A *flight to quality* is the act of moving capital away from "risky" investments and toward "safer" investments due to uncertainty about the overall economy.
[48] Such as dummy variables.

6.3.1 The Regime-Switching Framework

In the following sections, we discuss the Markov regime-switching model that forms the basis of our empirical analysis. We follow Perez-Quiros and Timmermann (2001) and Kosowski (2011), who use the original specification purposed by Hamilton (1989) but generalize the model to let the intercept term, regression coefficients and the variance term to be regime dependent. Specifically, we allow for two possible regimes, a recession and a non-recession regime, in the Carhart (1997) unconditional four-factor model, and let the identity of these regimes be determined by the data. This allows us to explicitly control for state dependencies in fund's alpha and beta risk and link fund performance to recession and non-recession states. The regime-switching framework combines a set of equations into one system. The system allows the regression coefficients in each equation to be state dependent. For the unconditional four-factor model of Carhart (1997), the system can be specified as follows:

$$r_{i,t} - r_{f,t} = \alpha_{i,S_t} + \beta_{1i,S_t} \cdot MKT_t + \beta_{2i,S_t} \cdot SMB_t + \beta_{3i,S_t} \cdot HML_t + \beta_{4i,S_t} \cdot PR1YR_t + \varepsilon_{i,t} \quad (9)$$

where S_t is a latent state variable that can take on two possible values (i.e. $S_t = 1, 2$). The state variable S_t is unobserved and is assumed to follow a first order Markov process so that:

$$\Pr(S_t = i | S_{t-1} = j) = \pi_{i,j}, \qquad i, j = 1, 2. \quad (10)$$

The value $\pi_{i,j}$ is the *transition probability* of moving to state i from state j. These transition probabilities are usually assumed to stay constant over time, but it is also possible to allow them to be time-varying (i.e. depend upon some predetermined variables). In the model, α_{i,S_t} and β'_{i,S_t} represent the mutual funds' performance and beta risk, respectively. The alphas and betas are discrete random variables that can take two values, and is dependent on the state of the economy. Since fund managers may have different information sets in recessionary and non-recessionary states, alphas and betas may be different in the two states.

6.3.2 Transition Probabilities

Previous empirical evidence with Markov regime-switching models argue that the flexibility gained by allowing the state transition to vary over time as a function of a vector of predetermined variables, z_{t-j}, can be very substantial (See e.g. Filardo, 1994; Diebold et al., 1994; Durland and McCurdy, 1994; Gray, 1996; Perez-Quiros and Timmermann, 2000). Following Kosowski (2011), we thus establish our regime-switching model by using time-varying transition probabilities (TVTP), and make the usual assumption that the state transition probabilities follow a first-order Markov chain:

$$p_t = \Pr(S_t = 1|S_{t-1} = 1, z_{t-j}) = p(z_{t-j}) \tag{11}$$

$$1 - p_t = \Pr(S_t = 2|S_{t-1} = 1, z_{t-j}) = 1 - p(z_{t-j}) \tag{12}$$

$$q_t = \Pr(S_t = 2|S_{t-1} = 2, z_{t-j}) = q(z_{t-j}) \tag{13}$$

$$1 - q_t = \Pr(S_t = 1|S_{t-1} = 2, z_{t-j}) = 1 - q(z_{t-j}) \tag{14}$$

The candidate series for the predetermined information variables are those considered to be useful to predict business cycles. We follow Filardo (1994), Perez-Quiros and Timmermann (2000), and Kosowski (2011) who use the Composite Leading Indicator as the variable for the state transition in the model. The CLI measure is designed to provide qualitative information on short-term economic movements and as such represent a forward-looking indicator about the future evolution of the economy.[49] The logistic index function for the state transitions can be expressed as follows:

$$p_t = \Pr(S_t = 1|S_t = 1) = \phi(c_1 + d_1(\Delta CLI_{t-2})) \tag{15}$$

$$q_t = \Pr(S_t = 2|S_t = 2) = \phi(c_2 + d_2(\Delta CLI_{t-2})) \tag{16}$$

where $\phi(\cdot)$ is the cumulative density function of a standard normal variable, ΔCLI_{t-2} is the two month lagged change in the composite leading indicator and c_s is a constant. For completeness, we also estimate the Markov regime-switching model with constant transition probabilities, and briefly comment on the differences. The logistic index function for the model with constant transitions only contains a constant:

$$p_t = \Pr(S_t = 1|S_t = 1) = \phi(c_1) \tag{17}$$

$$q_t = \Pr(S_t = 2|S_t = 2) = \phi(c_2) \tag{18}$$

The transition probabilities together with the other parameters in Equation 9 can be estimated recursively by maximum likelihood if assumptions are made on the density of the innovations $\varepsilon_{i,t}$.[50] Appendix I provides a description of the likelihood function used to estimate the parameters. This approach estimates means of both alpha and beta coefficients, variance and the transition probabilities for both recessionary and non-recessionary states. Hence, it allows us to connect fund performance explicitly to the state of the business cycle, and provide answers on how Norwegian equity mutual funds perform in recession and non-recessionary states of the economy.

[49]Return-series on the OECD CLI for Norway from 1983 to 2014 is obtained from OECD Statistics webpage.

[50]In this study, we assume that the innovations are Gaussian.

6.3.3 Markov Regime-Switching Models of Mutual Fund Alpha

Given the advantages of the regime-switching model outlined in Section 6.3, we now examine the performance of Norwegian funds in recessions and non-recession periods using this forward-looking latent state approach and find that the results are mostly in line with our previous findings. As previously discussed, we estimate two models. The first model assumes constant state transition probabilities, whereas the second model allows the state transition probabilities to vary over time with changes in the Composite Leading Indicator. Table XIV below reports maximum likelihood estimates of the Markov-switching models in addition to the single-state model.

Model I (fixed transition probabilities) and Model II (time-varying transition probabilities) in Table XIV confirm our findings based on the Norwegian recession periods that the recession alpha is higher than the non-recession alpha. The estimated yearly alpha is positive of 0.33% and 0.43% in recession periods and negative of -1.19% and -1.46% in non-recession periods in Model I and Model II, respectively. The alpha estimates are, however, not statistically significant in any state in any of the models. From the log-likelihood values, we see that by allowing the state transition probabilities to vary over time with the two-month lagged change in CLI, the in-sample fit of the model increases.

The recession state is identified by three pieces of evidence. First, the transition probability coefficients, d_1 and d_2, have opposite signs in the time-varying transition probability specification (Model II). Hence, the transition probabilities, p_t and q_t move in the opposite direction when CLI_{t-2} fluctuates. Although not significant, the signs in front of d_1 and d_2 are intuitively plausible. The economic interpretation of the coefficients is that as CLI decreases the probability of being and staying in State 1 (non-recessions) decreases while the probability of being and staying in State 2 (recessions) increases. Second, we find that the residual volatility is higher in State 2 for both models. The higher residual volatility in State 2 represents the higher uncertainty regarding the predictive power of the model in State 2. From economic theory, one could expect the recessionary state to be more volatile than the non-recessionary state, and consequently we can expect a higher residual volatility in recession periods compared to non-recession periods. This premise is supported by Schwert (1990) and Hamilton and Lin (1996) who find that stock return volatility is highest during economic recessions. The higher residual volatility is consistent with our findings in Table XIII in Section 6.2.2. Third, in the fixed transition probability specification (Model I), State 2 is less persistent ($p_{22} = 0.84$) than State 1 ($p_{11} = 0.96$).[51] [52] This is consistent with Perez-Quiros and Timmermann (2000) and Ang and Bekaert (2002) who find that the high volatility state is also the less persistent state.

[51] p_{11} is the probability that the process will stay in State 1 at time $t+1$ given that the process was in State 1 at time t. The same intuition applies to p_{22}.

[52] $p_{11} = \frac{1}{1+exp(-3.196)} = 0.96$; $p_{22} = \frac{1}{1+exp(-1.625)} = 0.84$

Table XIV
Regime-Switching Estimates of Carhart's (1997) Four-Factor Alpha

This table reports regime-switching estimates of fund performance based on Carhart's (1997) four-factor model. The table includes estimates when regime-switching is applied on two-state models using both fixed transition probabilities (Model I) and time-varying transition probabilities (Model II). Rows 1 to 5 report mean parameters, which includes alpha (in percent per year) and the four risk factors in Carhart's model, MKT, SMB, HML, and PR1YR. Row 6 reports variance parameters in percent per month, and Rows 7 and 8 report the transition probability parameters. Row 9 show the log-likelihood estimates of each three models. The t-statistics are reported in parentheses below each estimate. The sample period is 1983 to 2014. The following regime-switching Carhart (1997) four-factor model is estimated for the equally-weighted portfolio of all funds:

$$r_{i,t} - r_{f,t} = \alpha_{i,S_t} + \beta_{1i,S_t} \cdot MKT + \beta_{2i,S_t} \cdot SMB_t + \beta_{3i,S_t} \cdot HML_t + \beta_{4i,S_t} \cdot PR1YR_t + \varepsilon_{i,t}$$

where $\varepsilon_t \sim N(0, \sigma_{s_t})$

Column 1 reports single state estimates, whereas Columns 2 (Model I) and 3 (Model II) report estimates in the two states. The transition probabilities in model I are fixed, and the time-varying transition probabilities in Model II depend on the two month lagged change in the Composite Leading Indicator ΔCLI_{t-2}:

Model I: p_t=Pr(S_t=1|S_{t-1}=1)=$\phi(c_1)$, q_t=Pr(S_t=2|S_{t-1}=2)=$\phi(c_2)$
Model II: p_t=Pr(S_t=1|S_{t-1}=1)=$\phi(c_1+\pi_1\Delta CLI_{t-2})$, q_t=Pr(S_t=2|S_{t-1}=2)=$\phi(c_2+\pi_2\Delta CLI_{t-2})$

where $\phi(.)$ is the cumulative density function of a standard normal variable

	Single State Model	Model I		Model II	
Mean Parameters		State 1	State 2	State 1	State 2
Alpha (in pct p.a.)	-0.428	-1.193	0.332	-1.457	0.429
	(-0.33)	(-1.13)	(0.63)	(-1.51)	(0.87)
MKT	0.995***	1.002***	0.967***	1.001***	0.970***
	(-66.44)	(66.08)	(14.97)	(69.36)	(15.89)
SMB	0.101***	0.074***	0.180**	0.078***	0.163**
	(4.69)	(3.56)	(2.16)	(3.90)	(2.31)
HML	-0.050	-0.063***	-0.004	-0.064***	-0.005
	(-2.88)	(-3.88)	(-0.04)	(-3.89)	(-0.07)
PR1YR	-0.029	-0.017	-0.067	-0.010	-0.082
	(-1.61)	(-0.69)	(-0.87)	(-0.52)	(-1.27)
Variance Parameters					
S.D. (in pct. p.m.)	1.678***	1.181***	2.935***	1.175***	2.910***
	(113.14)	(52.78)	(28.73)	(68.22)	(31.73)
Transition Probability Parameters					
c_s		3.196***	1.625***	3.301***	1.694***
		(4.77)	(2.96)	(5.69)	(3.31)
$\pi_s\Delta CLI_{t-2}$				0.295	-0.051
				(1.30)	(0.22)
Log Likelihood	1022.2	1062.8		1065.7	

Moreover, consistent with our previous findings, the regime-switching estimates show that actively managed Norwegian mutual funds, on average, reduces their exposure to the market factor in recessionary periods.[53] The regime-switching estimates also confirm our previous findings regarding the funds' loading on the momentum factor and the value factor, which is higher (lower) on the value (momentum) factor in recessions. Hence, these results provide further evidence of positive timing regarding the market, value and momentum factor. However, the regime-Switching estimates imply that Norwegian mutual funds increase their exposure to small firms during recessionary periods. This is in contrast to our regression results in Table XII using Norwegian recession and non-recession sub-periods, and in contrast to a *flight to quality* tendency often observed in bear market.[54] This may indicate bad timing in regards to the size factor. However, as stated before, these results should be interpreted with caution given the fact that portfolio betas can change as a result of changes in the betas of the assets comprising the portfolio.

Our findings are robust to different factor models as Appendix J shows. The alpha remains insignificant and positive in recessions, and insignificant and negative in non-recessions regardless of whether alpha is defined with respect to the Carhart (1997) four-factor model, the Fama and French (1993) three-factor model or the single-factor CAPM.

6.3.4 Statistical Tests for Asymmetries

One of the crucial tests in Markov Regime-Switching models concerns the number of states that best characterize the data. Unfortunately, tests for the number of states cannot be performed with the common likelihood ratio test (LRT) because, under the null of a single state, state transition probabilities are unidentified nuisance parameters and standard results on the distribution of LRT no longer apply. Various methods to overcome this problem have been proposed in the literature (See e.g. Boldin, 1990; Hansen, 1992; Garcia, 1998). Hansen (1992) suggest a Monte Carlo method to derive the asymptotic distribution of the regression coefficients under the null. However, given the number of parameters that needs to be estimated in our model, this approach will quickly run into computational limitations. We therefore follow Perez-Quiros and Timmermann (2000) and Kosowski (2011) and test for asymmetries in the conditional mean condition on the existence of two states in the conditional volatility and vice versa. The resulting LRT follows a standard chi-squared distribution. Panel A and Panel B in Table XV below reports the outcome of the tests for identical mean and identical variance using Model II, respectively.[55] We reject the null of symmetry in both the conditional mean and conditional volatility. These results confirm that asymmetries are statistically significant.

[53]Implied by a lower beta loading on the market proxy in recession periods.

[54]This is a result of letting the Markov Swithcing Model define the high-volatility recessionary states, which is different from that of Aastveit et al. (2014).

[55]These results also hold for Model I.

Table XV
Likelihood Ratio Tests for Identical Mean and Variance across States in the Markov Regime-Switching Model

The table reports the results of the likelihood ratio tests for equality of parameters across the two states in the Markov Regime-Switching model (Model II). Panel A reports the outcome of the test for identical mean parameters, which assumes that there are two states in the conditional variance. Panel B reports the outcome of the test for identical variance parameters, which assumes that there are two states in the conditional mean. In both panels, Row 1 (2) reports the unrestricted (restricted) log-likelihood value, whereas Rows 3 and 4 report the likelihood ratio test statistic (LRT) and p-value. The LRT statistic is based on the chi-squared distribution with m degrees of freedom, where m is the number of restrictions. The sample period is 1983 to 2014. The following regime-switching Carhart (1997) four-factor model is estimated for the equally-weighted portfolio of all funds:

$$r_{i,t} - r_{f,t} = \alpha_{i,S_t} + \beta_{1i,S_t} \cdot MKT + \beta_{2i,S_t} \cdot SMB_t + \beta_{3i,S_t} \cdot HML_t + \beta_{4i,S_t} \cdot PR1YR_t + \varepsilon_{i,t}$$

$$\text{where } \varepsilon_t \sim N(0, \sigma_{s_t})$$

$$p_t{=}\Pr(S_t{=}1|S_{t-1}{=}1){=}\phi(c_1{+}\pi_1\Delta CLI_{t-2}), \quad q_t{=}\Pr(S_t{=}2|S_{t-1}{=}2){=}\phi(c_2{+}\pi_2\Delta CLI_{t-2})$$

Panel A: Test for Identical Means

	EW (All Funds)
Unrestricted Log-Likelihood	1065.697
Restricted Log-Likelihood with ($\alpha_{s_t=1}{=}\alpha_{s_t=2}$ and $\beta_{q,s_t=1}{=}\beta_{q,s_t=2}$, $q{=}\{1,2,3,4\}$)	1060.026
LRT Statistic	11.342
p-value (LRT)	0.045

Panel B: Test for Identical Variance

	EW (All Funds)
Unrestricted Log-Likelihood	1065.697
Restricted Log-Likelihood with ($\sigma_{s_t=1}{=}\sigma_{s_t=2}$)	1053.999
LRT Statistic	23.396
p-value (LRT)	0.000

6.4 Interpretation of Results and Economic Importance

Our results above suggest some evidence of asymmetries in the performance of actively managed Norwegian equity funds. Specifically, based on our Markov regime-switching model with time-varying transition probabilities we find a difference in risk-adjusted performance (alpha) of 1.89% between recession and non-recession periods. Differences in risk-adjusted performance between recession and non-recession periods are robust to the binary classification of Norwegian recession dates based on Aastveit et al. (2014).[56] However, even though statistical tests provide evidence of asymmetries in Norwegian mutual fund returns, we cannot reject the hypothesis that the alpha in recession and non-recession periods is individually statistically different from zero. Despite the lack of statistical sig-

[56] We use Aastveit et al. (2014)'s binary classification of recession dates to construct two separate sub-samples for recession and non-recession periods and estimate the Carhart (1997) four-factor model on the separate sub-samples.

nificance, we find that the signs on the alpha estimates are in line with Kosowski (2011) who conduct the same analysis on US data. That is, Norwegian mutual fund managers show some tendencies to perform better in recessions when performance matters the most to investors, and that the documented underperformance appears to stem from non-recessionary periods when information are relatively symmetrically distributed in the market.

7 Conclusion

This paper investigates the performance, persistence, and business cycle asymmetries in Norwegian mutual fund returns using a dataset free of survivorship bias containing 98 active Norwegian mutual funds with net monthly returns from 1983 to 2014. We use the unconditional four-factor model of Carhart (1997) as our primary performance model and examine mutual fund performance both at the aggregate and the individual level. Due to non-normalities in the return residuals, we adapt a bootstrap approach similar to Kosowski et al. (2006), Fama and French (2010) and Cuthbertson et al. (2008) to evaluate the statistical significance of our results.

We find that actively managed Norwegian mutual funds on aggregate produce returns that underperform the four-factor benchmark net of costs. This implies that if the average fund manger in fact possesses superior skill, they do little beside collect fees. When we study individual funds, our bootstrap simulations suggest no evidence of superior fund mangers. On the contrary, we find credible evidence of inferior performing funds in the left tail of the performance spectrum. We reject the hypothesis that the bad performance among the worst performing funds is due to bad skill.

To evaluate whether there exists persistence in the performance, we employ several reputable parametric and non-parametric statistical tests adapted from previous literature. Tests for persistence reveals no evidence of risk-adjusted performance persistence among the top quintile portfolio of funds. The lack of evidence of persistence in performance among top-performing funds is bad news for Norwegian mutual fund investors, as investors cannot look at past performance to earn positive risk-adjusted returns. However, among the bottom quintile portfolio of funds, we find that performance strongly persists for time horizons up to six months before it largely disappears. These results are robust to rankings based on returns, alpha and t-statistic of alpha. Additionally, we find positive abnormal performance on a hypothetical spread portfolio, which consist of a long position in the top quintile portfolio and a short position in the bottom quintile portfolio. Specifically, buying last years top quintile portfolio and selling last years bottom quintile portfolio yields a yearly alpha of 3.63% if funds are ranked on the past 12 months return and held for one month before the portfolios are rebalanced. The economic implication of this finding is, however, questionable, as investors cannot easily short mutual funds. Additionally, the strategy entails frequent portfolio rebalancing, which would most certainty impose substantial transaction costs. Our series of non-parametric two-period tests reveals evidence of persistence amongst both winner and loser funds relative to other funds in the Norwegian market for time windows up to six months. Overall, our results are in line with Bollen and Busse (2005) who argue that post-ranking abnormal returns largely disappear when funds are evaluated over longer holding periods.

Finally, we investigate potential asymmetries in the performance of the aggregated portfolio of actively managed Norwegian mutual funds. Specifically, we aim to link fund performance to bad and good states of the economy, using two different methodologies. The first approach is based on a binary classification of recession periods in Norway. The second approach involves a Markov regime-switching framework, where we let the data determine the indicator of the different states in the economy. We find that managers of active Norwegian funds indeed show some indications to perform better in recessionary periods when performance matters the most to investors. Statistical tests show evidence of asymmetries in the returns of actively managed mutual funds, but we cannot explicitly reject the hypothesis that the recession alpha is statistically significantly different from zero. In economical terms, we find that the difference in alpha between recession and non-recession periods is 1.89% based on our Markov regime-switching model with time-varying transition probabilities. Differences in alpha between recession and non-recession periods are robust to the binary classification of recessions in Norway.

The general consensus in the existing literature on mutual fund performance is confirmed by the empirical results in our study. As previously emphasized, active management is not justified from an academic point of view because it would violate the efficient market hypothesis. Our results make it hard to disagree. The puzzle why investors keep investing in actively managed mutual funds still remain unanswered, but our results on asymmetric performance provide some important insight into a possible answer.

All taken together, it seems like Norwegian mutual fund investors are better of by investing in passively managed low-cost funds, as managers of active Norwegian mutual funds do little besides collecting fees. At the very least, our results show that Norwegian mutual fund investors should stay away from past losers.

Interesting avenues for further research could be to investigate the performance of Norwegian funds with international investment mandates. As illustrated in Figure 1, this category of mutual funds has experienced considerable growth during the past twenty years. To our knowledge, there has been conducted no research on the performance of this category of mutual funds. A major challenge on this subject, however, is to choose appropriate benchmarks. Additionally, previous research suggests that in- and outflows of funds could have severe impact on mutual fund performance (See e.g. Berk and Green, 2004; Lückoff, 2011). We leave these questions for further research.

References

Aastveit, K. A., Jore, A. S., and Ravazzolo, F. (2014). Forecasting recessions in real time. Norges Bank Working Paper 2014/02.

Agarwal, V. and Naik, N. Y. (2000). Multi-period performance persistence analysis of hedge funds. *Journal of Financial and Quantitative Analysis*, 35(3):327–342.

Ang, A. and Bekaert, G. (2002). Regime switches in interest rates. *Journal of Business and Economic Statistics*, 20(2):163–182.

Ang, A. and Chen, J. (2007). CAPM over the long run: 1926-2001. *Journal of Empirical Finance*, 14(1):1–40.

Artis, M. J., Kontolemis, Z. G., and Osborn, D. R. (1995). Classical business cycles for G7 and European countries. CEPR Discussion Paper No. 1137.

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

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

Berk, J. B. and Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, 112(1):1269–1295.

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

Bjørnland, H. C. (2000). Detrending methods and stylized facts of business cycles in Norway - an international comparison. *Empirical Economics*, 25(3):369–392.

Bjørnland, H. C., Brubakk, L., and Jore, A. S. (2008). Forecasting inflation with an uncertain output gap. *Empirical Economics*, 35(3):413–436.

Blake, D. and Timmermann, A. (1998). Mutual fund performance: Evidence from the UK. *European Finance Review*, 2(1):57–77.

Boldin, M. (1990). Characterizing business cycles with a Markov switching model: Evidence of multiple equilibria. Federal Reserve Bank of New York, Research Paper no. 9037.

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

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7(3):265–296.

Brown, S. J. and Goetzmann, W. N. (1995). Performance persistence. *Journal of Finance*, 50(2):679–698.

Brown, S. J., Goetzmann, W. N., Ibbotson, R. C., and Ross, S. A. (1992). Survivorship bias in performance studies. *Review of Financial Studies*, 5(4):553–580.

Bry, G. and Boschan, C. (1971). Cyclical analysis of time series: Selected procedures and computer programs. Technical Paper 20, National Bureau of Economic Research, Columbia University Press.

Burns, A. F. and Mitchell, W. C. (1946). Measuring business cycles. NBER Books. National Bureau of Economic Research.

Busse, J. A., Goyal, A., and Wahal, S. (2010). Performance and persistence in institutional investment management. *Journal of Finance*, 65(2):765–790.

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

Carpenter, J. N. and Lynch, A. W. (1999). Survivorship bias and attrition effects in measures of performance persistence. *Journal of Financial Economics*, 54(3):337–374.

Chang, E. C. and Lewellen, W. G. (1984). Market timing and mutual fund investment performance. *Journal of Business*, 57(1):57–72.

Chauvet, M. and Potter, S. (2000). Coincident and leading indicators of the stock market. *Journal of Empirical Finance*, 7(1):87–111.

Christensen, M. (2005). Danish mutual fund performance-selectivity, market timing and persistence. Working Paper, Aarhus School of Business, Finance Research Group.

Christensen, M. (2013). Danish mutual fund performance. *Applied Economics Letters*, 20(8):818–820.

Christensen, R. (1990). *Log-linear models*. Springer-Verlag, New York.

Christoffersen, P. (2000). Dating the turning points of Nordic business cycles. EPRU Working Paper No. 00/13.

Christopherson, J. A., Ferson, W. E., and Glassman, D. A. (1998). Conditioning manager alphas on economic information: Another look at the persistence of performance. *Review of Financial Studies*, 11(1):111–142.

Connor, G. and Korajczyk, R. A. (1991). The attributes, behavior, and performance of U.S. mutual funds. *Review of Quantitative Finance and Accounting*, 1(1):5–26.

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

Dahlquist, M., Engström, S., and Söderlind, P. (2000). Performance and characteristics of Swedish mutual funds. *Journal of Financial and Quantitative Analysis*, 35(3):409–423.

Daniel, K., Grinblatt, M., Titman, S., and Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance*, 52(3):1035–1058.

Diebold, F. X., Lee, J.-H., and Weinbach, G. C. (1994). Regime switching with time-varying transition probabilities. In *Nonstationary Time-Series Analyses and Cointegration*, pages 283–302. Oxford: Oxford University Press.

Durland, M. J. and McCurdy, T. H. (1994). Duration-dependent transitions in a Markov model of U.S. GNP growth. *Journal of Business and Economic Statistics*, 12(3):279–288.

Edelen, R. M. (1999). Investor flows and the assessed performance of open-end mutual funds. *Journal of Financial Economics*, 53(3):439–466.

Eika, T. and Lindquist, K. G. (1997). Konjunkturimpulser fra utlandet. Statistics Norway.

Eitrheim, Ø., Klovland, J. T., and Qvigstad, J. F. (2004). Historical monetary statistics for Norway 1819-2003. *Norges Bank Occasional Papers*, 35(1):1–408.

Elton, E. J., Gruber, M. J., and Blake, C. R. (1995). Fundamental economic variables, expected returns, and bond fund performance. *Journal of Finance*, 50(4):1229–1256.

Elton, E. J., Gruber, M. J., and Blake, C. R. (1996a). The persistence of risk-adjusted mutual fund performance. *Journal of Business*, 69(2):133–157.

Elton, E. J., Gruber, M. J., and Blake, C. R. (1996b). Survivor bias and mutual fund performance. *Review of Financial Studies*, 9(4):1097–1120.

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

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

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

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

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

Ferreira, M. A., Keswani, A., Miguel, A. F., and Ramos, S. B. (2012). The determinants of mutual fund performance: A cross-country study. *Review of Finance*, 17(2):483–525.

Ferson, W. E. and Schadt, R. W. (1996). Measuring fund strategy and performance in changing economic conditions. *Journal of Finance*, 51(2):425–461.

Filardo, A. J. (1994). Business-cycle phases and their transitional dynamics. *Journal of Business and Economic Statistics*, 12(3):299–308.

Gallefoss, K., Hansen, H. H., Haukaas, E. S., and Molnár, P. (2015). What daily data can tell us about mutual funds: Evidence from Norway. *Journal of Banking and Finance*, 55(1):117–129.

Garcia, R. (1998). Asymptotic null distribution of the likelihood ratio test in Markov switching models. *International Economic Review*, 39(3):763–788.

Gjerde, Ø. and Sættem, F. (1991). Performance evaluation of Norwegian mutual funds. *Scandinavian Journal of Management*, 7(4):297–307.

Glode, V., Hollifield, B., Kacperczyk, M. T., and Kogan, S. (2011). Time-varying predictability in mututal fund returns. Working Paper, University of Pensylvania.

Glosten, L. R. and Jagannathan, R. (1994). A contingent claim approach to performance evaluation. *Journal of Empirical Finance*, 1(2):133–160.

Goetzmann, W. N. and Ibbotson, R. G. (1994). Do winners repeat? *Journal of Portfolio Management*, 20(2):9–18.

Goetzmann, W. N., Jr, J. I., and Ivkovic, Z. (2000). Montly measurement of daily timers. *Journal of Financial and Quantitative Analysis*, 35(3):257–290.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42(1):27–62.

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

Grinblatt, M. and Titman, S. (1989b). Portfolio performance evaluation: Old issues and new insights. *Review of Financial Studies*, 2(3):393–421.

Grinblatt, M. and Titman, S. (1992). The persistence of mutual fund performance. *Journal of Finance*, 47(5):1977–1984.

Grinblatt, M. and Titman, S. (1993). Performance measurement without benchmarks: An examination of mutual fund returns. *Journal of Business*, 66(1):47–68.

Grossman, S. J. and Shiller, R. J. (1981). The determinants of the variability of stock market prices. *American Economic Review*, 71(2):222–227.

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

Gruber, M. J. (1996). Another puzzle: The growth in actively managed mutual funds. *Journal of Finance*, 51(3):783–810.

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

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384.

Hamilton, J. D. and Lin, G. (1996). Stock market volatility and the business cycle. *Journal of Applied Econometrics*, 11(5):573–593.

Hansen, B. E. (1992). The likelihood ratio test under non-standard conditions: Testing the Markov switching model of GNP. *Journal of Applied Econometrics*, 7(1):61–82.

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

Henriksson, R. D. (1984). Market timing and mutual fund performance: An empirical investigation. *Journal of Business*, 57(1):73–96.

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

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

Huij, J. and Verbeek, M. (2007). Cross-sectional learning and short-run persistence in mutual fund performance. *Journal of Banking and Finance*, 31(3):973–997.

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

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

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

Kacperczyk, M., Nieuwerburgh, S. V., and Veldkamp, L. (2010). Rational attention allocation over the business cycle. Working Paper, New York University.

Kon, S. J. (1983). The market-timing performance of mutual fund managers. *Journal of Business*, 56(3):323–347.

Kon, S. J. and Jen, F. C. (1979). The investment performance of mutual funds: An empirical investigation of timing, selectivity, and market efficiency. *Journal of Business*, 52(2):263–289.

Korkeamaki, T. and Smythe, T. (2004). Effects on market segmentation and bank concentration on mutual fund expenses and returns: evidence from Findland. *European Financial Management*, 10(3):413–38.

Kosowski, R. (2011). Do mutual funds perform when it matters most to investors? US mutual fund performance and risk in recessions and expansions. *Quarterly Journal of Finance*, 1(3):607–664.

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.

Kothari, S. P., Shu, S., and Wysocki, P. D. (2009). Do managers withhold bad news? *Journal of Accounting Research*, 47(1):241–276.

Krueger, T. M. and Callaway, R. F. (1995). The persistence of three-year mutual fund performance. *Journal of Financial Planning*, 8(3):136–141.

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

Lehmann, B. N. and Modest, D. M. (1987). Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons. *Journal of Finance*, 42(2):233–265.

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

Lückoff, P. (2011). *Mutual Fund Performance and Performance Persistence: The impact of fund flows and manager changes.* Springer Gabler, Wiesbaden.

Malkiel, B. G. (1995). Returns from investing in equity mutual funds 1971 to 1991. *Journal of Finance*, 50(2):549–572.

Mamaysky, H., Spiegel, M., and Zhang, H. (2008). Estimating the dynamics of mutual fund alphas and betas. *Review of Financial Studies*, 21(1):233–264.

Merton, R. C. (1981). On market timing and investment performance. I. An equilibrium theory of value for market forecasts. *Journal of Business*, 54(3):363–406.

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

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

Næs, R., Skjeltorp, J. A., and Ødegaard, B. A. (2009). What factors affect the Oslo Stock Exchange? Norges Bank and University of Stavanger (UiS).

Ødegaard, B. A. (2015). Empirics of the Oslo Stock Exchange. Basic, descriptive, results 1980-2014. Working Paper.

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

Otten, R. and Bams, D. (2004). How to measure mutual fund performance: economic versus statistical relevance. *Accounting and Finance*, 44(2):203–222.

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

Pástor, L. and Stambaugh, R. F. (2002b). Mutual fund performance and seemingly unrelated assets. *Journal of Financial Economics*, 63(3):315–349.

Perez-Quiros, G. and Timmermann, A. (2000). Firm size and cyclical variations in stock returns. *Journal of Finance*, 55(3):1229–1262.

Perez-Quiros, G. and Timmermann, A. (2001). Business cycle asymmetries in stock returns: Evidence from higher order moments and conditional densities. *Journal of Econometrics*, 103(1-2):259–306.

Pesaran, M. H. and Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance. *Journal of Finance*, 50(4):1201–1228.

Qiu, J., Faff, R. W., and Benson, K. L. (2011). Business cycles and mutual fund timing performance: An application of regime switching and GARCH modelling. *SSRN Electronic Journal*, 1913935.

Roll, R. (1977). A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. *Journal of Financial Economics*, 4(2):129–176.

Roll, R. (1978). Ambiguity when performance is measured by the securities market line. *Journal of Finance*, 33(4):1051–1069.

Roll, R. (1980). Orthogonal portfolios. *Journal of Financial and Quantitative Analysis*, 15(5):1005–1023.

Rubinstein, M. (1976). The valuation of uncertain income streams and the pricing of options. *Bell Journal of Economics*, 7(2):407–425.

Schwert, G. W. (1990). Stock returns and real activity: A century of evidence. *Journal of Finance*, 45(4):1237–1257.

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

Sharpe, W. F. (1992). Asset allocation: Management style and performance measurement. *Journal of Portfolio Management*, 18(2):7–19.

Sørensen, L. Q. (2009a). Mutual fund performance at the Oslo Stock Exchange. *SSRN Electronic Journal*, 1488745.

Sørensen, L. Q. (2009b). Oil price shocks and stock return predictability. *SSRN Electronic Journal*, 1341013.

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

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

Veit, T. E. and Cheney, J. M. (1982). Are mutual funds market timers? *Journal of Portfolio Management*, 8:35–42.

Vidal-Garcia, J. (2013). The persistence of European mutual fund performance. *Research in International Business and Finance*, 28:45–67.

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

Appendix A

Table A.I
Mutual Fund Database Summary Statistics

This table provides summary statistics of our mutual fund database. Column 1 shows the number of funds available to investors at the end of each year. The second (third) columns show how many funds that are born (liquidated) each year. Column 4 reports annualized returns on an equally weighted portfolio comprising all funds each year, whereas Column 5 reports returns on the OSEAX benchmark. Returns are reported in percent.

	Number of funds			Returns	
Year	Year-end	Born	Liquidated	EW (All)	OSEAX
2014	54	0	6	7.86	3.33
2013	60	3	4	21.78	21.19
2012	61	1	1	15.29	11.23
2011	62	4	0	-19.86	-8.17
2010	58	2	0	22.00	16.71
2009	55	0	0	58.16	46.54
2008	55	0	0	-66.05	-63.95
2007	55	0	2	12.04	13.53
2006	57	3	6	27.12	30.41
2005	60	1	4	39.23	44.46
2004	63	1	2	32.26	34.59
2003	64	2	3	45.14	41.74
2002	65	7	4	-38.88	-26.48
2001	62	4	6	-14.63	-11.86
2000	64	7	3	4.36	3.57
1999	60	3	3	45.05	42.09
1998	60	12	0	-32.73	-27.97
1997	48	9	0	29.44	26.74
1996	39	7	0	35.25	24.24
1995	32	7	0	15.97	11.75
1994	25	8	0	4.96	8.72
1993	17	1	0	57.64	52.84
1992	16	4	0	-13.37	-7.11
1991	12	2	0	-8.24	-6.85
1990	10	3	0	-17.12	-11.60
1989	7	0	0	49.02	46.90
1988	7	0	0	27.97	33.87
1987	7	1	0	-3.01	-4.82
1986	6	0	0	-5.49	-8.21
1985	6	1	0	19.26	28.66
1984	5	0	0	37.19	24.47
1983	5	3	0	73.50	68.11

Appendix B

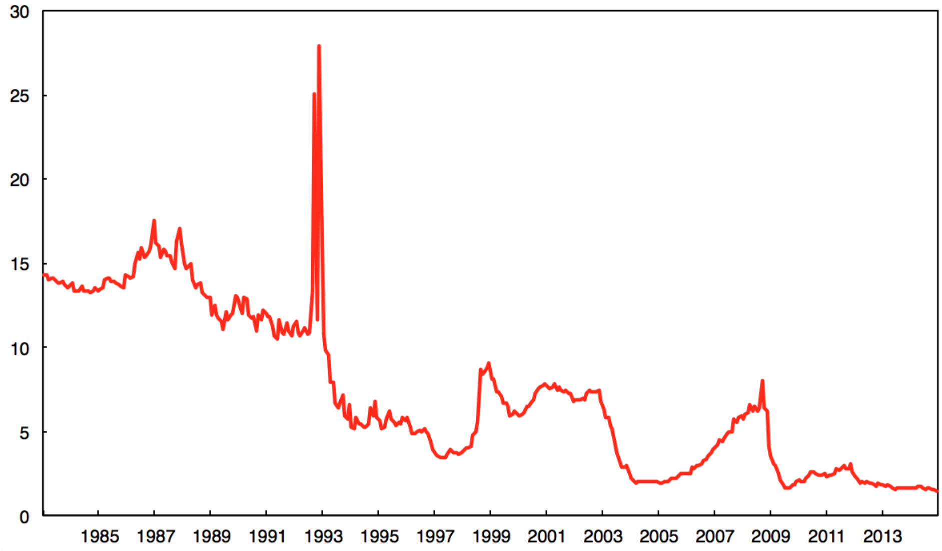

Figure B.1. Monthly risk-free interest rate from 1983 to 2014. The figure plots the annualized one-month interest rate for the period 1983 to 2014. For the period between 1983 and 1986, the overnight NIBOR rate is used. Between 1986 and 2014, the figure plots the one-month NIBOR interbank rate.

Appendix C

Following Elton et al. (1996b), we estimate the magnitude of survivorship bias in the performance (i.e. alpha) between surviving funds and all funds in our sample. If survivorship bias is present, excluding non-surviving (i.e. dead) funds in the sample data will yield unrealistically high estimates of the aggregate mutual fund performance. Specifically, we divide the sample into all funds (98), surviving funds (54) and non-surviving funds (44). Non-surviving funds are included in the sample until their last monthly observation, upon which the fund was liquidated or merged with another fund. Then, we estimate the potential bias imposed by restricting our sample to include surviving funds only by estimating the difference between the performance of surviving and non-surviving funds. It is evident from Table C.I below that the aggregate performance is biased upwards if we were to include surviving funds only, regardless of which performance model we use. The table shows that the aggregate fund portfolio delivers neutral performance by only including surviving funds, which indicates a survivorship bias of 0.428% per year when measuring against Carhart's (1997) four-factor model.

Table C.I
Survivorship Bias

The table presents alphas for all funds (Column 1), surviving funds (Column 2), and non-surviving funds (Column 4) in our sample. The alphas are estimated by using Jensen's (168) single-factor CAPM, Fama-French's (1993) three-factor model, and the four-factor model of Carhart (1997). The estimated bias in Column 3 is the difference in the average performance between surviving funds and all funds. Alphas are annualized and reported in percent. The sample period is 1983 to 2014.

	All Funds	Surviving Funds	Bias	Dead Funds
CAPM	0.391	0.801	0.410	-0.810
Fama-French	-0.569	-0.214	0.355	-1.633
Carhart	-0.428	0.000	0.428	-1.257

Appendix D

Table D.I

Aggregate Fund Performance Using the Conditional Four-Factor Model

The table report regression results on the equally weighted portfolio comprising all funds in our sample using different versions of the conditional four-factor model. Column 1 reports the annualized alpha, whereas Columns 2 - 5 reports the four unconditional variables included in the models, MKT, SMB, HML, and PR1YR. Columns 6 - 8 reports regression coefficients of the three information variables included in the conditional models (i.e. Models I, II, and III), being the yield on the oil price, industrial production growth, and market return, respectively. The conditional variables are lagged one month. The final column reports the adjusted R-squared. Row 4 reports estimates from the unconditional model for comparative purposes. The t-statistics are reported in parentheses below each coefficient. Numbers assigned stars, ***, **, and * indicates significance at the 1%, 5%, and 10% levels, respectively. The sample period is 1983 to 2014.

		Unconditional Variables				Conditional Variables			
	α	MKT	SMB	HML	PR1YR	Oil Price$_{t-1}$	Industrial Production$_{t-1}$	MKT$_{t-1}$	R^2_{Adj}
Model I	-0.186	0.996***	0.101***	-0.051***	-0.028*	-0.025*			0.929
	(-0.17)	(66.04)	(4.68)	(-2.91)	(-1.65)	(-1.73)			
Model II	-0.509	0.998***	0.102***	-0.050***	-0.029*		0.005		0.929
	(-0.46)	(65.89)	(4.71)	(-2.84)	(-1.67)		(1.07)		
Model III	-0.614	0.999***	0.099***	-0.051***	-0.029*			0.022*	0.929
	(-0.56)	(65.77)	(4.60)	(-2.88)	(-1.65)			(1.65)	
Unconditional Carhart	-0.428	0.995***	0.101***	-0.050***	-0.029*				0.929
	(-0.33)	(66.01)	(4.66)	(-2.86)	(-1.69)				

Appendix E

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

The table shows descriptive statistics on all individual Norwegian mutual funds in our mutual fund database. Column 1 reports the number of monthly return observations in each fund, whereas yearly excess returns, skewness and kurtosis are reported in Columns 2 - 4. Column 5 reports outputs from the Jarque-Bera test, where a value of 1 (0) implies rejection (no rejection) of normality on a 5% level. The estimates are constructed by running regressions on each individual fund from 1983 to 2014.

Name	Obs	Return	Skewness	Kurtosis	Jarque-Bera
ABIF Norge ++	55	3.01	-0.59	4.50	1
Alfred Berg Aksjefond Norge	112	1.34	-0.47	5.98	1
Alfred Berg Aksjespar	105	0.75	-0.22	3.42	0
Alfred Berg Aktiv	228	9.70	0.24	4.65	1
Alfred Berg Aktiv II	180	4.80	0.21	4.12	1
Alfred Berg Gambak	251	11.46	0.10	3.21	0
Alfred Berg Humanfond	180	4.44	0.31	3.28	0
Alfred Berg N. Pensjon	204	6.44	-0.13	3.55	0
Alfred Berg Norge Classic	195	6.72	-0.26	4.02	1
Alfred Berg Norge Etisk	290	5.20	0.18	2.78	0
Alfred Berg Norge Inst	144	8.41	-0.01	3.75	0
Alfred Berg Norge +	51	7.71	0.14	3.35	0
Alfred Berg Vekst	69	1.54	1.18	6.89	1
Arctic Norwegian Equities Class A	48	7.79	-0.23	2.43	0
Arctic Norwegian Equities Class B	49	10.58	-0.22	2.49	0
Arctic Norwegian Equities Class D	22	15.46	0.62	2.95	0
Arctic Norwegian Equities Class I	49	10.51	-0.19	2.49	0
Atlas Norge	202	7.29	1.43	11.43	1
Banco Norge	36	8.83	0.03	2.31	0
Carnegie Aksje Norge	233	9.71	0.05	7.03	1
Danske Invest Aktiv Formuesforvaltning A	19	23.21	-0.77	3.69	0
Danske Invest Norge Aksj. Inst 1	176	10.51	0.48	5.13	1
Danske Invest Norge Aksj. Inst 2	97	8.61	0.42	4.83	1
Danske Invest Norge I	251	7.16	0.04	3.95	1
Danske Invest Norge II	251	6.61	-0.02	4.29	1
Danske Invest Norge Vekst	251	9.69	2.00	14.17	1
Delphi Norge	245	12.04	-0.18	4.34	1
Delphi Vekst	191	6.67	-0.19	3.70	0
DNB Norge	232	6.06	0.25	5.56	1
DNB Norge (Avanse I)	285	5.76	-0.27	3.75	1
DNB Norge (Avanse II)	374	5.23	0.14	4.72	1
DNB Norge (I)	284	3.48	-1.03	13.22	1
DNB Norge (III)	226	6.48	-0.24	4.61	1
DNB Norge (IV)	145	13.04	-0.47	5.50	1
DNB Norge Selektiv	224	8.62	0.14	3.61	0
DNB Norge Selektiv (II)	155	9.32	0.04	3.12	0
DNB Norge Selektiv (III)	246	7.52	-0.19	5.51	1
DnB Real-Vekst	150	-4.50	-0.16	4.11	1
DNB SMB	165	10.33	0.26	3.04	0
Eika Norge	135	12.94	-0.47	4.52	1
Eika SMB	185	4.08	0.09	3.29	0
Fokus Barnespar	30	-8.15	-1.03	4.58	1
Fondsfinans Aktiv II	46	-7.05	-0.21	2.95	0
Fondsfinans Norge	144	15.96	0.15	5.32	1

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

Name	Obs	Return	Skewness	Kurtosis	Jarque-Bera
FORTE Norge	46	0.61	-0.13	2.55	0
FORTE Trønder	20	11.23	0.34	2.33	0
GAMBAK Oppkjøp	17	-3.40	-0.17	2.21	0
GJENSIDIGE AksjeSpar	150	1.02	0.65	7.23	1
GJENSIDIGE Invest	102	7.95	-0.20	4.25	1
Globus Aktiv	88	7.19	-0.27	3.84	0
Globus Norge	103	-0.10	-0.34	4.11	1
Globus Norge II	95	4.28	-0.33	4.10	1
Handelsbanken Norge	238	8.78	-0.15	4.86	1
Holberg Norge	35	7.50	0.37	2.62	0
K-IPA Aksjefond	168	7.18	0.15	2.87	0
KLP Aksjeinvest	95	-0.85	0.07	4.43	1
KLP AksjeNorge	189	8.39	0.18	3.41	0
Landkreditt Norge	103	6.65	0.43	4.81	1
Landkreditt Utbytte	22	10.02	-1.52	5.02	1
NB-Aksjefond	205	4.96	-0.24	4.60	1
Nordea Avkastning	375	4.02	-0.73	7.80	1
Nordea Barnespar	45	-11.09	0.59	5.39	1
Nordea Kapital	237	9.86	-0.27	5.56	1
Nordea Kapital II	82	8.92	-0.85	6.65	1
Nordea Kapital III	68	7.81	-0.90	6.27	1
Nordea Norge Pluss	44	5.46	-0.52	3.56	0
Nordea Norge Verdi	226	8.73	-0.31	3.98	1
Nordea SMB	211	1.15	0.01	4.72	1
Nordea SMB II	68	-23.54	-0.44	3.74	1
Nordea Vekst	377	4.20	-0.25	7.35	1
ODIN Norge	270	12.03	1.17	7.72	1
ODIN Norge II	127	6.04	-0.02	3.15	0
Orkla Finans 30	155	11.61	0.21	3.90	1
Pareto Aksje Norge	159	12.30	0.30	3.27	0
Pareto Aktiv	147	11.69	0.32	3.36	0
Pareto Investment Fund A	360	6.21	0.40	6.92	1
Pareto Verdi	216	4.05	0.38	6.11	1
PLUSS Aksje (Fondsforvaltning)	108	6.92	0.02	2.77	0
PLUSS Markedsverdi (Fondsforvaltning)	239	8.61	-0.28	4.64	1
Postbanken Aksjevekst	95	1.82	0.86	4.91	1
RF Aksjefond	114	6.02	-0.49	3.62	1
RF Plussfond	52	15.21	-0.05	2.49	0
SEB Norge LU	66	-12.24	-0.10	3.25	0
Skandia Horisont	95	7.98	1.42	7.36	1
Skandia SMB Norge	95	-9.52	-0.29	5.22	1
Storebrand Aksje Innland	221	4.63	-0.92	5.96	1
Storebrand AksjeSpar	219	6.09	0.28	6.49	1
Storebrand Norge	375	6.06	0.00	9.18	1
Storebrand Norge A	41	19.11	-1.04	4.70	1
Storebrand Norge I	176	5.31	-0.62	5.52	1
Storebrand Norge Institusjon	168	4.23	-0.58	4.92	1
Storebrand Optima Norge	37	6.22	0.00	2.29	0
Storebrand Vekst	267	11.65	0.70	6.98	1
Storebrand Verdi	204	7.76	0.44	5.82	1
Swedbank Generator	51	12.79	0.28	2.91	0
Terra Norge	185	5.09	0.22	4.02	1
Terra Vekst	41	-14.11	0.16	2.60	0
VÅR Aksjefond	37	0.38	0.09	3.16	0

Appendix F

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

The table shows individual fund characteristics on all Norwegian funds in our mutual fund database. Columns 1 - 6 report alpha, alpha t-statistic, and individual factor loadings on the risk factors used in the unconditional four-factor model of Carhart (1997), MKT, SMB, HML and PR1YR, respectively. Column 7 reports the R-squared. The estimates are constructed by running regressions on each individual fund from 1983 to 2014. Alphas are reported in percent per year.

Name	α	$t_{\hat{\alpha}}$	β_{MKT}	β_{SMB}	β_{HML}	β_{PR1YR}	R^2
ABIF Norge ++	0.14	0.05	1.04	-0.11	-0.03	-0.11	0.95
Alfred Berg Aksjefond Norge	-4.31	-2.52	1.00	0.07	0.01	-0.03	0.95
Alfred Berg Aksjespar	-6.20	-2.61	1.07	0.10	0.00	0.03	0.92
Alfred Berg Aktiv	-1.58	-0.77	1.15	0.29	-0.19	0.00	0.88
Alfred Berg Aktiv II	-2.95	-1.16	1.09	0.30	-0.18	-0.04	0.87
Alfred Berg Gambak	-0.85	-0.37	1.17	0.44	-0.32	0.05	0.84
Alfred Berg Humanfond	-1.08	-0.67	1.01	-0.01	-0.13	-0.09	0.93
Alfred Berg N. Pensjon	0.51	0.38	1.03	-0.01	-0.07	-0.05	0.95
Alfred Berg Norge Classic	-1.26	-1.08	1.05	0.01	-0.02	-0.02	0.94
Alfred Berg Norge Etisk	-0.52	-0.30	1.05	0.02	-0.14	-0.15	0.95
Alfred Berg Norge Inst	1.86	1.37	1.03	-0.02	-0.07	-0.03	0.95
Alfred Berg Norge +	-2.53	-0.77	1.07	0.06	0.01	0.04	0.93
Alfred Berg Vekst	-7.68	-1.29	1.14	0.29	-0.11	0.18	0.79
Arctic Norwegian Equities Class A	-2.98	-1.01	1.06	0.19	-0.16	0.27	0.86
Arctic Norwegian Equities Class B	-2.47	-0.85	1.05	0.19	-0.16	0.27	0.88
Arctic Norwegian Equities Class D	-3.59	-0.79	0.90	0.06	-0.12	0.40	0.83
Arctic Norwegian Equities Class I	-2.53	-0.87	1.05	0.19	-0.16	0.27	0.88
Atlas Norge	-0.89	-0.36	1.13	0.17	-0.31	-0.02	0.87
Banco Norge	-3.64	-0.85	1.09	0.18	-0.27	-0.19	0.96
Carnegie Aksje Norge	1.45	0.86	0.99	-0.03	-0.16	0.08	0.90
Danske Invest Aktiv Formuesforvaltning A	-8.81	-0.97	0.76	0.07	0.19	0.30	0.84
Danske Invest Norge Aksj. Inst 1	5.03	3.03	0.96	-0.04	-0.04	-0.12	0.93
Danske Invest Norge Aksj. Inst 2	6.58	2.70	0.98	-0.01	-0.02	-0.06	0.92
Danske Invest Norge I	1.39	0.88	1.00	0.03	-0.04	-0.13	0.90
Danske Invest Norge II	0.85	0.55	1.00	0.02	-0.04	-0.13	0.91
Danske Invest Norge Vekst	-1.45	-0.54	1.08	0.43	-0.27	0.01	0.77
Delphi Norge	1.12	0.49	1.18	0.32	-0.23	-0.06	0.85
Delphi Vekst	-1.79	-0.64	1.10	0.35	-0.29	-0.09	0.84
DNB Norge	-0.47	-0.50	1.00	-0.06	-0.03	-0.05	0.97
DNB Norge (Avanse I)	-1.39	-1.20	0.98	-0.01	-0.06	-0.06	0.94
DNB Norge (Avanse II)	-0.61	-0.46	0.93	0.03	-0.03	-0.05	0.89
DNB Norge (I)	-1.44	-1.38	0.98	-0.02	-0.04	-0.03	0.95
DNB Norge (III)	-0.45	-0.49	1.00	-0.04	-0.03	-0.05	0.97
DNB Norge (IV)	0.99	0.86	1.01	-0.03	-0.06	-0.05	0.97
DNB Norge Selektiv	2.05	1.33	1.04	0.01	-0.06	-0.12	0.93
DNB Norge Selektiv (II)	0.39	0.32	1.01	-0.04	-0.05	-0.06	0.97
DNB Norge Selektiv (III)	-0.54	-0.42	1.02	0.05	-0.06	-0.02	0.94
DnB Real-Vekst	-3.42	-2.41	0.97	0.07	-0.03	-0.03	0.95
DNB SMB	0.20	0.07	1.26	0.58	-0.16	-0.23	0.85
Eika Norge	1.68	0.67	1.05	0.17	-0.05	-0.15	0.88
Eika SMB	-0.82	-0.34	0.98	0.17	-0.03	-0.21	0.86
Fokus Barnespar	-13.29	-1.29	0.99	0.04	0.01	-0.24	0.87
Fondsfinans Aktiv II	-3.80	-0.88	0.95	-0.10	0.00	-0.16	0.91
Fondsfinans Norge	4.07	1.85	1.02	0.08	-0.11	-0.16	0.90

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

Name	α	$t_{\hat{\alpha}}$	β_{MKT}	β_{SMB}	β_{HML}	β_{PR1YR}	R^2
FORTE Norge	-5.95	-1.41	1.09	-0.07	-0.16	0.07	0.83
FORTE Trønder	4.35	0.48	0.57	-0.19	-0.22	0.12	0.52
GAMBAK Oppkjøp	-6.43	-0.60	0.58	0.04	-0.35	0.37	0.90
GJENSIDIGE AksjeSpar	-4.14	-2.17	0.95	0.04	0.03	0.03	0.92
GJENSIDIGE Invest	-5.30	-2.65	0.99	0.16	0.12	0.08	0.93
Globus Aktiv	-5.44	-1.03	1.08	0.08	-0.26	-0.41	0.79
Globus Norge	-7.06	-1.52	1.07	0.15	-0.26	-0.39	0.81
Globus Norge II	-7.72	-1.52	1.07	0.10	-0.27	-0.42	0.78
Handelsbanken Norge	0.64	0.38	1.03	0.03	-0.06	-0.02	0.90
Holberg Norge	3.02	0.53	0.99	0.10	0.07	0.07	0.91
K-IPA Aksjefond	0.58	0.25	1.01	0.25	-0.10	-0.16	0.86
KLP Aksjeinvest	-1.41	-0.54	0.94	0.00	-0.04	-0.08	0.89
KLP AksjeNorge	1.31	0.78	1.02	0.02	-0.04	-0.11	0.92
Landkreditt Norge	4.90	1.80	0.97	0.08	0.02	-0.16	0.88
Landkreditt Utbytte	-1.68	-0.18	0.66	0.05	0.18	0.17	0.50
NB-Aksjefond	-1.72	-1.03	1.00	0.05	-0.02	-0.14	0.92
Nordea Avkastning	-1.66	-1.14	0.99	0.03	-0.03	-0.04	0.89
Nordea Barnespar	-6.71	-1.93	0.97	0.07	-0.02	0.01	0.93
Nordea Kapital	2.05	1.48	1.02	0.02	-0.07	-0.06	0.93
Nordea Kapital II	-2.72	-1.15	1.04	0.00	-0.07	-0.10	0.94
Nordea Kapital III	-2.71	-0.98	1.04	-0.01	-0.09	-0.16	0.94
Nordea Norge Pluss	-1.19	-0.36	1.08	0.02	-0.15	0.07	0.89
Nordea Norge Verdi	1.60	0.92	0.95	0.16	-0.04	-0.13	0.88
Nordea SMB	-7.84	-3.20	1.10	0.48	-0.08	-0.12	0.83
Nordea SMB II	-19.32	-3.19	1.04	0.53	-0.13	-0.08	0.79
Nordea Vekst	-1.69	-1.22	1.00	0.06	-0.04	-0.06	0.89
ODIN Norge	1.30	0.57	1.02	0.30	0.06	-0.11	0.80
ODIN Norge II	-4.11	-1.36	0.97	0.31	-0.04	-0.07	0.80
Orkla Finans 30	-0.94	-0.48	1.04	0.10	-0.05	-0.07	0.91
Pareto Aksje Norge	1.98	0.88	0.97	0.20	0.00	-0.04	0.86
Pareto Aktiv	-1.39	-0.59	0.95	0.21	-0.05	-0.02	0.85
Pareto Investment Fund A	0.79	0.49	1.00	0.05	-0.08	-0.04	0.87
Pareto Verdi	1.10	0.62	0.99	0.00	-0.05	-0.10	0.90
PLUSS Aksje (Fondsforvaltning)	0.82	0.32	0.95	0.22	-0.08	-0.17	0.87
PLUSS Markedsverdi (Fondsforvaltning)	2.98	2.37	0.95	-0.10	-0.02	-0.08	0.94
Postbanken Aksjevekst	-2.84	-1.12	1.02	0.02	-0.18	-0.08	0.92
RF Aksjefond	-1.09	-0.52	0.95	0.00	-0.06	-0.11	0.92
RF Plussfond	-6.34	-1.28	1.11	0.17	-0.29	-0.21	0.88
SEB Norge LU	-0.01	0.00	1.09	0.01	-0.07	-0.03	0.94
Skandia Horisont	4.75	1.50	1.04	0.18	-0.11	0.03	0.86
Skandia SMB Norge	-14.25	-3.59	1.03	0.40	-0.14	-0.10	0.81
Storebrand Aksje Innland	-2.61	-2.37	1.02	-0.04	-0.01	-0.03	0.96
Storebrand AksjeSpar	-1.34	-0.71	0.95	0.09	0.05	0.02	0.88
Storebrand Norge	0.31	0.19	0.99	0.02	-0.04	-0.02	0.86
Storebrand Norge A	-3.12	-0.76	1.06	-0.16	-0.06	-0.22	0.94
Storebrand Norge I	-0.93	-0.59	1.05	-0.01	-0.03	-0.12	0.94
Storebrand Norge Institusjon	-0.27	-0.14	1.05	0.03	-0.04	-0.13	0.92
Storebrand Optima Norge	-2.09	-0.97	1.08	0.05	-0.07	0.00	0.96
Storebrand Vekst	0.33	0.11	1.07	0.29	-0.43	-0.06	0.75
Storebrand Verdi	2.05	1.52	0.99	-0.06	0.12	0.02	0.94
Swedbank Generator	1.35	0.30	1.25	0.13	-0.20	0.05	0.82
Terra Norge	-0.84	-0.44	1.06	0.10	-0.16	-0.09	0.92
Terra Vekst	-10.36	-1.26	1.01	0.31	-0.41	0.20	0.84
VÅR Aksjefond	1.74	0.36	1.09	0.06	0.19	0.04	0.94

Appendix G

Table G.I
Sensitivity Analysis

Panel A (B) of this table shows the results of the cross-sectional bootstrap of Norwegian mutual fund performance when the sample is restricted to only include funds that have at least 36 (24) monthly net return observations during the sample period from 1983 to 2014. Panel C reports the results for the full sample. The first row in each panel reports the actual (estimated) t-statistics of alpha for various points and percentiles of the performance distribution, ranging from worst fund (bottom) to best fund (top). The second row reports the associated alpha (annualized) for these t-statistics. Row 3 reports the bootstrapped p-values of the t-statistics of alpha based on 10,000 bootstrap resamples. Funds are ranked according to their unconditional Carhart (1997) four-factor t-statistic of alpha from the model as follows:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{1i} \cdot MKT_t + \beta_{2i} \cdot SMB_t + \beta_{3i} \cdot HML_t + \beta_{4i} \cdot PR1YR_t + \varepsilon_{i,t}$$

Panel A: 36 $\geq$ Observations										
	Bottom Fund	2^{nd}	3^{rd}	Bottom 5%	Bottom 10%	Top 10%	Top 5%	3^{rd}	2^{rd}	Top Fund
t-alpha	-3.59	-3.20	-3.19	-2.61	-2.17	1.37	1.80	2.37	2.70	3.03
Alpha	-14.25	-7.84	-19.32	-6.20	-4.32	1.86	4.90	3.98	6.58	5.03
p-tstat	0.026	0.003	0.003	0.000	0.000	0.303	0.249	0.063	0.053	0.136
Panel B: 24 $\geq$ Observations										
	Bottom Fund	2^{nd}	3^{rd}	Bottom 5%	Bottom 10%	Top 10%	Top 5%	3^{rd}	2^{nd}	Top Fund
t-alpha	-3.59	-3.20	-3.19	-2.61	-2.17	1.37	1.80	2.37	2.70	3.03
Alpha	-14.25	-7.84	-19.32	-6.20	-4.32	1.86	4.90	3.98	6.58	5.03
p-tstat	0.026	0.004	0.004	0.000	0.000	0.306	0.255	0.071	0.057	0.140
Panel C: Full Sample										
	Bottom Fund	2^{nd}	3^{rd}	Bottom 5%	Bottom 10%	Top 10%	Top 5%	3^{rd}	2^{nd}	Top Fund
t-alpha	-3.59	-3.20	-3.19	-2.61	-1.93	1.33	1.80	2.37	2.70	3.03
Alpha	-14.25	-7.84	-19.32	-6.20	-6.71	2.05	4.90	3.98	6.58	5.03
p-tstat	0.036	0.005	0.005	0.000	0.000	0.415	0.261	0.091	0.075	0.175

Appendix H

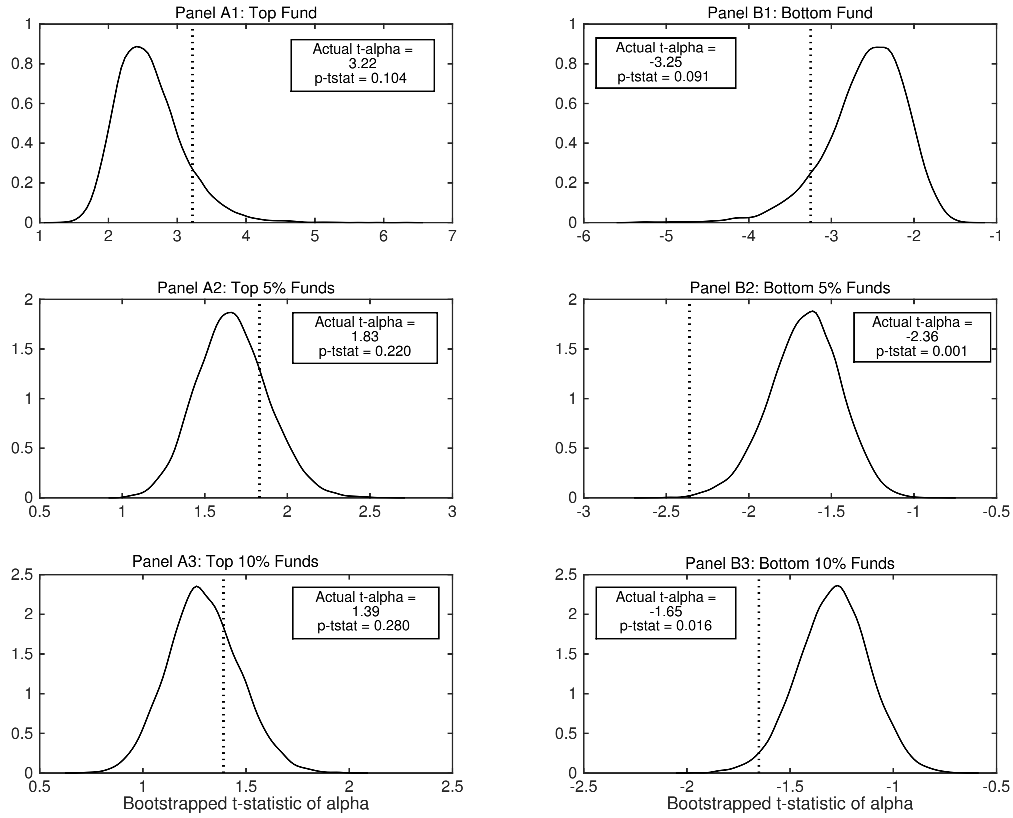

Figure H.1. Actual (estimated) t-statistics of alpha vs. bootstrapped t-statistics of alpha distributions for individual funds using the conditional four-factor model. This figure plots kernel density estimates of the bootstrapped conditional four-factor t-statistic of alpha distribution (solid line) for various percentile points in the cross-section. The x-axis shows the t-statistic of alpha (performance measure) and the y-axis the kernel density estimate. The vertical dotted line shows the actual (estimated) fund t-statistic of alpha. Panel A1-A3 (B1-B3) show marginal funds in the right (left) tail of the performance distribution. For example, “Top 5%” in Panel A2, refers to the marginal alpha t-statistic at the top 5 percentile of the performance distribution. The bootstrapped distributions of t-statistics of alpha (under $H_0 : \hat{t}_{\hat{\alpha}} = 0$) is based on 10,000 bootstrap resamples. Funds are ranked on their conditional four-factor t-statistic of alpha, where the highest ranked fund has the highest t-statistic over the sample period.

Appendix I

This appendix describes the maximum likelihood derivation of the coefficients in the Markov regime-switching model following Perez-Quiros and Timmermann (2000) and Kosowski (2011). We start by denoting the mixture density of $r_{i,t}$ in the model as $\phi(r_{i,t}|\Omega_{t-1};\theta)$. $r_{i,t}$ is conditional on Ω_{t-1}, where Ω_{t-1} represents an *information set* of the respective variables X_{t-1}, r_{t-1}, and z_{t-1} in addition to lagged values of these variables, i.e. $\Omega_{t-1} = \{X_{t-1}, r_{t-1}, z_{t-1}, \Omega_{t-2}\}$. Further, the respective state density functions is conditional on the state, i.e. $\eta(r_t|\Omega_{t-1}; S_t = j; \theta)$. By using the respective state probabilities as weights, and letting θ represent a vector of parameters in the likelihood function of the data, the mixture density can be obtained as follows:

$$\phi(r_{i,t}|\Omega_{t-1};\theta) = \sum_{j=1}^{k} \eta(r_t|\Omega_{t-1}; S_t = j; \theta) Pr(S_t = j|\Omega_{t-1};\theta) \tag{I1}$$

where $\Pr(S_t = j|\Omega_{t-1};\theta)$ is the conditional probability of being in state j at time t, given the available information at time $t-1$. Under assumptions about the state densities of the innovations, ϵ_t, and a law specifying how the state evolves over time, the parameters of this model can be obtained by maximum likelihood estimation. The remainder of this appendix describes in detail the specifications in arriving at the final maximum likelihood. For now, let us decompose the log-likelihood (LL) function as follows:

$$LL(r_T, r_{T-1}, \cdots, r_1; \theta) \tag{I2}$$

We assume that the state densities, $\eta(r_t|\Omega_{t-1}; S_t = j; \theta)$, are Gaussian (i.e. follows a continuous probability distribution). Within each state, we enter (linearly) the variable depending on the performance measure (X_t) in the excess return equation. However, we allow the coefficients to vary between states for $j = 1, \cdots, k$ so that:

$$\eta(r_t|\Omega_{t-1}; S_t = j; \theta) = \frac{1}{\sqrt{2\pi\sigma_j}} exp \frac{-(r_t - \alpha_j - \beta_j' X_t)^2}{2\sigma_j^2} \tag{I3}$$

That is, we estimate the likelihood by weighting across the states by multiplying the likelihood with the probability of the respective state. Since mixtures of normal can be approximate a very broad set of density families, this assumption is not very restrictive. Following the law of total probability, the conditional state probabilities can be obtained recursively:

$$\Pr(S_t = i|\Omega_{t-1};\theta) = \sum_{j=1}^{k} \Pr(S_t = i|S_{t-1} = j, \Omega_{t-1};\theta) \Pr(S_{t-1} = j|\Omega_{t-1};\theta) \tag{I4}$$

According to Bayes' rule, the conditional state probabilities can then be derived as follows:

$$\Pr(S_{t-1} = j|\Omega_{t-1};\theta) = \Pr(S_{t-1} = j|r_{t-1}, X_{t-1}, z_{t-1}, \Omega_{t-2};\theta) = \tag{I5}$$

$$\frac{\eta(r_{t-1}|S_{t-1} = j, X_{t-1}, z_{t-1}, \Omega_{t-2};\theta)\ \Pr(S_{t-1} = j|X_{t-1}, z_{t-1}, \Omega_{t-1};\theta)}{\sum_{j=1}^{k} \eta(r_{t-1}|S_{t-1} = j, X_{t-1}, z_{t-1}, \Omega_{t-2};\theta)\ \Pr(S_{t-1} = j|X_{t-1}, z_{t-1}, \Omega_{t-2};\theta)}$$

Equations I4 and I5 can be iterated to derive the state probabilities $\Pr(S_t|\Omega_{t-1};\theta)$ by using recursive estimation, and the parameters of the likelihood function can finally be obtained.

Appendix J

Table J.I

Regime-Switching Estimates of Fama-French's (1993) Three-Factor Alpha

This table reports regime-switching estimates of fund performance based on Fama and French's (1993) three-factor model. The table includes estimates when regime-switching is applied on two-state models using both fixed transition probabilities (Model I) and time-varying transition probabilities (Model II). Rows 1 to 5 report mean parameters, which includes alpha (in percent per year) and the three risk factors in Fama and French's model, MKT, SMB, and HML. Row 6 reports variance parameters in percent per month, and Rows 7 and 8 report the transition probability parameters. Row 9 show the log-likelihood estimates of each three models. The t-statistics are reported in parentheses below each estimate. The sample period is 1983 to 2014. The following regime-switching Fama-French (1993) three-factor model is estimated for the equally-weighted portfolio of all funds:

$$r_{i,t} - r_{f,t} = \alpha_{i,S_t} + \beta_{1i,S_t} \cdot MKT + \beta_{2i,S_t} \cdot SMB_t + \beta_{3i,S_t} \cdot HML_t + \varepsilon_{i,t}$$

where $\varepsilon_t \sim N(0, \sigma_{s_t})$

Column 1 reports single state estimates, whereas Columns 2 (Model I) and 3 (Model II) report estimates in the two states. The transition probabilities in model I are fixed, and the time-varying transition probabilities in Model II depend on the two month lagged change in the Composite Leading Indicator ΔCLI_{t-2}:

Model I: $p_t{=}\Pr(S_t{=}1|S_{t-1}{=}1){=}\phi(c_1)$, $q_t{=}\Pr(S_t{=}2|S_{t-1}{=}2){=}\phi(c_2)$

Model II: $p_t{=}\Pr(S_t{=}1|S_{t-1}{=}1){=}\phi(c_1{+}\pi_1\Delta CLI_{t-2})$, $q_t{=}\Pr(S_t{=}2|S_{t-1}{=}2){=}\phi(c_2{+}\pi_2\Delta CLI_{t-2})$

where $\phi(.)$ is the cumulative density function of a standard normal variable

	Single State Model	Model I		Model II	
Mean Parameters		State 1	State 2	State 1	State 2
Alpha (in pct p.a.)	-0.569	-1.352	0.292	-1.535	0.359
	(-0.54)	(-1.34)	(0.53)	(-1.61)	(0.72)
MKT	0.997***	1.001***	0.979***	1.001***	0.980***
	(66.41)	(66.03)	(15.68)	(69.50)	(15.78)
SMB	0.097***	0.074***	0.168**	0.077***	0.156**
	(4.54)	(3.58)	(2.07)	(3.90)	(2.24)
HML	-0.049	-0.062***	-0.012	-0.063***	-0.009
	(-2.81)	(-3.79)	(-0.15)	(-3.95)	(-0.12)
Variance Parameters					
S.D. (in pct. p.m.)	1.683***	1.174***	2.954***	1.175***	2.960***
	(113.04)	(53.48)	(28.14)	(68.73)	(31.47)
Transition Probability Parameters					
c_s		3.142***	1.595***	3.287***	1.672***
		(5.29)	(3.13)	(5.81)	(3.34)
$d_s\Delta(CLI)_{t-2}$				0.289	-0.049
				(1.30)	(-0.22)
Log Likelihood	1020.9	1061.6		1063.9	

Table J.II
Regime-Switching Estimates of Jensen's (1968) Single-Factor Alpha

This table reports regime-switching estimates of fund performance based on Jensen's (1968) single-factor model. The table includes estimates when regime-switching is applied on two-state models using both fixed transition probabilities (Model I) and time-varying transition probabilities (Model II). Rows 1 to 5 report mean parameters, which includes alpha (in percent per year) and the risk factor in Jensen's model, MKT (i.e $r_M - r_f$). Row 6 reports variance parameters in percent per month, and Rows 7 and 8 report the transition probability parameters. Row 9 shows the log-likelihood estimates of each three models. The t-statistics are reported in parentheses below each estimate. The sample period is 1983 to 2014. The following regime-switching Jensen (1968) single-factor model is estimated for the equally-weighted portfolio of all funds:

$$r_{i,t} - r_{f,t} = \alpha_{i,S_t} + \beta_{1i,S_t} \cdot MKT + \varepsilon_{i,t}$$

where $\varepsilon_t \sim N(0, \sigma_{s_t})$

Column 1 reports single state estimates, whereas Columns 2 (Model I) and 3 (Model II) report estimates in the two states. The transition probabilities in model I are fixed, and the time-varying transition probabilities in Model II depend on the two month lagged change in the Composite Leading Indicator ΔCLI_{t-2}:

Model I: $p_t = \Pr(S_t=1|S_{t-1}=1) = \phi(c_1)$, $q_t = \Pr(S_t=2|S_{t-1}=2) = \phi(c_2)$
Model II: $p_t = \Pr(S_t=1|S_{t-1}=1) = \phi(c_1 + \pi_1 \Delta CLI_{t-2})$, $q_t = \Pr(S_t=2|S_{t-1}=2) = \phi(c_2 + \pi_2 \Delta CLI_{t-2})$

where $\phi(.)$ is the cumulative density function of a standard normal variable

	Single State Model	Model I		Model II	
Mean Parameters		State 1	State 2	State 1	State 2
Alpha (in pct p.a.)	0.391	-0.066	0.474	-0.063	0.499
	(0.28)	(-0.63)	(0.95)	(-0.62)	(0.93)
MKT	0.965***	0.978***	0.925***	0.979***	0.920***
	(67.78)	(66.53)	(15.21)	(66.26)	(14.20)
Variance Parameters					
S.D. (in pct. p.m.)	1.751***	1.254***	3.045***	1.275***	3.114***
	(111.94)	(57.48)	(26.79)	(58.90)	(25.90)
Transition Probability Parameters					
c_s		3.259***	1.753***	3.473***	1.821***
		(5.31)	(3.20)	(4.66)	(3.01)
$d_s \Delta(CLI)_{t-2}$				0.164	-0.042
				(0.53)	(0.14)
Log Likelihood	1005.97	1042.3		1043.9	