Norwegian School of Economics
Bergen, Fall 2020

Morningstar Ratings and Norwegian Mutual Fund Performance

An empirical study of the Morningstar rating system as a predictor of performance for mutual funds investing primarily in Norwegian equity

Sara Jahr Kirkeby

Supervisor: Jørgen Haug

Master thesis, Economics and Business Administration

Major: Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This thesis would not have been possible to write without valuable insights from several contributors. First and foremost, I would like to thank my master thesis supervisor, Jørgen Haug, for providing excellent guidance, insights, and constructive criticism throughout the process of writing the thesis. I would also like to direct a thank you to Steinar Eikeland at Industrifinans for his valuable insights on fund investing from a practitioner's point of view. Furthermore, I would like to thank the people at Morningstar for providing data and answering my questions about the Morningstar rating methodology. Lastly, I would like to thank NHH for motivational and educational courses on the topics of finance and programming, providing the necessary academic background for this thesis. I am grateful that NHH, in recent years, has started offering multiple courses in programming. Based on these courses, I was able to conduct my analysis in R and write the thesis in Latex.

Norwegian School of Economics

Bergen, December 2020

Sara Jahr Kirkeby

Sara Jahr Kirkeby

Abstract

This thesis examines the Morningstar rating system as a predictor of mutual fund performance for mutual funds investing primarily in Norwegian equity. The predictive abilities are examined using a data set free of survivorship bias, which contains data on 136 mutual funds from January 2002 to December 2019. Mutual fund performance is evaluated using estimates of alpha from three different factor models of performance measurement: the Capital Asset Pricing Model (CAPM), Fama and French 3-factor model, and Carhart 4-factor model. To comprehensively test for predictive abilities in the rating system, both random effects panel data regressions and strategies of buying historically top-rated ("winners") versus low-rated ("losers") funds are employed. The results indicate findings that are robust across different performance measures and styles of funds. First, for the period before the financial crisis, from January 2002 to March 2008, low ratings from Morningstar generally indicate relatively poor future performance. Second, for the period during and after the financial crisis, from April 2008 to December 2019, low ratings from Morningstar, on the contrary, indicate relatively high future performance. Hence, in the period before the financial crisis, investors could invest in past winners to generate higher returns than an investment in past losers would have yielded. Contrary, during and after the financial crisis, the loser strategies outperformed the winner strategies. Third, there is little statistical evidence that Morningstar's highest-rated funds outperform the next-to-highest, median, and next-to-lowest rated funds in both periods.

Contents

List of Figures

List of Tables

1 Introduction

Norwegians are increasingly using mutual funds as their financial investment of choice. According to a recent survey conducted by Opinion for the Norwegian Fund and Asset Management Association (VFF, 2020), 40% of all Norwegians have money invested in mutual funds. This corresponds to approximately 1.7 million Norwegians over the age of 18 and is the highest proportion since the annual survey started in 2001, when only 26% of all Norwegians invested in mutual funds. This dramatic increase has lead to a corresponding increase in the demand for information on mutual funds. Several rating agencies provide such information and investment advice for mutual funds, and one of the most popular, if not the most popular, is Morningstar. Morningstar ranks mutual funds on a scale of one to five stars, where one star is the worst rating and five stars are the best. Because of the rating system's simplicity, the ratings have been given increased attention since their inception in Norway in 2001. The ratings are heavily used in mutual fund advertising, suggesting that mutual fund companies believe that investors care about Morningstar ratings. In fact, in some cases, the only mention of return performance in the mutual fund advertisement is the Morningstar rating. Moreover, the importance of Morningstar ratings has been highlighted by several high-profile publications (e.g., Blume (1998) and Sharpe (1998)) that have investigated the underlying properties of the Morningstar rating system.

Evidence from academic research also indicates that investors care much about Morningstar's star ratings. Both Sirri and Tufano (1998) and Del Guercio and Tkac (2008) find that the Morningstar rating itself has a significant effect on fund flows. They find evidence that higher-rated funds experience cash inflows that are far greater in size than the cash outflows experienced by lower-rated funds. More specifically, Del Guercio and Tkac (2008) find that a fund's initial five-star rating produces inflows of 53% above the ordinarily expected inflow. In contrast, funds with rating downgrades experience significant outflows beyond what would typically be expected.

Given the inherent interest in the Morningstar ratings by investors, this thesis examines the following key question:

Can the Morningstar ratings be used to predict future risk-adjusted performance for mutual funds investing primarily in Norwegian equity?

This is an important question because many investors use Morningstar's rating system to decide which funds to invest in. Also, examining performance across funds grouped by Morningstar rankings will indicate if the large documented cash inflows for top-rated funds are justified by subsequent relative performance. The research question will be answered using a data set free of survivorship bias, which contains data on 136 mutual funds[1] from January 2002 to December 2019.

An outline of the thesis follows. First, Section 2 presents literature that is related to the research question of the thesis. Section 3 presents the Morningstar rating methodology, and Section 4 describes the data. Section 5 further presents the methodology used to measure the Morningstar rating's ability to predict risk-adjusted mutual fund performance, employing estimates of alpha from common factor models of performance measurement. Section 6 describes the empirical analysis and results. Finally, the main findings are summarized in a final conclusion in Section 7.

[1]The number of 136 mutual funds includes the number of share classes of each fund in the sample, and is the total number of fund observations in the sample.

2 Related literature

This section presents literature that is related to the research question of the thesis. First, I present related literature conducted in the U.S. Second, I describe related Norwegian literature.

2.1 U.S. related literature

Given the importance of Morningstar ratings to investors, a considerable amount of research has been conducted in the U.S. on whether the star ratings have predictive ability. These papers include Khorana and Nelling (1998), Blake and Morey (2000), Morey (2002), and Morey and Gottesman (2006). The findings of these papers suggest that low-rated funds generally indicate relatively poor future performance, measured by various risk-adjusted return measures. However, the papers are more ambiguous regarding associating top-rated funds with superior future performance.

Khorana and Nelling (1998) examine the determinants and predictive ability of the Morningstar ratings in the U.S. by assessing the degree of persistence in the fund ratings. They analyze the persistence by comparing a group of funds in December 1992 to the June 1995 rating of those same funds and find evidence of persistence over these thirty months. In other words, they find that highly rated funds tend to continue to be successful in the future, and poorly rated funds continue to underperform relative to their peers.

Blake and Morey (2000) further examine the Morningstar rating system as a predictor of mutual fund performance for U.S. domestic equity funds using a data set free from survivorship bias. The main result of their investigation is that low Morningstar ratings generally indicate relatively poor future performance. At the same time, they find little statistical evidence that Morningstar's highest-rated funds outperform the next-to-highest and median funds. The results are robust across different samples, ages and styles of funds, and performance measures. Blake and Morey's (2000) results suggest that investors should be very cautious about associating a highly rated fund with superior future performance. Morey (2002) supports Blake and Morey's findings of poor predictability.

The Morningstar rating system is revised periodically. Morey and Gottesman (2006) perform an extensive examination of how the 2002-revised Morningstar rating system predicts future performance. The results are quite different from those of Blake and Morey (2000) and Morey (2002). Morey and Gottesman (2006) find support for the notion that the 2002-revised Morningstar rating system can predict future performance, at least in the first three years out of sample. Specifically, they find that higher-rated funds significantly outperforms lower-rated funds. After 2002, the Morningstar rating methodology was revised in 2006 and 2016. However, the building blocks of the rating methodology remains the same. To my knowledge, no extant studies have examined the predictability of Morningstar's ratings with the newest 2016 methodology. This thesis is thus the first to examine predictability with the newest methodology.

2.2 Norwegian related literature

Limited research has been conducted on the predictability of the Morningstar ratings for mutual funds investing primarily in Norwegian equity. This thesis is trying to fill that void.

To my knowledge, the only extant academic study on the predictive abilities of the Morningstar ratings for Norwegian equity mutual funds is the study by Aasheim (2013). However, the setup of Aasheim's (2013) study differs from this thesis. First, Aasheim's (2013) study is not restricted to funds which primarily invest in Norwegian equity, but rather an examination of all equity mutual funds available for sale in Norway. Also, Aasheim (2013) examines the Morningstar ratings' predictive abilities for fund performance measured by raw returns, which are not adjusted for risk. In contrast, this thesis examines risk-adjusted returns. Nevertheless, Aasheim (2013) finds that the Morningstar ratings lack predictive power based on raw returns for this set of data.

Dahl and Madsen (2017) examine another rating system's predictability: Dine Penger's Norwegian mutual fund ratings. They find evidence that investors were more likely to benefit from the rating system before the financial crisis than after, as the rating system possessed better predictive abilities measured by risk-adjusted return measures. Using a somewhat similar methodology as Dahl and Madsen (2017), it remains interesting to examine whether this also applies to the Morningstar ratings.

3 Morningstar rating methodology

Since this thesis studies the Morningstar ratings, it is important with an understanding of how the Morningstar ratings are calculated. Hence, this section describes the Morningstar rating methodology.

3.1 Morningstar Inc.

Morningstar Inc., established in 1984, is a world-leading provider of investment research. The company collects and provides comprehensive data and insights on stocks, bonds, funds, and general market data to individual and institutional investors, financial advisors, retirement plan providers, and asset managers.

3.2 Morningstar rating system

Morningstar is known for its rating system based on rating funds on a one- to five-star scale. The stars reflect how funds have performed historically on a risk-adjusted basis relative to other funds in the same investment category (Morningstar, 2016). The one-star funds are considered the worst performers on a risk-adjusted basis, and the five-star funds are the best performers.

As of the 2016-revised Morningstar rating methodology, the Morningstar rating calculation can be divided into the following three steps.

3.2.1 Peer groups

First, the funds ranked by Morningstar are grouped into peer groups. These groups are generally based on a fund's geographical investment exposure: which country or countries a fund primarily invests in, and the types of securities a fund primarily invests in. Furthermore, category membership can, in addition, in some geographical areas, be based on a fund's long-term or "normal" style profile. However, all the mutual funds studied in this thesis have "Norway Equity" as their Morningstar Category or peer group. The Morningstar Category Index for this group of funds is the Oslo Stock Exchange Mutual Fund Index (OSEFX). This suggests that Morningstar does not consider factor risk-exposures, relative to, e.g., a Fama-French 3-factor model, to determine peer groups for mutual funds investing primarily in Norwegian equity.

3.2.2 Morningstar risk-adjusted return

The second step of the Morningstar rating calculation is to compute the Morningstar Risk-Adjusted Return (MRAR). To compute the MRAR, Morningstar first calculates monthly total returns for the funds, as shown in equation 3.1.

$$TR_t = \left\{ \frac{P_e}{P_b} \prod_{i=1}^{n} \left(1 + \frac{D_i}{P_i} \right) \right\} - 1 \tag{3.1}$$

where

TR_t	=	total return for the fund for month t
P_e	=	end of month net asset value (NAV) per share
P_b	=	beginning of month NAV per share
D_i	=	per share distribution at time i
P_i	=	reinvestment NAV per share at time i
n	=	number of distributions during the month

Next, Morningstar adjusts the funds' monthly returns for the risk-free rate (1-month NIBOR for Norwegian equity funds) to get Morningstar Return, as shown in equation 3.2.

$$MorningstarReturn = \left[\prod_{t=1}^{T} \left(1 + \left(\frac{1+TR_t}{1-RF_t} - 1 \right) \right) \right]^{\frac{12}{T}} - 1 = \left[\prod_{t=1}^{T} (1+ER_t) \right]^{\frac{12}{T}} - 1 \tag{3.2}$$

where

ER_t	=	geometric excess return for the fund for month t
TR_t	=	total return for the fund for month t
RF_t	=	total return for the risk-free rate for month t
T	=	number of months in the period (for example three, five or ten years)

Further, the Morningstar Return is adjusted for risk to get MRAR, as shown in equation 3.3. In the MRAR calculation, Morningstar uses expected utility theory to model how investors trade off return and risk. In this framework, a return distribution with high expected return and low risk is preferable to low expected return and high risk. "Gamma" (γ) represents the degree of risk aversion in the framework. When $\gamma = 0$, the investor is indifferent between a riskless choice and a risky choice as long as the average geometric expected return is the same. When γ is less than zero, the investor is risk-loving. However, when γ is larger than zero, the investor is risk-averse. Morningstar has concluded that $\gamma = 2$ results in fund rankings consistent with the risk tolerances of typical investors. Hence, Morningstar applies a γ equal to two in the calculation of its star ratings. Applying $\gamma = 2$ to the rating system means that the rating system provides a penalty for risk-taking. For example, when $\gamma = 2$, two funds with identical average returns can receive different ratings because of different risk (volatility). The fund with the highest risk will receive a lower rating than the fund with the lowest risk. If the rating system were based solely on returns, when $\gamma = 0$, both funds would be assigned the same star rating.

$$MRAR = \left[\frac{1}{T}\sum_{t=1}^{T}(1+ER_t)^{-\gamma}\right]^{-\frac{12}{\gamma}} - 1 = \left[\frac{1}{T}\sum_{t=1}^{T}(1+ER_t)^{-2}\right]^{-\frac{12}{2}} - 1 \qquad (3.3)$$

where

$MRAR_t$	=	Morningstar risk-adjusted return for the fund for month t
ER_t	=	geometric excess return for the fund for month t
T	=	number of months in the period (for example three, five or ten years)
γ	=	degree of risk aversion

MRAR is calculated differently than many other methods. For example, Modern Portfolio Theory uses standard deviation as a unit of risk. In contrast, MRAR gives more weight to downside risk and does not make any assumptions about the distribution of excess returns. With this Morningstar methodology, funds that experience fewer heavy losses tend to look better than they would by merely using standard deviation. Moreover, compared with asset pricing models like the Fama-French 3-factor model, the MRAR calculation does not account for risk factors included in such models.

3.2.3 Fund ranking

In the third and last step of the Morningstar rating calculation, Morningstar ranks all funds in a category or peer group using MRAR. The funds with the highest MRAR scores receive the most stars. Funds in the top 10% of risk-adjusted scores receive a five-star rating, the next 22.5% receive a four-star rating, the middle 35% receive three stars, and those in the last two categories (two-star and one-star rating) represent the next 22.5%

and 10%.

Morningstar requires a fund to have at least three years of return data to receive a rating. The rating is calculated for three years, five years, and ten years, and the overall rating is a weighted average of the time-period ratings. Table 3.1 shows the weightings for different time horizons. For example, for funds with a ten-year return history, the ten-year rating is given a 50% weighting, the five-year rating is given a 30% weighting, and the three-year rating is given a 20% weighting in the overall rating calculation. Another example is a fund with a four-year return history, where the rating is solely based on the last three years.

Table 3.1: Overall rating weightings (Morningstar, 2016)

Table 3.1 shows the weightings of the Morningstar ratings for different time horizons.

Months of Total Returns	Overall (Weighted) Morningstar Rating
36-59	100% three-year rating
60-119	60% five-year rating 40% three-year rating
120 or more	50% ten-year rating 30% five-year rating 20% three-year rating

4 Data

In order to answer the question of whether the Morningstar ratings can be used to predict future mutual fund performance measured by risk-adjusted return, data on Norwegian equity mutual funds, benchmark indexes, risk-free interest rates, and factors are required. This section presents the data obtained.

4.1 Norwegian equity mutual fund data

The data in this thesis are obtained from Morningstar Direct, which has one of the most comprehensive databases on mutual funds internationally. The sample contains data on 136 Norwegian equity mutual funds from January 2002 to December 2019[2]. Morningstar started to rate Norwegian equity mutual funds in October 2001. Therefore, to analyze the data on a yearly basis, the data sample is limited to the 2002-2019 period. The data sample can be characterized as panel data, as it includes observations of multiple funds over multiple periods. It includes monthly data on rating, returns, and fund size for each fund in the sample from January 2002 to December 2019. Some basic descriptive statistics are reported in table 4.1. The second and third column of table 4.1 shows the number of fund observations at the beginning (January) and the end (December) of each year. Furthermore, the fourth and fifth column shows the number of new funds that emerged each year and the number of funds that ceased to exist, those that "died", during each year. Column six through eight show average rating, return, and fund size per year. Table 4.1 suggests that the average fund size grew significantly during the early 2000s, indicating a positive growth of Norwegian equity funds. In 2003 and 2008, there was a massive decline in fund size compared to 2002 and 2007. The average fund size decreased by approximately 40% from 2002 to 2003 and 30% from 2007 to 2008. As table 4.1 shows, the reductions were due to the large negative returns after the dot-com bubble and during the financial crisis, not because of a massive withdrawal of funds. The average rating of the funds is roughly equal to three for all the years. This is consistent with Morningstar's rating methodology, where most of the funds should receive a three-star rating (the middle 35% of risk-adjusted scores receive three stars).

[2]See table A1.1 in the appendix for a detailed overview of the fund list. The fund list contains in total 136 fund observations included all share classes of each fund.

Table 4.1: Descriptive statistics for the sample funds

Table 4.1 shows descriptive statistics per year for the mutual funds in the data sample. The sample period is January 2002 to December 2019. Column two shows the number of funds in the sample that existed at the beginning (January) of each year, while column three shows the number of funds that existed at the end (December) of each year. Column four and five shows the number of new funds that emerged each year and the number of funds that ceased to exist, those that "died", during each year. Column six through eight show average rating, return, and fund size per year. The return is expressed in percentage terms, while the fund size is reported in million NOK.

Year	Number funds beg. year	Number funds end year	Number new funds	Number funds "died"	Avg. rating	Avg. return (%)	Avg. fund size (mill NOK)
2002	40	42	2	0	3.18	-3.11	747
2003	42	46	4	0	2.91	3.81	427
2004	46	48	2	0	2.95	2.87	623
2005	50	52	4	2	3.03	3.38	1,347
2006	52	52	1	1	3.13	2.35	1,618
2007	52	52	1	1	3.22	1.00	1,818
2008	52	54	2	0	3.19	-5.53	1,278
2009	54	56	2	0	3.20	4.79	1,305
2010	56	58	2	0	3.19	1.87	1,717
2011	58	61	3	0	3.05	-1.64	1,745
2012	61	63	4	2	2.97	1.35	1,594
2013	63	66	7	4	3.02	1.82	1,766
2014	66	61	2	7	3.15	0.65	2,090
2015	59	62	5	2	3.17	0.60	2,064
2016	63	63	2	2	3.19	1.47	2,057
2017	63	67	9	5	3.22	1.24	2,959
2018	69	72	4	1	3.10	-0.32	3,349
2019	73	82	13	4	3.11	1.50	3,349

The database's fund returns are computed each month by taking the change in monthly net asset value (NAV), reinvesting all income and capital-gains distributions during that month, and dividing by the starting NAV. The monthly returns between time t_0 and t_1, denoted $R(t_0 \to t_1)$, are thus computed according to equation 4.1. Morningstar does not adjust the returns for broker commissions, such as front-end or back-end loads, or other costs associated with purchases and sales of mutual funds. However, the returns account for management and administrative fees and other costs automatically taken out of fund assets.

$$R(t_0 \to t_1) = \frac{NAV(t_1)}{NAV(t_0)} \left(1 + \frac{D_i}{NAV_i}\right) - 1 \tag{4.1}$$

where

$NAV(t_1)$	=	current month-end NAV
$NAV(t_0)$	=	previous month-end NAV
D_i	=	amount of distribution during month i
NAV_i	=	reinvestment NAV per share at time i

4.1.1 Criteria for including funds

It is essential to ensure that the fund sample is consistent with the intention of the thesis. For this, the sample of mutual funds is completed through various filters in Morningstar's mutual fund database. First, all mutual funds are open-end funds, meaning that shares in the funds can be issued and redeemed at any time. Second, the sample is restricted to mutual funds that invest primarily in Norwegian equity. Thus, the Morningstar Category has to be "Norway Equity". The funds in this category are required to have at least 75% of total assets in equities and at least 75 % of equity assets in Norwegian equities (Morningstar, 2019). By restricting the sample to funds that invest primarily in Norwegian equities, more accurate benchmark returns can be computed when assessing the fund's risk-adjusted performance. Furthermore, the base currency is set to Norwegian Krone (NOK), for the full sample, to exclude funds listed multiple times with different currencies. Finally, the sample funds have to be available for sale in Norway, as all the funds in the sample should be available for investors to purchase at the time of the Morningstar evaluation.

4.1.2 Sample adjustments

It is further necessary to discuss whether the sample needs some adjustments to ensure the analysis's preciseness.

First, the fund list contains funds with different share classes. Mutual fund companies can have multiple classes of shares for a particular fund. The difference between the share classes is the fee structure. One example from the data sample is Pareto Investment Fund A, B, and C. Pareto Investment Fund A, B, and C require a minimum investment of 500 NOK, 10 million NOK, and 50 million NOK. The funds further have fixed fees of respectively 1.80%, 0.95%, and 0.50%. Since the share classes have different expense ratios, the ratings of the different share classes of mutual funds can differ. Therefore, to account for all rating variations, all share classes of each fund are included in the fund sample.

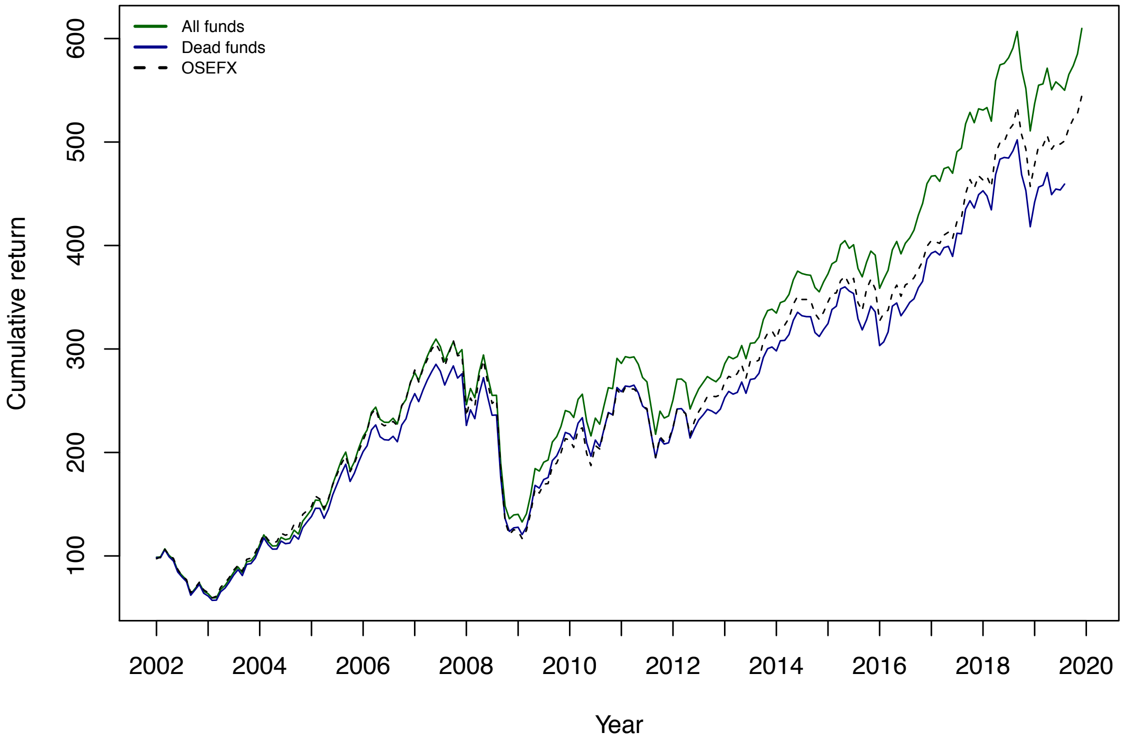
Second, previous studies like Kosowski et al. (2006) exclude funds that only have existed for a short period to avoid a large number of funds with short return histories. According to Sørensen (2009), unreported results show that such a filter induces a survivorship bias in the Norwegian data. The advantage of not removing any funds is thus that the data sample is free of survivorship bias. However, including funds with few observations can have consequences for the preciseness of the estimated regression. It is a trade-off between a long return history and mitigating survivorship bias. Nevertheless, including returns of short-lived funds is essential to gain an accurate and survivorship bias free understanding of fund performance. Therefore, no short-lived funds are excluded from the analysis. However, funds without any return history at all are excluded.

4.1.3 Survivorship bias

A commonly known problem when analyzing fund performance is the survivorship bias, described by Brown et al. (1992), Elton et al. (1996), and others. Survivorship bias occurs when the returns of funds that no longer exist, those that have "died" during the sample period, are removed from the sample. The "dead" funds have typically had poor performance, so their removal produces an unrealistically high estimate of aggregate mutual fund performance. Therefore, such exclusion would yield misleading results. Figure 4.1 illustrates the problem of survivorship bias for the data used in this thesis. The figure shows the cumulative return on 100 NOK invested in equal-weighted indexes of fund returns from 2002 to 2019. The solid green line shows the cumulative return of all funds that existed at some point over the sample period, and the solid blue line shows the return of the funds that no longer exist. As expected, the graph for the "dead" funds is everywhere below the graph for all the funds, suggesting that the surviving funds' returns are higher than the returns of those funds that "died" during the sample period. This suggests the presence of survivorship bias. To avoid the problem of survivorship bias, all the "dead" funds are included in the analysis.

Figure 4.1: Equal-weighted cumulative returns for all funds and dead funds

Figure 4.1 plots the cumulative return on 100 NOK invested in equal-weighted indexes of all the sample funds from 2002 to 2019. The solid green line plots the cumulative indexed return of all Norwegian equity mutual funds that existed at some point over the sample period. The solid blue line plots the cumulative indexed return of the funds that no longer exist. As a reference, the dotted line plots the OSEFX benchmark index.



4.2 Benchmark indexes

In order to measure relative mutual fund performance, an appropriate benchmark index is required. This thesis employs one single benchmark index as the market return for all the sample funds. Another possibility could have been to use each fund's benchmark index, as stated in the fund's prospectus, as the market return for the respective fund. The main reason for using the same index for all funds is to achieve comparability of the funds' performance, which is essential in this paper's analysis. However, when measuring funds with a specific benchmark index, e.g., index funds tracking a specific index, against another index, the relative fund performance could represent an alternative estimation.

The key question is which benchmark index should be used to measure all the sample funds' performance against. Most mutual funds in the data sample have the Oslo Stock Exchange Mutual Fund Index (OSEFX) as their benchmark index. However, some of the funds in the sample use the Oslo Stock Exchange Benchmark Index (OSEBX), the Oslo

Stock Exchange Small Cap Index (OSESX), or the OBX Total Return Index (OBX) as their benchmark index. The "Landkreditt" funds[3] in the sample do not have a benchmark index. However, these funds are measured against both OSEFX and OSEBX in their fund presentations. The Oslo Stock Exchange All Share Index (OSEAX) can also be considered a proxy for the market return, as it consists of all shares listed on the Oslo Stock Exchange.

First, the OBX index is not an alternative because it is a price index that does not account for dividends. Morningstar's return calculation accounts for dividends, and one cannot compare a return accounting for dividends against a return not accounting for dividends. Second, the OSESX is neither an alternative because it spans an entirely different universe, in which only small-cap stocks are considered. Thus, only the OSEFX, OSEBX, and OSEAX remains. Table 4.2 shows descriptive statistics for these benchmark returns, where the return series are retrieved from Morningstar Direct. The table shows that the OSEAX had the highest mean return and lowest standard deviation during the sample period from 2002 to 2019. Even though the mean returns and standard deviations are different, the cross-correlations in table 4.2 show that the various benchmark returns are closely correlated. The correlations are close to 1, suggesting that the different benchmark indexes approximately follow the same pattern. Thus, the choice of benchmark index will probably not drastically impact the results of this paper.

Table 4.2: Descriptive statistics for benchmark returns

Table 4.2 shows the descriptive statistics for the OSEFX, OSEAX, and OSEBX during the sample period from January 2002 to December 2019. The left hand side of the table shows the mean, standard deviation, minimum return and maximum return for the benchmark indexes, all expressed in percentage terms. The mean and standard deviation have been annualized, while the minimum return and maximum return are monthly. The right hand side of the table shows the cross-correlations over the sample period.

					Cross-correlations		
	Mean	St.dev.	Min	Max	OSEFX	OSEAX	OSEBX
OSEFX	11.69	20.66	-27.17	16.52	1		
OSEAX	12.19	18.78	-23.93	15.05	0.97	1	
OSEBX	11.60	19.78	-25.22	15.83	0.99	0.99	1

The OSEFX index is a natural choice because most of the sample funds have the OSEFX as their benchmark index. The OSEFX index is designed to meet specific diversification requirements and comply with the UCITS directives to regulate investments in mutual

[3] Landkreditt Norge, Landkreditt Utbytte A and Landkreditt Utbytte I

funds. More specifically, the OSEFX index is a capped version of the OSEBX index. The capping rules state that the mutual funds' weight in any company must not exceed 10% of the index's total market value, and securities exceeding 5% must not exceed 40% combined. Furthermore, Morningstar uses the OSEFX as their Morningstar Category Index for Norwegian equity mutual funds. This highlights the index's applicability in the study of mutual funds. I thus choose the OSEFX index to measure all sample funds' performance against.

It is worth mentioning that choosing the OSEFX index may impact the interpretation of the other factors that go into the factor models for returns, described in section 5.2. Since the factors are constructed as value-weighted averages (Ødegaard, 2020), the capping of the OSEFX index may not comply with the factors' construction. Thus, the OSEBX could serve as a better proxy for market return. However, with a high correlation of 0.99 between the OSEBX and OSEFX (table 4.2), the choice between these indexes is not paramount. Thus, the choice of using the OSEFX remains.

4.3 Risk-free interest rate

Mutual fund performance in the subsequent sections is evaluated in excess of the risk-free rate. A proxy for the risk-free rate is thus required. Equation 4.2 shows how the excess returns are computed, where r_i is the return of fund i, and r_f is the risk-free rate. Theoretically, an investor should be compensated for bearing risk and earn a positive excess return, i.e., a higher return than the risk-free rate. Otherwise, the investor could invest in a risk-free asset to earn a higher and more certain return.

$$ExcessReturn = r_i - r_f \tag{4.2}$$

For the risk-free interest rate, Nibor is used. Nibor reflects the rate of return banks can borrow from each other. As an approximation, the 1-month Nibor rate is used. Since Morningstar also uses the 1-month Nibor as a proxy for the risk-free interest rate in their rating calculations, this ensures consistency. The Nibor rate in this thesis is retrieved from Norges Bank (2014) before 2013 and the Oslo Stock Exchange (2020) after 2013.

4.4 Factor returns

This thesis measure relative fund performance by comparing a fund's return to the return implied by three different factor models for returns; the Capital Asset Pricing Model (CAPM), the Fama-French 3-factor model, and the Carhart 4-factor model, described in section 5.2. For this, data on the factor returns used in the different factor models are required. The factor returns for the Norwegian market are acquired from Professor Bernt Arne Ødegaard (2020).

5 Methodology

This section presents the methodology used to measure the Morningstar ratings' ability to predict mutual fund performance measured by risk-adjusted return. First, two different sample periods for the analysis are introduced, where the period split is the beginning of the financial crisis. Second, I present three different factor models used to measure performance or risk-adjusted return. These factor models are (1) CAPM, (2) Fama and French 3-factor model, and (3) Carhart 4-factor model. I further describe two methods used to evaluate whether the rating system contains any predictive abilities: (1) the random effects panel data regression and (2) the buying winners versus losers strategy inspired by Jegadeesh and Titman (1993).

5.1 Sample periods

In the subsequent sections, I split the study into two different sample periods, where the split is the beginning of the financial crisis of 2008. The reason for this split is to illustrate the effects of the financial crisis, and because the prediction results before and after the financial crisis significantly differs. Thus, the first sample period is before the financial crisis from January 2002 to March 2008, while the second sample period is during and after the financial crisis from April 2008 to December 2019. The market returns for the two sample periods, here represented by the OSEFX, are shown in figure 5.1. The first sample period is characterized by considerable market growth after the 2000 dot-com bubble until the stock market top around March 2008 before the financial crisis. The second sample period is characterized by a large drop in returns during the financial crisis and growth after the financial crisis, with some market fluctuations throughout the growth period. Figure 5.2 shows the rating distribution of all funds over the two sample periods. This figure yields an indication of whether the ratings are normally distributed. Normality plays a role in the analysis as it is an essential assumption for the random effects panel data regression explained in section 5.3. The figure suggests that the distribution of funds in the rating groups is approximately normal in both periods, with most funds ranked in the three-star category and subsequently fewer funds in the other rating groups. However, slightly more funds are rated as four-star funds and five-star funds than two-star and one-star funds, suggesting a small negative skew. However, I conclude that this is not a significant issue, so there is no need to change the rating groups to achieve normality.

Figure 5.1: Market returns over the sample periods

Figure 5.1 plots the cumulative return on 100 NOK invested in the market (OSEFX) over the first and second sample period. Figure A shows the cumulative return for the first period from January 2002 to March 2008, and figure B for the second period from April 2008 to December 2019.

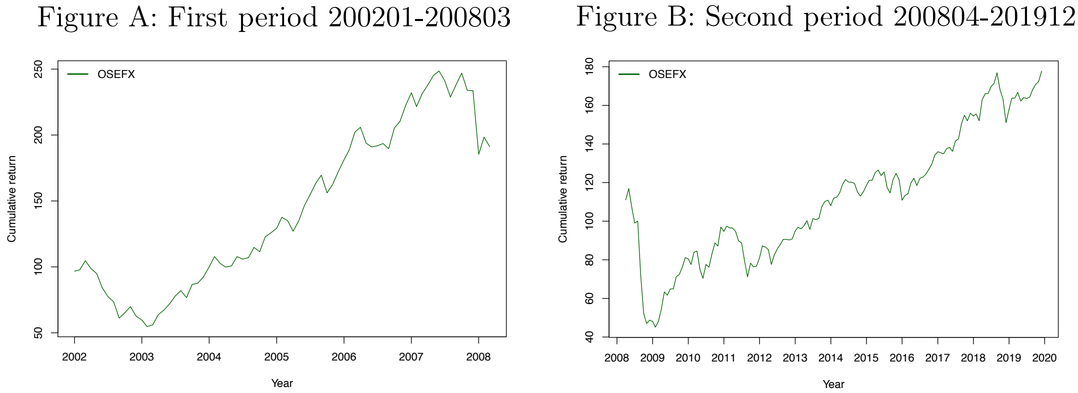
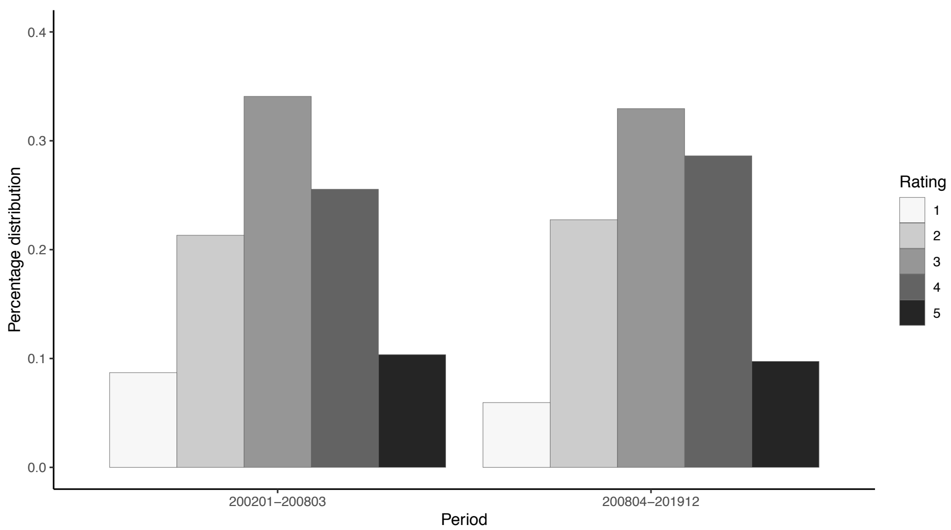


Figure 5.2: Distribution of ratings over the sample periods

Figure 5.2 shows the percentage distribution of funds within each rating group over the first (200201-200803) and second (200804-201912) sample period.



5.2 Factor models

Extensive literature is available that tries to explain the cross-section of returns, and various factor models have been developed to evaluate risk-adjusted fund performance. The convention in fund performance evaluation literature is to compare a fund's return to the return implied by factor models for returns. Performance evaluation is based on the alpha in the factor models. The alpha is the rate of return above or below the model's prediction, and thus a measure of fund performance. For this study, to be consistent with the fund performance evaluation literature, and to increase robustness because different models can produce different results, I employ estimates of alpha from three different factor models of performance measurement: the Capital Asset Pricing Model (CAPM) described in Sharpe (1964) and Lintner (1965), Fama and French (1993) 3-factor model, and Carhart (1997) 4-factor model.

As explained in section 4.4, I acquire the factor returns used in the different models for the Norwegian market from Professor Bernt Arne Ødegaard (2020). The acquired data contains time-series of monthly factor returns for a size factor (SMB), a value factor (HML) and a momentum factor (PR1YR) during the sample period 2002-2019. The SMB (*small minus big*) factor reflects the return of a portfolio with a long position in small-capitalization stocks and a short position in large-capitalization stocks. The HML (*high minus low*) factor represents a portfolio that is long in value stocks with high book-to-market ratio and short in growth stocks with low book-to-market ratio. The PR1YR (*prior one-year*) factor is constructed by going long the stocks with the highest returns over the past year and shorting the stocks with the lowest returns over the past year.

Figure 5.3 plots time-series of the factor returns for the Norwegian market for the two sample periods. The figure shows that over the first sample period, the accumulated return on all factors is positive. SMB and PR1YR have the highest mean returns. However, only the SMB factor shows a general upward trend over the entire first sample period. For the second sample period, the accumulated return of the HML factor is negative, while the accumulated return of the SMB is slightly positive. PR1YR has the absolute highest mean return over the second period and displays a general upward trend over the entire second period.

Figure 5.3: Time-series of factor returns

Figure 5.3 plots the cumulative return on 100 NOK invested in the size (SMB), value (HML), and momentum (PR1YR) factors. Figure A shows the cumulative return for the first period from January 2002 to March 2008, and figure B for the second period from April 2008 to December 2019. The solid green line plots the cumulative indexed return for the SMB factor, the solid blue line for the HML factor and the dotted black line for the PR1YR factor.

Figure A: First period 200201-200803

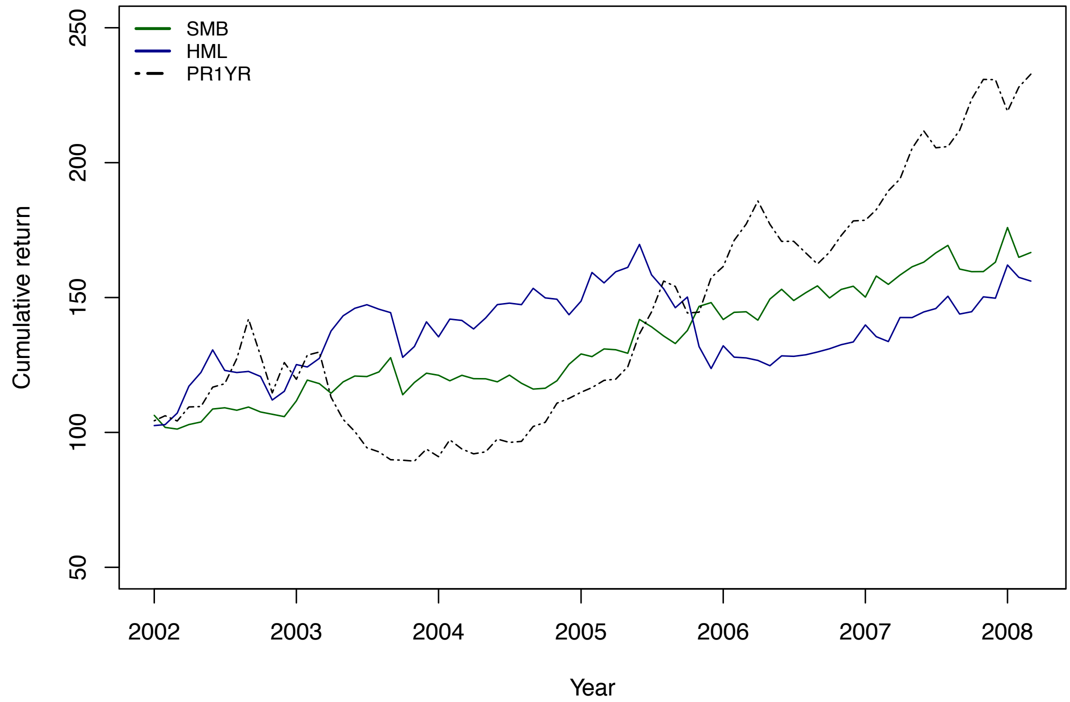
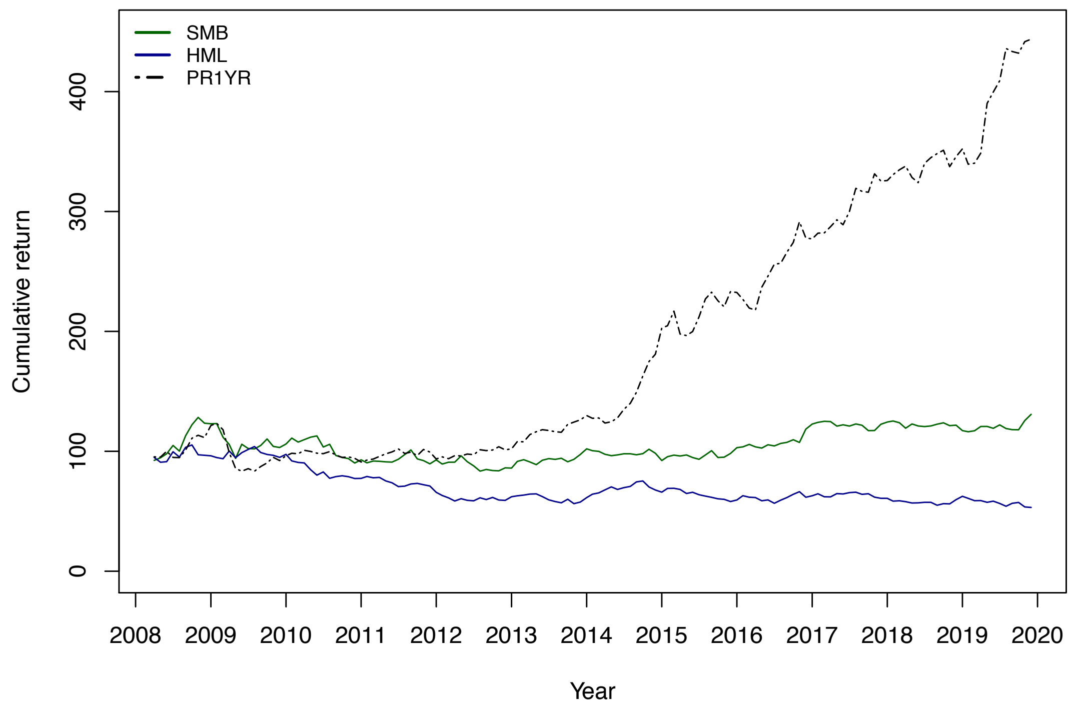


Figure B: Second period 200804-201912



Descriptive statistics for the factors in the two sample periods are reported in table 5.1. The table shows that the equity premium estimate, computed using OSEFX, is not statistically significantly different from zero for either sample periods. The momentum factor (PR1YR) is the only factor statistically significantly different from zero for both sample periods. Panel D reports the correlation matrix between the factors over the entire sample period from 2002 to 2009. SMB, HML, and PR1YR are all negatively correlated with the market premium with correlation coefficients of respective -0.57, -0.15, and -0.25. SMB and HML are positively correlated with a correlation coefficient of 0.11.

Table 5.1: Descriptive statistics for factors

Table 5.1 shows descriptive statistics for the factor returns for the two sample periods (200201-200803 and 200804-201912). Panel A and B show the monthly average return and standard deviation expressed in percentage terms, while panel C shows the t-statistics. t-statistics above 2 or below -2 suggests significance. Panel D reports the cross-correlations over the entire sample period from 2002 to 2019.

	$R_m - R_f$	SMB	HML	PR1YR
Panel A: Average returns				
200201-200803	1.08	0.74	0.69	1.25
200804-201912	0.58	0.26	-0.39	1.14
Panel B: Standard deviations				
200201-200803	6.5	3.34	4.27	4.83
200804-201912	5.74	3.80	3.51	4.03
Panel C: t-statistics				
200201-200803	1.44	1.91	1.39	2.24
200804-201912	1.21	0.82	-1.30	3.37
Panel D: Cross-correlations over entire sample				
$R_m - R_f$	1			
SMB	-0.57	1		
HML	-0.15	0.11	1	
PR1YR	-0.25	0.11	-0.0083	1

5.2.1 The Capital Asset Pricing Model (CAPM)

The CAPM describes the relationship between systematic risk, or market risk, and expected returns. According to CAPM, a measure of risk-adjusted performance is given by the alpha in equation 5.1, referred to as Jensen's (1968) alpha. A positive Jensen's alpha implies that the fund has outperformed compared to the market return, in this case the OSEFX, while a negative Jensen's alpha suggests underperformance.

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{m,i}(r_{m,t} - r_{f,t}) + \epsilon_{i,t} \tag{5.1}$$

where $r_{i,t}$ is fund i's return in period t, $r_{f,t}$ is the risk-free rate (one-month NIBOR) in period t and $r_{m,t}$ is the market return (OSEFX) in period t. α_i represent fund i's alpha over the sample period and $\beta_{m,i}$ is fund i's sensitivity to the market return over the sample period.

Table 5.2: Top and bottom funds based on Jensen's alpha

Table 5.2 shows the top and bottom five funds in the fund sample based on Jensen's alpha. The average ratings are included in the table to show the ratings' applicability. Note that the reported alphas have been annualized. Panel A shows the top and bottom funds over the first period from January 2002 to March 2008, while panel B shows the top and bottom funds over the second period from April 2008 to December 2019.

Panel A: First period 200201-200803

Top five funds			**Bottom five funds**		
Name	Alpha	Avg. rating	Name	Alpha	Avg. rating
Pareto Aksje Norge I	0.0910	4.65	Globus Norge II Acc	-0.1280	1.14
DNB SMB A	0.0899	4.88	Globus Aktiv Acc	-0.1181	1.26
Eika Norge	0.0876	4.94	Storebrand Vekst A	-0.0591	1.49
Holberg Norge A	0.0718	4.96	Nordea 1 - Norwegian Equity BP	-0.0470	1.64
Pareto Aksje Norge A	0.0506	3.45	RF Plussfond Acc	-0.0458	1.47
Average:	0.0782	4.58		-0.08	1.40

Panel B: Second period 200804-201912

Top five funds			**Bottom five funds**		
Name	Alpha	Avg. rating	Name	Alpha	Avg. rating
Landkreditt Utbytte A	0.0065	4.66	Nordea SMB	-0.0085	2.17
FORTE Trønder	0.0057	4.81	DNB Norge R	-0.0035	1.00
Storebrand Vekst A	0.0056	3.43	SSgA Norway Index Equity Fund I	-0.0023	2.49
SEB Norway Focus Fund IC	0.0038	5.00	ODIN Norge II	-0.0018	1.53
SEB Norway Focus Fund HNWC	0.0034	5.00	First Norway Delta Kl.IV (LAMP)	-0.0016	1.48
Average:	0.0050	4.58		-0.0035	1.73

By running the regression in equation 5.1 on the Norwegian equity mutual funds in the data sample, I find that the sample funds on average realize a yearly Jensen's alpha of approximately 0.12% in the first period from January 2002 to March 2008 and 0.08% in the second period from April 2008 to December 2019. Thus, the average fund has obtained a positive alpha in both periods, with a larger alpha in the first period than the second period. Furthermore, from running individual regressions on each fund (table 5.2), I find that there exist fund managers that can generate a high alpha. For instance, the top five funds in the first period have generated an average yearly alpha of 7.82%, while the bottom five funds have generated an average yearly alpha of -8.0%. This can be an indication of stock-picking abilities for some fund managers. Table 5.2 further shows that the highest-performing funds (lowest-performing funds) in the first period is not the highest-performing funds (lowest-performing funds) in the second period. This suggests that stock-picking abilities rarely sustain for longer periods and that a market shock like the financial crisis can have a large impact. Furthermore, table 5.2 suggests that the top funds have higher average ratings than the bottom funds, indicating the ratings' applicability.

5.2.2 Fama-French 3-factor model

The Fama and French (1993) 3-factor model expands on the CAPM by adding a size risk factor (SMB) and a value risk factor (HML) to the market risk factor in CAPM. By including these two additional factors, the Fama-French model accounts for the tendency that value and small-cap stocks outperform markets on a regular basis. According to Fama and French (1993), a measure of risk-adjusted performance is given by the alpha in equation 5.2.

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{m,i}(r_{m,t} - r_{f,t}) + \beta_{SMB,i}SMB_t + \beta_{HML,i}HML_t + \epsilon_{i,t} \tag{5.2}$$

where SMB_t and HML_t are the returns of the respective SMB and HML factors in period t. $\beta_{SMB,i}$ and $\beta_{HML,i}$ represent fund i's exposure to the SMB and HML factors over the sample period.

5.2.3 Carhart 4-factor model

Carhart (1997) augments the Fama-French 3-factor model further by a momentum factor (PR1YR). Momentum in a stock is described as the tendency for the stock price to continue rising if it is going up and to continue declining if it is going down. Carhart (1997) presents evidence that high-momentum stocks outperform markets on a regular basis. Based on this, the Carhart 4-factor model is commonly used as an active management and mutual fund evaluation model. The Carhart model evaluates performance according to the alpha in equation 5.3.

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{m,i}(r_{m,t} - r_{f,t}) + \beta_{SMB,i}SMB_t + \beta_{HML,i}HML_t + \beta_{PR1YR,i}PR1YR_t + \epsilon_{i,t} \tag{5.3}$$

where $PR1YR_t$ is the return of the PR1YR factor in period t. $\beta_{PR1YR,i}$ represent fund i's exposure to the PR1YR factor over the sample period.

Table 5.3 shows descriptive statistics for the factor loadings (β) of the factors included in the Carhart 4-factor model. The factor loadings are estimated from the regression in equation 5.3. Each loading explains how a fund is exposed to each of the four risk factors in the Carhart model. A factor loading equal to one indicates a perfect correlation between a fund's return and the respective factor return. In contrast, a factor loading of zero means no exposure to the risk factor. A negative factor loading indicates a negative exposure to the respective factor. The "Mean" columns in table 5.3 show that the Norwegian equity mutual funds in the sample are nearly perfectly correlated with the excess market return ($OSEFX - R_f$) in both sample periods. Hence, the average fund moves together with the market. The table further shows that the average sample fund is somewhat exposed to size risk (SMB), in which the exposure is highest for the first period. However, the range between the minimum and maximum exposure is relatively large for the SMB factor, indicating that there exists large differences between the sample funds regarding size risk exposure. Furthermore, the table shows that the average fund is only slightly exposed to the HML and PR1YR factor. The slightly negative loadings for HML in the first period and PR1YR for both periods indicate that the average fund has a small negative exposure to value risk in the first period and momentum risk in both periods. However, a slightly positive loading in the HML factor in the second period suggests that the average fund is slightly positive exposed to value risk in the second period. Moreover, since the market is the most prominent risk factor, Jensen's alpha is a suitable evaluation metric to consider in addition to the Fama-French 3-factor alpha and Carhart 4-factor alpha.

Table 5.3: Descriptive statistics for Carhart 4-factor loadings in fund sample

Table 5.3 reports descriptive statistics for the factor loadings of the factors in the Carhart 4-factor model for both sample periods. The first sample period is from January 2002 to March 2008 and the second sample period is from April 2008 to December 2019. The "Mean" columns show the average loading, while the rest of the columns show the respective standard deviation, minimum load and maximum load in the fund sample.

	First period: 200201-200803				Second period: 200804-201912			
Factor	Mean	St.dev.	Min	Max	Mean	St.dev.	Min	Max
$R_m - R_f$	1.0188	0.0845	0.8102	1.2973	0.9785	0.1148	0.6106	1.5350
SMB	0.1548	0.1905	-0.1224	0.6558	0.1117	0.1520	-0.2236	0.8933
HML	-0.0492	0.0972	-0.3175	0.1992	0.0287	0.0644	-0.1950	0.2874
PR1YR	-0.0098	0.0981	-0.3272	0.2476	-0.0083	0.0737	-0.2556	0.2063

5.3 Random effects panel data regression

The first method I use to examine whether the Morningstar ratings can predict risk-adjusted mutual fund performance is a random effects panel data regression with monthly lagged dummy variables for star-ratings. The idea behind this regression is to find out whether or not the highest-rated funds systematically outperform the lower-rated funds. The regression is estimated for the two sample periods split by the financial crisis and for each of the three factor models presented in section 5.2 (CAPM, Fama-French 3-factor model, Carhart 4-factor model).

The choice of the random effects model (RE) is based on intuition and the Hausman test. Contrary to the fixed effects model (FE), RE assumes no individual fixed effects, captured in the model's unique error, correlated with the model's explanatory variables. That is, RE assumes that the variation across funds is random and uncorrelated with the other explanatory variables. In this case, I expect most of the fund characteristics to not be constant across funds. For example, fund characteristics like fund size, investment strategy, fee structure, and fund managers are expected to vary across funds. The Hausman test, shown in table A2.1 in the appendix, supports this intuition. The null hypothesis in the Hausman test is that the assumption that makes random effects both consistent and efficient holds, i.e., zero correlation between unique errors and explanatory variables. Thus, if the test fails to reject the null, random effects should be used. If the test rejects the null, fixed effects should be used. For all models, I can not reject the null, and I thus prefer random effects over fixed effects.

The random effects model is estimated as reported in the regression equation 5.4. The RE estimator can be viewed as a feasible generalized least squares (FGLS) procedure. In the estimation, various feasible GLS estimators have been put forth to tackle serial correlation induced by the group-invariant random effect. This has been proven to be equivalent to ordinary least squares (OLS) on partially demeaned data.

$$\alpha_{i,t} = \beta_0 + \beta_1 D1_{i,t-l} + \beta_2 D2_{i,t-l} + \beta_3 D3_{i,t-l} + \beta_4 D4_{i,t-l} + u_i + \epsilon_{i,t} \tag{5.4}$$

where $\alpha_{i,t}$ represent fund i's alpha at time t, estimated for each factor model: CAPM, Fama-French 3-factor model, and Carhart 4-factor model. $D1, D2, D3$, and $D4$ are the funds' dummy variables for respectively one-star, two-star, three-star and four-star lagged rating, where $l = 1, 3, 6, 12, 24$ is the monthly lag period length. For example, $D1 = 1$ if fund i is a one-star fund at time $t - l$, otherwise 0. u_i is the unobserved individual unique error term and $\epsilon_{i,t}$ is the error term.

The choice of lag periods is based on an assessment of a fund's characteristics over its life cycle. A lag period of more than 24 months (2 years) can be problematic because, even if the fund survives, the fund manager is often replaced, and the investment strategy is often changed. Therefore, I examine the predictive performance only using lags up to 24 months.

The five-star funds are the reference group in the regressions. Hence, when measuring performance using the different factor models, the β_0 constant represents the five-star funds' expected alpha. The β_1 through β_4 coefficients represent the difference between the dummy variables and the reference group (five-star funds). If the star ratings accurately forecast performance, in the sense that the five-star funds outperform the lower-rated funds, the coefficients should be increasingly negative and significant when moving from β_4 to β_1.

5.4 Strategy of buying winners vs. losers

The second method I use to examine the predictive abilities of the Morningstar ratings on risk-adjusted fund performance is a strategy of buying past winners versus losers, inspired by Jegadeesh and Titman (1993). The paper by Jegadeesh and Titman (1993) documents that strategies that take a long position in stocks that have performed well in the past and short stocks that have performed poorly in the past generate significant positive returns over 3- to 12-month holding periods. Jegadeesh and Titman's (1993) strategy is referred to as a J/K strategy. Securities are ranked in ascending order based on their returns in the past J months, and the strategy goes long in the equal-weighted top decile ("winner") portfolio and short in the equal-weighted bottom decile ("loser") portfolio, holding the portfolio for K months. This strategy can be applied to the mutual fund environment in order to examine whether investors can benefit from the Morningstar mutual fund ratings. I thus apply Jegadeesh and Titman's (1993) strategy with some modifications.

The first modification I introduce is to create the portfolios based on past Morningstar ratings instead of past returns, like Jegadeesh and Titman (1993) do.

The second modification is to eliminate the short sell element of Jegadeesh and Titman's (1993) strategy. The reason for this elimination is that mutual fund shares are not traded in a functioning market to borrow or sell fund shares, so the classical shorting strategy does not work. Therefore, mutual funds can not be shorted the same way that stocks can. To short mutual funds, one would need to build a short position that could be close to the fund if one shorts all of the fund's assets individually in the right proportions. However, this requires ongoing and up to date information about the funds' transactions. Even very transparent funds will probably not supply such information in sufficient form or within a reasonable time. Even then, it would only work with funds investing only in deeply liquid shares. Consequently, to create a feasible strategy for all fund investors, I eliminate the short sell element of Jegadeesh and Titman's (1993) strategy. Instead, I study a strategy of buying winners and compare this to a strategy of buying losers.

The third modification is related to the portfolio holding periods of the strategy. Jegadeesh and Titman (1993) study holding periods of 3, 6, 9, and 12 months. To make the strategy consistent with the panel data regression in this thesis, I study holding periods of 1, 3, 6, 12, and 24 months.

With these modifications introduced, the J/K strategy used in this paper is constructed as follows: At the beginning of each month, the funds are ranked in ascending order based on their average Morningstar ratings in the past J months, where $J = 1, 3, 6, 12, 24$. Based on these rankings, ten decile portfolios are formed that equally weigh the funds in the top decile, the second decile, etc. The top decile portfolio (top 10% rated funds) is called the "winners" decile, and the bottom decile (bottom 10% rated funds) is called the "losers" decile. Each month, one strategy buys and holds the winner portfolio for K months, and another strategy buys the loser portfolio holding this position for K months, where $K = 1, 3, 6, 12, 24$. These strategies are then compared. To obtain robust results, the explained strategy will also be computed using quartile portfolios (top quartile: top 25% rated funds; bottom quartile: bottom 25% rated funds) instead of decile portfolios. Applying quartile portfolios, the winner and loser portfolios will consist of funds with a larger range of average ratings. For example, the top decile portfolio could one month include funds with an average rating ranging from 4.5 to 5. In contrast, the top quartile portfolio could include funds with an average rating ranging from 4 to 5. Using both decile and quartile portfolios, I can thus examine whether there is a different in results for the top (bottom) 10% rated funds and the top (bottom) 25% rated funds. Note that the J/K strategy used in this paper is calculated for both sample periods split by the financial crisis, like the rest of the paper.

6 Empirical analysis and results

In this section, I present the analysis and corresponding results of the two methods used to examine the Morningstar ratings' predictive ability. First, I report the analysis and results using the random effects panel data regression. Second, I present the analysis and results using the strategy of buying winners versus losers, inspired by Jegadeesh and Titman (1993). The analyses make use of the alphas in the factor models from section 5.2: CAPM, Fama-French 3-factor model and Carhart 4-factor model. Furthermore, the analyses are split into the two sample periods explained in section 5.1. The first period is before the financial crisis from January 2002 to March 2008, while the second period is during and after the financial crisis from April 2008 to December 2019.

6.1 Random effects panel data regression

In this subsection, I discuss the random effects panel data regression results. I divide the discussion into an analysis of the two sample periods[4]. The regressions examine how well the Morningstar ratings predict fund performance, as measured by the alphas of the CAPM, Fama-French (FF) 3-factor model, and Carhart 4-factor model. With dummy variable rating lags ranging from one month to 24 months, I assess the predictive abilities for different prediction horizons. In addition to estimating regressions for all funds for both sample periods, I examine whether there is a "style effect" on fund performance predictability by separating funds according to investment style. The investment styles are based on the three style categories of Morningstar's equity style box: "value", "growth", and "blend". Two of these categories, "value" and "growth," are common to stocks and funds. However, the blend category for funds represents a mixture of growth and value stocks, or mostly core stocks. Random effect panel data regressions are estimated separately for each investment style category. Since the regression results for each investment style yield approximately the same results as the regression results for all funds, I only present and discuss the regression results for all funds in this subsection. The estimated regressions for funds separated by investment style are reported in section A4 in the appendix.

[4]The regression for the full sample period from January 2002 to December 2019 is also estimated, shown in section A3 in the appendix. The results for the full period are less significant than for the other periods, and are therefore not discussed further.

Table 6.1: Predictive regression for first period: 200201-200803

Table 6.1 shows the random effects panel data regression for the first period from January 2002 to March 2008. The regression is estimated for three performance metrics: Jensen's alpha (CAPM), Fama-French three-factor alpha (FF) and Carhart four-factor alpha (Carhart). For each model, in order to measure future performance, dummy variables for lagged 4-star rating, 3-star rating, 2-star rating and 1-star rating are included. The 5-star rating is the reference group. The lag periods are 1, 3, 6, 12 and 24 months. Standard errors are in parenthesis.

	Lag = 1 month			Lag = 3 months			Lag = 6 months			Lag = 12 months			Lag = 24 months		
	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart
β_0 (Constant)	0.0001	−0.001	−0.001	0.0005	−0.001	−0.0004	0.001	−0.0002	−0.00003	0.003***	0.001	0.001	0.001	−0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_4 (4-star)	0.0001	0.0001	0.0001	0.001	0.001	0.001	0.001	0.001	0.001	−0.002	−0.001	−0.001	−0.0002	−0.0002	−0.0002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_3 (3-star)	−0.00001	0.0001	0.00001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.002**	−0.002*	−0.002*	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_2 (2-star)	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.002*	−0.002*	−0.002*	0.0002	−0.00001	0.00001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_1 (1-star)	−0.001	−0.001	−0.001	−0.001	−0.002	−0.002	−0.002	−0.002	−0.002	−0.003*	−0.003*	−0.002*	−0.005***	−0.005***	−0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Observations	3,519	3,519	3,519	3,405	3,405	3,405	3,243	3,243	3,243	2,925	2,925	2,925	2,286	2,286	2,286
Adjusted R^2	0.921	0.926	0.926	0.922	0.927	0.927	0.919	0.923	0.923	0.912	0.916	0.916	0.910	0.915	0.916

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

6.1.1 First period: 200201-200803

Table 6.1 reports the random effects panel data regression for all funds for the first period.

First, table 6.1 shows that the β_0 constants differs for the different prediction horizons, and factor models. For all prediction horizons except the 12-month horizon, the constants are close to zero and insignificant, suggesting that the five-star funds generally obtain alphas which are not significantly different from zero. The CAPM shows slightly positive constants for all horizons, while the FF and Carhart models show constants just below zero. The 12-month horizon constants are positive for all factor models but only significant for the CAPM. These results indicate that the reference group (the five-star funds) performs quite different for different prediction horizons. The five-star funds only obtain a positive and significant Jensen's alpha (CAPM) over a 12-month horizon.

Second, the results show weak statistical evidence that the four-, three- and two-star funds diverge from the five-star funds in terms of future performance. In many cases, even the signs of the coefficients are the opposite of what one would expect. However, the coefficients for the three- and two-star funds are negative and significant for the 12-month horizon, indicating that they underperform relative to the five-star funds after the specific period of 12 months. Nevertheless, in terms of 24-month future performance, the signs of the three- and two-star coefficients are the opposite of what one would expect, suggesting that one should be careful associating three- and two-star funds with underperformance.

Third, there is statistical evidence that the Morningstar ratings predict low-performing funds. The one-star funds significantly diverge from the five-star funds in terms of future 12 months and 24 months performance, indicating that the one-star funds' performance is significantly worse than the five-star funds' performance over longer prediction horizons of one and two years. The difference in results for different prediction horizons may be related to an intuition of random walk in the market prices for short time horizons. Whether a fund's portfolio value increases or decreases in the short term could be related to "luck". However, for longer horizons, a fund's portfolio value could be more related to skills, e.g., fund manager skills, or a good investment strategy.

The investment style regressions in section A4 in the appendix show similar but weaker results.

Table 6.2: Predictive regression for second period: 200804-201912

Table 6.2 shows the random effects panel data regression for the second period from April 2008 to December 2019. The regression is estimated for three performance metrics: Jensen's alpha (CAPM), Fama-French three-factor alpha (FF) and Carhart four-factor alpha (Carhart). For each model, in order to measure future performance, dummy variables for lagged 4-star rating, 3-star rating, 2-star rating and 1-star rating are included. The 5-star rating is the reference group. The lag periods are 1, 3, 6, 12 and 24 months. Standard errors are in parenthesis.

	Lag = 1 month			Lag = 3 months			Lag = 6 months			Lag = 12 months			Lag = 24 months		
	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart
β_0 (Constant)	0.001	0.0003	0.0004	0.001	0.0002	0.0004	0.001	0.0004	0.001	−0.0001	−0.001	−0.0004	0.0004	−0.0001	0.0001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_4 (4-star)	0.00000	0.0001	0.0001	−0.001	−0.0005	−0.0005	−0.001	−0.001	−0.001	0.0003	0.0004	0.0004	−0.001	−0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_3 (3-star)	−0.001	−0.001	−0.001	−0.0002	−0.0002	−0.0002	−0.0004	−0.0004	−0.0005	0.001	0.001	0.001	0.0004	0.0003	0.0003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_2 (2-star)	0.0003	0.0001	0.0001	0.0004	0.0004	0.0004	−0.0001	0.0001	−0.00001	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_1 (1-star)	0.002**	0.002*	0.002*	0.0002	0.0003	0.0003	0.001	0.001	0.001	0.003***	0.003***	0.003***	0.002**	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	8,750	8,750	8,750	8,643	8,643	8,643	8,490	8,490	8,490	8,192	8,192	8,192	7,685	7,685	7,685
Adjusted R^2	0.904	0.907	0.907	0.904	0.907	0.907	0.904	0.907	0.907	0.906	0.909	0.909	0.909	0.912	0.912

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

6.1.2 Second period: 200804-201912

Table 6.2 presents the random effects panel data regression for all funds for the second sample period from April 2008 to December 2019. This is the period during and after the financial crisis.

First, the results show that the β_0 constants are positive for most prediction horizons and factor models during and after the financial crisis. However, after 12 months and 24 months, some of the constants turns negative. None of the constant's are significantly different from zero. These results indicate that the five-star funds' performance differ for the shorter and longer prediction horizons and that there exists little evidence of significance of the constant's. That is, the five-star funds do not obtain alphas significantly different from zero over this second sample period.

Second, consistent with the regression results for the first sample period, the results for the second period show weak statistical evidence that the four-, three- and two-star funds diverge from the five-star funds in terms of future performance. The coefficients of the four-, three and two-star funds are insignificant for all prediction horizons and factor models. Furthermore, the signs of the coefficients are positive for some horizons and negative for others. Consequently, these results indicate no significant difference in future performance of next-to-highest, median, and next-to-lowest rated funds and top-rated funds.

Third, contrary to the results for the period before the financial crisis, there is evidence that the lowest-rated funds outperform the top-rated funds in the period during and after the financial crisis. The coefficients of the one-star funds are positive for all prediction horizons and significant for the 1-, 12-, and 24-month horizons. The significance is strongest for the 12-month horizon. Thus, these results indicate that the one-star funds outperform the five-star funds, and the effect is strongest for longer time horizons. The results suggest mean-reversion in fund returns, in which fund prices converge on an average value over time. Experienced investors, who have seen many market ups and downs, often take the view that the market over time will even out. This can be related to the regression results during and after the financial crisis, as the performance of the five-star and one-star funds are evened out.

The regressions for the growth style and blend style for the second sample period (reported in section A4 in the appendix) also show similar results, in that the one-star funds outperform the five-star funds. However, the value style regression suggests no statistical evidence that the lowest-rated funds outperform the top-rated funds in the period during and after the financial crisis.

6.1.3 Main results of regression

The regression results for the two sample periods can be summarized in several main findings. First, for the sample period before the financial crisis, Morningstar is able to "predict" low-performing funds for one- and two-year horizons. That is, the lowest-rated funds (one-star funds) generally underperform relative to the top-rated funds (five-star funds) for these longer time horizons. Second, on the contrary, for the sample period during and after the financial crisis, there exist evidence that the lowest-rated funds outperform the top-rated funds. Third, there is only weak statistical evidence for both sample periods that the five-star funds (top-rated funds) outperform the four-, three- and two-star funds (next-to-highest, median and next-to-lowest rated funds).

The substantial difference in results for the periods split by the financial crisis highlights the importance of market timing in buying mutual funds and may be related to the funds' average factor exposures and overall risk-taking over these periods.

Table 6.3 reports the average factor loadings for each rating group over both sample periods. The table shows that the one-star funds have negative exposure to value risk (HML) over the second sample period during and after the financial crisis, while the five-star funds are positively exposed to value risk over this second period. In section 5.2, figure 5.3 shows that the HML factor has a negative cumulative return over the second period. Isolated, since the five-star funds are positively exposed to a negative cumulative return factor, and vice versa for the one-star funds, these results indicate lower performance of the five-star funds than the one-star funds over the second period. Moreover, table 6.3 shows that the one-star funds were contrarians in the first period to a large degree. The one-star funds were highly negatively exposed to the momentum factor (PR1YR) in the first period and much less so in the second period. This change in exposure can be a contributor to the improvement in the one-star funds' performance relative to the five-star funds from the first period to the second period. Morningstar does not account for these discussed factor exposures when assigning ratings. These factor exposures can thus be one part of the explanation for the remarkable results after the financial crisis.

Table 6.3: Factor loadings for each rating group

Table 6.3 shows the differences in average factor loadings of the factors in Carhart 4-factor model for each rating group and sample period. The first sample period is from January 2002 to March 2008 and the second sample period is from April 2008 to December 2019.

	First period: 200201-200803				Second period: 200804-201912			
Rating group	$R_m - R_f$	SMB	HML	PR1YR	$R_m - R_f$	SMB	HML	PR1YR
1-star funds	1.1054	0.2790	-0.1526	-0.1865	1.0157	0.1749	-0.0332	-0.0932
2-star funds	1.0405	0.2202	-0.0759	-0.0370	1.0107	0.1247	0.0203	-0.0299
3-star funds	1.0106	0.0879	-0.0233	-0.0086	0.9605	0.0921	0.0329	-0.0109
4-star funds	0.9973	0.0955	-0.0210	0.0102	0.9434	0.0845	0.0147	0.0078
5-star funds	1.0140	0.2612	-0.0513	0.0870	0.9604	0.1133	0.0186	0.0408

Furthermore, the difference in results between the two periods can also be related to the overall risk-taking of the different rating groups of mutual funds. Table 6.4 presents the standard deviation (risk) for each rating group over both sample periods. The table suggests that the one-star funds are the largest risk takers over both periods. Also, the five-star funds are one of the fund groups with the lowest risk over both periods. This can, to some extent, explain the large difference in results between the two sample periods. With considerable risk, the one-star funds can have either superior or inferior performance. For the second period, the first was true (superior performance), while the second was true for the first period (inferior performance). Table 6.4 further suggests that the five-star funds takes on more risk relative to the one-star funds in the second period than in the first period. In the first period, the five-star funds' risk was approximately 15% lower than the one-star funds. However, in the second period, this percentage was approximately 12%. Consequently, with a higher relative risk of the five-star funds, there are higher risks of performing poorly (or superior) in the second period.

Table 6.4: Risk for each rating group

Table 6.4 reports the risk (standard deviation) of the different rating groups over the first sample period (200201-200803) and second sample period (200804-201912). The standard deviations have been annualized.

	First period: 200201-200803	Second period: 200804-201912
Rating group	St.dev.	St.dev.
1-star funds	0.2535	0.2012
2-star funds	0.2280	0.1880
3-star funds	0.2198	0.1747
4-star funds	0.2262	0.1868
5-star funds	0.2166	0.1767

6.2 Strategy of buying winners vs. losers

In this subsection, I implement the buying winners versus losers strategy (J/K-strategy), inspired by Jegadeesh and Titman (1993). This strategy is another way of examining the predictive abilities of the Morningstar fund ratings. I present the main findings and results when implementing the J/K-strategy for decile portfolios for the two sample periods. The first period is from January 2002 to March 2008 and the second period is from April 2008 to December 2019. The quartile portfolio J/K-strategy yields similar but less profitable results and is thus only reported in section A5 of the appendix.

6.2.1 First period: 200201-200803

Table 6.5 documents the results in terms of average annualized returns of the decile portfolio strategy for the first sample period before the financial crisis, from January 2002 to March 2008. Panel A reports the results for the top ("winner") decile, while panel B reports the results for the bottom ("loser") decile. The returns of all the portfolios are positive, and most returns are statistically significantly different from zero. Furthermore, the top decile portfolios' returns are greater than the bottom decile portfolios' returns for most evaluation (J) and holding (K) periods. Consequently, investors can implement top decile J/K strategies to generate higher returns than the bottom decile J/K strategies. The magnitude and significance of the effect that the top decile funds outperforms relative to the bottom decile funds are greater for longer evaluation and holding periods. These results suggest that the Morningstar ratings have greater predictive abilities for longer time horizons. This complies with the regression results in section 6.1 of the paper. The most successful decile portfolio strategy for the first period is the $J = 12/K = 6$ top decile ("winner") strategy, which selects funds based on their ratings over the previous 12 months and then holds the portfolio for six months. This strategy yields an annualized average return of 27.13%, as reported in table 6.5. The quartile portfolio strategy (table A5.1 in the appendix) yields similar but less profitable results than the decile portfolio strategy. The most profitable quartile portfolio strategy yields an annualized average return of 25.99%, in contrast to 27.13% for the decile portfolio.

Table 6.6 documents the decile portfolio strategy's performance in the first period further in terms of annualized alphas of the three factor models: Jensen's alpha (CAPM), Fama-French 3-factor alpha, and Carhart 4-factor alpha. In general, the table reports higher alphas for the top decile portfolios (panel A) than for the bottom decile portfolios (panel B). This suggests that the historically top-rated funds outperforms the historically low-rated funds. Furthermore, the difference between the top and bottom decile portfolio is larger and more significant for longer evaluation (J) and holding (K) periods. These results are consistent with the results in the above paragraph and the previous regression results, where the top-rated funds outperform the lower-rated funds and the effect is stronger for longer prediction horizons. The top decile $J = 12$ months holding period yields the highest alphas, with the highest for the top decile $J = 12/K = 12$ strategy. This strategy yields a Jensen's alpha of approximately 4.60%. The quartile portfolio strategy, reported in table A5.2 in the appendix, yields similar results. However, the quartile portfolio strategy's highest alpha is less than the highest alpha of the decile portfolio strategy: alpha of 2.73% against 4.60%. Thus, the decile portfolio strategy is more profitable than the quartile portfolio strategy.

Table 6.5: Returns of decile portfolios for first period: 200201-200803

Table 6.5 reports the average annualized returns of the decile portfolios for the first sample period from January 2002 to March 2008. The portfolios are formed based on J-month lagged ratings and held for K months. The values of the J and K for the different strategies are indicated in the first column and row, respectively. Panel A shows the returns for the top decile, while Panel B shows the returns for the bottom decile. The t-statistics are reported in parantheses.

Panel A: Top decile

	K=1	K=3	K=6	K=12	K=24
J=1	0.1648*	0.1696*	0.1613*	0.1585*	0.1603*
	(1.85)	(1.9)	(1.84)	(1.73)	(1.76)
J=3	0.1804*	0.1742*	0.1723*	0.1639*	0.1723*
	(1.98)	(1.93)	(1.93)	(1.8)	(1.89)
J=6	0.1915**	0.1907**	0.1947**	0.1875*	0.1848*
	(2.07)	(2.05)	(2.07)	(1.98)	(1.98)
J=12	0.2670***	0.2706***	0.2713***	0.2653***	0.2638***
	(3.15)	(3.19)	(3.19)	(3.26)	(3.22)
J=24	0.2357***	0.2347***	0.2348***	0.2334***	0.2412***
	(2.73)	(2.72)	(2.72)	(2.71)	(2.83)

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Panel B: Bottom decile

	K=1	K=3	K=6	K=12	K=24
J=1	0.1689*	0.1663*	0.1610	0.1557*	0.1361
	(1.72)	(1.69)	(1.66)	(1.68)	(1.45)
J=3	0.1643	0.1708*	0.1726*	0.1820*	0.1302
	(1.61)	(1.67)	(1.67)	(1.82)	(1.41)
J=6	0.1859*	0.1805*	0.1788*	0.1911*	0.1610*
	(1.78)	(1.73)	(1.73)	(1.91)	(1.68)
J=12	0.2390**	0.2368**	0.2339**	0.2231**	0.2240**
	(2.41)	(2.39)	(2.4)	(2.27)	(2.28)
J=24	0.2034**	0.2011**	0.1976**	0.2005**	0.2055**
	(2.08)	(2.06)	(2.02)	(2.09)	(2.22)

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Table 6.6: Alphas of decile portfolios for first period: 200201-200803

Table 6.6 reports the annualized performance of the decile portfolios for three performance measures: Jensen's alpha, Fama-French 3-factor alpha and Carhart 4-factor alpha. The sample period is the first period from January 2002 to March 2008. The portfolios are formed based on J-month lagged ratings and held for K months. The values of the J and K for the different strategies are indicated in the first column and row, respectively. Panel A shows the alphas for the top decile, while Panel B shows the alphas for the bottom decile.

Panel A: Top decile

	Performance measure	K=1	K=3	K=6	K=12	K=24
	Jensen's alpha	-0.00129	0.00381	-0.00244	-0.01054	-0.00682
J=1	Fama-French 3-factor alpha	-0.02459	-0.01966	-0.02382	-0.03724***	-0.04341**
	Carhart 4-factor alpha	-0.03938**	-0.0339***	-0.03744***	-0.03266	-0.0306
	Jensen's alpha	0.00854	0.00438	0.00401	-0.00619	0.00253
J=3	Fama-French 3-factor alpha	-0.01194	-0.01857	-0.01948	-0.03813	-0.03727
	Carhart 4-factor alpha	-0.02436	-0.02918	-0.02805	-0.03625	-0.02522
	Jensen's alpha	0.01338	0.01225	0.015	0.00737	0.00692
J=6	Fama-French 3-factor alpha	-0.01359	-0.01533	-0.01385	-0.02613	-0.03288
	Carhart 4-factor alpha	-0.02454	-0.02491	-0.01946	-0.02672	-0.02378
	Jensen's alpha	0.04104	0.04456***	0.04485***	0.04595**	0.04245***
J=12	Fama-French 3-factor alpha	0.00898	0.01401	0.01309	0.00697	0.00406
	Carhart 4-factor alpha	-0.00494	0.0025	0.00188	-0.00216	-0.00677
	Jensen's alpha	0.02329	0.02281	0.02235	0.02127	0.0293
J=24	Fama-French 3-factor alpha	-0.01227	-0.01303	-0.01252	-0.01377	0.00095
	Carhart 4-factor alpha	-0.01531	-0.01614	-0.01657	-0.01832	-0.01121

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Panel B: Bottom decile

	Performance measure	K=1	K=3	K=6	K=12	K=24
	Jensen's alpha	-0.00751	-0.01082	-0.01395	-0.01474	-0.03496
J=1	Fama-French 3-factor alpha	-0.02789	-0.0378	-0.03631	-0.03196	-0.05039**
	Carhart 4-factor alpha	-0.00454	-0.01375	-0.0153	-0.02196	-0.04058***
	Jensen's alpha	-0.01902	-0.01341	-0.01325	0.00087	-0.04245***
J=3	Fama-French 3-factor alpha	-0.04741	-0.04131	-0.04203	-0.02182	-0.05871*
	Carhart 4-factor alpha	-0.02006	-0.01513	-0.01857	-0.00818	-0.05553**
	Jensen's alpha	-0.00701	-0.01211	-0.01248	0.00363	-0.02173
J=6	Fama-French 3-factor alpha	-0.04105	-0.04529	-0.04544	-0.02277	-0.05118**
	Carhart 4-factor alpha	-0.015	-0.02076	-0.02325	-0.01354	-0.05211**
	Jensen's alpha	-0.01514	-0.01688	-0.0175	-0.02937	-0.03103
J=12	Fama-French 3-factor alpha	-0.05418	-0.05663***	-0.05443***	-0.07363**	-0.07558*
	Carhart 4-factor alpha	-0.03244	-0.03687	-0.03924	-0.06303***	-0.07061**
	Jensen's alpha	-0.02682	-0.02844	-0.03189	-0.0273	-0.01755
J=24	Fama-French 3-factor alpha	-0.06202***	-0.06196***	-0.06531***	-0.0588***	-0.04509
	Carhart 4-factor alpha	-0.06984***	-0.07365***	-0.07599***	-0.06093***	-0.04984

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

6.2.2 Second period: 200804-201912

Table 6.7 documents the results in terms of average annualized returns of the decile portfolio strategy for the second sample period during and after the financial crisis, from April 2008 to December 2019. Panel A reports the results for the top ("winner") decile, while panel B reports the results for the bottom ("loser") decile. First, the average returns of all the portfolios are positive for this second period, but lower compared to the period before the financial crisis. Second, contrary to the period before the financial crisis, the bottom decile portfolios' returns are, for the second period, higher than the top decile portfolios' returns. This suggests the remarkable results that the bottom-rated funds outperforms the top-rated funds over the second period during and after the financial crisis. Furthermore, table 6.7 shows that the magnitude and significance of the effect that bottom-rated funds outperforms are greater for longer evaluation periods (J), consistent with previous results. The most successful decile portfolio strategy for the second period is the $J = 12/K = 1$ bottom decile ("loser") strategy, which selects funds based on their ratings over the previous 12 months and then holds the portfolio for one month. This strategy yields an annualized average return of 16.50%, which is lower than the first period's most profitable strategy (27.13%). The quartile portfolio strategy (table A5.3 in the appendix) yields similar results.

Table 6.8 documents the decile portfolio strategy's performance in the second period further in terms of annualized alphas of the three factor models: Jensen's alpha (CAPM), Fama-French 3-factor alpha, and Carhart 4-factor alpha. The table shows that most alphas are negative in panel A (top decile), while the majority is positive in panel B (bottom decile). Hence, the reported alphas are higher for the bottom decile portfolios than for the top decile portfolios. This further suggests the effect that the historically low-rated funds outperforms the top-rated funds in the second period during and after the financial crisis. However, few of the alphas in table 6.8 are significant. Nevertheless, the top decile Fama-French 3-factor alphas (and Carhart 4-factor alphas) are significant for some evaluation and holding periods, e.g., for the $J = 12/K = 1, 3, 6$ strategy. None of the alphas in the bottom decile portfolio are significant. The most profitable strategy based on the alphas is the bottom decile $J = 12/K = 1$ strategy, with a Jensen's alpha of approximately 1.80%. However, this alpha is not significantly different from zero. The quartile portfolio strategy, reported in table A5.4 in the appendix, yields similar but even less significant results in which none of the quartile portfolio alphas are significant.

Table 6.7: Returns of decile portfolios for second period: 200804-201912

Table 6.7 reports the average annualized returns of the decile portfolios for the second sample period from April 2008 to December 2019. The portfolios are formed based on J-month lagged ratings and held for K months. The values of the J and K for the different strategies are indicated in the first column and row, respectively. Panel A shows the returns for the top decile, while Panel B shows the returns for the bottom decile. The t-statistics are reported in parantheses.

Panel A: Top decile

	K=1	K=3	K=6	K=12	K=24
J=1	0.0816	0.0778	0.0750	0.0660	0.0814
	(1.51)	(1.45)	(1.41)	(1.24)	(1.54)
J=3	0.0669	0.0665	0.0624	0.0498	0.0677
	(1.24)	(1.25)	(1.17)	(0.95)	(1.29)
J=6	0.0743	0.0739	0.0706	0.0735	0.0858
	(1.37)	(1.38)	(1.31)	(1.38)	(1.63)
J=12	0.1267***	0.1253***	0.1300***	0.1343***	0.1426***
	(2.78)	(2.77)	(2.87)	(3.05)	(3.3)
J=24	0.0910**	0.0913**	0.0917**	0.0943**	0.0921**
	(2.1)	(2.11)	(2.13)	(2.21)	(2.15)

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Panel B: Bottom decile

	K=1	K=3	K=6	K=12	K=24
J=1	0.1050*	0.0999*	0.1000*	0.0916	0.0938*
	(1.87)	(1.77)	(1.77)	(1.65)	(1.68)
J=3	0.0896	0.0854	0.0852	0.0784	0.0692
	(1.57)	(1.52)	(1.5)	(1.41)	(1.23)
J=6	0.0968*	0.0955*	0.0960*	0.0975*	0.0874
	(1.69)	(1.68)	(1.68)	(1.73)	(1.54)
J=12	0.1650***	0.1618***	0.1564***	0.1477***	0.1555***
	(3.54)	(3.48)	(3.35)	(3.12)	(3.3)
J=24	0.1054**	0.1085**	0.1124**	0.1179**	0.1112**
	(2.3)	(2.35)	(2.46)	(2.56)	(2.39)

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Table 6.8: Alphas of decile portfolios for second period: 200804-201912

Table 6.8 reports the annualized performance of the decile portfolios for three performance measures: Jensen's alpha, Fama-French 3-factor alpha and Carhart 4-factor alpha. The sample period is the second period from April 2008 to December 2019. The portfolios are formed based on J-month lagged ratings and held for K months. The values of the J and K for the different strategies are indicated in the first column and row, respectively. Panel A shows the alphas for the top decile, while Panel B shows the alphas for the bottom decile.

Panel A: Top decile

	Performance measure	K=1	K=3	K=6	K=12	K=24
	Jensen's alpha	-0.00124	-0.00472	-0.00731	-0.01603	0.00003
J=1	Fama-French 3-factor alpha	-0.01011	-0.01248	-0.01402	-0.02325**	-0.01093
	Carhart 4-factor alpha	-0.00704	-0.00937	-0.01164	-0.02198***	-0.0091
	Jensen's alpha	-0.00325	-0.0032	-0.00704	-0.01916	-0.00075
J=3	Fama-French 3-factor alpha	-0.01117	-0.01181	-0.01533	-0.02703**	-0.01277
	Carhart 4-factor alpha	-0.00513	-0.00623	-0.00877	-0.02328***	-0.00765
	Jensen's alpha	-0.00789	-0.0079	-0.01133	-0.00792	0.00544
J=6	Fama-French 3-factor alpha	-0.01588	-0.01552	-0.01945	-0.01825	-0.0058
	Carhart 4-factor alpha	-0.00734	-0.00953	-0.01235	-0.01145	-0.00103
	Jensen's alpha	-0.01936	-0.01954	-0.01555	-0.00726	0.00187
J=12	Fama-French 3-factor alpha	-0.0296**	-0.02961**	-0.02573**	-0.01762	-0.00588
	Carhart 4-factor alpha	-0.01905	-0.01945	-0.01447	-0.00455	0.00073
	Jensen's alpha	-0.01217	-0.01142	-0.01062	-0.00731	-0.0101
J=24	Fama-French 3-factor alpha	-0.01994***	-0.01965	-0.01907	-0.01455	-0.01627
	Carhart 4-factor alpha	-0.00424	-0.00324	-0.00435	-0.0005	-0.00349

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Panel B: Bottom decile

	Performance measure	K=1	K=3	K=6	K=12	K=24
	Jensen's alpha	0.01942	0.01416	0.014	0.00649	0.00862
J=1	Fama-French 3-factor alpha	0.00966	0.00395	0.00348	-0.00244	0.00074
	Carhart 4-factor alpha	0.00858	0.00142	0.00242	-0.00302	0.00078
	Jensen's alpha	0.01696	0.01353	0.01301	0.00693	-0.00322
J=3	Fama-French 3-factor alpha	0.00238	-0.00049	-0.00019	-0.0053	-0.01178
	Carhart 4-factor alpha	0.00405	-0.00143	0.00493	-0.00137	-0.00867
	Jensen's alpha	0.01153	0.01035	0.01037	0.01235	0.00179
J=6	Fama-French 3-factor alpha	-0.00082	-0.00148	0.00056	0.00208	-0.00585
	Carhart 4-factor alpha	0.00007	0.00247	0.00107	0.00289	-0.00434
	Jensen's alpha	0.01803	0.015	0.00856	-0.00172	0.00471
J=12	Fama-French 3-factor alpha	0.00376	0.00092	-0.00432	-0.01235	-0.00367
	Carhart 4-factor alpha	0.00542	0.00456	-0.0014	-0.01068	-0.0009
	Jensen's alpha	-0.00062	0.00194	0.00664	0.01131	0.00292
J=24	Fama-French 3-factor alpha	-0.01077	-0.00819	-0.00285	0.00137	-0.00662
	Carhart 4-factor alpha	-0.0113	-0.00966	-0.00542	-0.00107	-0.00685

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

6.2.3 Main results of strategy

The main results of the buying winners versus losers strategy are summarized in figure 6.1. Figure A shows that the top decile portfolio outperforms the bottom decile portfolio over the first sample period before the financial crisis. However, figure B shows that the bottom decile portfolio performs better than the top decile portfolio over the second sample period during and after the financial crisis. These results are consistent with the regression results, and suggests that investors before the financial crisis could invest in historically top-rated funds to generate higher returns than an investment in historically low-rated funds would have yielded. Contrary, during and after the financial crisis, an investment in historically low-rated funds would have generated the highest returns.

Figure 6.1: Cumulative returns of most profitable J/K decile portfolio strategies

Figure 6.1 plots the cumulative return on 100 NOK invested in equal-weighted indexes of the most profitable decile portfolio strategies for the two sample periods (200201-200803 and 200804-201912). Panel A shows the profitable $J = 12/K = 6$ strategy for the first period, while Panel B shows the profitable $J = 12/K = 1$ strategy for the second period. J is the evaluation period and K is the holding period. The solid green line plots the cumulative indexed return for the top decile portfolio. The solid blue line plots the cumulative indexed return for the bottom decile portfolio.

Figure A: $J = 12/K = 6$ strategy for first period 200201-200803

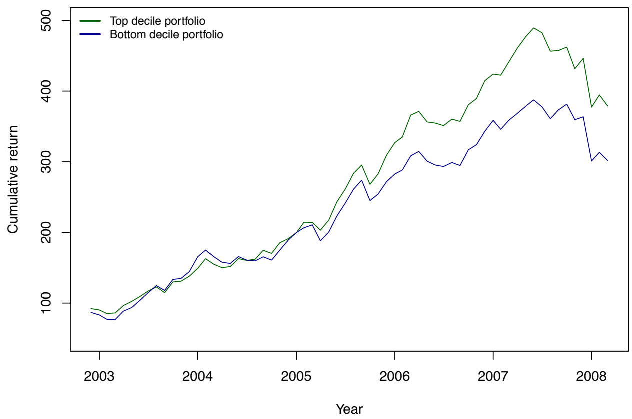
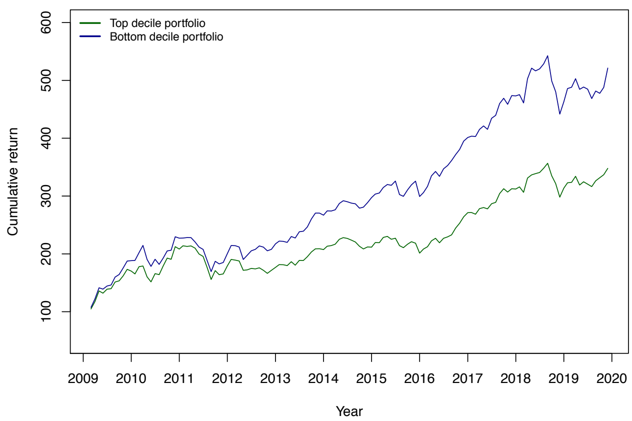


Figure B: $J = 12/K = 1$ strategy for second period 200804-201912



6.3 Limitations

There exist some limitations of the above results worth mentioning.

First, the analysis could experience some issues regarding minimum investment amounts in some funds. There exist funds in the sample that require an investor to invest a minimum amount in the funds. If an investor were to follow the J/K strategy with a limited investment amount available below the investment limit of these funds, some funds could be unavailable. If the funds with these limitations are included in the J/K strategy, the strategy would not be feasible for such investors. Examining the sample of funds included in the J/K strategy, I find that approximately 27% of the included funds require a high minimum investment amount of 1 million NOK or more. This can affect the strategy's feasibility. However, the majority of the high minimum investment funds are available without this high minimum investment restriction in a different share class of the same fund, indicating that the issue is not paramount. Hence, investors with limited capital available can replicate the strategy replacing high minimum investment funds with other share classes of the same fund. Nevertheless, even though the strategy is feasible for such investors, this replacement can yield a somewhat lower overall return because of the different share classes' fee structure.

Second, broker commissions, such as costs associated with purchases and sales of mutual funds, are not considered in the analyses. To examine whether broker commissions would change the prediction results, complex and comprehensive data from multiple sources would be required. However, these commissions are rarely present when investors buy and sell mutual funds. Also, since Morningstar does not account for broker commissions when rating the funds, this thesis does not account for it either, ensuring that the rating methodology and analysis are closely connected.

Third, if the assumptions of the random effects model turn out to be violated in some way, this could affect the credibility of the results. For example, if there exists individual or time fixed effects captured in the model's unique error correlated to the explanatory variables of the model, fixed effects should have been used instead. However, the Hausman test in section A2 in the appendix gives clear indications of the opposite. Therefore, this paper assumes that the assumptions of the random effects model hold.

7 Conclusion

This thesis examines the well-known Morningstar rating system's ability to predict future mutual fund performance, measured by risk-adjusted return, for mutual funds investing primarily in Norwegian equity. This is an important issue because many investors use Morningstar's rating system when deciding upon which funds to invest in. Also, several studies like Sirri and Tufano (1998) and Del Guercio and Tkac (2008) have shown that highly ranked funds attract the greatest investor cash inflow. Examining performance across funds grouped by Morningstar rankings will thus indicate if these cash inflows are justified by subsequent relative performance.

To investigate the Morningstar ratings predictive ability, I implement a data set free of survivorship bias for 136 mutual funds investing primarily in Norwegian equity from January 2002 to December 2019. This allows for an examination of the predictive abilities of the rating system for different periods, prediction horizons, fund investment styles, and with different performance metrics. In general, two periods are examined in this paper: the period before the financial crisis (200201-200803) and the period during and after the financial crisis (200804-201912). Mutual fund performance is evaluated using estimates of alpha from three different factor models of performance measurement: CAPM, Fama-French 3-factor model, and Carhart 4-factor model. Furthermore, both random effects panel data regressions and J/K strategies of buying historically top-rated ("winners") versus low-rated ("losers") funds are employed to investigate the predictive abilities of the Morningstar ratings.

My investigation results in several main findings. First, for the period before the financial crisis, from January 2002 to March 2008, Morningstar is able to "predict" low-performing funds. That is, low ratings from Morningstar generally indicate relatively poor future performance. Second, for the period during and after the financial crisis, from April 2008 to December 2019, low ratings from Morningstar, on the contrary, indicate relatively high future performance. Hence, in the period before the financial crisis, investors could implement top decile ("winner") J/K strategies to generate higher returns than the bottom decile ("loser") strategies. Contrary, during and after the financial crisis, the loser J/K strategies outperformed the winner J/K strategies. Third, there is only weak statistical evidence that Morningstar's highest-rated funds (five-star funds) outperform the next-to-highest, median-rated, and next-to-lowest rated funds (four-, three- and two-star funds) in both periods. The results are relatively robust across the different performance measures and investment styles of funds.

The results have similarities with the paper of Dahl and Madsen (2017). However, my results differ in that the rating system possesses some predictive abilities during and after the financial crisis, but opposite of what one would expect. Furthermore, the results for the first period before the financial crisis are broadly consistent with much of the mutual fund performance persistence literature, in which it is much more difficult to predict superior future performance than to predict poor future performance. However, the results for the second period during and after the financial crisis differ from extant academic studies. I relate the substantial difference in results between the two periods to funds' overall risk-taking and different factor exposures. For instance, the value (HML) factor shows a negative cumulative return over the second period. One-star funds have experienced a negative exposure to this factor over the second period, and vice versa for five-star funds. I argue that this suggests greater performance for the one-star funds relative to the five-star funds in the second period.

Furthermore, consistent with Blake and Morey's (2000) results, my results suggest that investors should be cautious about associating highly rated funds with superior future performance. Although several studies have shown that highly ranked funds attract the greatest investor cash inflow, my results suggest that those cash inflows are not necessarily justified by subsequent performance. However, Morningstar highlights that the rating system is not a predictor of future performance but rather an evaluation of history. Thus, my results do not refute the Morningstar rating system. Nevertheless, many investors use the ratings as indicators of future performance, and the ratings are well used in marketing mutual funds to the public.

In summary, this thesis answers an essential question that mutual fund investors should ask: Do the Morningstar star ratings predict future performance?

References

Aasheim, L. K. (2013). The effect of the morningstar rating on fund flows in the norwegian mutual fund market. Master's thesis, ISCTE Business School, Lisboa, Portugal.

Blake, C. R. and Morey, M. R. (2000). Morningstar ratings and mutual fund performance. *Journal of financial and Quantitative Analysis*, 35(3):451–483.

Blume, M. E. (1998). An anatomy of morningstar ratings. *Financial Analysts Journal*, 54(2):19–27.

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

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

Elton, E. J., Gruber, M. J., and Blake, C. R. (1996). Survivor bias and mutual fund performance. *The review of financial studies*, 9(4):1097–1120.

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

Guercio, D. D. and Tkac, P. A. (2008). Star power: The effect of morningstar ratings on mutual fund flow. *Journal of Financial and Quantitative Analysis*, pages 907–936.

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

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

Khorana, A. and Nelling, E. (1998). The determinants and predictive ability of mutual fund ratings. *The Journal of Investing*, 7(3):61–66.

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

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

Madsen, H. M. and Dahl, O. A. S. (2017). Performance prediction in dine penger's norwegian mutual fund ratings: evidence from ratings in the financial media. Master's thesis, Norwegian School of Economics (NHH), Bergen, Norway.

Morey, M. R. (2002). Mutual fund age and morningstar ratings. *Financial Analysts Journal*, 58(2):56–63.

Morey, M. R. and Gottesman, A. A. (2006). Morningstar mutual fund ratings redux. *Pace University Finance Research Paper*, (2006/03).

Morningstar (2016). *The Morningstar Rating for Funds*. Morningstar Research Report.

Morningstar (2019). *Morningstar Category Definitions: Europe, Asia and South Africa*. Morningstar Research Report.

Norges Bank (2014). Short term interest rates. https://www.norges-bank.no/en/topics/Statistics/Interest-rates/. [Online; accessed 15-September-2020].

Norwegian Fund and Asset Management Association (2020). Fondsundersøkelsen 2020. https://www.vff.no/news/2020/ny-undersokelse-rekordmange-sparer-i-fond-i-2020. [Online; accessed 03-December-2020].

Oslo Børs (2020). Statistics. https://www.oslobors.no/ob_eng/Oslo-Boers/Statistics. [Online; accessed 15-September-2020].

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

Sharpe, W. F. (1998). Morningstar's risk-adjusted ratings. *Financial Analysts Journal*, 54(4):21–33.

Sirri, E. R. and Tufano, P. (1998). Costly search and mutual fund flows. *The journal of finance*, 53(5):1589–1622.

Sørensen, L. Q. (2009). Mutual fund performance at the oslo stock exchange. *Available at SSRN 1488745*.

Ødegaard, B. A. (2020). Asset pricing data at ose. http://finance.bi.no/~bernt/financial_data/ose_asset_pricing_data/index.html. [Online; accessed 19-October-2020].

Appendix

A1 Norwegian equity mutual fund sample

Table A1.1: List of mutual funds in data sample

Table A1.1 displays an overview of the 136 fund observations in the data sample. Please note that some of the funds have changed name during the sample period from 2002 to 2019. Also note that some of the funds are new funds that have emerged during the sample period and some of the funds have "died" during the sample period, while other funds lived throughout the whole sample period.

Alfred Berg Aktiv	FORTE Norge
Alfred Berg Aktiv II	FORTE Trønder
Alfred Berg Gambak	Globus Aktiv Acc
Alfred Berg Humanfond	Globus Norge II Acc
Alfred Berg Indeks C	Handelsbanken Norge
Alfred Berg Indeks I	Handelsbanken Norge (A1)
Alfred Berg Norge +	Handelsbanken Norge (A10)
Alfred Berg Norge C	Handelsbanken Norge Index (A1)
Alfred Berg Norge Etisk	Handelsbanken Norge Index (A9)
Alfred Berg Norge I	Holberg Norge A
Arctic Norwegian Equities A	KLP AksjeNorge
Arctic Norwegian Equities B	KLP AksjeNorge Indeks
Arctic Norwegian Equities D	KLP AksjeNorge Indeks II
Arctic Norwegian Equities E	Landkreditt Norge
Arctic Norwegian Equities I	Landkreditt Utbytte A
Arctic Norwegian Value Creation A	Landkreditt Utbytte I
Arctic Norwegian Value Creation B	NB Aksjefond
Arctic Norwegian Value Creation C	Nordea 1 - Norwegian Equity BP
Arctic Norwegian Value Creation D	Nordea Avkastning
C WorldWide Norge	Nordea Kapital
C WorldWide Norge III	Nordea Norge Pluss
Carnegie Norge Indeks	Nordea Norge Verdi
Danske Invest Engros Norske A I R	Nordea Norwegian Stars A
Danske Invest Norge I	Nordea Norwegian Stars Fund C growth
Danske Invest Norge II	Nordea Norwegian Stars I
Danske Invest Norge Vekst	Nordea SMB
Danske Invest Norske Aksjer Inst I	Nordea Vekst
Danske Invest Norske Aksjer Inst II	Nordnet Indeksfond Norge
Delphi Norge A	ODIN Norge A
Delphi Vekst	ODIN Norge B
DIX Norway Restr	ODIN Norge C
DIX Norway Restr W	ODIN Norge D
DNB Norge (Avanse I)	ODIN Norge II
DNB Norge (Avanse II)	Pareto Aksje Norge A
DNB Norge (I)	Pareto Aksje Norge B
DNB Norge (III)	Pareto Aksje Norge C
DNB Norge A	Pareto Aksje Norge D
DNB Norge C	Pareto Aksje Norge I
DNB Norge D	Pareto Investment Fund A
DNB Norge Indeks A	Pareto Investment Fund B
DNB Norge N	Pareto Investment Fund C
DNB Norge R	PLUSS Aksje
DNB Norge Selektiv	PLUSS Indeks
DNB Norge Selektiv (II)	PLUSS Markedsverdi
DNB Norge Selektiv A	RF Aksjefond Acc
DNB Norge Selektiv C	RF Plussfond Acc
DNB Norge Selektiv E	Sbanken Framgang Sammen
DNB Norge Selektiv N	SEB Norway Focus Fund C
DNB Norge Selektiv R	SEB Norway Focus Fund HNWC
DNB SMB A	SEB Norway Focus Fund IC
DNB SMB N	SR-Bank Norge A
DNB SMB R	SR-Bank Norge B
Eika Norge	SR-Bank Norge C
Eika SMB	SR-Bank Norge D
Equinor Aksjer Norge	SSgA Norway Index Equity Fund I
FIRST Aksjer Norge	Storebrand Aksje Innland
FIRST Aksjer Norge Kl.I	Storebrand Indeks - Norge A
FIRST Aksjer Norge Kl.III	Storebrand Norge A
FIRST Generator A	Storebrand Norge B
FIRST Generator S	Storebrand Norge Fossilfri A
FIRST Norge Verdi	Storebrand Norge H
FIRST Norway	Storebrand Norge I
First Norway Delta Kl.IV (LAMP)	Storebrand Norge Institusjon
FIRST Opportunities	Storebrand Optima Norge B
FIRST SMB	Storebrand Vekst A
Fondsfinans Norge	Storebrand Verdi A
Fondsfinans Utbytte	Storebrand Verdi N
Formue Diversifiserte Norske Aksjer	Terra Norge

A2 Hausman test

Table A2.1: Hausman test

Table A2.1 shows the Hausman test for the CAPM, Fama-French and Carhart models. The Hausman test can be used to differentiate between fixed effects model and random effects model in panel analysis. The null hypothesis in the Hausman test is that the assumption that makes random effects both consistent and efficient holds, i.e. zero correlation between unique errors and independent variables. Thus, if the test fails to reject the null, random effects should be used. If the test rejects the null, fixed effects should be used. For all models, I can not reject the null, and I thus prefer random effects over fixed effects.

Panel A: CAPM

data: $\alpha \sim$ 4-star + 3-star + 2-star + 1-star
chisq = 3.821, df = 4, p-value = 0.431
alternative hypothesis: one model is inconsistent

Panel B: Fama-French

data: $\alpha \sim$ 4-star + 3-star + 2-star + 1-star
chisq = 0.653, df = 4, p-value = 0.957
alternative hypothesis: one model is inconsistent

Panel C: Carhart

data: $\alpha \sim$ 4-star + 3-star + 2-star + 1-star
chisq = 0.448, df = 4, p-value = 0.978
alternative hypothesis: one model is inconsistent

A3 Regression for full sample period

Table A3.1: Predictive regression for full sample period: 200201-201912

The table shows the random effects panel data regression for the full sample period from January 2002 to December 2019. The regression is estimated for three performance metrics: Jensen's alpha (CAPM), Fama-French three-factor alpha (FF) and Carhart four-factor alpha (Carhart). For each model, in order to measure future performance, dummy variables for lagged 4-star rating, 3-star rating, 2-star rating and 1-star rating are included. The 5-star rating is the reference group. The lag periods are 1, 3, 6, 12 and 24 months. Standard errors are in parenthesis.

	Lag = 1 month			Lag = 3 months			Lag = 6 months			Lag = 12 months			Lag = 24 months		
	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart
β_0 (Constant)	0.001	−0.0002	−0.0001	0.001	−0.0001	−0.00002	0.001**	0.0002	0.0003	0.001	−0.0001	−0.0001	0.001	−0.0001	−0.0001
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.001)	(0.0005)	(0.0005)	(0.001)	(0.0005)	(0.0005)	(0.001)	(0.001)	(0.001)
β_4 (4-star)	0.00002	0.0002	0.0001	−0.0003	−0.0002	−0.0002	−0.001	−0.0004	−0.0005	−0.0002	−0.0001	−0.0001	−0.0005	−0.0005	−0.0005
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_3 (3-star)	−0.001	−0.001	−0.001	−0.0004	−0.0004	−0.0004	−0.001	−0.001	−0.001	−0.0001	−0.0002	−0.0002	0.001	0.0004	0.0004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_2 (2-star)	−0.00003	−0.0001	−0.0001	0.00005	0.00001	0.00000	−0.0003	−0.0002	−0.0002	0.00002	0.00002	0.00001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_1 (1-star)	0.001	0.001	0.001	−0.0003	−0.0003	−0.0003	−0.0002	−0.0002	−0.0002	0.002**	0.002*	0.002*	−0.0002	−0.0003	−0.0003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	12,269	12,269	12,269	12,048	12,048	12,048	11,733	11,733	11,733	11,117	11,117	11,117	9,971	9,971	9,971
Adjusted R^2	0.909	0.913	0.913	0.910	0.913	0.913	0.909	0.913	0.913	0.908	0.911	0.911	0.909	0.913	0.913

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

A4 Regressions for investment styles

A4.1 First period: 200201-200803

Table A4.1: Predictive regression for value style funds for first period: 200201-200803

Table A4.1 shows the random effects panel data regression for value style funds for the first period from January 2002 to March 2008. The regression is estimated for three performance metrics: Jensen's alpha (CAPM), Fama-French three-factor alpha (FF) and Carhart four-factor alpha (Carhart). For each model, in order to measure future performance, dummy variables for lagged 4-star rating, 3-star rating, 2-star rating and 1-star rating are included. The 5-star rating is the reference group. The lag periods are 1, 3, 6, 12 and 24 months. Standard errors are in parenthesis.

	Lag = 1 month			Lag = 3 months			Lag = 6 months			Lag = 12 months			Lag = 24 months		
	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart
β_0 (Constant)	0.003	0.001	−0.0004	0.002	0.0005	0.0005	0.002	0.001	0.001	0.004*	0.002	0.002	−0.002	−0.003	−0.003
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
β_4 (4-star)	−0.002	−0.001	−0.0001	−0.002	−0.001	−0.001	−0.002	−0.001	−0.001	−0.004	−0.003	−0.003	0.001	0.002	0.002
	(0.003)	(0.003)	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
β_3 (3-star)	−0.003	−0.002	−0.0003	−0.002	−0.001	−0.001	−0.001	−0.001	−0.001	−0.004*	−0.004*	−0.004*	0.004	0.004	0.004
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
β_2 (2-star)	−0.003	−0.002	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	0.002	0.002	0.002	0.004	0.004	0.004
	(0.003)	(0.002)	(0.001)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
β_1 (1-star)	−0.007	−0.004	−0.002	−0.003	−0.001	−0.001	0.002	0.002	0.002	−0.005	−0.008	−0.008	0.008	0.008	0.008
	(0.006)	(0.006)	(0.001)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	625	625	3,366	605	605	605	575	575	575	515	515	515	395	395	395
Adjusted R^2	0.923	0.928	0.925	0.924	0.929	0.929	0.921	0.926	0.926	0.915	0.921	0.921	0.907	0.914	0.914

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Table A4.2: Predictive regression for growth style funds for first period: 200201-200803

Table A4.2 shows the random effects panel data regression for growth style funds for the first period from January 2002 to March 2008. The regression is estimated for three performance metrics: Jensen's alpha (CAPM), Fama-French three-factor alpha (FF) and Carhart four-factor alpha (Carhart). For each model, in order to measure future performance, dummy variables for lagged 4-star rating, 3-star rating, 2-star rating and 1-star rating are included. The 5-star rating is the reference group. The lag periods are 1, 3, 6, 12 and 24 months. Standard errors are in parenthesis.

	Lag = 1 month			Lag = 3 months			Lag = 6 months			Lag = 12 months			Lag = 24 months		
	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart
β_0 (Constant)	−0.003	−0.003	−0.002	−0.002	−0.003	−0.002	0.001	−0.001	−0.0001	0.003	−0.0002	−0.0001	0.0005	−0.001	−0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
β_4 (4-star)	0.004	0.001	0.001	0.004	0.003	0.002	0.002	0.001	0.001	−0.001	−0.0001	−0.0001	−0.003	−0.004	−0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
β_3 (3-star)	0.003	−0.00002	−0.0003	−0.001	−0.002	−0.002	−0.004	−0.004	−0.004	−0.003	−0.002	−0.002	0.003	0.003	0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
β_2 (2-star)	0.001	−0.001	−0.001	0.001	−0.0004	−0.001	−0.001	−0.002	−0.001	−0.003	−0.003	−0.003	0.003	0.001	0.001
	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
β_1 (1-star)	0.002	−0.001	−0.002	−0.0002	−0.001	−0.002	−0.002	−0.002	−0.002	−0.004	−0.004	−0.004	−0.007*	−0.009**	−0.009**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Observations	799	799	799	771	771	771	738	738	738	672	672	672	531	531	531
Adjusted R^2	0.860	0.881	0.883	0.864	0.884	0.885	0.857	0.876	0.878	0.837	0.857	0.857	0.847	0.869	0.870

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Table A4.3: Predictive regression for blend style funds for first period: 200201-200803

Table A4.3 shows the random effects panel data regression for blend style funds for the first period from January 2002 to March 2008. The regression is estimated for three performance metrics: Jensen's alpha (CAPM), Fama-French three-factor alpha (FF) and Carhart four-factor alpha (Carhart). For each model, in order to measure future performance, dummy variables for lagged 4-star rating, 3-star rating, 2-star rating and 1-star rating are included. The 5-star rating is the reference group. The lag periods are 1, 3, 6, 12 and 24 months. Standard errors are in parenthesis.

	Lag = 1 month			Lag = 3 months			Lag = 6 months			Lag = 12 months			Lag = 24 months		
	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart
β_0 (Constant)	0.001	−0.0001	−0.00004	0.001	0.0002	0.0003	0.001	−0.0003	−0.0003	0.002	0.001	0.001	0.002	0.0001	−0.0003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_4 (4-star)	−0.001	−0.0005	−0.0005	0.00003	0.00005	0.00004	0.001	0.001	0.001	−0.001	−0.001	−0.001	−0.0001	0.0002	0.0003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_3 (3-star)	−0.001	−0.0005	−0.0005	−0.001	−0.001	−0.001	−0.0003	−0.0003	−0.0003	−0.001	−0.001	−0.001	−0.0001	0.0002	0.0003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_2 (2-star)	−0.001	−0.001	−0.001	−0.002*	−0.002*	−0.002*	−0.001	−0.001	−0.001	−0.003*	−0.003**	−0.003**	−0.002	−0.002	−0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_1 (1-star)	−0.002	−0.002	−0.002	−0.002	−0.002	−0.002	−0.003*	−0.003*	−0.003*	−0.0002	−0.0005	−0.0005	−0.003	−0.003	−0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	2,021	2,021	2,021	1,957	1,957	1,957	1,861	1,861	1,861	1,675	1,675	1,675	1,309	1,309	1,309
Adjusted R^2	0.949	0.950	0.950	0.949	0.950	0.950	0.946	0.948	0.948	0.943	0.945	0.945	0.938	0.941	0.941

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

A4.2 Second period: 200804-201912

Table A4.4: Predictive regression for value style funds for second period: 200804-201912

Table A4.4 shows the random effects panel data regression for value style funds for the second period from April 2008 to December 2019. The regression is estimated for three performance metrics: Jensen's alpha (CAPM), Fama-French three-factor alpha (FF) and Carhart four-factor alpha (Carhart). For each model, in order to measure future performance, dummy variables for lagged 4-star rating, 3-star rating, 2-star rating and 1-star rating are included. The 5-star rating is the reference group. The lag periods are 1, 3, 6, 12 and 24 months. Standard errors are in parenthesis.

	Lag = 1 month			Lag = 3 months			Lag = 6 months			Lag = 12 months			Lag = 24 months		
	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart
β_0 (Constant)	−0.002	−0.002	−0.001	−0.001	−0.002	−0.001	−0.0005	−0.001	−0.0003	−0.001	−0.001	−0.0004	0.001	−0.0001	0.0002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
β_4 (4-star)	0.003	0.002	0.002	0.001	0.001	0.001	−0.0003	0.0003	−0.0003	−0.0002	−0.0001	−0.00003	−0.0004	0.0002	0.0003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
β_3 (3-star)	−0.001	−0.002	−0.002	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	0.0004	0.0003	−0.0001	−0.002	−0.002	−0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
β_2 (2-star)	0.0001	−0.0003	−0.0002	0.001	0.001	0.001	−0.0004	−0.0004	−0.001	−0.001	−0.001	−0.001	−0.002	−0.002	−0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
β_1 (1-star)	0.003	0.002	0.002	−0.001	−0.002	−0.002	−0.002	−0.002	−0.002	0.001	0.001	0.001	−0.001	−0.001	−0.0001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Observations	1,305	1,305	1,305	1,289	1,289	1,289	1,265	1,265	1,265	1,216	1,216	1,216	1,132	1,132	1,132
Adjusted R^2	0.874	0.881	0.882	0.873	0.880	0.882	0.873	0.881	0.882	0.878	0.885	0.887	0.888	0.895	0.896

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Table A4.5: Predictive regression for growth style funds for second period: 200804-201912

Table A4.5 shows the random effects panel data regression for growth style funds for the second period from April 2008 to December 2019. The regression is estimated for three performance metrics: Jensen's alpha (CAPM), Fama-French three-factor alpha (FF) and Carhart four-factor alpha (Carhart). For each model, in order to measure future performance, dummy variables for lagged 4-star rating, 3-star rating, 2-star rating and 1-star rating are included. The 5-star rating is the reference group. The lag periods are 1, 3, 6, 12 and 24 months. Standard errors are in parenthesis.

	Lag = 1 month			Lag = 3 months			Lag = 6 months			Lag = 12 months			Lag = 24 months		
	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart
β_0 (Constant)	0.002	0.001	0.0003	0.002	0.001	0.0001	0.003**	0.002	0.001	0.0005	−0.001	−0.001	0.001	0.0002	−0.0001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_4 (4-star)	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.002	−0.002	−0.002	0.0004	0.001	0.001	−0.002	−0.001	−0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
β_4 (3-star)	−0.001	−0.001	−0.001	−0.001	−0.001	−0.0003	−0.002	−0.002	−0.001	0.00004	0.001	0.001	0.0004	0.001	0.0003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
β_2 (2-star)	−0.001	−0.001	−0.0001	−0.001	−0.001	−0.0001	−0.001	−0.001	−0.0002	0.001	0.002	0.002	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
β_1 (1-star)	0.007***	0.007***	0.007***	0.008***	0.008***	0.009***	0.007***	0.007***	0.007***	0.009***	0.010***	0.010***	0.003	0.003	0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	1,767	1,767	1,767	1,749	1,749	1,749	1,722	1,722	1,722	1,666	1,666	1,666	1,582	1,582	1,582
Adjusted R^2	0.883	0.888	0.888	0.884	0.889	0.889	0.884	0.888	0.889	0.885	0.889	0.890	0.884	0.889	0.889

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Table A4.6: Predictive regression for blend style funds for second period: 200804-201912

Table A4.6 shows the random effects panel data regression for blend style funds for the second period from April 2008 to December 2019. The regression is estimated for three performance metrics: Jensen's alpha (CAPM), Fama-French three-factor alpha (FF) and Carhart four-factor alpha (Carhart). For each model, in order to measure future performance, dummy variables for lagged 4-star rating, 3-star rating, 2-star rating and 1-star rating are included. The 5-star rating is the reference group. The lag periods are 1, 3, 6, 12 and 24 months. Standard errors are in parenthesis.

	Lag = 1 month			Lag = 3 months			Lag = 6 months			Lag = 12 months			Lag = 24 months		
	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart	CAPM	FF	Carhart
β_0 (Constant)	0.001*	0.001	0.001	0.001	0.0005	0.001	0.001	0.0004	0.001	−0.0002	−0.0004	−0.0003	0.0002	0.00002	0.0002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_4 (4-star)	−0.0005	−0.0003	−0.0003	−0.001	−0.0005	−0.0005	−0.001	−0.0004	−0.0004	0.0004	0.0004	0.0004	−0.0002	−0.0004	−0.0004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_3 (3-star)	−0.001	−0.001	−0.001	−0.0002	−0.0003	−0.0002	0.00004	−0.0001	−0.00003	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_2 (2-star)	0.0005	0.0004	0.0004	0.001	0.001	0.001	0.001	0.001	0.001	0.001*	0.001*	0.001*	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_1 (1-star)	−0.0002	0.0001	0.0001	−0.001	−0.0003	−0.0002	0.001	0.001	0.001	0.002**	0.002**	0.002**	0.002	0.002	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	5,307	5,307	5,307	5,246	5,246	5,246	5,162	5,162	5,162	5,000	5,000	5,000	4,718	4,718	4,718
Adjusted R^2	0.927	0.929	0.929	0.927	0.929	0.929	0.927	0.929	0.929	0.928	0.930	0.930	0.930	0.931	0.931

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

A5 Quartile portfolio J/K-strategy

A5.1 First period: 200201-200803

Table A5.1: Returns of quartile portfolios for first period: 200201-200803

Table A5.1 reports the average annualized returns of the quartile portfolios for the first sample period from January 2002 to March 2008. The portfolios are formed based on J-month lagged ratings and held for K months. The values of the J and K for the different strategies are indicated in the first column and row, respectively. Panel A shows the returns for the top quartile, while Panel B shows the returns for the bottom quartile. The t-statistics are reported in parantheses.

Panel A: Top quartile

	K=1	K=3	K=6	K=12	K=24
J=1	0.1680*	0.1717*	0.1693*	0.1685*	0.1732*
	(1.89)	(1.93)	(1.91)	(1.88)	(1.94)
J=3	0.1798*	0.1777*	0.1792*	0.1806*	0.1852**
	(1.97)	(1.95)	(1.97)	(1.97)	(2.01)
J=6	0.1897**	0.1918**	0.1941**	0.1896**	0.1882**
	(2.04)	(2.06)	(2.09)	(2.04)	(2.03)
J=12	0.2595***	0.2579***	0.2562***	0.2599***	0.2510***
	(3)	(2.99)	(2.97)	(3.04)	(2.96)
J=24	0.2423***	0.2403***	0.2374**	0.2378**	0.2465***
	(2.73)	(2.7)	(2.66)	(2.65)	(2.76)

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Panel B: Bottom quartile

	K=1	K=3	K=6	K=12	K=24
J=1	0.1658*	0.1639*	0.1642*	0.1655*	0.1601*
	(1.82)	(1.78)	(1.78)	(1.8)	(1.75)
J=3	0.1637*	0.1675*	0.1684*	0.1676*	0.1621*
	(1.72)	(1.76)	(1.76)	(1.77)	(1.73)
J=6	0.1763*	0.1786*	0.1749*	0.1834*	0.1793*
	(1.81)	(1.83)	(1.79)	(1.89)	(1.87)
J=12	0.2383**	0.2376**	0.2372**	0.2381**	0.2416**
	(2.55)	(2.55)	(2.54)	(2.55)	(2.58)
J=24	0.2213**	0.2224**	0.2219**	0.2194**	0.2116**
	(2.38)	(2.38)	(2.37)	(2.33)	(2.3)

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Table A5.2: Alphas of quartile portfolios for first period: 200201-200803

Table A5.2 reports the annualized performance of the quartile portfolios for three performance measures: Jensen's alpha, Fama-French 3-factor alpha and Carhart 4-factor alpha. The sample period is the first period from January 2002 to March 2008. The portfolios are formed based on J-month lagged ratings and held for K months. The values of the J and K for the different strategies are indicated in the first column and row, respectively. Panel A shows the alphas for the top quartile, while Panel B shows the alphas for the bottom quartile.

Panel A: Top quartile

	Performance measure	K=1	K=3	K=6	K=12	K=24
	Jensen's alpha	0.00055	0.00445	0.00273	0.00038	0.00569
J=1	Fama-French 3-factor alpha	-0.0138	-0.00925	-0.01044	-0.01329	-0.01299
	Carhart 4-factor alpha	-0.01915	-0.01542	-0.01475	-0.01113	-0.00871
	Jensen's alpha	0.00588	0.00415	0.00612	0.00644	0.01113
J=3	Fama-French 3-factor alpha	-0.00469	-0.00662	-0.00539	-0.00863	-0.00768
	Carhart 4-factor alpha	-0.0103	-0.01168	-0.00844	-0.00688	-0.00178
	Jensen's alpha	0.00814	0.0102	0.01276	0.0088	0.00847
J=6	Fama-French 3-factor alpha	-0.00545	-0.00397	-0.0011	-0.009	-0.01347
	Carhart 4-factor alpha	-0.0089	-0.00599	-0.00151	-0.00814	-0.00761
	Jensen's alpha	0.02427	0.02341	0.02202	0.02732***	0.02012
J=12	Fama-French 3-factor alpha	0.0061	0.00429	0.00275	0.00277	-0.00559
	Carhart 4-factor alpha	-0.00237	-0.00473	-0.0067	-0.00338	-0.01005
	Jensen's alpha	0.01964	0.0174	0.01403	0.0134	0.0224
J=24	Fama-French 3-factor alpha	-0.0032	-0.00494	-0.00887	-0.00968	0.00469
	Carhart 4-factor alpha	-0.00701	-0.0086	-0.01241	-0.01407	-0.00161

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Panel B: Bottom quartile

	Performance measure	K=1	K=3	K=6	K=12	K=24
	Jensen's alpha	-0.00365	-0.00639	-0.00678	-0.00497	-0.01019
J=1	Fama-French 3-factor alpha	-0.02116	-0.02651	-0.02549	-0.02043	-0.02193
	Carhart 4-factor alpha	-0.00868	-0.01392	-0.01497	-0.01613	-0.02027
	Jensen's alpha	-0.01449	-0.01084	-0.01105	-0.01052	-0.01523
J=3	Fama-French 3-factor alpha	-0.03418***	-0.02998	-0.02921	-0.02738	-0.02989***
	Carhart 4-factor alpha	-0.02239	-0.01867	-0.01898	-0.02024	-0.02658***
	Jensen's alpha	-0.01072	-0.00883	-0.01204	-0.00348	-0.00574
J=6	Fama-French 3-factor alpha	-0.03463***	-0.03171	-0.0338***	-0.02234	-0.02346
	Carhart 4-factor alpha	-0.02286	-0.02017	-0.02341	-0.01638	-0.01905
	Jensen's alpha	-0.00994	-0.01041	-0.01139	-0.0099	-0.00784
J=12	Fama-French 3-factor alpha	-0.03652***	-0.03628***	-0.03816***	-0.03745***	-0.03628***
	Carhart 4-factor alpha	-0.02678	-0.02682	-0.029	-0.02825	-0.03088
	Jensen's alpha	-0.0072	-0.00687	-0.00749	-0.01092	-0.01547
J=24	Fama-French 3-factor alpha	-0.03038	-0.02977	-0.03043	-0.03524	-0.03334
	Carhart 4-factor alpha	-0.03535	-0.03473	-0.03639	-0.04126	-0.03746***

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

A5.2 Second period: 200804-201912

Table A5.3: Returns of quartile portfolios for second period: 200804-201912

Table 6.5 reports the average annualized returns of the quartile portfolios for the second sample period from April 2008 to December 2019. The portfolios are formed based on J-month lagged ratings and held for K months. The values of the J and K for the different strategies are indicated in the first column and row, respectively. Panel A shows the returns for the top quartile, while Panel B shows the returns for the bottom quartile. The t-statistics are reported in parantheses.

Panel A: Top quartile

	K=1	K=3	K=6	K=12	K=24
J=1	0.0878 (1.63)	0.0866 (1.62)	0.0872 (1.63)	0.0850 (1.59)	0.0862 (1.63)
J=3	0.0736 (1.36)	0.0732 (1.36)	0.0747 (1.4)	0.0699 (1.31)	0.0765 (1.44)
J=6	0.0825 (1.51)	0.0836 (1.54)	0.0816 (1.5)	0.0838 (1.55)	0.0912* (1.69)
J=12	0.1408*** (3.14)	0.1415*** (3.16)	0.1417*** (3.17)	0.1425*** (3.24)	0.1452*** (3.31)
J=24	0.0990** (2.28)	0.1003** (2.32)	0.1009** (2.34)	0.1037** (2.42)	0.1019** (2.37)

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Panel B: Bottom quartile

	K=1	K=3	K=6	K=12	K=24
J=1	0.0988* (1.8)	0.0992* (1.81)	0.0978* (1.78)	0.0960* (1.75)	0.0917* (1.67)
J=3	0.0827 (1.49)	0.0787 (1.42)	0.0852 (1.53)	0.0844 (1.53)	0.0773 (1.4)
J=6	0.0965* (1.73)	0.0938* (1.68)	0.0959* (1.72)	0.0953* (1.72)	0.0933* (1.68)
J=12	0.1567*** (3.45)	0.1557*** (3.46)	0.1540*** (3.43)	0.1551*** (3.41)	0.1562*** (3.44)
J=24	0.1087** (2.46)	0.1093** (2.48)	0.1108** (2.5)	0.1147** (2.57)	0.1198*** (2.69)

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Table A5.4: Alphas of quartile portfolios for second period: 200804-201912

Table A5.4 reports the annualized performance of the quartile portfolios for three performance measures: Jensen's alpha, Fama-French 3-factor alpha and Carhart 4-factor alpha. The sample period is the second period from April 2008 to December 2019. The portfolios are formed based on J-month lagged ratings and held for K months. The values of the J and K for the different strategies are indicated in the first column and row, respectively. Panel A shows the alphas for the top quartile, while Panel B shows the alphas for the bottom quartile.

Panel A: Top quartile

	Performance measure	K=1	K=3	K=6	K=12	K=24
	Jensen's alpha	0.00421	0.00326	0.00404	0.00199	0.00372
J=1	Fama-French 3-factor alpha	-0.00063	-0.00133	-0.00067	-0.00241	-0.00205
	Carhart 4-factor alpha	0.00331	0.00213	0.00302	0.00143	0.00182
	Jensen's alpha	0.00238	0.00219	0.00417	-0.00053	0.00624
J=3	Fama-French 3-factor alpha	-0.0022	-0.00267	-0.00085	-0.00523	-0.00013
	Carhart 4-factor alpha	0.00201	0.00135	0.00247	-0.00162	0.00284
	Jensen's alpha	-0.00139	0.000002	-0.00201	0.00053	0.00807
J=6	Fama-French 3-factor alpha	-0.00617	-0.00494	-0.00624	-0.00464	0.0023
	Carhart 4-factor alpha	-0.00103	0.00085	-0.00051	0.00019	0.00507
	Jensen's alpha	-0.00636	-0.00537	-0.00487	-0.00203	0.00075
J=12	Fama-French 3-factor alpha	-0.0138***	-0.01287	-0.01233	-0.00979	-0.00593
	Carhart 4-factor alpha	-0.00817	-0.00739	-0.00709	-0.00412	-0.00163
	Jensen's alpha	-0.00614	-0.00462	-0.00376	-0.00031	-0.00255
J=24	Fama-French 3-factor alpha	-0.01161	-0.00991	-0.00874	-0.00528	-0.00837
	Carhart 4-factor alpha	-0.00521	-0.00333	-0.00224	0.00015	-0.0039

Signif. codes: *p<0.1; **p<0.05; ***p<0.01

Panel B: Bottom quartile

	Performance measure	K=1	K=3	K=6	K=12	K=24
	Jensen's alpha	0.01422	0.01455	0.01304	0.01123	0.00679
J=1	Fama-French 3-factor alpha	0.00774	0.00767	0.00578	0.00415	0.00129
	Carhart 4-factor alpha	0.00575	0.00599	0.00517	0.00352	0.00078
	Jensen's alpha	0.01031	0.00637	0.01273	0.01225	0.00501
J=3	Fama-French 3-factor alpha	0.00239	-0.00134	0.00449	0.00298	-0.00187
	Carhart 4-factor alpha	0.0023	-0.00168	0.00419	0.0001	-0.00292
	Jensen's alpha	0.01142	0.00881	0.01098	0.01079	0.00856
J=6	Fama-French 3-factor alpha	0.00315	0.00108	0.00313	0.00312	0.00192
	Carhart 4-factor alpha	0.00395	0.00147	0.0028	0.0013	0.00196
	Jensen's alpha	0.00939	0.00951	0.00808	0.008	0.00844
J=12	Fama-French 3-factor alpha	-0.00128	-0.00137	-0.00235	-0.00196	0.00035
	Carhart 4-factor alpha	-0.00117	-0.0019	-0.00244	-0.00109	0.00124
	Jensen's alpha	0.00349	0.00428	0.00547	0.0088	0.01427
J=24	Fama-French 3-factor alpha	-0.00439	-0.00429	-0.00332	0.00056	0.00545
	Carhart 4-factor alpha	-0.0093	-0.0091	-0.00788	-0.00504	-0.00096

Signif. codes: *p<0.1; **p<0.05; ***p<0.01