Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
Dept. of Industrial Economics and Technology Management

Lyngedal, Magnus
Pedersen, Ludvig

Incentive Fee Fund Performance Evaluation and Multi-Dimensional Risk Classification using Self-Organizing Maps

Master's thesis in Industrial Economics and Technology Management

Supervisor: Belsom, Einar

June 2020

NTNU
Norwegian University of Science and Technology

Lyngedal, Magnus
Pedersen, Ludvig

Incentive Fee Fund Performance Evaluation and Multi-Dimensional Risk Classification using Self-Organizing Maps

Master's thesis in Industrial Economics and Technology Management
Supervisor: Belsom, Einar
June 2020

Norwegian University of Science and Technology
Faculty of Economics and Management
Dept. of Industrial Economics and Technology Management

NTNU
Norwegian University of
Science and Technology

Preface

This master's thesis is written during the spring of 2020 at the Norwegian University of Science and Technology (NTNU), and it marks the end of our Master of Science degree in Industrial Economics and Technology Management. The purpose of this thesis is dual. First, we examine a particular class of mutual funds, and second, we explore methodology that has seen little use in financial literature.

We would like to thank our supervisor, Einar Belsom, for invaluable support and advice throughout the work with this thesis. We also thank those who have assisted in the collection of data, namely Sjur Resaland of the Oslo Stock Exchange and the team behind TITLON at the University of Tromsø.

Magnus Lyngedal

Ludvig Pedersen

Trondheim, June 11, 2020

Abstract

We examine the performance and map the risk-taking behaviour of incentive fee equity funds registered on the Oslo Stock Exchange in the period 2000-2018. We map risk properties by a self-organizing map, which is an unsupervised neural network, and cluster its output using the k-means and the hierarchical clustering algorithms. Some incentive fee fund managers increase volatility and market beta in attempt to earn positive fees. In contrast, others take nuanced exposure to systematic or unsystematic factors to beat their benchmark, perhaps in the belief of possessing skill. Assessing risk-changing behaviour, we find indication that incentive contracts add to the convexity of flow-related incentives. For performance, we find no verification of any theoretical inclination that incentive fee funds attract the best or hardest-working managers.

Sammendrag

Vi undersøker prestasjon og kartlegger måten aksjefond med resultatbasert forvaltningsgodtgjørelse som er registrerte på Oslo Børs i perioden 2000-2018 tar risiko på. Vi kartlegger risikoegenskaper ved et selvorganiserende kart, som er et uovervåket nevralt nettverk, og klynger utdataen fra kartet ved k-means- og hierarkisk gruppering. Noen forvaltere av insentivfond øker volatiliteten og markedsbetaen i forsøk på å tjene positive resultatavhengige honorarer. I kontrast til det, tar andre nyansert eksponering for systematiske eller usystematiske faktorer for å slå sin referanseindeks, kanskje i troen på egen dyktighet. Når vi vurderer atferd for risikoendring, finner vi indikasjon på at insentivkontrakter forsterker konveksiteten fra tegningsinsentiver. Når det gjelder prestasjon, finner vi ingen verifikasjon av noen teoretisk tilbøyelighet for at insentivfond tiltrekker seg de beste eller hardest arbeidende forvalterne.

Contents

List of Figures

List of Tables

1 Introduction

Mutual funds manage a pool of assets on behalf of their investors for a fee. These fees are usually paid yearly and given as a fraction of assets under management, but fee structures vary across funds. The structure of the fee has implications for the way fund managers take on risk. On the other side of the relationship, the mutual fund investor is interested in the maximization of risk-adjusted net returns on their assets regardless of any fees accrued. In this thesis, we study the implication that the contracts of so-called incentive fee mutual funds have on the relationship between these parties.

Incentive fee funds are mutual funds that charge a fee given as a function of the difference in the performance of the fund and its reference index. This fee supplements the conventional fraction fee structure. The reference index is chosen by the mutual fund to reflect the general movements in the investment universe of the fund, such as the geographical or industrial sector. Incentive fee funds can generally be partitioned into two subgroups by the structure of their variable fee, namely symmetric and asymmetric incentive fee funds. For symmetric incentive fee funds, the variable component is positive when the fund return is higher than that of the index and negative in the opposite. This usually applies within some range of the fixed component of the fund fee. For asymmetric incentive fee funds, the variable component is positive when the fund return is higher than the index return but zero otherwise, resembling the payoff of a call option. For instance, the yearly fees earned by the incentive fee fund Skagen Kon-tiki A is 2.00% of its total assets plus an asymmetric incentive fee of 10% of the excess performance compared to the index MSCI Emerging Markets.

The structure of mutual fund fee schedules does, in concept, leave room for a disparity in the incentives of the mutual fund investor and their manager. As a result, in order to maximize their payoff function, the fund manager might use their private information to take actions that deviate from those that would maximize the payoff for the investor. Throughout this thesis, we write as if the incentive fees obtained by the fund mirrors the employment contract of the fund manager. That is, the fund manager is not trying to optimize with respect to their *employment* contract or any incentives in that regard. As a consequence, a mutual fund manager should invest to maximize the value of future fund fees.

In theory, there are various common-sensical arguments for the existence of incentive fee funds. From the perspective of an investor, the variable payoff structure should, to some extent, align the investor and fund manager's incentives. The fund manager should increase efforts to maximize the return of their portfolio, as predicted by agency theory (e.g. Jensen and Meckling, 1976). Besides, the fund should attract the best managers, as any manager capable of generating excess return would be best renumerated for their services in such a fund. From the perspective of a manager, the same argument implies that an incentive structure would induce a signalling effect, which, coupled with any excess performance of the fund, leads to positive net investment flows (e.g. Berk and Green, 2004). In practice, the intricacies of an incentive contract add to any existing agency effects in delegated portfolio management.

There is a considerable body of theoretical work on how different contract structures in mutual funds impact the behaviour of the fund manager, and as a consequence, how well investor and fund manager interests are aligned. Starks (1987) compare different contract structures among incentive fee funds. Despite not eliminating agency costs, they find that the symmetric contract dominates the asymmetric. Ou-Yang (2003) arrives at the same conclusion, provided that the fund benchmark is appropriate. Notably, this is a significant assumption in the eyes of Admati and Pfleiderer (1997), who find that commonly used benchmark schemes among mutual funds are inconsistent with optimal risk-sharing. Regardless, equitable distribution of risk seems to remain the major advantage that symmetric contracts have over asymmetric ones. For example, Grinblatt and Titman (1989) show, given that managers can hedge their compensation, that the convex schedule in asymmetric contracts induce managers to increase leverage as much as possible. They note that while greater leverage also increases the probability for liquidation of the fund, the net effect remains positive. While Carpenter (1989) reports ambiguous findings when removing the manager opportunity to hedge compensation, she agrees that managers with asymmetric fees weigh the gains of overperforming heavier than the implicit impact of underperformance. While other researchers agree that asymmetric incentive contracts lead to managers engaging in strategies with high variance around their benchmark, they show for their configurations that option-like contracts align interests better than both linear (Stoughton (1993); Li and Tiwari (2009)) and symmetric contracts (Das and Sundaram, 2002). The model of (e.g. Stoughton, 1993) emphasizes the benefits of the contractual incentives that managers have to make an effort in acquiring private information. While they identify considerable agency costs, they argue that the asymmetric contract remains dominant as long as the investor is risk-tolerant.

Turning away from the purely theoretical side, the empirical side of contract structure evaluation has also received attention in the literature of financial economics. Notably, there has been considerable interest in the topic among U.S. researchers. In underlying terms, this is due to both the maturity and size of the U.S. mutual fund market. The interest for research on incentive fee funds in delegated portfolio management was however sparked in 1971 with legislation that prohibited the use of asymmetric incentive fees in U.S. mutual funds.

There is some evidence that managers of incentive fee funds outcompete fraction fee funds in generating excess returns. Massa and Patgiri (2009) find that greater incentives not only increase the risk-adjusted return of the fund but that the performance is persistent. In line with this, Elton et al. (2003) show that managers of symmetric incentive fee funds exhibit significantly better ability in generating risk adjusted net returns than their fraction fee counterparts, even when adjusting for incentive fee funds in their sample charging lower fees. Importantly, they find that even the incentive fee funds do not on average outperform their benchmark indices. The idea that greater incentives motivate managers to generate superior returns is backed by Ibert (2018), by showing a correlation between abnormal fund returns and amount of personal wealth in the fund (that is, they have great incentives).

Even if incentive fee fund managers deliver better results net of fees and seem to act solely in the interests of their investors, the agency issues of imperfect contracts persist. There are signs that the investment strategies that incentive fee fund managers employ are designed to maximize the income of the fund, and not necessarily to maximize the return of the investor. Before discussing the strategies that incentive fee fund managers are found to use, it is vital to remember that they are also affected by any existing agency effects, such as flow incentives (e.g. Gruber, 1996). For example, even though an option-like contract increases strictly in value with volatility when we consider one period in isolation, no manager is of infinite risk tolerance when they regard multiple periods. Still, researchers have found differences in how incentive and fraction fee fund managers invest.

First, an incentive contract should lead managers to seek greater volatility around their reference index. The fund manager reward structure does not penalize large market beta and the option-like payoff of the fee increases in value from volatility. As long as reference index returns are non-negative, a beta greater than one would earn positive incentive fees. Elton et al. (2003) argue that this logic holds for symmetric in addition to asymmetric funds, as they can be shown to have mathematically equivalent convexity features to a capped asymmetric fee fund, given that the incentive component of the fund fee only holds within the capped range. They find that U.S. funds with symmetric fee structures do take on more systematic risk than fraction fee mutual funds, although the average incentive fee fund has a beta below unity. Cuoco and Kaniel (2010) find similar results in terms of market beta. They point to risk-aversion among managers in funds with symmetric incentive contracts as a possible explanation for the tendency for symmetric incentive fee funds to follow their benchmarks more closely.

Second, incentive fee fund managers should allocate more money outside of their reference index. They should especially take more exposure to assets they believe yield positive differential expected return, such as e.g. small-capitalization stocks. The reason is that the incentive fee reward structure does not penalize the priced increase in risk. Elton et al. (2003) find that U.S. symmetric incentive fee funds act on this logic and do employ a higher tracking error to their benchmarks. Golec and Starks (2004) find that managers of asymmetric incentive fee funds deviate more from their benchmarks than those of symmetric incentive fee funds. They show that a sample of U.S. funds that were forced to change the structure from asymmetric to symmetric in 1971 reduced their tracking error in the subsequent period. An alternative method of achieving benchmark deviation is to assign a reference index that takes exposure to different risk factors than those of the fund. For example, Ervik and Qvale (2017) point to the equity mutual fund Pareto Global A comparing its performance to an international bond index.

Third, incentive fee fund managers should be more willing to change the level of risk within an evaluation period. That is because incentive structures add to the convexity of the fee schedule. While this evidently holds for asymmetric incentive fee funds, Elton et al. (2003) similarily to above argue that this also holds for capped symmetric incentive fee funds. Depending on the performance in the first segment of the evaluation period, convexity in the fund income schedule may lead fund managers to increase the risk to finish ahead of their benchmark or decrease risk to lock in gains, as documented by Grinblatt and Titman (1992), Basak et al. (2007) and Kempf et al. (2009). Furthermore, the adverse effects of spurious risk-changing to beat a benchmark are shown by Chen and Pennachi (2005) and Huang et al. (2011), as funds that increase risk intra-period tend to perform worse than others. As an extension, Massa and Patgiri (2009) find that the risk-seeking nature of funds with greater incentives lead to them having a lower chance of survival.

These insights provide a clear connection between the findings of those that study mathematical contracts and those that have observed the empirical characteristics of incentive fee funds. However, the broader literature in the strategic characterization of funds explores more general approaches.

As a start, the concept of style analysis was brought to the forefront by Sharpe (1992), who characterized mutual fund investment styles by linear regression against a set of benchmarks. In a similar vein, others have used extended CAPM and other multi-factor models to theorize on the strategies employed by fund managers (Blake et al. (1999); Gruber (2001)). In parallell with the continued pervasiveness of linear multi-factor models as a vehicle to describe the behaviour of mutual funds since the work of Fama and French (1993), increased efforts have been made to delve deeper into statistics to explore useful methodology for classification of fund behaviour. The techniques used for factor identification range from principal component analysis (Brown and Goetzmann, 1997), option-like return representative strategies (Fung and Hsieh (2001); Agarwal and Naïk (2000)), cluster analysis (Marate and Shawky (1999); Gruber (2001); Lisi and Otranto (2010); Sun et al. (2012)), hierarchical tree (Mantegna, 1998), to genetic algorithms (Pattarin et al., 2004) and network-like approaches such as self-organizing maps (Maillet and Rousset, 2003).

Research that characterizes the behaviour and performance of incentive fee funds registered on the Oslo Stock Exchange (OSE) is a relevant issue. In 2017, the Norwegian Financial Supervisory Authority repealed 2001 regulation that prohibited registration of asymmetrical incentive mutual funds in Norway. Since, there has been an uptick in the number of such funds marketed towards Norwegian investors. For example, DNB, the largest Norwegian consumer bank, opened ten asymmetrical incentive fee funds in 2019. Previous work with a similar sample and focus is to our best knowledge sparse. Ervik and Qvale (2017) find that a sample of Norwegian incentive fee funds charges higher fees than a small sample of large Norwegian fraction fee funds. They do not address risk-adjusted performance or risk characteristics.

In this thesis, we study performance and risk-taking aspects for 409 fraction fee funds and 13 incentive fee funds registered on the OSE in the period 2000-2018. We investigate the existence of well-known risk characteristics for a new sample and map their risk-taking characteristics in the mutual fund universe.

We generally do not distinguish between symmetric and asymmetric incentive fee funds. The size of our data set does not allow such granularity. Our focus is thus to study incentive fee funds as a gross class compared to fraction fee funds. We find solace in the fact that asymmetric incentive fee funds that cap their payoff are mathematically equivalent to symmetric incentive fee funds.

We investigate the risk-adjusted performance of aggregate portfolios of incentive fee funds by the measure of Jensen's alpha (Jensen, 1968). We employ the regression framework of Carhart (1997) and assign factor data based on geographical investment regions in the spirit of Fama and French (2012). We further explore our data set for the existence of skilled individual managers using the bootstrapping methodology of Fama and French (2010), which adjusts for sampling variation and non-normality in the aggregated distribution of cross-sectional regression intercepts.

As a first step for understanding the risk properties of the incentive fee funds in our sample, we investigate their systematic exposure to the market portfolio and tracking error to benchmark indices. As a second step, we test whether explicit incentive fee contracts add to the convexity of the payoff schedule. In our tests, we make assumptions similar to those of Chevalier and Ellison (1997). For the second analysis, we consider a fund's tracking error to their stated benchmark index.

In a third effort to gauge to the risk-taking behaviour of incentive fee funds, our approach is both more explorative and comprehensive. We use a self-organizing map, which is a two-layer neural network, to map the patterns in risk properties of the funds in our sample. We further cluster the output of the map to obtain fund classes in an objective manner.

Self-organizing maps have been used in various problem domains[1], while application in finance is sparse and focused on style analysis of mutual and hedge funds[2]. Our approach differs from previous work in finance in two ways. First, we cluster the output of our self-organizing map, instead of taking the map as the final output. Second, we train our network purely on risk measures, and not on all available fund features.

A notable consequence of deviating from previous literature on feature selection is that we surrender external means for direct comparison of results. If we chose to map funds on investment style, we would be able to measure the validity of our methodology externally by comparison to the pre-defined investment style classes from a commonly used data provider (e.g. Refinitiv Eikon). While we employ a range of criteria to test the internal validity of our model, the lack of a direct external comparison to our results implies that we alter

[1]Recent examples: Robotics (Zhu et al., 2017); geology (Huang et al., 2017); natural language processing (Lokesh et al., 2019); image recognition (Chen et al., 2017).

[2]Noteable examples: Deboeck (1998); Maillet and Rousset (2003); Baghai-Wadji et al. (2006).

our perspective slightly. Our work in this area hence serves a dual purpose. For one, we categorize incentive fee funds in risk property classes and compare them to the tendencies found in previous work. Secondly, we give an example of the use of a non-linear, robust and intuitive tool for exploring patterns in fund behaviour.

The rest of the thesis is structured as follows. In Chapter 2, we describe our data sources and the steps taken to construct our sample of funds. In Chapter 3, we compare the performance of funds as groups and individuals. For individual funds, we run tests to distinguish fund manager skill from luck. In Chapter 4, we explore the risk-taking characteristics of the funds in our sample by methods that follow previous empirical work with a focus on incentive fee funds. In Chapter 5, we train a neural network for mapping the funds registered on the OSE by risk characteristics. Chapter 6 concludes.

2 Data

In this chapter, we describe the data sources we use, the filtering and the processing of the data to obtain our final sample.

2.1 Data sources

We download monthly time series for Net Asset Value (NAV), Total Net Assets (TNA), Technical Indicators (TI) and Fund Manager Benchmark (FM) from Refinitiv Eikon for 1462 equity mutual funds that were present in the registries on the Oslo Stock Exchange (OSE) at any time from 2000 to 2018. We retrieve all time series in USD. We focus on equity funds as incentive fee structures are most prominent in this segment. Table A.1 shows explicit fund selection criteria.

By including liquidated, merged and active funds, we limit the amount of survivorship bias in our sample. A sample solely consisting of funds that existed at a certain in time would likely lead to overestimation of risk-adjusted performance, as there is a correlation between underperformance and discontinuation of mutual funds Brown (1992).

OSE has provided an incomplete list of 54 equity funds that employ incentive fees.[1] These records contain funds that have been de-listed or changed their fee structure in the period considered. We further categorize 12 funds as incentive fee funds by reading fund prospectus, resulting in a gross sample of 66 incentive fee funds. In the sample of 1462 funds retrieved from Refinitiv Eikon, 56 of these are present. In the cases where funds change the fee structure, we treat the fund time series as two separate funds, split on the date of the change.

We collect pre-computed monthly Fama-French regression factors and risk-free rates for funds that invest in Norwegian and various international equity categories from the web pages of Ødegaard (2020) and French (2020), respectively. The latter source characterizes international funds by the categories Asia Pacific (hereinafter referred to as Asia), Emerging, Europe, Global, Japan, Norway and U.S.[2] Our approach of assigning factor data by fund classification is motivated by Fama and French (2012). They find that locally adapted models have greater explanatory power of returns and that patterns in risk anomalies vary between the international markets. The risk-free rate of Ødegaard is a one-month forward-looking rate constructed from a combination of the NIBOR and government securities, while the one of French is the one month T-bill rate. Both sources follow Fama and French (2015) in creation of risk factor returns.

[1]These records are the most comprehensive overview of the use of incentive fees among mutual funds traded on the OSE. Other sources for such overviews are not known to the authors at the time of writing.

[2]A full list of countries included in each factor set can be found in Table A.4.

To complete the factor data for Norwegian funds, we download time series data for the Oslo Stock Exchange Mutual Fund Index (OSEFX) from TITLON.[3] Norwegian law requires mutual funds to invest in at least 16 different equities, where the weight of each asset cannot exceed 10%, and the OSEFX reflects these requirements. Using a different reference for the funds in this category could cause misleading results. For example, the performance of Norwegian mutual funds as a gross group would likely seem weaker in comparison to a Norwegian index with fewer constraints.

2.2 Data filtering and processing

In line with our ambition of measuring how individual fund managers behave, we filter the fund sample to ensure fair comparisons across time series of funds. For the following steps, we note that we base many exclusions on fund names. We argue that it does not lead to any systematic bias aside from what is pointed out below, as all funds are filtered through the same set of rules.

First, we exclude passive funds such as index funds, as we aim to compare actions taken by active managers. In the same step, in a similar vein, we exclude fund-of-funds, as their performance is derivative of decisions made by other fund managers. Second, we exclude funds that require an initial purchase of at least USD 100 000 or more. Our focus is the perspective of all Norwegian investors, and funds with large buy-ins are outside of their investment universe. Third, for each set of share classes, we exclude all but the oldest, to ensure that manager decisions are counted only once. We note that although the time series net of fees for different share classes of the same fund are slightly different, they are the result of the same risk exposure. Fourth, to ensure that we compare fund time series to factor portfolios that represent their investment universe, we exclude any fund that we cannot assign regression factor data based on the fund classification by Refinitiv Eikon.[4] For completeness in data, we make sure that all funds have an assigned TI.[5] Finally, we exclude any funds that have less than 24 months of observations between 2000 and 2018.

For the Norwegian funds, we compute their time series for NAV, TI and FM in NOK to match the NOK-denominated time series for both the OSEFX and the Norwegian factor data. For one fund that lacked one observation in their time series, we interpolate NAV linearly.

We calculate monthly arithmetic returns from the NAV of each fund. Despite the smoothing and symmetric properties of logarithmic returns, we use arithmetic returns, as the time series for the risk pricing factors we use are derived in the framework of French (2020) and of Ødegaard (2020), who both employ arithmetic returns.[6]

[3] The TITLON database provides data reported from OSE. The University of Tromsø manages the database.
[4] Criteria for factor assignment is shown in Table A.3.
[5] A full list of reference indexes can be found in Table A.2.
[6] This has been verified through direct communication with Mr Ødegaard.

Table 2.1 shows an overview of the filtering process. The process is intended to return funds that are available to the majority of investors, make decisions with just an incentive share class in mind and that has time series fit for analysis. The filtering steps reduce the sample size for the incentive fee fund category remarkably. Including more funds would perhaps make it easier to obtain statistically significant figures in various analyses. Those figures would however be a less precise description of the issues we focus on.

Table 2.1: Data filtering process.

The table shows an overview of the data filtering process. The second and third columns show the number of incentive fee funds and the total number of funds in each step, respectively. The fourth to seventh columns show the first four moments of the Compounded Annual Growth Rate (CAGR). Numbers are in per cent per annum. Kurtosis follows Fisher's definition (standard of 0). The last two rows show the final sample split into fraction and incentive fee funds.

Filter	N_{inc}	N_{tot}	Mean	St. dev.	Skewness	Kurtosis
Initial sample	56	1462	4.97	6.1	1.98	15.96
Passive funds	53	1394	4.96	6.17	2.01	15.98
Institutional funds	46	1194	4.89	6.14	2.16	18.06
One share class	20	511	4.09	5.47	1.84	9.82
Factor portfolios	15	463	4.31	5.37	2.03	11.54
Time series length	13	422	3.82	4.05	-0.21	-0.06
Fraction fee	-	409	3.79	4.02	-0.2	-0.05
Incentive fee	-	13	4.58	4.31	-0.46	-0.11

3 Performance

In this chapter, we investigate manager performance. We examine our data set for signs of incentive fee fund managers differing from their fraction fee competitors in generating excess returns for investors. We study differences across the groups as a whole and later explore the cross-section of funds for skilled individual fund managers.

3.1 Risk-adjusted return

Through a linear regression approach that has been the standard in the literature since Fama and French (1993), we here explore the risk-return characteristics of our fund sample.

3.1.1 Regression framework

Linear regression models that explain the return of some asset by some set of systematic risk factors have long been the standard way of measuring risk-adjusted return within financial literature. The regression takes the form

$$r^e_{i,t} = \alpha_i + \boldsymbol{\beta}^{'}_i \boldsymbol{X}_{i,t} + \varepsilon_{i,t}, \tag{3.1}$$

where for some asset i at time t, $r^e_{i,t}$ is the return in excess of the risk-free rate, α_i is the intercept, $\boldsymbol{\beta}^{'}_i$ is the vector of factor loadings, $\boldsymbol{X}_{i,t}$ is the vector of returns on a set of systematic risk factors and $\varepsilon_{i,t}$ is the residual. Assuming that the set of risk factors explain the movement of dependent variable well, one interprets the intercept as the risk-adjusted abnormal performance of the asset (e.g. a mutual fund). The magnitude of each beta coefficient represents the systematic risk exposure to the respective risk factor.

Sharpe (1964) introduced the Capital Asset Pricing Model (CAPM). It was further developed by Lintner (1965) and Mossin (1966). They found that much of the risk of an asset can be explained by the returns on a broad market portfolio. Assuming that investors diversify away idiosyncratic risk, they argue that only systematic risk should affect asset prices. Building on this, Jensen (1968) was the first to describe α as a performance measure. Fama and French (1993) extended the CAPM to by adding factors that adjust for risk by firm size (*SMB*) and book-to-market (*HML*). As so-called small capitalization and value stocks historically outperformed large capitalization and growth stocks respectively, they argue that this risk too should be priced to determine to what extent the performance of a portfolio was attributable to these factors.

Since Fama and French (1993), there has been much research in pursuit of identifying additional systematic risk factors. Perhaps most notably, Carhart (1997) extended the three-factor model by adding the momentum factor (*MOM*) of Jegadeesh and Titman (1993). The momentum factor has its empirical reasoning in the short-term overperformance of those assets that have performed well in the previous months. Recently, Fama and French (2015) extended their three-factor model by adding both a profitability factor (*RMW*) and an investment factor (*CMA*). They find increased explanatory power on data from the New York Stock Exchange (NYSE). In the Norwegian setting, the most recent extensive research is by Grimeland (2018). He finds that a combination of the Fama-French three-factor model and a liquidity factor (*LIQ*) outperforms even the Fama-French five-factor model for Norwegian stocks.

An unfavourable aspect of these extensions to the CAPM is that their inclusion has little theoretical foundation. The origins of the risk factors are not very clear, aside from their empirical existence. Fama and French (1993) argue that the components are proxies for common risk factors and that they may appropriately account for risk despite their uncertain source.

While we note recent development in multivariate regression configuration, we face practical limitations in accessing pre-computed data for every relevant factor, such as *RMW*, *CMA* and *LIQ*. In this thesis, we opt for the regression of Carhart (1997). The setup may be written as

$$r^e_{i,t} = \alpha_i + \beta_{i,MKT} MKT_{i,t} + \beta_{i,SMB} SMB_{i,t} + \beta_{i,HML} HML_{i,t} + \beta_{i,MOM} MOM_{i,t} + \varepsilon_{i,t}, \tag{3.2}$$

where $r^e_{i,t}$ is the return of fund i at time t in excess of the risk free rate, and the risk factors are denoted by MKT (market portfolio), SMB (size portfolio), HML (book-to-market portfolio) and MOM (momentum portfolio). All the risk factors represent investable strategies that are structured as zero-investment portfolios. α is the intercept, $\beta_{i,MKT}, ..., \beta_{i,MOM}$ are the factor loadings, and $\varepsilon_{i,t}$ is the residual.

We employ the Carhart four-factor regression assuming that it is suited to finding risk-adjusted mutual fund returns, as many before us. However, financial data from a cross-section of mutual funds is often hard to reconcile with standard Gauss-Markov assumptions (e.g. Bickel and Freedman (1984); Hall and Martin (1988)). As we are working with a self-constructed sample of funds, we take steps to ensure that our parameter estimates and thus, potential inferences are reliable.

We test for non-constant variance in residuals using the heteroskedasticity test of Breusch and Pagan. We find that 40.0% of the funds in our sample exhibit heteroskedasticity with a confidence of 95% or higher. We further find signs of autocorrelation in a number of time series using the test of Durbin and Watson. Test results in form of a histogram of the test statistics are listed in Figure B.1. To account for heteroskedasticity and autocorrelation, we perform our regressions using Newey-West heteroscedasticity- and autocorrelation-consistent standard errors. Testing for normality in residuals with the Shapiro-Wilk test, we reject normality in residuals with a confidence of 95% for 49.5% of the funds.

In order to alleviate the fact that residuals from regressions on individual funds are non-normal, we employ a bootstrapping procedure for generating confidence intervals for each parameter estimate. We account for anomalies by not imposing an ex-ante parametric distribution for our parameters. We use the chosen regression framework to simulate an empirical distribution for each parameter, where we assume each simulated observation to be equally likely. To perform the bootstrap, we first estimate the Carhart four-factor model for the time series of each portfolio i.[1] We save the coefficient estimates $\{\hat{\alpha}_i, \hat{\beta}_{i,MKT}, \hat{\beta}_{i,SMB}, \hat{\beta}_{i,HML}, \hat{\beta}_{i,MOM}\}$ and the estimated residuals $\hat{\varepsilon}_i = [\hat{\varepsilon}_{i,1}, ..., \hat{\varepsilon}_{i,T}]$, for each portfolio $i \in I, I$ denoting the set of portfolios, and T denoting the set of months the portfolio has registered data. For every portfolio, we draw a sample with replacement from the portfolio residuals saved from the original regression, creating a pseudo time series of resampled residuals, $[\varepsilon^b_{i,1} ..., \varepsilon^b_{i,T}]$, where b is the bootstrap index. We use sampled residuals to construct a new time series of pseudo monthly excess returns $\tilde{r}^e_{i,t}$:

$$\tilde{r}^e_{i,t} = \hat{\alpha}_i + \hat{\beta}_{i,MKT} MKT_{i,t} + \hat{\beta}_{i,SMB} SMB_{i,t} + \hat{\beta}_{i,HML} HML_{i,t} + \hat{\beta}_{i,MOM} MOM_{i,t} + \hat{\varepsilon}_{i,t}. \quad (3.3)$$

We further regress the Carhart four-factor model on the pseudo time series and save the parameters estimated for each portfolio. Repeating this for all bootstrap iterations, $b = 1, ..., B$, we build an empirical distribution for each parameter. We use the empirical distribution to construct confidence intervals.

We use this bootstrapping method to gauge the significance of our parameter estimates when we compare the performance of fund groups in Section 3.1.2. In Section 3.1.3, we shift the focus to comparing the performance of individual funds in the tails of the cross-sectional distribution of alphas. Before we proceed, we note that we there extend the described bootstrap method to distinguish manager skill from luck. First, however, we explore the performance of funds in groups.

3.1.2 Equally weighted regressions

We regress the Carhart four-factor model for equally weighted portfolios of funds with and without incentive fees for seven geographical regions from 2000 to 2018. Regression results are presented in Table 3.1.

For fraction fee funds, although magnitude and significance varies, the regressions show that alphas are negative across all geographies. For the equally weighted portfolio of Japanese funds, the alpha is significantly different from zero at the 5% level. For Asian, Global and U.S. fraction fee funds, the results are even stronger with significance at the 1% level. In contrast, most equally weighted portfolios of incentive fee funds show positive, albeit insignificant, alphas across the geographies where they are present in our sample.

[1] Such a portfolio may represent an equally weighted portfolio of a selection of funds, while for regression on just one fund it will consist of just the individual fund.

Table 3.1: Carhart four-factor regression.

The table shows regression parameters for the funds in the categories Asia, Emerging, Europe, Global, Japan, Norway and U.S., separated on fraction and incentive fee funds. For each category, the Carhart four-factor model is computed from an equally weighted portfolio of funds. R^2 denotes fit, and N denotes the number of funds in each portfolio. The regression is performed using Newey-West heteroscedasticity- and autocorrelation-consistent standard errors. Alphas are annualized by multiplication.

	$\alpha(\%)$	β_{MKT}	β_{SMB}	β_{HML}	β_{MOM}	R^2	N
Panel A: Asia.							
Fraction fee	-4.42***	0.95***	0.09**	0.18***	0.03	0.92	34
	(1.43)	(0.02)	(0.04)	(0.04)	(0.03)		
Incentive fee	-	-	-	-	-	-	-
	-	-	-	-	-	-	-
Panel B: Emerging.							
Fraction fee	-1.28	0.99***	-0.04	0.11**	-0.02	0.96	94
	(1.15)	(0.02)	(0.05)	(0.05)	(0.03)		
Incentive fee	1.37	1.04***	0.11	-0.04	-0.03	0.9	2
	(2.38)	(0.03)	(0.11)	(0.13)	(0.07)		
Panel C: Europe.							
Fraction fee	-1.13	1.01***	0.25***	-0.14***	-0.02	0.98	99
	(0.7)	(0.01)	(0.03)	(0.02)	(0.01)		
Incentive fee	1.28	1.17***	0.35***	-0.29***	-0.11***	0.9	5
	(1.83)	(0.03)	(0.08)	(0.06)	(0.04)		
Panel D: Global.							
Fraction fee	-2.76***	1.02***	0.11***	0.0	0.02	0.97	88
	(0.67)	(0.01)	(0.03)	(0.02)	(0.01)		
Incentive fee	1.63	1.25***	0.43***	0.06	-0.03	0.9	5
	(1.6)	(0.03)	(0.07)	(0.06)	(0.03)		
Panel E: Japan.							
Fraction fee	-2.36**	0.94***	0.08**	-0.16***	0.05**	0.93	24
	(1.07)	(0.02)	(0.03)	(0.03)	(0.02)		
Incentive fee	-	-	-	-	-	-	-
	-	-	-	-	-	-	-
Panel F: Norway.							
Fraction fee	-0.98	0.96***	0.12***	-0.04***	0.01	0.97	33
	(0.84)	(0.01)	(0.02)	(0.02)	(0.02)		
Incentive fee	-5.3	0.99***	0.1	-0.02	-0.06	0.8	1
	(6.15)	(0.12)	(0.18)	(0.16)	(0.17)		
Panel G: USA.							
Fraction fee	-2.23***	0.98***	0.16***	-0.06***	-0.03**	0.97	37
	(0.68)	(0.01)	(0.03)	(0.02)	(0.01)		
Incentive fee	-	-	-	-	-	-	-
	-	-	-	-	-	-	-

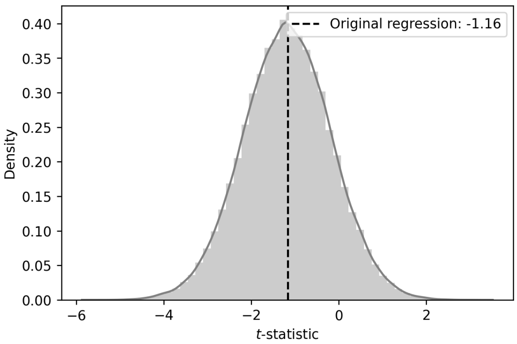
Note: *p<0.1; **p<0.05; ***p<0.01

Figure 3.1: Bootstrapped distribution of t-statistics.

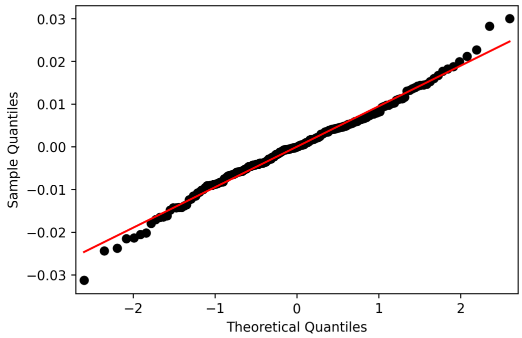
Subfigure (a) shows the empirical distribution for the t-statistic of alpha for Norwegian fraction fee funds. Subfigure (b) displays the corresponding quantile-quantile-plot.[a] $B = 100\,000$.

[a]Bootstrapping on funds with other geographical focus shows similar results and may be provided upon request.

(a) Norway: Empirical distribution of t-statistics.



(b) Norway: QQ-plot for t-statistics.



To ensure the validity of our coefficient estimates, we build empirical distributions for each parameter using a bootstrap method as described in Section 3.1.1. We show the empirical distribution for the t-statistic of alpha and its quantile-quantile plot for Norwegian fraction fee funds in Figure 3.1. We list 95% bootstrapped confidence intervals for the fund groups for each of the coefficients in Table B.2. Despite evidence of non-normality from the Jarque-Bera test, and deviations in the tails from the QQ-plot, the distribution shares the shape with a normal distribution to the extent that it yields the very same conclusions as those we come to from Table 3.1.

The measures computed in Table 3.1 are from the time series for fund NAV, which is stated after fees.[2] The implication is that a negative alpha means that a fund manager is not able to generate an excess return from the perspective of the investor. The results are thus generally in line with the fundamental theory of equilibrium accounting (Sharpe, 1991), where funds participate in a zero-sum alpha game pre-fees. For fraction fee funds, this means that the alpha is negative by the magnitude of the fund fees. For reference, Gallefoss et al. (2015) find that the average fee for Norwegian funds is 1.7% annually. For incentive fee funds, the interpretation of pre-fee performance depends on a variable fund fee, and we do not have access to its historical size. As a result, the post-fee performance of incentive fee funds may look better if the incentive fee funds charge a smaller fraction fee than the average of that of pure fraction fee funds while simultaneously being beaten by their benchmark index. This was the case for the sample studied by Elton et al. (2003).

[2]Irregular fees such as for front-end and back-end loads are exceptions to this. We do not have access to data that incorporates this. The discussion is thus on the implied assumption that investors of corresponding funds follow a buy-and-hold strategy.

The exception of our results being in line with Sharpe (1991) is arguably the evidence from the Asian and Global fraction fee fund groups, as these groups show more negative alphas than one likely may explain purely by fund fees. We touch on this in the following paragraphs.

For Asian fraction fee funds, the 14 of 34 funds that exist before 2006 perform especially poorly. In this period, this minority of funds alone make up the monthly observations that are taken as input into regression of the equally weighed portfolio of Asian funds. As a result of the skewness in existence for the Asian funds in our data set, these 14 funds impact the regression disproportionally. In unreported tests, we regress the portfolio of equally weighed Asian fraction fee funds from 2006 to 2018 and find alpha at -1.44%. We hypothesize that there may have been a skewness in which Asian funds were registered on the OSE in the early 2000's and that these funds had a tendency to invest in assets that underperformed the local market in that period.

For the Global fraction fee funds, the slightly low alpha estimate is due to a subgroup of funds that both deviate from and underperform the global factor set. The low explanatory power of global factors on this subgroup is a symptom of two underlying issues. First, the assignment of the global fund category by Refiniv Eikon has some inconsistencies.[3] Second, French (2020) includes only developed countries when constructing the global factor portfolios (Table A.4). Funds that correctly invest globally may have broader exposure, for example to emerging or frontier markets.

The comparison of Global fraction fee funds that deviate from the factor set to the factors themselves is punishing. In unreported tests, we perform individual regressions on every Global fraction fee fund and find that the low-fit funds systematically invest differently than those with high fit. The 28 funds with regression fit R^2 below 0.8 produce an average post-fee alpha of -4.78%, while the 60 remaining funds yield -2.3%. To see this tendency in a broader context, we compare the performance of widely used equity indices for developed countries (MSCI World) and emerging markets (MSCI Emerging Markets). When we consider the period from 2000 to 2018, the connection is not apparent, as the latter outperforms the first. However, taking into account that more funds are present in the latter half of our sample than our first, the link is evident. Counting from any year post-2005, MSCI World outperforms MSCI Emerging Markets over the remaining years in our considered period.

To test the robustness of our regression choice, we exclude the momentum factor from Equation 3.2 and re-run the regression for the three-factor model of Fama and French (1993). The results are listed in Table B.1. For fraction fee funds, we find very similar results to those listed in Table 3.1. For the equally weighted portfolios of incentive fee funds, the alphas are reduced as their negative exposure to the momentum factor no longer is explained. For European incentive fee funds, this change sees incentive fee fund alpha go below zero. Al-

[3] Examples are "Nomura Funds Ireland-NEWS EM Small Cap Eq A EUR" ($R^2 = 0.6$, $\alpha = -11.7$) and "Odin Maritim" ($R^2 = 0.69$, $\alpha = -8.9$). These belong among funds that invest in emerging markets and the maritime sector, respectively.

phas for all incentive fee fund categories remain insignificantly different from zero at the 10% significance level. For all fund groups, we note slight changes in coefficient estimates, indicating that the various factor portfolios are not entirely orthogonal to the momentum portfolio.

We test for any currency issues from data processing of the Norwegian fund set. In unreported tests, we download fund data for Norwegian funds from a source that offers time series in NOK (TITLON) and re-run the regression in Table 3.1. The differences in coefficient estimates are negligible.

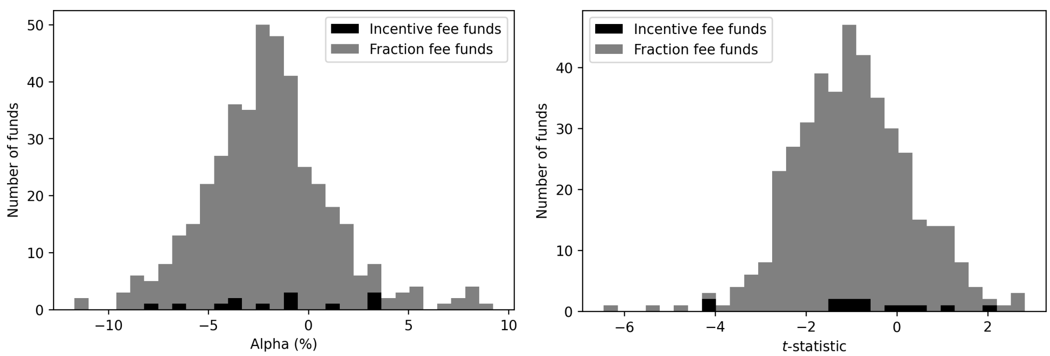
In summary, we find no evidence of any fund portfolio being able to generate significantly positive alpha net of fees. The tendency holds across both incentive and all geographical categories. While various fraction fee fund portfolios produce significantly negative alpha, we are for the incentive fee fund portfolios not able to reject the hypothesis of zero post-fee alpha. Thus, we cannot say certainly whether groups of incentive fee funds over- or underperformed in risk-adjusted terms. In the next subsection, we explore the individual performance of incentive fee funds. This setting allows us to compare funds across geographical segments.

3.1.3 Individual regressions

Figure 3.2: Distribution of individual fund alphas and t-statistics.

The figure shows histograms of regressed alphas and t-statistics from the Carhart four-factor model for individual funds in the sample. Incentive fee funds are highlighted.

(a) Alpha for individual funds. **(b)** t-statistic of alpha for individual funds.



Even if aggregated groups of funds perform on the norm (that is, in the vicinity of zero pre-fee alpha), equally weighted portfolios tell us little about the distribution of fund alphas within each group. We present the distributions of alpha and the t-statistic of alpha for regressions on individual funds in Figure 3.2. The t-statistic can be interpreted as a normalized coefficient, as it scales inversely by the standard error. Importantly, this lets us compare coefficients more reliably, as we account for different levels of idiosyncratic risk and number of observations between funds (Brown, 1992).

Before we put the performance of individual funds under the magnifying glass, we test for the similarity between the distributions for the t-statistic of alpha between fraction fee and incentive fee funds. We have already seen that various equally weighted portfolios of fraction fee funds rejected the hypothesis of zero post-fee alpha, while incentive fee funds gave insignificant results. Here, the question is whether the geographically aggregated distributions of individual fund t-statistics share statistical properties. For this, we use the two-sided Kolmogorov-Smirnov test, for which the null hypothesis is that the two sample sets are drawn from the same distribution. The test returns a p-value of 0.43. In words, we cannot conclude that incentive and fraction fee fund t-statistics are drawn from different distributions. Elton et al. (2003) find that samples of the incentive and fraction fee fund alphas are significantly different at the 10% level. The higher level of significance may both be due to different properties of the funds in their sample, and that the small size of our incentive fee fund sample makes it hard to obtain statistical significance from such tests even when a similar trend is present in the data.

Gauging the distributions in Figure 3.2 we note that there are both incentive and fraction fee funds in the right and left tails of either distribution. If one were to assume that alpha is drawn from a distribution that closely resembles a normal distribution, t-statistics crossing 1.96 (-1.96) would indicate that a fund manager generates positive (negative) alpha for their investors significant at the traditional threshold of 95% confidence. In the right tail of the t-statistics, we find one incentive fee fund and four fraction fee funds, while we in the left tail find one incentive fee fund in company with 86 fraction fee funds.

In summary, from analyzing regressions of equally weighted portfolios and individual funds, we find slight but insignificant indications that incentive fee fund managers outperform fraction fee fund managers. If these indications are symptoms of a broader trend, our results would be in line with the empirical work on incentive fee funds of ((Elton et al., 2003); (Massa and Patgiri, 2009); (Ibert, 2018)) and the theoretical work of those that find incentive contracts to best align investor-manager interests (e.g. Stoughton, 1993).

Even if there is a weak trend of incentive fee fund managers outperforming fraction fee managers in generating alpha for their investors, we do not know from this whether that outperformance is due to skill or luck. In the following section, we follow Fama and French (2010) and investigate the existence of skill among the individual fund managers in our sample in a more robust manner.

3.2 Fund manager ability

Even if some managers produce significant post-fee alphas at various thresholds under normality assumptions, we cannot yet conclude that they are *skilled* in generating excess returns for their investors. In this section, we test for this by employing a bootstrap method adapted by Fama and French (2010).

3.2.1 Bootstrapping to distinguish skill from luck

Good performance does not unequivocally equate skill. For a setup with long individual regressions where we measure performance by regression intercept and its t-statistic as in Section 3.1.3, Kosowski et al. (2006) point out two reasons why. First, when we examine a sufficiently large sample of funds, we increase the chance of recording significant performance due to sampling variation (luck). That is, even if the null hypothesis of no significant fund manager ability is correct, we increase the chance of drawing a sample in the tails of the distribution (*Type I* error). Second, the regressions assume above the aggregate distribution of fund alphas is normal, while it likely is not. If we e.g. draw fund alphas from an aggregate distribution that has fat tails, as seems to be the case in the alpha distribution of Figure 3.2, we overestimate the extremity of those observations when we compare it to the quantiles of a normal distribution. For the second reason, non-normality in the aggregate distribution of alphas can be a result of individual fund returns not being normally distributed (83% were not at a significance level of 5%) or due to different levels of idiosyncratic risk between funds.

To account for these issues, Kosowski et al. (2006) propose a bootstrap method to distinguish skill from luck. They simulate empirical distributions of alpha for each fund while imposing true alpha equal to zero. They then compare every alpha from the original regression with the correspondingly ranked alphas from each of the simulated runs. For example, we compare the best fund from the original regression to the distribution consisting of the highest alphas from each simulation. Similarly, we compare the worst fund to the distribution of worst-performing funds across the simulations. If a fund performs well (poorly) in comparison to the distribution of equally ranked alphas, we conclude that the fund manager is skilled (incompetent).

Fama and French (2010) modify the procedure slightly. Kosowski et al. (2006) sample only the residuals and use the historical sequence of explanatory returns in each simulation. Fama and French jointly sample factor and fund returns. This way, they take into account cross-correlation of alpha between funds that arise when a benchmark model does not capture all common variation in fund returns. A second benefit of joint sampling of the sample fund and explanatory returns is capturing correlation in heteroskedasticity of the explanatory returns and disturbances of a benchmark model.

The alternate procedure has drawbacks. First, while the method by Kosowski et al. (2006) generates pseudo time series with the same length as the original series, the length varies in the modified method. Fama and French (2010) sample random dates, and for each fund, include data points present at the sampled dates, resulting in varying length of the sampled time series. When considering the alpha, the result depends on the number of data points in regression, meaning that funds with shorter time series risk producing thicker tails. Fama and French (2010) argue that the use of t-statistics mitigate this issue. Secondly, the random sampling of dates ignores the potential effects of autocorrelation. Third, random sampling of results risk losing the effects of variation through time.

We implement both methods. Results from the Fama French procedure are presented in Section 3.2.3. Results from the method of Kosowski et al. (2006) are discussed briefly in Section 3.2.4 and listed in Table B.3.

3.2.2 Bootstrapping procedure

The bootstrapping procedure proceeds as follows. As in Section 3.1.1, we first estimate the Carhart four-factor regression for the time series of each fund. We save the coefficient estimates $\{\hat{\alpha}_i, \hat{\beta}_{i,MKT}, \hat{\beta}_{i,SMB}, \hat{\beta}_{i,HML}, \hat{\beta}_{i,MOM}\}$ and the estimated residuals $\hat{\varepsilon}_i = [\hat{\varepsilon}_{i,1}, ..., \hat{\varepsilon}_{i,t}]$, for each fund $i \in I$, I denoting the set of funds, and t denoting each month the fund has registered data. Starting from the sampling method we use to generate pseudo time series in Section 3.1.1, this approach is different. Here, for each simulation $b = 1, ..., B$, we build pseudo time series that are of the same length for each fund i. We do so by random sampling (with replacement) from all the months in our timeframe and then jointly sampling factor returns and residual for each fund at that time.[4] The funds that do not span the entire timeframe only extend their pseudo time series when they have data for the chosen month. We require that each pseudo time series is at least 24 months, and re-run the simulation if not. Together with the estimated betas, we construct pseudo time series of monthly excess returns. We impose the null hypothesis of $\alpha = 0$ by construction:

$$\tilde{r}^e_{i,t} = \hat{\beta}_{i,MKT} MKT_{i,t} + \hat{\beta}_{i,SMB} SMB_{i,t} + \hat{\beta}_{i,HML} HML_{i,t} + \hat{\beta}_{i,MOM} MOM_{i,t} + \hat{\varepsilon}_{i,t}, \tag{3.4}$$

where the meaning of the parameters is the same as in equation 3.3, noted for each fund i. Setting alpha to zero when using time series net of fees is equivalent of imposing a null hypothesis that the fund managers are able to generate abnormal returns that cover all investment-related costs for the investor, such as investment fees and transaction cost. We further regress Carhart four-factor model on the pseudo time series and save the alpha estimated for the cross-section of individual funds $i = 1, ..., N$. We repeat this for B bootstrap iterations, which yields B cross-sections of N alphas. We rank each simulated cross-section, as well as the cross-section of original alphas. We then compare each original alpha with its corresponding vector of B simulated alphas. To avoid ambiguity, we emphasize that this means that the highest real alpha competes with the highest simulated alpha from each bootstrap simulation. For the top (bottom) performers, the fraction of simulated alphas for which the original is higher (smaller) in absolute value is equivalent to a p-value. For the right (left) tail of the original alpha distribution, we use this to infer chances of skill (inability) in our sample of mutual fund managers. We repeat the procedure of ranking and comparing using the t-statistic of alpha, due to its property of controlling for the varying precision of alpha estimates across funds (due to different length of pseudo time series or different idiosyncratic risk levels) Kosowski et al. (2006).

[4]Note that the factor returns are not sampled in historical order, and that we allow for cross-correlation of alpha by not separating residuals from their factor returns.

3.2.3 Results from the bootstrap of Fama and French (2010)

We present the results from the bootstrap in Table 3.2. We focus on t-statistics in interpretation, and include alpha for completeness.

We find that only a handful of the top performers produce t-statistics in the vicinity of the cross-section of equally ranked t-statistics. Only the top funds exhibit skill that is average or better in post-fee terms. In order to reject the null hypothesis of managers not generating abnormal post-fee performance for the top performers at, e.g. the 5% significance level, we would require a win rate of 95%. There are no funds in the sample, regardless of the incentive fee structure, where this is the case. This also holds when considering the alpha measure directly. When we consider the bottom performers, the story is less nuanced. In terms of t-statistics, every fund descending from the 80^{th} percentile record a win rate below 1%. For the alpha measure, the worst funds are not beaten as decisively by the equally ranked simulated distributions. We hypothesize that this is due to a subset of funds have traits in their return series that may generate poor alpha when we are unlucky in sampling months for building pseudo-time series. If these fund alphas average high standard deviation, it explains why the effect is slight when considering t-statistics.

We plot the t-statistics for the most, third and fifth extreme funds against their simulated distribution of cross-sectionally ranked equivalents in Figure 3.4. Under the assumption of zero post-fee alpha, the probability of the best fund t-statistic being drawn from the distributions we compare them with is likely. The t-statistic of the worst-performing funds lie far to the left of the probability mass, indicating that the distributions they are drawn from represent funds that do not perform to the standard of zero post-fee alpha.

The results achieved here are generally in line with those of Fama and French (2010). They investigate the performance of U.S. mutual funds from 1984 to 2006 and find that net fund returns have the same characteristics as they do for our sample. In the 80^{th} percentile in terms of t-statistics of alpha, Fama and French (2010) find win rates above 1%, whereas we find the same for the for the 90^{th} percentile for our sample. Hence, our results too contradict the claims of Berk and Green (2004) that most fund managers are skilled enough to generate positive risk-adjusted returns for their investors. For the Norwegian fund universe in isolation, previous studies find the same patterns that we find in our results. For Norwegian funds, both Sørensen (2009) and Børsheim and Eilertsen (2016) show that it is easier to detect inability among the poor performers than skill among the top performers.

Figure 3.3 compares the cumulative distribution functions (CDFs) for actual and simulated t-statistics. For the simulated t-statistics, we plot the mean of each ranked distribution from the bootstrap. The CDF for actual fund statistics lies to the left of the simulated mean for nearly every quantile with exceptions only in the tails. While this tells much of the same story as the numbers presented in Table 3.2, it offers visual intuition for the fact that using the null hypothesis of zero post-fee alpha is a strong and perhaps unrealistic assumption. While the median of the simulated CDF lies near zero, the median of the actual CDF is drawn to the left by a combination of management fees and incompetence in making

Table 3.2: Results from the bootstrapping procedure of Fama and French (2010).

The table shows original regression values and simulated means displayed as *Act* and *Sim* for the alphas and their *t*-statistic. The leftmost columns list the five best and worst values for alphas and their *t*-statistics, as well as deciles. The top performer in terms of alpha is not necessarily the same fund as the top performer in terms of *t*-statistic. The fourth and seventh columns show the *win-rate* of the original regression values to the distribution of 1000 simulations of alphas and *t*-statistics. In the upper (lower) part of the table, high (low) win rates translate to low p-values. Monthly alphas are annualized.

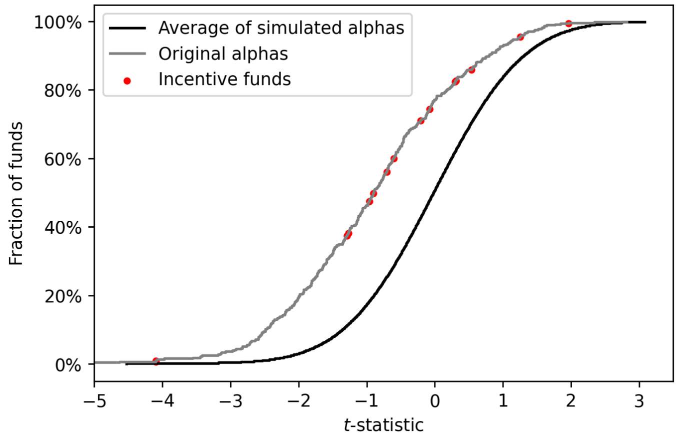
	Alpha			*t*-statistic		
	Act(%)	*Sim*(%)	%<*Act*	*Act*	*Sim*	%<*Act*
Best	8.61	12.48	11.3	2.82	3.08	32.8
2	8.48	10.32	24.1	2.34	2.74	16.0
3	8.35	9.22	39.3	1.96	2.57	4.6
4	8.23	8.55	48.5	1.85	2.45	3.5
5	7.9	8.01	52.6	1.77	2.35	3.8
90 %	2.15	3.3	7.2	0.8	1.31	2.3
80 %	0.46	1.96	0.2	0.18	0.85	0.0
70 %	-0.65	1.14	0.0	-0.26	0.53	0.0
60 %	-1.31	0.52	0.0	-0.61	0.25	0.0
50 %	-2.0	-0.03	0.0	-0.9	-0.02	0.1
40 %	-2.56	-0.59	0.1	-1.23	-0.28	0.1
30 %	-3.4	-1.22	0.1	-1.59	-0.56	0.1
20 %	-4.44	-2.04	0.3	-1.98	-0.89	0.1
10 %	-5.97	-3.47	1.6	-2.49	-1.35	0.1
5	-8.9	-9.49	54.7	-4.09	-2.53	0.5
4	-9.28	-10.31	60.2	-4.1	-2.66	0.8
3	-9.46	-11.56	73.6	-4.63	-2.85	0.4
2	-11.34	-13.66	70.8	-5.4	-3.18	0.6
Worst	-11.69	-17.68	96.0	-6.45	-4.53	4.5

investment decisions that generate positive risk-adjusted returns.

These results bring nuance to those we find for incentive fee funds in Section 3.1.3. One incentive fee fund manager beats the correspondingly ranked simulated mean, and the remaining incentive fee funds lie to the left of the simulated CDF. As mentioned previously, we are not able to separate the figurative data-generating processes for individual *t*-statistics of alpha between incentive and fraction fee funds by statistical tests, perhaps due to too small a sample. With these results, we can state that incentive fee funds, in addition, share the property of generally not having the skill to cover costs. We again note that management fees in incentive fee funds are variable and that they may unknowingly distort the performance of incentive fee funds in our analysis. The notion that incentive fee fund managers do not have enough *skill* to generate positive risk-adjusted return for their investors conflicts slightly with the findings of positive and significant difference in alphas between fraction and incentive fee funds by Elton et al. (2003). A potential explanation may be that our

Figure 3.3: Cumulative distribution functions for actual and simulated t-statistics.

This figure shows the empirical cumulative distribution functions for the actual and simulated t-statistics. Incentive fee funds are highlighted.



analysis extends further. Elton et al. (2003) do not account for luck, non-normality in alpha distributions nor the size of standard errors in individual alpha estimates.

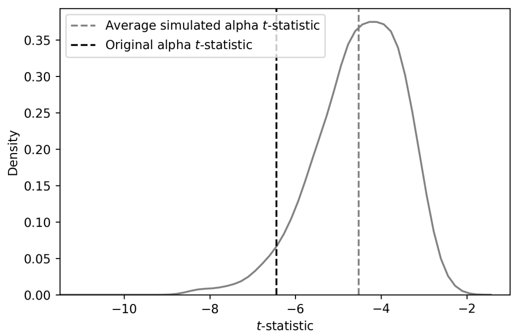
We re-run our analysis for the Fama-French three-factor model. Results are listed in Table B.1. Fama and French (2010) find that this configuration yields a stronger indication of skill amongst managers in this bootstrap procedure. The difference stems from the excess return generated by momentum exposure is transferred to the alpha. Here, the results are very similar, as there is no clear trend for the entire sample of funds having positive or negative exposure to the momentum factor. We also test for robustness by altering our bootstrap procedure to match that of Kosowski et al. (2006). We discuss our findings in Section 3.2.4.

The results in this section are more nuanced than those in Section 3.1.3 because we take luck and non-normality of alpha distributions into account. We find that only a few top performers may exhibit skill to cover costs, while the majority of fund managers decisively are unable to produce risk-adjusted returns that exceed their management fees. For our data set, fraction and incentive fee funds generally share these properties. These results are in slight contrast to those of the equally weighted regressions in Section 3.1.2 which do not reject the null hypothesis of zero post-fee alpha for incentive fee fund categories. This disputes any trend in incentive fee fund managers outperforming fraction fee fund managers due to skill.

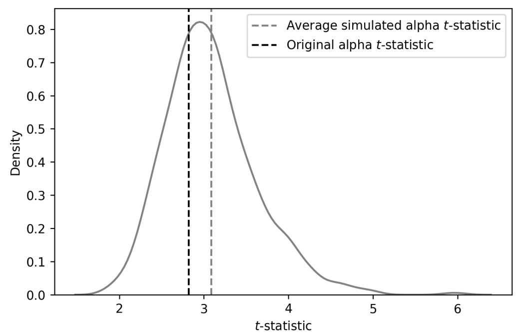
Figure 3.4: Parametric t-statistics compared to equally ranked simulated distributions.

The figures show actual t-statistics plotted against the empirical distributions of equally ranked t-statistics by the bootstrapping procedure of Fama and French (2010). An actual t-statistic being far to the right (left) in its distribution indicates skill (inability).

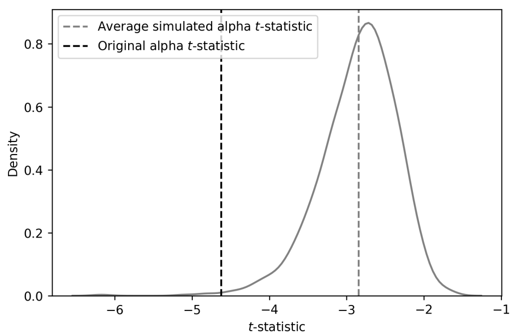
(a) Worst fund performance.



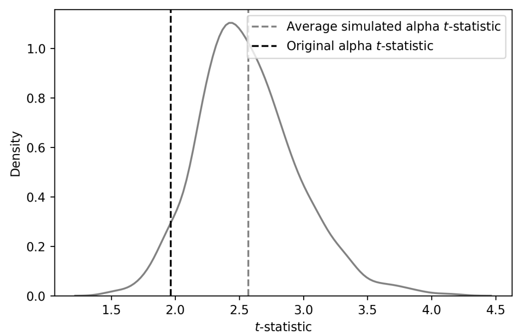
(b) Best fund performance.



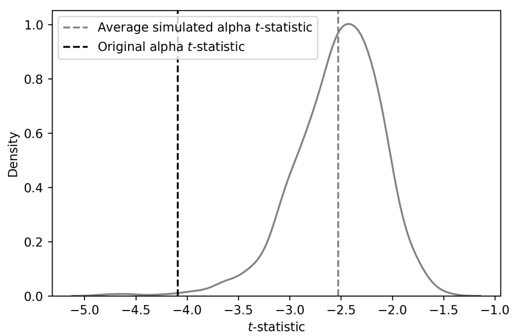
(c) Third worst fund performance.



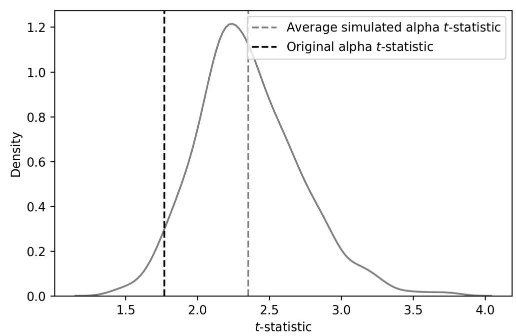
(d) Third best fund performance.



(e) Fifth worst fund performance.



(f) Fifth best fund performance.



3.2.4 Comparison to the procedure of Kosowski et al. (2006)

To account for differences in the simulation method, we adjust our bootstrap procedure to match that of Kosowski et al. (2006). We present the results and distribution plots in Table B.3 and Figure B.2. The results are nearly identical, and the conclusions from the previous chapter hold here as well.

Kosowski et al. (2006) find different results than we do when adjusting our simulation method to match theirs. They similarly reject that the worst performers have the ability to generate zero post-fee alpha, but they differ in that they find evidence that top performers exhibit skill. As pointed by Fama and French (2010), there are likely two reasons for this. First, Kosowski et al. (2006) remove funds from the data sample containing less than 60 months of returns, possibly implying a survivorship bias. Second, the time series investigated by Kosowski et al. (2006) ranges from 1975 to 2002, for which large subperiods the investment environment was less professionalized than it is today. Interestingly, the second argument does not separate our results from those of Fama and French (2010), although our data set is more recent than theirs.

For both simulation methods, the conclusion suggested is thus that all funds but the very top performers with near certainty lack skill in covering their management fees. Incentive fee funds share those properties with fraction fee funds. For an investor that wishes to maximize the risk-adjusted returns on their assets through investments in actively managed funds, this gives little guidance. In the following sections, we switch our focus from investigating risk-adjusted performance to exploring the risk implications that an investor faces when pooling their money in incentive fee funds.

4 Risk

As discussed in Chapter 1, previous theoretical and empirical work has lead to various hypotheses about the risk-taking behaviour of incentive fee funds. In this section, we examine whether those hypotheses hold for our data set. We first consider a selection of descriptive measures in Section 4.1, and secondly investigate intra-period risk changing behaviour in Section 4.2.

4.1 Risk over time

In this section, we address the differences in Carhart market beta, regression fit and tracking error to Technical Indicators and Fund Manager Benchmarks between fraction fee and incentive fee funds. Tracking error to an index is the standard deviation of the difference in returns. We present the statistics in Table 4.1. While some tendencies are clear, we interpret the results with caution. Our sample of incentive fee funds is small, and there is thus a chance that we have unintentionally cherry-picked funds with certain characteristics. For some reassurance, we run column-wise Kolmogorov-Smirnov tests to quantify dissimilarity per measure between samples. We also note that calculating tracking error using post-fee data distorts the measures by the fees. For incentive fee funds, the fee level is furthermore more uncertain than for fraction fee funds.

Incentive fee funds invest with a higher average market beta than fraction fee funds. The difference is significantly positive at the 1% level. Average beta levels over unity for incentive fee funds is in line with theoretical literature (e.g. Grinblatt and Titman, 1989). For an option-like contract, the increasing volatility of the underlying (in this case, the market portfolio) is beneficial. There is also a more straightforward argument for incentive fee funds employing beta over unity. In rising markets, an incentive fee fund manager holding a portfolio with beta over unity would earn positive incentive fees. Our findings are in contrast to those of Elton et al. (2003). They study a sample of symmetric incentive fee funds in the period 1990-2000 and are surprised to find levels of market beta under one.

R^2 is a measure of unsystematic risk in the sense that it measures the percentage of variability in the dependent variable for which the regression accounts.[1] For our sample, incentive fee funds recorded an average value of R^2 that is higher than that of fraction fee funds and significantly different at the 10% level.[2] Importantly, our regression accounts for returns that correlate with any of the four-factor portfolios, including that of the general market. High regression fit for incentive fee funds may be due to their returns are explained well by SMB-, HML- or MOM-portfolios. The compensation of an incentive fee fund is not

[1] R^2 is referred to as tracking error by others (e.g. Elton et al., 2003). While the set-up is not different in our case, we refer to it as unsystematic risk to avoid ambiguity.

[2] The value for all funds might be affected by the misassignment of factor sets as discussed in Chapter 2.

Table 4.1: Risk over time.

The table showns risk estimates for each fund category. $\bar{\beta}_{MKT}$ is the average market beta and R^2 the average fit from the Carhart four-factor regression (Equation 3.2). TE_{TI} and TE_{FM} denote the average tracking error to Technical Indicators and Fund Manager Benchmarks, respectively. The estimate for FM applies only to the 293 funds for which Refinitiv Eikon state such a benchmark. N is the total number of funds.

	$\bar{\beta}_{MKT}$	R^2	TE_{TI}	TE_{FM}	N
Fraction fee	0.95	0.79	2.83	2.97	409
	(0.19)	(0.16)	(2.36)	(2.38)	
Incentive fee	1.13	0.84	2.17	2.46	13
	(0.12)	(0.1)	(0.53)	(0.89)	
Difference	0.18***	0.05*	-0.67	-0.51	
Note:					*p<0.1; **p<0.05; ***p<0.01

measured directly in terms of exposure to such factors, and managers are free to use them in attempts to beat their benchmark indices.

For our sample, the average tracking error to the Technical Indicator is lower for incentive fee funds than for fraction fee funds. The difference is not significant, as the estimate for fraction fee funds has a high standard deviation. One might expect incentive fee funds to deviate more from their index in an attempt to beat it. From this, we are not able to draw any conclusions. The fourth column in Table 4.1 measures tracking error to the Fund Manager Benchmark, and the results are mostly similar to those for the Technical Indicator. The measures are slightly higher for both incentive and fraction fee categories. The fact that fund returns deviate more from the benchmarks that were chosen by the funds may be due to managers selecting benchmarks strategically. It is not uncommon for managers to select benchmarks that set the performance of the fund in a favourable light (Sensoy, 2009). Despite receiving fees based on differential performance, this jump is not markedly higher for the incentive fee fund group.

Across the categories considered, we see indications that the two fund groups take on risk in slightly different ways. We see a clear indication of incentive fee funds leveraging their convex contracts using the market beta. We find that incentive fee funds as a gross group owe more of their returns to sources inside of the regression portfolios assigned in Chapter 2. For tracking error to indices, the groups are hard to separate. We take these findings as a fundament for a slightly more nuanced approach to risk mapping of the fund universe in Chapter 5.

4.2 Risk changing

Particular effects influence the way a fund manager takes on risk based on the position within some time frame. These effects have in common that they are irrational for a manager with the simple goal of generating risk-adjusted returns for the investor. Chevalier and Ellison (1997) find that the relationship between yearly fund performance and yearly inflows is convex. The intuition is that funds which generate good performance statistics for a calendar year attract inflows from investors, which in turn increases the base for calculation of fraction fees. This results in a convex payoff schedule. A fund manager that near the end of the evaluation period is out-of-the-money, should increase tracking error relative to the benchmark. Similarly, a fund that is in-the-money will be less willing to hold risk outside of the benchmark to lock-in the achieved gains. Funds far in the right tail may even increase risk, however. Kempf and Ruenzi (2007) find that funds act as if they compete in a tournament to finish among the best performercer within a year. Previous work suggests that the existence of an incentive fee contract adds an extra layer of convexity of the payoff schedule (Elton et al., 2003). In this section, we test for this added convexity.

We measure risk change in terms of tracking error to the benchmark of the fund, following both Chevalier and Ellison (1997) and Elton et al. (2003). 293 of the 422 funds in our sample state such a benchmark. We assume that fraction fee funds invest on the assumption of yearly implicit incentive schedules, which is similar to the assumption made by Chevalier and Ellison (1997). We make this assumption on the grounds that yearly data is the most readily available to the investor, and thus often is used to make investment decisions. For our incentive fee funds, six from 13 funds do not load incentive fees yearly. We exclude those funds from our main results.

For each fund year, we only include it if the fund has return data for the full year. Moreover, we take as a proxy that the fund managers reexamine their position after exactly nine and twelve months and adjust the risk of their portfolio accordingly.[3] We present the results for testing for added risk-changing behaviour in Table 4.2.

For fraction fee funds, the top quintile performers after the first subperiod exhibit positive change in average tracking error, while the bottom quintile shows a decrease. The difference in tracking error change between top and bottom performers is negative and significant at the 10% level. It seems the yearly implicit flow incentives of Chevalier and Ellison (1997) are not the dominant force.

The increase in tracking error for the top funds may be due to managers gambling to finish among the very top performers (Kempf and Ruenzi, 2007). When we toggle the percentage threshold for being included in the top group, we find in unreported tests that the funds in top decile primarily drive the difference in average tracking error change to the entire fund set. This supports the tournament hypothesis. For the worst-performing funds, Basak et al. (2007) find that fund managers who shift risk to beat their benchmarks only do so to the

[3] In unreported tests, we find that the results are not sensitive to the intra-year split by toggling it two months in either direction. The material is available upon request.

Table 4.2: Intra-period change in tracking error.

The table shows an overview of the change in tracking error between the first nine and final three months of each fund calendar year for fraction and incentive fee funds in Panels A and B, respectively. Funds are ranked on their absolute differential return to their Technical Indicator in the first nine months. The rightmost column shows the average absolute change in tracking error. The fourth row of each subpanel displays the difference between the top and bottom quintiles.

	First subperiod		Second subperiod		
	$R_i - R_m(\%)$	$TE(\%)$	$R_i - R_m(\%)$	$TE(\%)$	ΔTE
Panel A: Fraction fee.					
Top 20%	9.87	2.74	1.46	3.1	0.36
All	-0.92	2.17	-0.18	2.34	0.16
Bottom 20%	-11.12	3.1	-1.22	2.72	-0.38
Difference	-20.99***	0.37	-2.68***	-0.38	-0.74*
Panel B: Incentive fee.					
Top 20%	15.84	2.93	2.15	2.33	-0.6
All	-0.29	2.34	-0.28	2.17	-0.18
Bottom 20%	-11.16	3.28	-2.7	2.83	-0.45
Difference	-27.0***	0.35	-4.85*	0.5	0.14

Note: *p<0.1; **p<0.05; ***p<0.01

limit of their risk tolerance. These funds may have operated above that risk tolerance in the first subperiod and be unwilling even to maintain it for the second subperiod.

For incentive fee funds, both top and bottom performers decrease tracking error to the benchmark in the second subperiod. The difference between the groups is small and non-significant. The fact that top quintile incentive fee funds seem to lock-in gains is in line with the results of Elton et al. (2003) and might be a sign that the incentive fee contract alters the shape of the payoff schedule. If one considers such behaviour hazardous, one would disagree with those that argue for the incentive alignment dominance of asymmetric contracts (Das and Sundaram (2002); Palomino and Prat (2003); Li and Tiwari (2009)). For the bottom quintile, a possible explanation for the similarity to the bottom fraction fee funds may be that incentive fee contracts lose importance when the fund is far out-of-the-money. Both incentive groups contain few observations, and we are thus hesitant to interpret this as more than indicators.

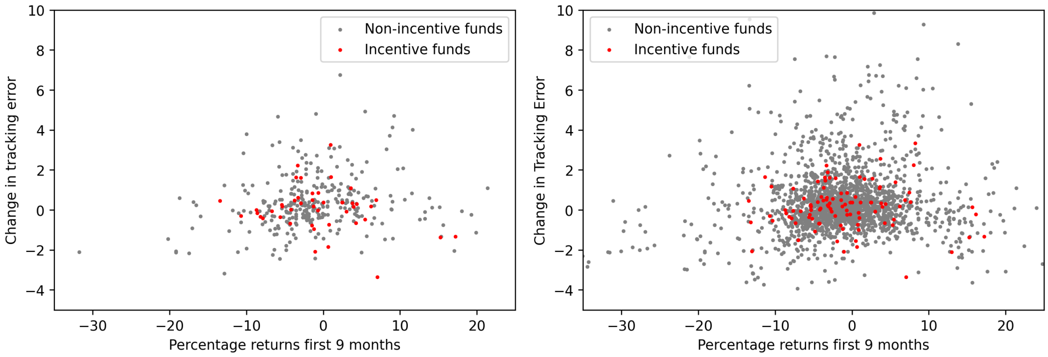
In contrast to Chevalier and Ellison (1997), we see little difference in overall results when filtering funds by size and age at various thresholds. Results filtered at seven years of prior existence and maximum size of USD 100 million can be found in Table C.1. The top quintile of incentive fee funds decrease risk more in this instance, but the sample is even smaller than in Table 4.2.

In unreported tests, we include the remaining six incentive fee funds and re-run the analysis for both the gross and filtered sample. The results broadly follow the same trends as those in Table 4.2. Barring any effects of a potentially unrepresentative sample, this indicates that incentive fee funds share features that motivate risk-changing behaviour and are not entirely explained by the explicit investor-fund contract. For instance, there may to a greater extent exist yearly explicit incentives in the employment contracts of incentive fee fund managers.

Figure 4.1: Intra-period change in tracking error.

The figure shows year-to-date absolute excess return against the benchmark for the first nine months plotted against changes in tracking error. The tracking error is computed against the benchmarks chosen by the funds (FM). Each data point represents a fund-year. Subfigure (a) shows the fund-years with less than 100 million dollars total net assets and fewer than seven years of prior existence. Subfigure (b) shows the entire fund-year sample. Incentive fee funds are highlighted.

(a) Small and young funds. **(b)** All funds.



Any inclination of risk-changing behaviour we find is of smaller magnitude and less systematic than those by the similar tests of Chevalier and Ellison (1997). As can be seen in Figure 4.1, any tendencies are not apparent. We hypothesize that funds have adapted to increasingly sophisticated investors. As investors have access to better data today than two decades ago, the assumption of implicit flow incentives matching the one-year time frame is weaker. Similarly, investors may more easily become aware that funds are toggling risk levels in the short term. For incentive fee funds, we do find indication that top performers lock-in gains and thus that the incentive contract adds to any convexity in the payoff schedule. For other incentive fee funds, results are ambiguous. It may be that incentive fee funds, as well as a majority of fraction fee funds, are more concerned about investing with some particular strategy than varying their risk levels in the short term to exploit their contracts. In the next section, we employ unconventional methodology to explore where the risk-taking behaviour of incentive fee funds place them in the universe of funds registered on the OSE.

5 Mapping funds on risk features

We employ self-organizing maps (SOM) to investigate clusters of risk-taking behaviour. Self-organizing maps were first proposed by Kohonen (1990). They are neural networks that belong to the class of unsupervised classification methods. From the time series from each fund, we calculate risk measures and attempt to discriminate incentive and fraction fee funds by mapping them in a SOM. As an extension, we perform both partitive and agglomerative cluster analysis on the map output.

The map analogy of a SOM comes from the fact that the neurons in its output layer usually are arranged in a two-dimensional grid. A SOM assigns input vectors to output vectors of the same dimensionality. Through the training process, the output neurons adjust their position relative to their neighbouring nodes dictated by a so-called neighbourhood function. The output layer neurons *self-organizes* in an unsupervised manner. In contrast to supervised learning, where a network learns from rewards and penalties from comparing its output to a known correct solution, unsupervised networks form groups of its input based on similar patterns in their attributes. The classification is non-linear. Intuitively, this allows for deviation from hyperplanar surface separation and instead facilitates a flexible fit on the data structure. We note that both the neighbourhood notion and inherent non-linearity of this approach separates it from other dimensionality reduction methods like e.g. principal component analysis.

A core property of the SOM is that the dimension reduction from d to two dimensions is designed to preserve the topology of the data. An implication of this is that points of data that land close on the output map lie close in the original data structure. While traditional clustering methods yield one fixed solution when grouping a set of high-dimensional features, the SOM yields a continuous mapping in dimensions that are conceivable to a human. With this, the SOM offers flexibility in choosing from multiple grouping alternatives. Applications of SOM are also more robust to outliers than those of traditional clustering methods, which increases in importance when training on high-dimensional feature sets (e.g. Kohonen, 1990).

While the SOM generates a continuous map of the input data where the topology is preserved, it does not offer a way to methodically place borders on the map to separate the data points into distinct groups. In addition, in most cases, as well as in ours, the desired number of clusters is smaller than the number of neurons in the output layer. For these reasons, we extend our procedure by a clustering step, as done by Vesanto and Alhoniemi (2000). In our two-step approach, the idea is first train the SOM and second cluster the neurons in the output layer. This gives us a way to generate candidate solutions for a sensible number of clusters consistently. Furthermore, as argued by e.g. Kiang (2001), clustering neurons that are topologically consistent still favours grouping those that are closely related. Moreover, because we have the map in the intermediate step, we gain intuition on the position of a fund within a cluster. We cluster by both partitive (k-means) and agglomerative (hierarchical) means and compare the results.

5.1 Method

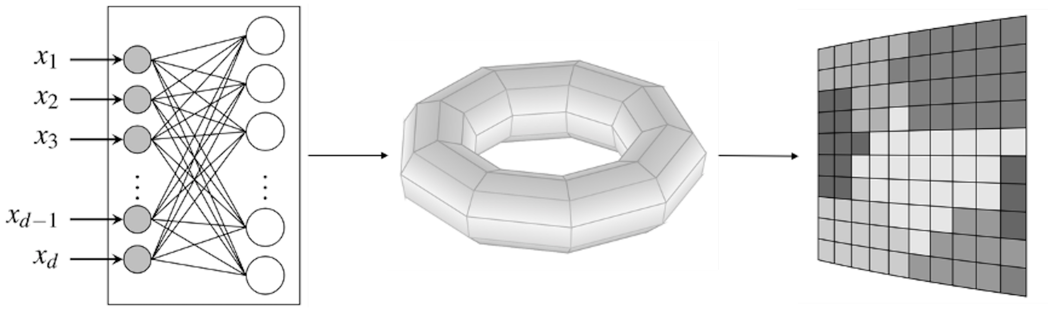
In this section, we lay out the SOM set-up and describe the k-means and hierarchical clustering algorithms. We later discuss selection of features for training, as well as the process of selection and tuning of hyperparameters.

5.1.1 Self-organizing maps

The SOM is a two-layer feed-forward neural network, visualized in Figure 5.1. The input layer takes as input a vector $x(t)$ of dimension d. Its output layer is a two-dimensional grid with a total number of neurons M, where each neuron is represented by a vector $w_j = [w_{j,1}, ..., w_{j,d}]$ of dimension d. The lattice of the grid is usually rectangular or hexagonal. We opt for the former. The input layer units are fully connected; that is, each input layer unit is connected to every neuron in the output layer. The training process is one of online stochastic learning, in that we iteratively present one randomly selected input.

Figure 5.1: Flattening of the SOM.

For each input vector $x(t)$ fed to the network, the weights (edges) in the network are adjusted for the output neurons to fit the data. The output layer forms a torus. The torus is projected onto a two-dimensional plane by cutting it open to first form a cylinder and cutting it again to form a sheet.



Each output neuron is initialized as a random vector. For each training iteration $t = 1, ..., T$, we draw a fund i randomly from our sample (with replacement). We calculate its set of features and construct its feature vector $x(t)$, as more thoroughly described in Section 5.1.4. We locate the *best-matching unit* (BMU) by finding the output-layer neuron c with the smallest distance:

$$c = \arg\min_j ||x(t) - w_j||, \; j \in \{1, ..., M\}. \tag{5.1}$$

The distance measure we use is the euclidean distance, which is standard in the literature (e.g. Kohonen, 1990). We define the kernel neighbourhood function $h_{j,c}(t)$ at time t to dictate the rate of weight updates for node $j \in \{1, ..., M\}$. The kernel is a non-increasing function of time. This means that the weight updates for neighbouring units of c too are decaying in time, aiding convergence of the map. We take the kernel function as a Gaussian function,

$$h_{j,c}(t) = e^{-(\frac{m(j,c)}{\sigma(t)})^2}, \tag{5.2}$$

where $\sigma(t)$ is the kernel width and is given by $\sigma(t) = \sigma_0 e^{\frac{t}{T}}$ at iteration t of total T iterations, σ_0 being the initial kernel width and $m(j,c)$ the manhattan distance between neuron j and c. The kernel function penalizes distance from the BMU exponentially, and thus the magnitude of change for neighbouring neurons is inverse of the exponential. Given the best-matching unit c, we write the weight update function as

$$w_j(t+1) = w_j(t) + \eta(t) h_{j,c}(t)(x(t) - w_j(t)), \tag{5.3}$$

where $\eta(t) = \eta_0 e^{\frac{t}{T}}$ is the learning rate at iteration t of total T iterations, and η_0 is the initial learning rate. Both the learning rate and the kernel width decay monotonically in time. Intuitively, this means that the changes we make to the output layer are of decreasing magnitude and width. Using the analogy of throwing a stone in a pond, regardless of where the stone hits the water (BMU) at time t, the stones become gradually lighter such that both the impact ($\eta(t)$) and the reach of its rings ($h_{j,c}(t)$) are smaller than in previous iterations. We comment further on our choice for hyperparameters in Section 5.1.5.

Schmidt et al. (2011) highlight the issue of the so-called edge effect in traditional SOM. They argue that having edge neurons with fewer neighbours set unnecessary constraints on the network. Motivated by this, we extend the vanilla SOM procedure by allowing neighbourhood relations to cross the edges of the rectangular map, illustrated in Figure 5.1. As such, our solution candidates live in a modified two-dimensional plane similar to a 'Pac-Man'-world. This modification of borders relax the optimization problem and potentially allows for more flexibility when folding a high-dimensional structure onto a plane. In our application, this extension improves our chances of finding similarity in patterns lying close to the edges of the map.

After training, the map has divided the hyperspace into M regions. The notion of neighbourhood relation is crucial for topology preservation. Whenever a neighbouring node is the best-matching unit, the current node is pulled in a similar direction. As a result, nodes that are close in the high-dimensional space tend to lie close in the two-dimensional projection. Distances between the neurons show similarity between items. The map is *topologically ordered* in the input space (Kohonen, 2014). As Kiang (2001) puts it, the SOM behaves like a net that folds onto a cloud.

For our purposes, interpretation of fund groups is the most interesting for a number of groups that is both flexible and smaller than the number of neurons in the map. The SOM does not suggest a way of methodically drawing borders. As mentioned, to produce consistent and smaller groups, we extend our procedure by both partitive and agglomerative clustering algorithms. In this setting, we interpret the neurons in the output-layer as proto-cluster centres, to again be clustered into larger groups. The task of the clustering algorithms is then to identify homogeneous disjoint sets of output from the SOM.

5.1.2 K-means clustering

We first briefly present the k-means clustering algorithm. It was introduced by MacQueen (1967) and is one of the most used clustering methods due to its simplicity and intuitiveness. K-means clustering is a partitive clustering algorithm; that is, we initially consider the data set as one group and iteratively split it into subgroups. The device for doing so is minimizing the euclidean distances from cluster centres to each of the data points that are members of the cluster. The procedure is described in Algorithm 1. As earlier, w_j, for $j \in \{1,...,M\}$, are the output neurons from the SOM.

Algorithm 1: K-means

Input: Number of clusters k

Output: k centroid positions

Data: Neurons w_j, $j \in \{1,...,M\}$, from the SOM

Initialize k clusters with random values for each feature in the feature range.

while *centroid change* **do**

1. Assign each data point to its most similar centroid, forming a cluster.
2. Calculate new centroid positions as mean of the fund features in each cluster.

While the k-means algorithm seems appealing for its simplicity, the approach has several drawbacks to keep in mind in application. First, the number of clusters k is defined ex-ante in Section 5.1.5. We tackle this with the widely used Davies-Bouldin index (Davies and Bouldin, 1979). Second, the random initialization of centroids may lead runs of the algorithm to different local optima. A popular way to mitigate this issue is to test for sensitivity by running the algorithm a number of times.

Because the k-means algorithm finds its new centroids per iteration by averaging the data points that belong to it, a third characteristic is that the procedure is sensitive to outliers in the data. The sensitivity is less of an issue when using the k-means algorithm as an extension to a SOM (Kiang, 2001), due to the neighbourhood weight updates. Even if certain output-layer neurons in the SOM that are the best-matching units for outlier feature vectors and thus adjust their weights starkly, neighbouring neurons will consequently pull them closer whenever non-outlier feature vectors are fed to the network.

A final drawback of the k-means approach is that considering just the distance of each data point to a centroid when assigning clusters makes the algorithm identify only hyper-spherical groups of neurons. While it is hard to gauge the validity of such clusters in isolation, hierarchical clustering methods lend some diagnostic tools for alleviating these concerns.

5.1.3 Hierarchical clustering

We briefly present hierarchical clustering. For a more extensive overview, we refer to Kaufman and Roussew (1990). Hierarchical is an agglomerative clustering approach; that is, we initially consider every data point as a cluster and recursively merge two groups until one group remains. In each step, the merging decision is a greedy function of the current state and the chosen merging criterion. Greediness, in this context, signifies that the decision of an optimal move is made only considering the current state. The procedure is described in Algorithm 2.

Algorithm 2: Hierarchical clustering

Input: Number of clusters k

Output: k sets of neurons

Data: Neurons w_j, $j \in \{1,...,M\}$, from the SOM

Let each data point w_j be a member of a singleton group such that there are $n = M$ groups.

while $n > k$ **do**

1. For each pair of groups A and B, compute the chosen merging criterion.
2. Merge the pair that optimizes the criterion.

As one may read from the description in Algorithm 2, both the initialization and iterations of hierarchical clustering are non-random. Given some data set as input, in our case the neurons w_j for $j \in \{1,...,M\}$, as well as the merging criterion, the algorithm is deterministic. We thereby mitigate the issue of varying results per simulation run. A noteworthy con of this method is that it is static (whereas k-means is dynamic), which means that data points cannot change membership between groups while the algorithm runs. This property may let outliers distort the final image.

We still have to choose the desired number of clusters k, but we here do so ex-post. There is little guidance on choice of k, outside the rule of thumb that is considering jumps in cluster distances for various steps. In Section 5.1.5, we use a dendrogram for this. The final degree of freedom is the merging criterion. For clustering of a SOM, Wu and Chow (2004) find that variance minimization criteria outperform minimization of various linkage and distance criteria. For our main results, we use the most commonly used variance minimization criterion, which is the Ward criterion. The cost of merging two clusters is expressed as the global increase in variance:

$$\begin{aligned}\Delta(A,B) &= \sum_{j \in A \cup B} ||w_j - \mu_{A \cup B}||^2 - \sum_{j \in A} ||w_j - \mu_A||^2 - \sum_{j \in B} ||w_j - \mu_B||^2 \\ &= \frac{m_A m_B}{m_A + m_B} ||\mu_A - \mu_B||^2,\end{aligned} \tag{5.4}$$

where m_A is the number of data points for cluster A, and μ_A is the cluster centre in cluster A. Δ is the merging cost for clusters A and B, which is to be minimized per iteration.

5.1.4 Feature extraction and selection

In the following section, we reason on feature selection for our purposes and discuss the set of features that we use as input for our network in Section 5.2.

As mentioned earlier, we cluster on risk features exclusively, and not on any full set of available fund features. The reason for this is our aim of understanding whether an incentive fee fund customer takes exposure to patterns of risk in a different way than one of a fraction fee fund. If we clustered on all available fund features, we would find groups per *investment style*, such as funds that invest in a particular sector or area. This restriction means that features that otherwise are commonly seen in various approaches for clustering funds (e.g. Deboeck, 1998) are not used as input here. One set of examples are performance measures such as absolute return against a fund's benchmark, the factor-regressed alpha, or the Treynor ratio.

Features used for training are listed in Table 5.1. For clarity, we first note that we include Carhart-regression coefficients as features, even though an important motivation by choice of clustering procedure is to detach the following analysis from the linear factor analysis paradigm that permeates the current financial literature. Our choice of including these coefficients as features means that we do not disregard the factor analysis paradigm completely, as factor analysis has a robust empirical track record and has proven a useful framework for analyzing fund returns. We instead shift focus and look for non-linear patterns between them and other features by unconventional means. In addition to the regression coefficients, we include regression goodness of fit, R^2, as a measure of systematic risk exposure to the assigned factor set. The regression outputs used as features are coefficient means from $T-24$ rolling window regressions of length 24, where T is the total number of months in the time series.

The remaining set of features are included because they describe some aspect of risk-taking behaviour in mutual fund returns. The standard deviation of returns gives absolute volatility, which is the most common measure of risk in fund returns. Semivariance is a measure of downside risk, defined as the average of the sum of squared differences for observations below the mean. We include tracking error from a fund's Technical Indicator as assigned by Refinitiv Eikon. Tracking error is the standard deviation of the difference in return between two time series. Motivated by the considerations in Section 4.2, we include the standard deviation of rolling estimates of tracking error to account for variation in risk level. As previously, we use monthly data intervals. We further preprocess the set by scaling each feature vector to zero mean and unit variance. This aids in dealing with outliers in the data.

Table 5.1: Input features to the SOM.

The table shows a list of the features used as input to the SOM.

Notation	Feature description
$t_{\beta_{MKT}}$	t-statistic of loading to market portfolio
$t_{\beta_{SMB}}$	t-statistic of loadingto small-minus-big portfolio
$t_{\beta_{HML}}$	t-statistic of loading to high-minus-low portfolio
$t_{\beta_{MOM}}$	t-statistic of loading to momentum portfolio
R^2	Goodness of fit
σ_r	Standard deviation of returns
γ	Semivariance
TE_{TI}	Tracking error to Technical Indicator
σ_{TE}	Standard deviation of tracking error

5.1.5 Choice and tuning of model parameters

In this subsection, we discuss our approach for tuning the hyperparameters of our SOM and ensuring the internal validity of the clustering of its output.

Training a self-organizing network requires selection of sensible hyperparameters for map size M (dimensionality of the neural network), initial learning rate η_0 and kernel width σ_0, as well as the number of training iterations T. Even then, the training procedure has stochastic elements, and networks trained with equal hyperparameters will vary slightly, as we randomly select input vectors in each training iteration. For the subsequent clustering we in addition select number of clusters k.

There exist several measures to quantify map validity, among which quantization error (QE) and topographic error (TE) are most commonly used in the literature (Pölzlbauer, 2004). Quantization error is a measure of map resolution, defined as the average euclidean distance between each data vector and its best-matching unit (BMU). For n data points in the training data and the mapping of fund feature vector x_i from the input space to the SOM, $\theta(x_i)$, we write:

$$QE = \frac{1}{n}\sum_{i=1}^{n} ||\theta(x_i) - x_i||. \tag{5.5}$$

Topographic error is a measure of topology preservation, defined as the proportion of all data vectors for which the first and second BMUs are not neighbouring units. We write this below using c as the first BMU and c' for the second:

$$TE = \frac{1}{n}\sum_{i=1}^{n} \rho(x_i) \tag{5.6}$$

$$\rho(x) = \begin{cases} 0 & \text{if } c \text{ and } c' \text{ are neighbouring nodes} \\ 1 & \text{otherwise} \end{cases} \tag{5.7}$$

While these measures are useful indicators for map validity, they should not be trusted blindly. Comparing values across data sets makes little sense as the mathematical possibility for a map to fit the data is a function of the high-dimensional data structure. Regarding one data set in isolation, the indicators generally give better results for large maps (Kohonen, 1990).

There is no definitive method for choice of map size M. Widely used software packages such as *SOM Toolbox* (Vatanen, 2015) suggest that M is a function of N, with $M = 5 * \sqrt{N}$ set as standard. The literature deviates substantially from this (Vesanto and Alhoniemi (2000), Wu and Chow (2004), Rajanen (Marghescu), Solidoro et al. (2007)), recognizing that M depends greatly on the structure and not only the number of elements of the data set. Maillet and Rousset (2003) analyze hedge funds and choose $M = 49$ for $N = 1358$. They argue that their choice of M is fitting in their implementation as there are few neurons for which no data points are mapped. For $N = 422$, we use $M = 100\,(10x10)$, which roughly matches the norm. Increasing the map size to $M = 144\,(12x12)$ we find that the fraction of unused neurons jump from 3% to 10%. Increasing the map size further to e.g. $M = 196\,(14x14)$ sets unused neurons of 14%. Correspondingly, we see drops in QE and TE as M increases. We present alternative configurations in Section 5.2.5.

We use QE and TE for guidance in choosing initial values for the learning rate $\eta(t)$ and the neighbourhood width $\sigma(t)$. The learning rate controls the magnitude of change per iteration. The neighbourhood width sets the radius for which the weight of neighbours to the BMU are affected. Throughout training, they both decay exponentially. In a SOM, the outcome in terms of QE and TE depends not only on the values of each of the hyperparameters, but on the relationship between them.

Setting the learning rate high and the neighbourhood width small, we observe a slight improvement in QE at the cost of worsening in TE. In such a situation, the output-layer neurons are less often pulled towards a winning neuron and sparsely updated when they are not the winning neuron themselves. This imbalance hurts topology preservation in the network. In the converse scenario, with a low learning rate and a wide initial neighbourhood, more neurons are pulled towards each other for a higher number of iterations. At the point when the pull on a neuron is more seldom, the learning rate might be too low for any meaningful perturbations away from the current local minima. The settings that lead to the second scenario mean that topology is better preserved, but hurts network resolution. Methods for determining optimal learning rate and sigma are not known (Deetz et al., 2009). We attempt to balance the two measures.

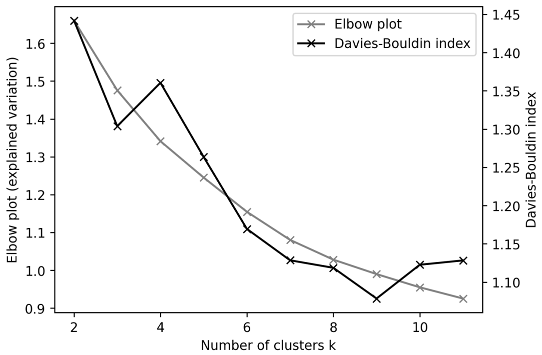
In our clustering extension, choosing a suitable number of clusters k is both a matter of the inherent structure of our data as well as the information value it gives in presentation. In other words, it depends on the number of meaningful fund groups that can be separated by clustering on risk features.

For k-means clustering, we measure validity by the Davies-Bouldin index (Davies and Bouldin, 1979), which optimizes for intra-cluster similarity and inter-cluster distance. We show the index for the SOM configuration presented in Chapter 5.2 in Figure 5.2.[1] To address the randomness in the initialization of the k-means algorithm, we iteratively toggle the computers' pseudo-randomness setting and re-run the clustering of the map configuration 100 times. The clusterings are largely similar, and we, therefore, do not present a comparison of them. A likely explanation for this is that the SOM mitigates the issue of outliers. The clusters presented in the following sections are from random initializations. For the hierarchical clustering algorithm, we select k from gauging a dendrogram (Figure 5.2b). There is no strictly correct procedure for this. We select an intersection that, in our opinion, seems clear and that partitions the map in a number of clusters that is in a similar range to that implied by the Davies-Bouldin index.

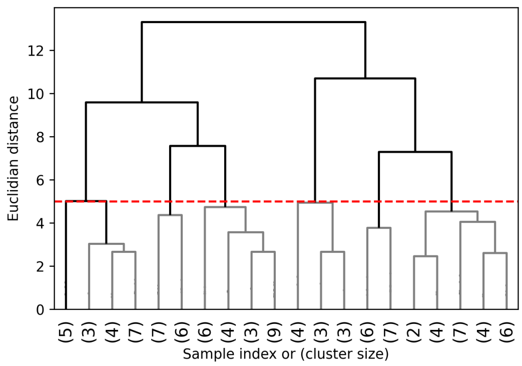
Figure 5.2: Choosing k.

Subfigure (a) shows the Davies-Bouldin index and elbow plot for k-means clustering. Subfigure (b) shows the dendrogram from hierarchical clustering using the Ward distance (Equation 5.4) as merging criteria. Vertical line lengths grow with merging cost. The input data is the neurons of the trained SOM in depicted in Figure 5.3.

(a) Davies-Bouldin index.



(b) Dendrogram.



[1]In addition, we display the elbow plot, which shows the explained variation among the clusters. A change in slope signifies good cluster validity. The latter does not suggest a specific value of k but is provided for completeness.

5.2 Results

In this section, we present and discuss results from clustering a SOM of mutual equity funds. We first describe the results from the SOM. We then present and compare inferences drawn on the strategy employed in delegated fund management from the clustering procedures. Finally, we discuss sensitivity in Section 5.2.5.

5.2.1 The self-organizing map

Figure 5.3: Self-organizing map.

The figure shows the distance plot of the trained SOM. Funds are plotted as dots in the square corresponding to their best-matching neuron. Incentive fee funds are plotted in blue. Randomness is added to each fund to show the varying number of funds that belong to each neuron. Cell colours indicate the normalized sum of distances from each neuron to all other neurons. $M = 100$, $\eta_0 = 0.01$, $\sigma_0 = 50$ and $I = 50\,000$.

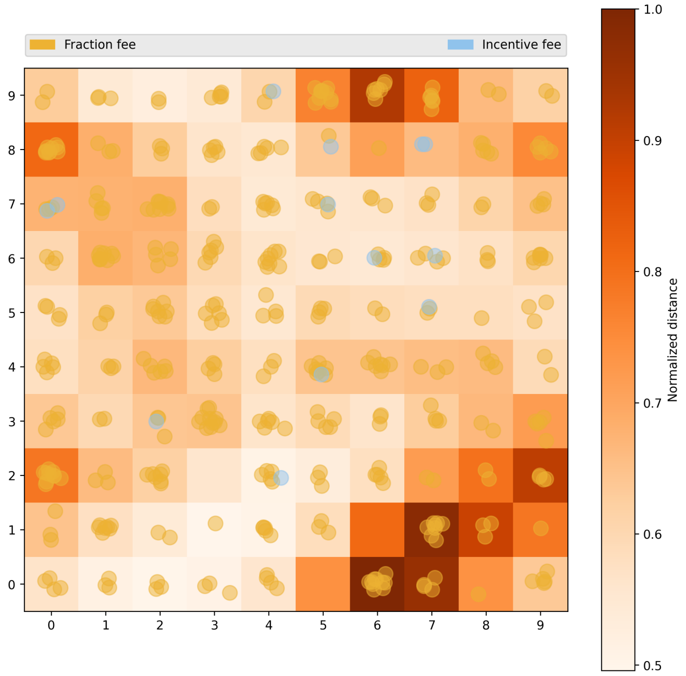
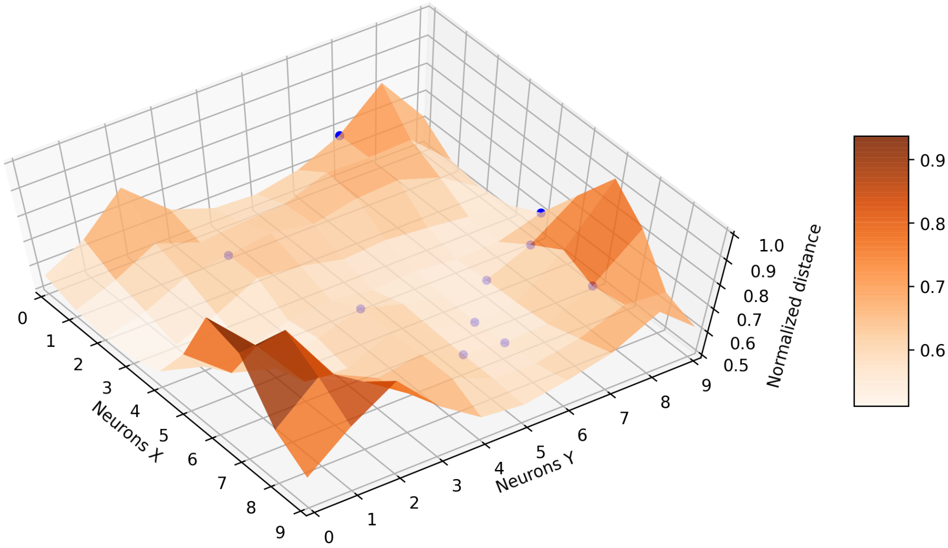


Figure 5.3 shows the normalized distance plot from the SOM trained on the features described in Table 5.1. Funds are scattered next to their best matching unit. The shade in each cell indicates the normalized distance from each neuron to all other neurons. Funds that lie in dark regions are likely far away from most funds in high-dimensional space. Funds positioned in lighter areas are closer to the average position of all funds.

We incorporate a third axis to visualize colour shade as height in Figure 5.4. The darker regions visibly form a mountain that stretches from the south-eastern corner and extends over borders to the other corners of the map. We observe a valley that stretches from the east and splits north and south in the central region of the map, where funds lie the closest to the weighted centre of fund features in high-dimensional space.

Figure 5.4: Visualization of the self-organizing map in three dimensions.

The figure shows the self-organizing map plotted in three dimensions using the normalized sum of distances from each neuron to all other neurons as height. Incentive fee funds are coloured blue. Incentive funds with the same BMU overlap.



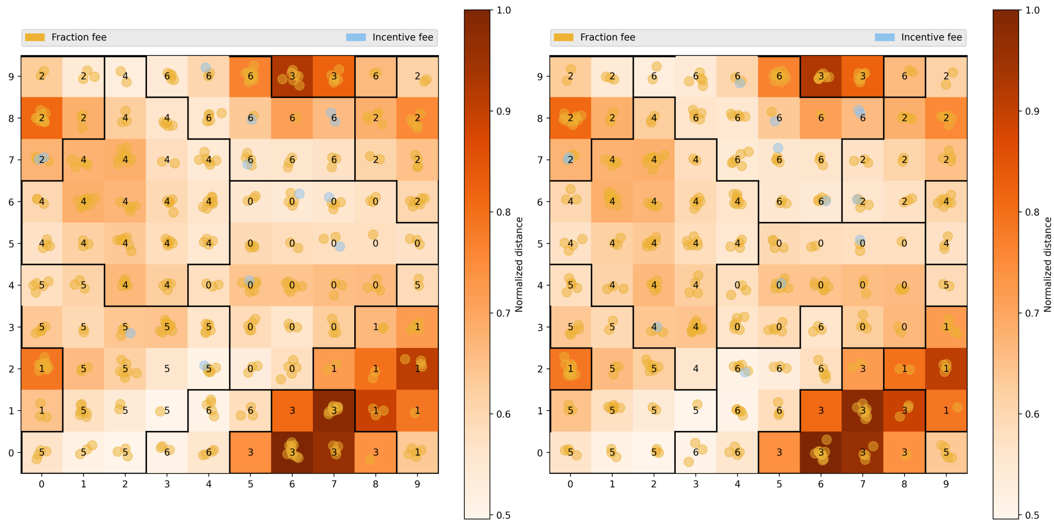
The fraction of funds for which the second-best matching neuron is not one of four direct neighbours to the best matching neuron (topographic error) for the map is 17%. Topology preservation means that funds that are close on the map lie close also in high-dimensional space. In a hypothetical case where the topology is not preserved, we could obtain maps where neurons could both lie close in the two-dimensional map and appear to have similar distances to the rest of the neurons, despite being dissimilar in the high-dimensional space. In such a case, interpreting a clustered map could yield spurious inferences.

The incentive fee funds in our sample are spread in both hilly and flat areas of the map. Incentive fee funds in the elevated regions of the map have feature sets which deviate from the most common patterns. To understand which feature sets lie the closest and how they can be characterized in groups, we employ clustering algorithms on the neurons in the map. We characterize clusters of mutual funds by the fraction of funds in that cluster which employs incentive fees.

Figure 5.5: Self-organizing maps with cluster borders.

The figures show the distance plot with cluster borders drawn by (a) the k-means algorithm and (b) the hierarchical algorithm. Incentive fee funds are coloured blue. Darker shades of orange signify higher normalized distance to all other neurons. $M = 100$, $\eta_0 = 0.01$, $\sigma_0 = 5$, $I = 50\,000$ and $k = 7$.

(a) K-means clustering. **(b)** Hierarchical clustering.



5.2.2 K-means clustering

We present the cluster borders from clustering the SOM in Figure 5.5a. The clusters consist of neurons that neighbour each other directly (with one exception). This is a further indication that topology is preserved. When the relative positions of neurons on the map reflect the those from high-dimensional spaces, a clustering algorithm will construct coherent groups also on the map. We also see a tendency for the cluster borders to follow the slopes around the hilly areas on the map. This tendency is a reflection of distant funds in tall terrain being separated from other groups.

Figure 5.6 gives an overview of the descriptive statistics for clusters $A0$-$A6$. The actual values are listed in Panel A in Table 5.2. Incentive fee funds appear in four of seven clusters ($A0$, $A2$, $A5$, $A6$). We discuss these, before we briefly describe the remaining groups ($A1$, $A3$, $A4$).

Cluster $A0$ contains four incentive fee funds. For this group, the best differentiator seems to be positive and negative loadings to the HML- and MOM-portfolios, respectively. For the HML-loading, we note that its significance stands much more out when mapping on t-statistics instead of directly on loading coefficients. A qualitative categorization of the strategy employed by this group could be mostly tracking the benchmark with a tendency of investing in undervalued and ill-performing stocks.

The cluster borders of $A2$ encircle a hill of neurons that are of higher relative distance. The clear separation from the pack of neurons is likely due to the significant loading to the

Figure 5.6: Radar plots of k-means cluster feature averages.

The radar plots show the linearly normalized cluster feature averages from Panel A in Table 5.2. 0% (100%) is equivalent to the lowest (highest) feature value average among the clusters. The fraction of incentive fee funds is shown in each figure header. Incentive fee fund averages are plotted as a grey outline.

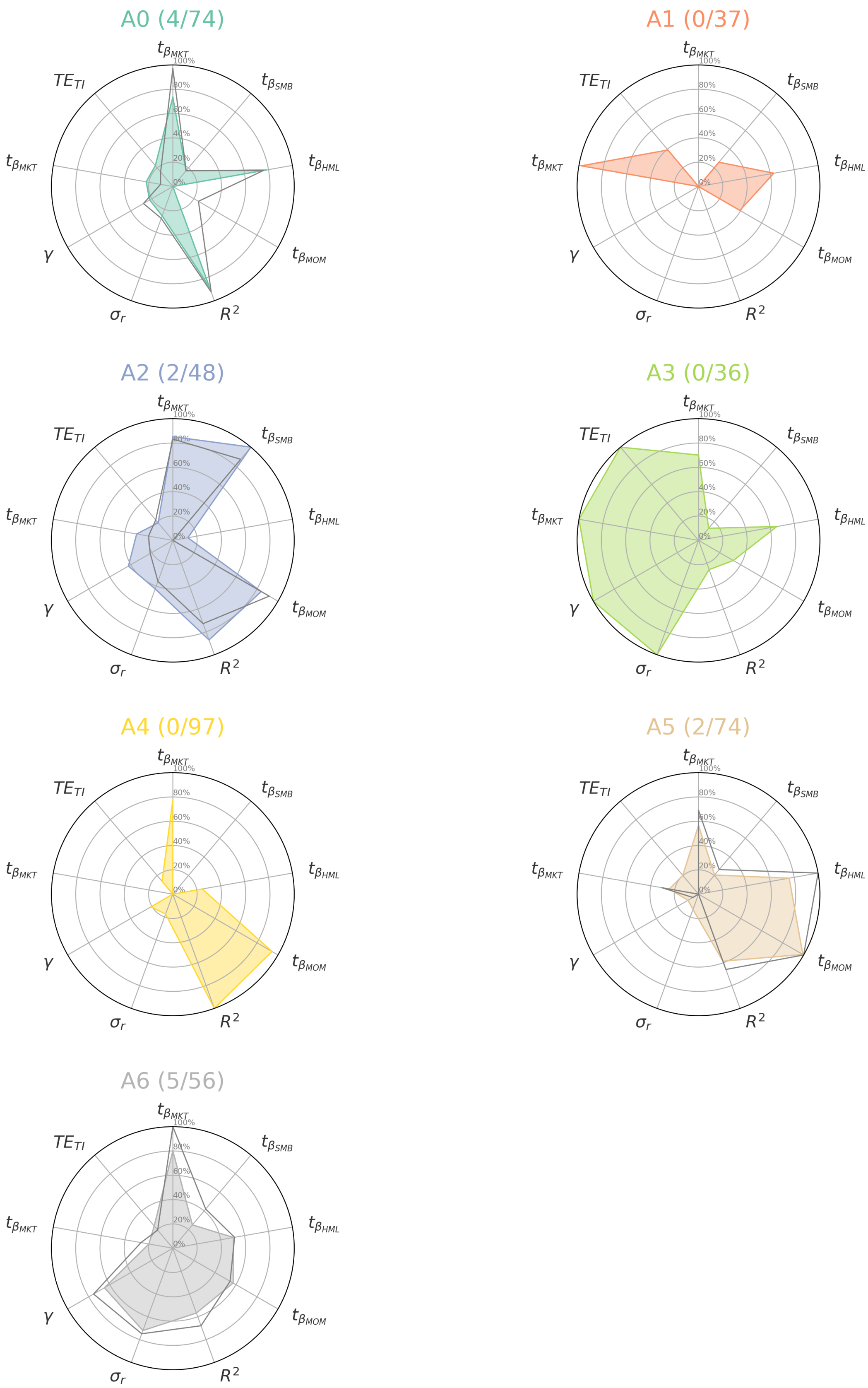
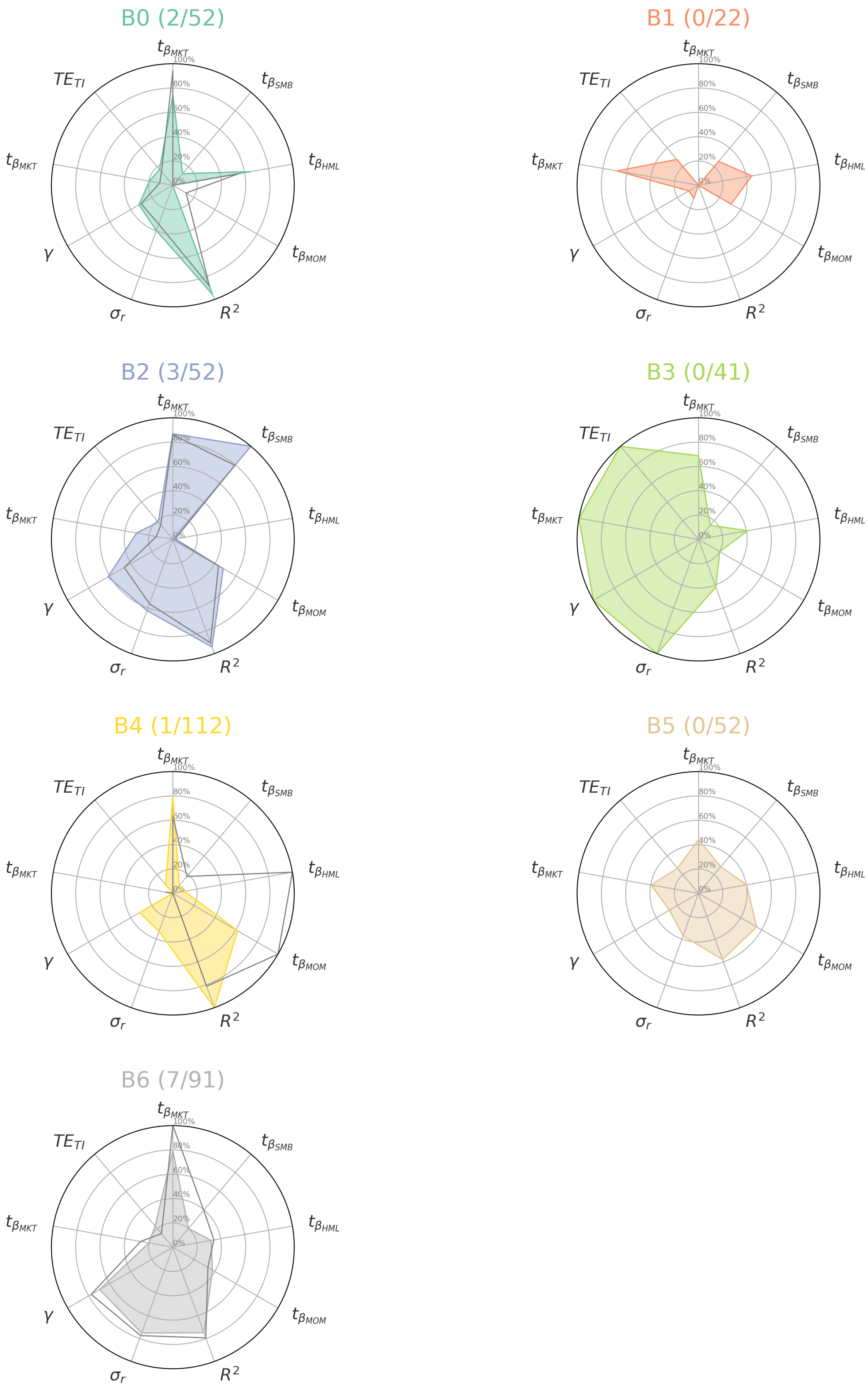


Figure 5.7: Radar plots of hierarchical cluster feature averages.

The radar plots show the linearly normalized cluster feature averages from Panel B in Table 5.2. 0% (100%) is equivalent to the lowest (highest) feature value average among the clusters. The fraction of incentive fee funds is shown in each figure header. Incentive fee fund averages are plotted as a grey outline.



SMB-portfolio for the funds in this area. The cluster contains two incentive fee funds, both of which are state in their prospects that they are small-capitalization funds. This group of funds moreover exhibits the most negative HML-loading and the highest market beta. It seems that the stocks selected by the small-capitalization funds in this cluster tend to have small book-to-market ratios and fluctuate strongly with movements in the overall market. While there may be other small-capitalization funds in bordering regions in $A5$ or $A6$, this pattern interestingly holds for a considerable fraction of the funds with large SMB-loadings in our sample.

$A5$ is situated on the south-western plains of Figure 5.5a and is characterized by market beta below unity, low regression fit and slight exposure to the HML- and MOM-portfolios. The combination of beta and fit could mean that managers prefer taking on considerable unsystematic risk to beat their benchmarks, which in theory is an expected characteristic among incentive fee fund managers (Carpenter, 1989).[2] Taking loadings to both HML- and MOM-portfolios indicate a strategy of investing in well-performing stocks that are cheap in terms in book-to-market ratio, perhaps in the hope of timing investment in a distressed company that is about to recover.

$A6$ covers the neurons on the slopes south and west of the mountain dominated by $A3$ in Figure 5.5a. From clustering by the k-means algorithm, this is the cluster that contains the largest fraction of incentive fee funds. While none of the single measures of risk is of abnormal size, the overall pattern is one that suits expectations one might have for incentive fee fund behaviour. The loading to the market beta is second-highest among the clusters, the regression fit is on the lower end, and both the unconditional volatility in returns and the semivariance measures are high. This pattern points to two a priori hypotheses on incentive fee funds. First, market beta and volatility levels indicate that managers leverage their returns against their benchmarks to take advantage of an option-like payoff function (Grinblatt and Titman, 1989). Secondly, low regression fit point to managers looking outside of the conventional factor portfolios for differential returns, perhaps in the belief of possessing skill in employing some unorthodox strategy. For both behaviours, the high level of semivariance shows that they present the fund investor with considerable downside risk.

From clustering the map by the k-means algorithm, groups $A1$ and $A3$ are two of three clusters that contain no incentive fee funds. Neurons in both groups share low regression fit. $A1$ is separated from the pack by loadings to the t-statistic of market beta far out in the left tail. The funds in $A3$ have the most volatile returns and very high tracking errors. While there is no doubt that there may be funds which invest far away from the index ($A1$) or with extreme volatility ($A3$), we fear that the clear separation may be a result of misassignment of factor portfolios and benchmark indices, respectively. The fact that $A3$ contains no incentive fee funds might be surprising, as an option-like payoff function motivates increased volatility around the benchmark. Gauging Figure 5.5a, we note that some of the incentive fee funds in $A6$ have patterns that border $A3$ also in high-dimensional space.

[2] Low regression fit may also be a sign of slight misassignment of factor portfolios. See Chapter 2.

The neurons in $A4$ exhibit the smallest and least varying tracking error together with market beta levels near unity and high regression fit. Together with an inclination to invest in large-capitalization stocks, it might seem that funds in this group are investing close to a large-capitalization stock index.[3] An investor that intends to pay for active management is hopefully aware of any tendency of 'closet-indexing'.

It seems that any contrast introduced by an incentive fee fund contract by no means is easily partitioned from a backdrop of fraction fee funds. Some incentive fee funds belong to groups with traits that are in line with conclusion patterns from contract theory, while others follow nuanced and non-extreme strategies. Before discussing implications for the investor, we explore the shape of the high-dimensional data set that the neurons of the map represent by an alternative clustering procedure.

5.2.3 Hierarchical clustering

We present the hierarchical clustering of the SOM in Figure 5.5b. Aside from hierarchical clustering being a bottom-up approach, the most notable difference between the clustering methods is that hierarchical clustering is more permissive to clusters non-hyperspherical shapes. While the underlying map remains the same, the alternative methodology of clustering thus draws borders in a different manner than the k-means algorithm. In other words, hierarchical clustering considers different partitions of neurons on the map to form natural groups.

The descriptive stats for clusters $B0$-$B6$ are listed in Panel B in Table 5.2, and visualized in Figure 5.7. The clusters broadly retain their feature patterns and are similar to the equally numbered clusters from $A0$-$A6$. Here, we focus on the regions where borders have changed and how they affect the interpretation of classes that contain incentive fee funds.

We first note that $B2$, in comparison with $A2$, gains neurons from $A0$ and $A6$. As a result, the cluster contains three incentive fee funds. The added funds slightly moderate the pronounced loadings to the SMB- and HML-portfolios that were present in $A2$. The new categorization of border neurons indicates that the added funds by the hierarchical algorithm share overall patterns. For the added incentive fee fund, the new categorization may have been helped by similarity in high market beta and tracking error. Nonetheless, the alternative clustering approach maintains that incentive fee funds are present in a high-beta small-capitalization fund group.

The hierarchical clustering algorithm considers $B6$ far larger than $A6$, pushing the cluster borders of $B6$ past neurons in the north and south of $A0$, in the east of both $A4$ and $A5$. The cluster expansion is primarily over low-lying areas of the map, which are the most similar to average feature patterns. As a result, the bulk of characteristics that separated $A6$ from other groups are diluted. $B6$ averages slight loadings to the market and SMB-portfolios, while regression fit and semivariance remain low and high, respectively.

[3]We remind the reader that we omit any funds with a stated strategy of replicating an index from the data set.

Interestingly, *B*6 contains seven incentive fee funds. Although the individual risk properties of each incentive fee fund remain static, the hierarchical clustering algorithm considers them similar in terms of feature pattern to a group that averages feature values that tend towards the generic. Important deviations from the norm for the group are the low regression fit and high semivariance measures, still. Interpretation of this configuration requires generalizing the hypotheses made from *A*6. From this, it seems a large portion of incentive fee funds belongs to a group that employs strategies which rely on subtle exposure outside of conventional factor portfolios to generate differential return. Common to those strategies is a market beta just above unity and chance for moderately high losses.

*B*0 changes in size but maintains the risk-taking profile of *A*0. *B*0 contains two incentive fee funds, in comparison to the four of *A*0.

Similarly, *B*5 is smaller than *A*5. The hierarchical clustering algorithm instead classifies the incentive fee funds in this region into *B*4 and *B*6. This clustering thus does not consider any incentive fee funds to be similar in pattern to a group that invests in well-performing and high book-to-market stocks. Although one incentive fee fund lies in *B*4, the fact that it lies on the edge of the cluster signals that it might be dissimilar from the broad group. This effect is illustrated in Figure 5.7.

5.2.4 Interpretation of incentive fee fund positions

Clustering of the self-organizing map gives some insight into the risk-taking classes and perhaps motivations of incentive fee fund managers. In our exploratory approach, we cluster by two methods for robustness, and some trends hold for both procedures.

For one, there are incentive fee funds that belong to groups which primarily are characterized by factor strategies, such as in *A*0, *A*5 and *A*6. The fact that incentive fee funds share risk properties with a large number of fraction fee funds might imply that the incentive fee fund managers are confident in their assumed stock-picking skills to beat the competition.

The two clustering methods, in addition, agree that incentive fee funds do not belong among the funds that most closely follow their index.[4] While one might have expected incentive fee funds to have more extreme positions on the map, there is little doubt that they rarely share risk properties with the group that invests the most passively. This is in line with e.g. Elton et al. (2003). For the incentive fee fund investor, this should alleviate any concern that the fund manager is not taking active bets.

For the classification of incentive fee funds in a group that relies on high market beta and unconditional volatility, the two clustering approaches disagree somewhat, and we draw inference with corresponding nuance. Class *A*6 by the k-means algorithm implies that a large portion of incentive fee funds belongs in a group that exhibits a tendency of high market beta

[4]The one exception in Figure 5.5b lies in the outskirts of cluster *B*4, indicating relative dissimilarity to the cluster averages.

and volatility. For an incentive fee fund manager, this is analogous to 'gaming' the option-like contract to earn positive fees. While such behaviour is expected from the literature (e.g. Grinblatt and Titman, 1989), it points to incentive contracts misaligning investor-manager interests.

In contrast, classification by the hierarchical algorithm categorizes the risk properties of the incentive fee funds in question with a broader group of fraction fee funds. From this clustering, these incentive fee funds belong to a strategic group that attempts to beat their benchmark mainly by deviating from factor portfolios. With less clear separation from sample averages, one would assume that the incentives aside from the explicit contract, such as, e.g. implicit flow incentives, weigh the heaviest in determining a fund's risk profile.

Another look at the map in Figure 5.5b gives nuance to the inferences drawn from the average feature values for neurons in *B*6. The incentive fee funds predominantly lie in the north-east of the cluster and are in effect, perhaps not well represented by average values from such a broad group. Comparing the feature means of the class to the outline formed by the incentive fee fund subsample in Figure 5.7, this seems evident. In any case, although a tendency to 'game' the contract is present for some of these incentive fee funds, it is not strong enough to separate those funds from a wider group of fraction fee funds by our chosen risk-taking features by hierarchical clustering.

From examining descriptive statistics for measures that were not clustered on in Table 5.2, we find it interesting to note that incentive fee funds tend to lie in the best-performing fund groups.[5] While results from Section 3.2.3 imply that incentive fee fund managers do not possess skill, we see here that many took risk in ways that generally outperformed other strategies in the period we consider. Simultaneously, it implies that incentive fee fund manager performance was less a result of stock-picking and more of general risk exposure. In addition, it seems that any investors who were subject to incentive fee fund managers levering contract pay-off coincidentally were exposed to a well-performing section of the risk map. This sets the work of Stoughton (1993) into perspective, who argues that incentive contracts share risk suboptimally but still are optimal for the investor.

We note with interest that these results are more nuanced regarding the metrics that were discussed in Section 4.1. First, even though incentive fee funds average high market beta as a gross group, it does not mean that it is the most potent descriptor of the risk properties for every fund in sample. Second, we observe that the high levels of tracking error in the fraction fee category stem from funds in clusters corresponding to *A*1 and *A*3.

[5] Notably *A*6, *B*2 and *B*6.

Table 5.2: Descriptive statistics for the clusters obtained from the 10x10 SOM.

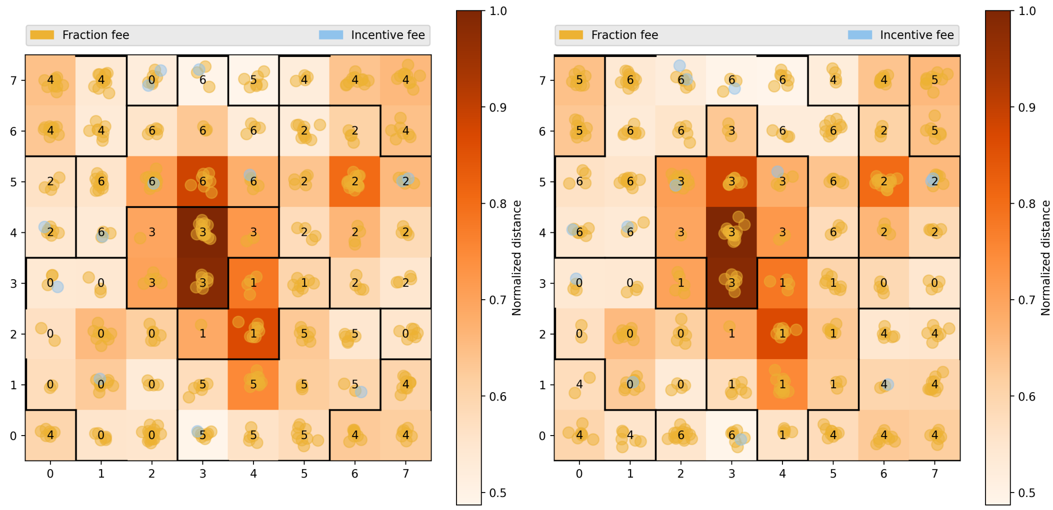
The tables show cluster statistics for each of the clusters obtained. The features used as input for the SOM are to the left of the dotted vertical line. $M = 100$, $\eta_0 = 0.01$, $\sigma_0 = 5$, $I = 50\,000$ and $k = 7$.

	$t_{\beta_{MKT}}$	$t_{\beta_{SMB}}$	$t_{\beta_{HML}}$	$t_{\beta_{MOM}}$	R^2	σ_r	γ	TE_{TI}	σ_{TE}	$\bar{\alpha}(\%)$	TNA	TE_{FM}	Flow	β_{MKT}	β_{SMB}	β_{HML}	β_{MOM}	N
Panel A: K-means clustering.																		
A0	-0.02	0.11	0.59	-0.67	0.89	0.05	0.29	2.09	0.66	-2.14	301.26	2.14	-0.93	1.0	0.04	0.12	-0.11	4/74
A1	-3.66	0.4	0.18	-0.18	0.59	0.04	0.18	3.79	1.05	-3.9	50.74	3.81	-0.26	0.62	0.13	0.08	-0.02	0/37
A2	0.55	2.72	-1.16	0.37	0.88	0.06	0.38	2.27	0.57	-1.92	247.33	3.02	-0.29	1.06	0.58	-0.24	0.08	2/48
A3	-0.19	-0.01	0.25	-0.26	0.68	0.08	0.65	3.83	2.47	-3.86	356.37	3.18	-0.42	0.97	-0.0	0.22	-0.13	0/36
A4	0.17	-0.42	-0.84	0.5	0.93	0.05	0.28	1.59	0.45	-2.32	249.65	1.9	0.62	1.02	-0.06	-0.13	0.08	0/97
A5	-0.86	0.23	0.52	0.56	0.79	0.05	0.23	2.14	0.6	-2.85	139.63	2.34	0.44	0.89	0.07	0.14	0.12	2/74
A6	0.3	0.37	-0.15	0.04	0.78	0.07	0.49	2.04	0.69	-1.74	405.28	2.18	-1.34	1.04	0.1	-0.07	0.0	5/56
Panel B: Hierarchical clustering.																		
B0	-0.14	-0.09	0.88	-0.63	0.9	0.05	0.25	2.07	0.57	-2.22	336.66	1.82	0.07	1.0	0.03	0.16	-0.09	2/52
B1	-4.14	0.33	0.32	-0.07	0.5	0.03	0.13	3.32	0.82	-3.46	60.85	3.49	-0.3	0.53	0.13	0.12	0.01	0/22
B2	0.55	2.64	-0.9	0.24	0.89	0.06	0.4	2.34	0.61	-1.79	246.61	3.22	-0.44	1.06	0.54	-0.19	0.06	3/52
B3	-0.43	-0.0	0.23	-0.27	0.67	0.07	0.6	4.14	2.47	-4.06	346.7	3.39	-0.42	0.94	0.0	0.21	-0.13	0/41
B4	0.16	-0.22	-0.64	0.48	0.91	0.04	0.25	1.56	0.4	-2.2	214.35	1.78	0.55	1.02	-0.01	-0.1	0.07	1/112
B5	-1.76	0.41	0.2	0.37	0.74	0.05	0.23	2.58	0.8	-3.7	59.18	3.0	-0.19	0.8	0.1	0.08	0.1	0/52
B6	0.14	0.15	-0.02	0.05	0.81	0.06	0.44	2.08	0.7	-1.93	374.81	2.28	-1.01	1.02	0.06	-0.0	-0.0	7/91

Figure 5.8: Self-organizing maps of size 8x8 with cluster borders.

The figures show the distance plot with cluster borders drawn by (a) the k-means algorithm and (b) the hierarchical algorithm. Incentive fee funds are coloured blue. Darker shades of orange signify higher normalized distance to all other neurons. $M = 64$, $\eta_0 = 0.01$, $\sigma_0 = 4$, $I = 50\,000$ and $k = 7$.

(a) K-means clustering. **(b)** Hierarchical clustering.



5.2.5 Sensitivity to network dimension

We plot the distance maps for self-organizing maps of size 8x8 for k-means and hierarchical clustering in Figures 5.8a and 5.8b, respectively. Corresponding maps for size 12x12 are plotted in Figures 5.9a and 5.9b. The tables describing the features for each of the fund clusters can be found in Appendix D. We keep the notation from the previous section and name the 8x8 cluster sets $C0 - C6$ and $D0 - D6$, and 12x12 sets $E0 - E6$ and $F0 - F6$. We assign cluster numbers by similarity to clusters in our main results.

The internal validity measures for the map of size 8x8 are inferior to those of the map of size 10x10. This is a result of having less two-dimensional space to unfold a high-dimensional data structure. By tweaking the hyperparameters, we are with this map unable to produce comparable values for QE or TE without sacrificing one. In order to achieve comparable QE, the value for TE shoots toward 50%, which means that half the second best-matching neurons are not adjacent to the first. In such a scenario of poor topology preservation, any spatial interpretation of the original data structure from the map is not reliable. In Figure 5.8a we opt for a configuration where TE is bearable (20%) at the cost of lower map resolution ($QE = 1.4$). Here, the validity measures point to a lack of granularity in representation consequently and that we attempt to model the original data set too sparsely.

Although the orientation and relative positioning of neurons are different, the 8x8 map shares the most recognizable properties with the map of size 10x10. The funds with the lowest regression fit and highest tracking error form the tallest peak. Funds where large loadings to the SMB-factor dominate form a smaller peak. The area that has the smallest relative distance to the rest of the map is dominated by funds that seem to manage the most passively. Incentive fee funds are, concerning the discussed map features, positioned broadly similarily in comparison with our main results. As in Figure 5.5a, we, for example, find that most of the incentive fee funds cluster around one side of the tallest area of the map (in this case, the north-western corner).

The cluster borders drawn from the k-means algorithm are similar to our main results, with each of clusters $C0-C6$ roughly matching one of $A0-A6$. A notable difference is that the borders around $C6$ traverse farther up towards the top of the mountain, yielding an interpretation of the cluster as more extreme in terms of unconditional volatility, semivariance and tracking error. Here, the clustering also considers one more incentive fee fund to be part of $C2$, which is characterized by a significant loading to the SMB-portfolio and some market beta.

For the output layer in the $8x8$ SOM, the hierarchical clustering algorithm draws borders that extend the tendencies we see for hierarchical clustering of the $10x10$ map in Figure 5.5b somewhat. The cluster borders differ the most for $D3$ and $D6$. In this configuration, more of the northern slope from the centre peak is considered as part of $D3$. As a result, we find two incentive fee funds in a group that is mainly characterized by very high volatility and activity. Examining $D6$, we see that the cluster has grown in similar directions as for $B6$, but to a greater extent. As a result, the feature averages of the cluster are less pronounced. Interpretation of incentive fee fund behaviour in this instance would require even more nuance than for the hierarchical clustering in our main results. We also note that $D4$ and $D5$ have redistributed a few neurons between the two groups.

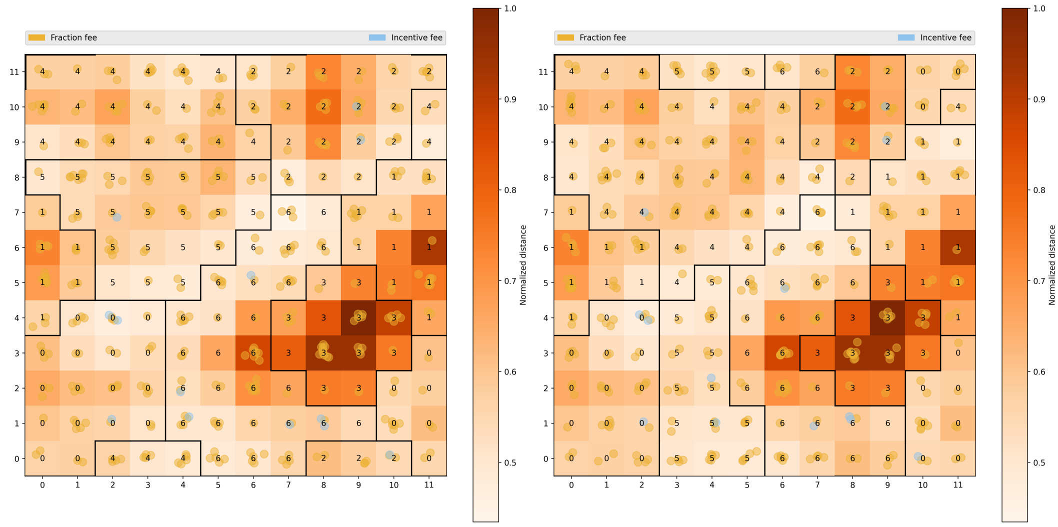
Increasing the map size beyond $10x10$, we find that the validity measures tend to decrease, but do so at a progressively slower rate. A larger amount of neurons on the map allows for increased granularity and makes it easier to preserve topology. In Figure 5.9a QE and TE decrease to 1.15 and 17%, respectively. Increasing map size to game these parameters is analogous to overfitting in supervised learning. The cost of large map size is that the distinction between pairs of neurons is too small and that more neurons are unused (Maillet and Rousset, 2003). Here, 6% of squares are empty, compared to 3% in Figure 5.5a.

As for the $8x8$ map, clustering the $12x12$ map by the k-means algorithm reveals that it shares core properties with the $10x10$ map. The map similarily contains areas and clusters for neurons that broadly are dominated by high tracking error ($E1$), SMB-loading ($E2$), low market beta and low regression fit ($E3$), positive HML-loadings and negative MOM-loadings ($E0$), and some loading to market beta and considerable volatility ($E6$). The most notable difference is that the borders in the flat regions of the map have shifted, yielding to two clusters that are characterized by quite passive management with an influx of various factor loadings ($E4$ and $E5$). Incentive fee funds belong to similar groups, as well. A considerable portion of the incentive fee funds land in $E6$, while other funds lie in $E0$ and $E2$, and one fund lands in $E5$.

Figure 5.9: Self-organizing maps of size 12x12 with cluster borders.

The figures show the distance plot with cluster borders drawn by (a) the k-means algorithm and (b) the hierarchical algorithm. Incentive fee funds are coloured blue. Darker shades of orange signify higher normalized distance to all other neurons. $M = 144$, $\eta_0 = 0.03$, $\sigma_0 = 6$, $I = 50\,000$ and $k = 7$.

(a) K-means clustering. **(b)** Hierarchical clustering.



Hierarchical clustering of the 12x12 map yields results that are broadly similar to the clusters $A0-A6$ and $B0-B6$ for the pronounced regions of the map. For $F0$ and $F5$, borders are shifted in comparison to the 10x10 map. Here, $F4$ is again a large area characterized by passive management. For $F5$, the more active neurons are given to $F0$ and $F6$, resulting in the cluster averaging at feature values that less distinctly point in a direction. As four incentive fee funds lie in $F0$ and two lie in $F5$, this configuration seems to argue that there are incentive fee fund managers who rely on strategies inside or outside of factor sets rather than market beta and volatility.

To summarize, clustering of the SOM seems robust in categorization for most areas of the map across grid sizes and clustering methods. The grouping of fraction fee and incentive fee funds into clusters that correspond to $A1$, $A2$, $A3$ and $A6$ are mostly consistent. Perhaps the most notable difference is the hierarchical algorithm used on a small map implies more strongly than other maps that incentive fee funds follow strategies that are only a nuance away from general fraction fee funds. While we note this tendency, we keep in mind that its presence diminishes with the improvement of internal validity measures.

For certain regions of the map that are predominantly flat, clustering of the SOM less consistently leads to invariable results. The neurons that belong to clusters $A0$, $A4$ and $A5$ are, for differing map sizes and clustering algorithms, shifted between groups. We hypothesize that this tendency is a function of the structure of the feature space. It may be the case that the data points often represented by central neurons lie in the feature space in a way that yields different shapes of neurons in the area when the number of neurons available for representation changes. While the topographic error measure tells us that relative positions of neurons on the map remain quite consistent across map sizes, changes in shapes formed by neurons in the high-dimensional space would explain the difference in clustering output seen for various configurations. Because of this, we have to be more cautious when drawing inference from the cluster affiliations for incentive fee funds that lie in the flattest areas of the map. In addition, for the cases when incentive fee funds land near the edge of such a cluster, orientation on feature averages of neighbouring clusters grows in importance.

6 Conclusion

Incentive fee funds are growing in importance in the Norwegian fund universe. The structure of the contracts inclines an investor to assume that managers of these funds are better and harder-working than other managers. Simultaneously, the contract structure may motivate the manager to take risk exposure that is against the interests of the investor. For the investor, a thorough understanding of these issues seems a priority. In this thesis, we study the performance and risk-taking behaviour of 409 fraction fee funds and 13 incentive fee funds using monthly data for the fund set registered on the OSE in the period from 2000 to 2018.

For groups of incentive fee funds, we are not able to reject the null hypothesis of zero post-fee alpha. If managers produce positive risk-adjusted returns, any significant alpha is accrued by the fund fees. For fraction fee funds, the results are slightly grimmer; several groups generate alphas for their investors that are significantly negative. The level of performance for funds, in general, is in line with seminal works in financial literature such as that of, e.g. Sharpe (1991). For incentive fee funds, Elton et al. (2003) and Massa and Patgiri (2009) conclude more positively that incentive fee fund managers outperform fraction fee managers.

A vast majority of fund managers decisively do not exhibit *skill* in producing positive alpha for their investors. In other words, nearly all fund managers that generate positive alpha are *lucky*. This holds for both fraction and incentive fee funds and is similar to what Fama and French (2010) find for a U.S. sample. From the perspective of a Norwegian investor, our results for fraction fee funds are similar to those of e.g. Sørensen (2009), although comparable studies generally disregard funds outside of those that invest in primarily Norwegian equities. In this setting, we find no verification of the theoretical inclination e.g. (Jensen and Meckling, 1976) that incentive fee fund managers as a general tendency exhibit skill.

We examine whether explicit incentive fee contracts add to the convexity of the payoff schedule in the spirit of Chevalier and Ellison (1997). For fraction fee funds, a risk-changing effect from annual implied incentives is not present. For incentive fee funds, we find indications that top performers, especially among young and small funds, lock in gains. This behaviour is a sign that investor-manager incentives are not aligned. For their sample in the period from 1990 to 2000, Elton et al. (2003), in addition, find evidence that the poorly performing incentive fee funds increase tracking error to their benchmark.

In an explorative approach, we map the risk properties of the funds in our sample by a self-organizing map (Kohonen, 1990). We cluster the output of the map by both a partitive and an agglomerative algorithm. By doing this, we not only study risk-taking characteristics of our set of funds but also give a thorough example within finance for the use of an intuitive, non-linear method for data exploration. The self-organizing map reveals a nuanced picture of the risk properties of incentive fee funds in relation to the fraction fee funds in our sample.

As expected from the literature (e.g. Grinblatt and Titman, 1989), we find that a considerable portion of incentive fee funds belongs in a group of funds that tend to have high market beta and volatility in returns. The average beta for the class as a whole is greater than unity and significantly different from that of fraction-fee funds. For an incentive fee fund manager, this is analogous to abusing the option-like contract to earn positive fees and likely a symptom of agency issues. Similarly, we see that the incentive funds in our sample generally are easily separated in risk properties from the subset of funds that manage the most passively. The incentive fee fund investor can feel confident that their manager is not replicating an index.

Interestingly, we find that there are many incentive fee fund managers on the spectrum between the two extremes of very active and passive management. This set of managers takes moderate exposure to factor portfolios or unsystematic risk to beat their benchmark, perhaps from stock-picking in the belief of possessing skill. For this group of incentive fee funds, investor-manager incentives seem to be better aligned than in the group that beats their contract by exploitation of its option-like element.

For another perspective on performance, we find that most incentive fee funds share risk properties with the fund groups that exhibit the best average four-factor regressed post-fee alpha. Even if incentive fee fund managers do not possess skill in picking stocks, for our sample and considered period, they show a tendency of outperforming the average fraction fee manager in selecting beneficial risk exposure. This insight helps in characterizing the manager but is of less relevance for the investor. As incentive fee funds seem to nonetheless generate post-fee alpha non-significantly different from zero, a risk-averse investor will likely achieve comparable returns by investing in a fund that replicates the corresponding benchmark index at a low cost.

Our work points to various avenues for future research. First, it would be of interest to run similar analyses to those in this thesis for a larger sample, given that incentive fee categorizations are available. For the funds registered on the OSE, for this we suggest the construction of suitable sets of factor portfolios for industrial sectors. Second, we believe that obtaining a comprehensive set of fund manager employment contracts would aid in shedding more light on the principal-agent relationship in question. Such a set would allow for removal of, or at the very least insight on, the assumption that managers act on fund fees. Third, we prompt future research, building on our work, to continue the exploration of the usefulness of self-organizing maps for mapping financial data.

A Data

Table A.1: Fund selection criteria.

The table shows the explicit search criteria we use to collect the initial sample of funds mentioned in Chapter 2. The criteria above the dashed line are used in the fund screener to retrieve the names and ISIN of the funds. The criteria below the dashed line are used to collect the time series for NAV and TNA from the formula builder in the Excel module provided by Refinitiv Eikon.

Parameter	Value
Countries Registered for sale	Norway
Lipper Global	Equity[a]
Fund	Active, Liquidated, Merged & Primary fund
Asset Universe	Mutual funds
Currency	United States Dollar (USD)
Exchange	The Oslo Stock Exchange (OSE)

[a] All equity categories.

Table A.2: Assignment of Technical Indicators (TI).

The table shows an overview of the Technical Indicators for the fund sample. *N* is the number of funds assigned to each respective Technical Indicator.

Technical Indicator (TI)	*N*	Technical Indicator (TI)	*N*
BOVESPA (Ibovespa) TR	3	Bombay Stock Exchange 100 Index	10
CAC 40 CR	1	DOW JONES U.S. SELECT dividend tot ret	1
Dow Jones US Select Dividend Total Return Index	1	EURO STOXX 50 TR EUR	1
FTSE 100 TR	6	FTSE AW/Industrials TR	1
FTSE AW/Oil & Gas TR	2	FTSE Bursa Malaysia KLCI TR	1
FTSE Singapore Straits Times TR	1	FTSE Turkey TR	2
Hang Seng CR	1	KOSPI Composite CR	3
MOEX Russia	5	MSCI AC ASEAN TR USD	1
MSCI AC Asia Pacific TR USD	3	MSCI AC Asia Pacific ex Japan TR USD	24
MSCI BRIC Daily TR	1	MSCI China TR USD	6
MSCI EM Small Cap NR USD	1	MSCI EM Small Cap TR USD	1
MSCI EM (Emerging Markets) TR USD	18	MSCI EMU Small Cap TR USD	1
MSCI Emerging Markets Eastern Europe TR	20	MSCI Emerging Markets Latin America TR	8
MSCI Europe High Dividend Yield TR	4	MSCI Europe Small Cap NR USD	1
MSCI Europe Small Cap TR USD	10	MSCI Europe Value NR USD	1
MSCI Europe ex UK TR USD	2	MSCI Golden Dragon TR US	1
MSCI Golden Dragon TR USD	11	MSCI Indonesia TR	2
MSCI Italy TR	1	MSCI Nordic Countries TR USD	21
MSCI Norway TR	35	MSCI Pacific Small Cap TR	2
MSCI Sweden Small Cap TR	2	MSCI World NR USD	1
MSCI World Small Cap TR USD	8	MSCI World TR USD	85
OMX Stockholm All Share CR	4	Russell 2000 TR	7
S&P 500 TR	25	S&P Africa 40 CR EUR	2
SBF 120 TR	1	STOXX Europe 50 CR EUR	17
STOXX Europe 50 TR EUR	27	STOXX Europe 600 NR	1
STOXX Nordic Small NR EUR	2	Swiss Performance Index TR	1
TAIEX TR	1	Tokyo SE 2nd Section CR	1
Tokyo SE 2nd Section TR	2	Topix TR	21

Table A.3: Criteria used to assign sets of regression factor data.

The table shows the criteria used for assigning sets of regression factor portfolios to funds in Chapter 2. The Lipper categorization of funds is required to contain one of the phrases in one of the lists in the second column to be assigned the respective factor set.[a] We select keywords by manual inspection of the Lipper categorization for our fund sample.

[a]The categorization is provided by Refinitiv Eikon.

Factor set	List of words to be contained in Lipper categorization
Asia ex. Japan factors	Asean, Asia, Ex Japan, Singapore, Hong Kong
Emerging markets factors	Brazil, China, Emerging Markets Global, Emerging Mkts Europe, Emerging Mkts Global, Emerging Mkts Latin, Emerzging Mkts Other, India, Indonesia, Korea, Malaysia, Russia, Taiwan, Thailand, Turkey
European factors	Nordic, Iberia, Equity Euro, France, UK, Germany, Sweden, Italy, Switzerland, Spain, Finland
Global factors	Equity Global
Japanese factors	Equity Japan
Norwegian factors	Norway
U.S. factors	US

Table A.4: Country data used in construction of factor portfolios.

The table shows the countries included by French (2020) when calculating factor returns for various developed regions. The table is adapted from French (2020).

Country	Global	Europe	Japan	Asia	U.S.
Australia	x			x	
Austria	x	x			
Belgium	x	x			
Canada	x				
Denmark	x	x			
Finland	x	x			
France	x	x			
Germany	x	x			
Great Britain	x	x			
Greece	x	x			
Hong Kong	x			x	
Ireland	x	x			
Italy	x	x			
Japan	x		x		
Netherlands	x	x			
New Zealand	x			x	
Norway	x	x			
Portugal	x	x			
Singapore	x			x	
Spain	x	x			
Sweden	x	x			
Switzerland	x	x			
United States	x				x

For calculation of factor portfolio returns in the Emerging category, French (2020) uses a set of countries that is disjoint from those in Table A.4. The countries included are Argentina, Brazil, Chile, China, Colombia, Czech Republic, Egypt, Greece, Hungary, India, Indonesia, Malaysia, Mexico, Pakistan, Peru, Philippines, Poland, Qatar, Russia, Saudi Arabia, South Africa, South Korea, Taiwan, Thailand, Turkey and the United Arab Emirates.

B Performance

Figure B.1: Distribution of Durbin-Watson test statistics.

The figure shows the histogram of Durbin-Watson test statistics for the funds in our sample. Values that deviate from 2 in the negative (positive) direction indicate negative (positive) autocorrelation in residuals.

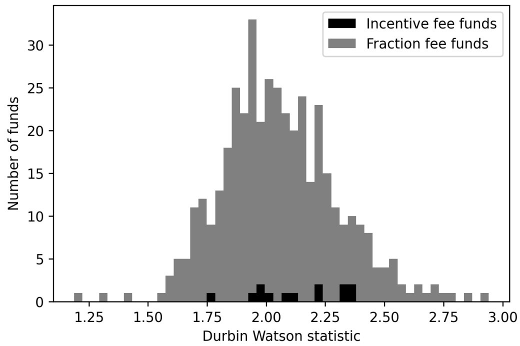


Table B.1: Fama-French three-factor regression.

The table shows regression parameters for the funds in the categories Asia, Emerging, Europe, Global, Japan, Norway and U.S., separated on fraction and incentive fee funds. For each category, the Fama-French three-factor model is computed from an equally weighted portfolio of funds. R^2 denotes fit, and N denotes the number of funds in each portfolio. The regression is performed using Newey-West heteroscedasticity- and autocorrelation-consistent standard errors. Alphas are annualized by multiplication.

	$\alpha(\%)$	β_{MKT}	β_{SMB}	β_{HML}	R^2	N
Panel A: Asia.						
Fraction fee	-4.03***	0.95***	0.1**	0.18***	0.91	34
	(1.37)	(0.02)	(0.04)	(0.04)		
Incentive fee	-	-	-	-	-	-
	-	-	-	-	-	-
Panel B: Emerging.						
Fraction fee	-1.49	0.99***	-0.04	0.12**	0.96	94
	(1.1)	(0.01)	(0.05)	(0.05)		
Incentive fee	1.1	1.05***	0.11	-0.03	0.9	2
	(2.28)	(0.03)	(0.11)	(0.12)		
Panel C: Europe.						
Fraction fee	-1.32*	1.02***	0.24***	-0.14***	0.98	99
	(0.68)	(0.01)	(0.03)	(0.02)		
Incentive fee	-0.1	1.21***	0.33***	-0.29***	0.9	5
	(1.78)	(0.03)	(0.08)	(0.06)		
Panel D: Global.						
Fraction fee	-2.62***	1.01***	0.12***	-0.0	0.97	88
	(0.66)	(0.01)	(0.03)	(0.02)		
Incentive fee	1.47	1.26***	0.42***	0.06	0.9	5
	(1.59)	(0.03)	(0.07)	(0.05)		
Panel E: Japan.						
Fraction fee	-2.42**	0.94***	0.1***	-0.16***	0.92	24
	(1.07)	(0.02)	(0.03)	(0.03)		
Incentive fee	-	-	-	-	-	-
	-	-	-	-	-	-
Panel F: Norway.						
Fraction fee	-0.89	0.96***	0.12***	-0.04***	0.97	33
	(0.81)	(0.01)	(0.02)	(0.02)		
Incentive fee	-6.07	0.99***	0.11	-0.03	0.8	1
	(5.63)	(0.12)	(0.18)	(0.15)		
Panel G: USA.						
Fraction fee	-2.35***	0.99***	0.15***	-0.05***	0.97	37
	(0.69)	(0.01)	(0.03)	(0.02)		
Incentive fee	-	-	-	-	-	-
	-	-	-	-	-	-

Note: *p<0.1; **p<0.05; ***p<0.01

Table B.2: Boostrapped confidence intervals.

The table shows bootstrapped confidence intervals for the coefficients from the Carhart four-factor regression. The bootstrap is described in Section 3.1.1. $B = 10\,000$.

	$\alpha(\%)$	β_{MKT}	β_{SMB}	β_{HML}	β_{MOM}
Panel A: Asia.					
Fraction fee	-4.42***	0.95***	0.09**	0.18***	0.03
	[-7.08, -1.76]	[0.91, 0.99]	[0.01, 0.17]	[0.1, 0.26]	[-0.03, 0.09]
Incentive fee	-	-	-	-	-
		-	-	-	-
Panel B: Emerging.					
Fraction fee	-1.28	0.99***	-0.04	0.11**	-0.02
	[-3.35, 0.88]	[0.96, 1.02]	[-0.13, 0.06]	[0.03, 0.2]	[-0.08, 0.04]
Incentive fee	1.38	1.04***	0.11	-0.04	-0.03
	[-2.79, 5.7]	[0.98, 1.11]	[-0.11, 0.32]	[-0.27, 0.19]	[-0.16, 0.11]
Panel C: Europe.					
Fraction fee	-1.13	1.01***	0.25***	-0.14***	-0.02
	[-2.43, 0.21]	[0.99, 1.03]	[0.19, 0.3]	[-0.19, -0.1]	[-0.05, 0.01]
Incentive fee	1.27	1.17***	0.34***	-0.31***	-0.11***
	[-2.16, 4.77]	[1.11, 1.23]	[0.2, 0.49]	[-0.43, -0.19]	[-0.18, -0.03]
Panel D: Global.					
Fraction fee	-2.76***	1.02***	0.11***	0.0	0.02
	[-3.98, -1.47]	[1.0, 1.05]	[0.05, 0.17]	[-0.04, 0.05]	[-0.01, 0.05]
Incentive fee	1.49	1.25***	0.47***	0.04	-0.02
	[-1.44, 4.58]	[1.19, 1.31]	[0.32, 0.61]	[-0.06, 0.15]	[-0.09, 0.04]
Panel E: Japan.					
Fraction fee	-2.36**	0.94***	0.08**	-0.16***	0.05**
	[-4.37, -0.32]	[0.9, 0.97]	[0.01, 0.15]	[-0.21, -0.1]	[0.0, 0.09]
Incentive fee	-	-	-	-	-
		-	-	-	-
Panel F: Norway.					
Fraction fee	-0.98	0.96***	0.12***	-0.04***	0.01
	[-2.6, 0.66]	[0.93, 0.98]	[0.08, 0.16]	[-0.07, -0.01]	[-0.02, 0.04]
Incentive fee	-5.3	0.99***	0.1	-0.02	-0.06
	[-16.21, 5.25]	[0.77, 1.2]	[-0.24, 0.41]	[-0.3, 0.27]	[-0.37, 0.24]
Panel G: USA.					
Fraction fee	-2.23***	0.98***	0.16***	-0.06***	-0.03**
	[-3.55, -0.98]	[0.95, 1.0]	[0.12, 0.2]	[-0.09, -0.03]	[-0.05, -0.01]
Incentive fee	-	-	-	-	-
		-	-	-	-

Note: *p<0.1; **p<0.05; ***p<0.01

Table B.3: Results from the boostrap of Kosowski et al. (2006).

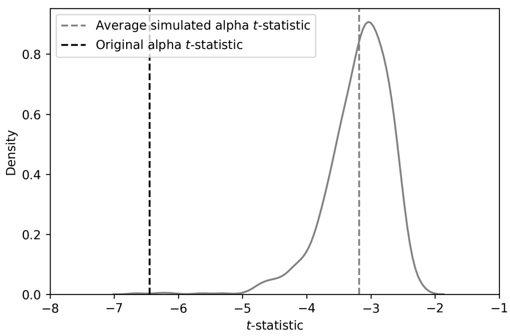
The table shows original regression values and simulated means displayed as *Act* and *Sim* for the alphas and their *t*-statistic. The leftmost columns list the five best and worst values for alphas and their *t*-statistics, as well as deciles. The top performer in terms of alpha is not necessarily the same fund as the top performer in terms of the *t*-statistic. The fourth and seventh columns show the *win-rate* of the original regression values to the distribution of 1000 simulations of alphas and *t*-statistics. In the upper (lower) part of the table, high (low) win rates translate to low p-values. Monthly alphas are annualized.

	Alpha			*t*-statistic		
	Act(%)	*Sim*(%)	%<*Act*	*Act*	*Sim*	%<*Act*
Best	9.23	17.04	8.4	2.82	3.04	32.3
2	8.48	10.45	17.5	2.76	2.72	60.8
3	8.35	9.07	35.2	2.73	2.54	79.3
4	8.23	8.29	52.8	2.34	2.42	37.4
5	7.9	7.71	62.1	1.96	2.33	0.9
90 %	1.78	3.04	0.0	0.76	1.28	0.0
80 %	0.35	1.78	0.0	0.14	0.84	0.0
70 %	-0.77	1.05	0.0	-0.38	0.53	0.0
60 %	-1.36	0.49	0.0	-0.69	0.25	0.0
50 %	-2.06	0.0	0.0	-0.96	0.0	0.0
40 %	-2.6	-0.49	0.0	-1.28	-0.25	0.0
30 %	-3.47	-1.05	0.0	-1.65	-0.53	0.0
20 %	-4.39	-1.77	0.0	-2.01	-0.85	0.0
10 %	-5.98	-3.06	0.0	-2.49	-1.3	0.0
5	-8.77	-7.33	6.6	-4.08	-2.28	0.0
4	-8.9	-7.82	13.7	-4.09	-2.37	0.0
3	-9.28	-8.44	23.4	-4.1	-2.47	0.0
2	-9.46	-9.34	40.7	-4.63	-2.6	0.0
Worst	-11.34	-11.01	36.6	-5.4	-2.8	0.0

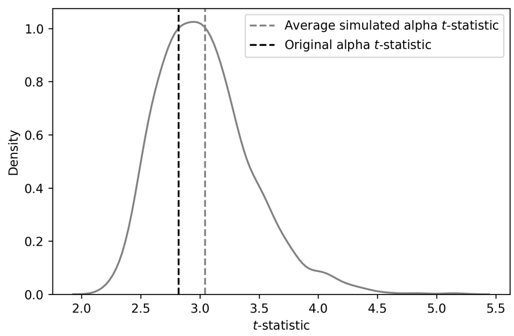
Figure B.2: Parametric t-statistics compared to equally ranked simulated distributions.

The figures show actual t-statistics plotted against the empirical distributions of equally ranked t-statistics by the bootstrapping procedure of Kosowski et al. (2006). An actual t-statistic being far to the right (left) in its distribution indicates skill (inability).

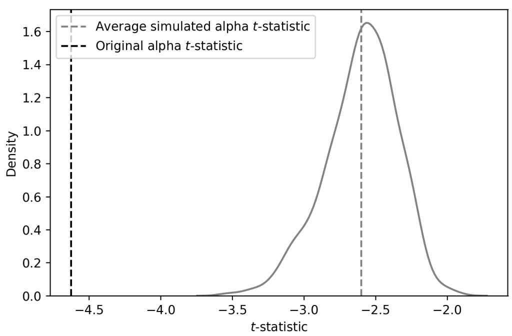
(a) Worst fund performance.



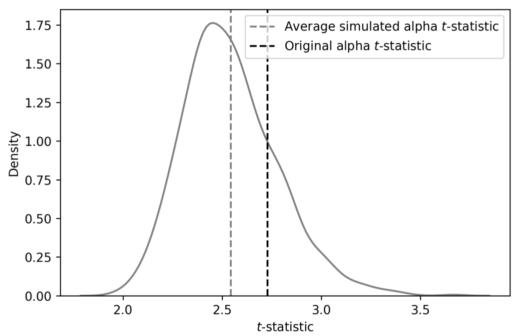
(b) Best fund performance.



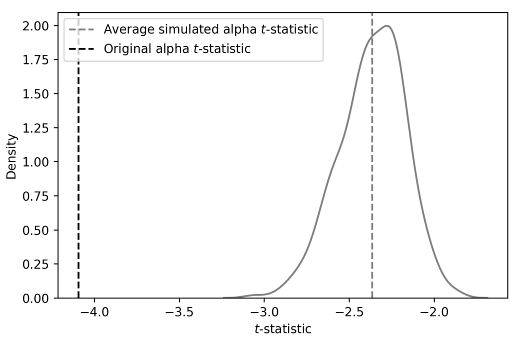
(c) Third worst fund performance.



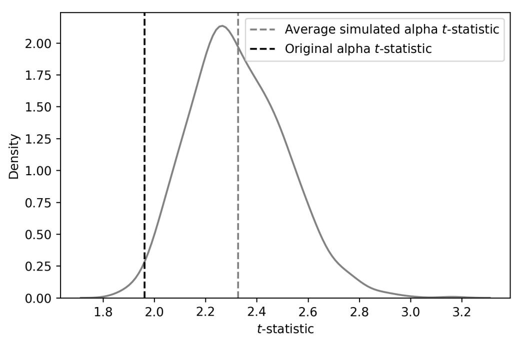
(d) Third best fund performance.



(e) Fifth worst fund performance.



(f) Fifth best fund performance.



C Risk

Table C.1: Intra-period change in tracking error for small and young funds.

The table shows an overview of the change in the tracking error between the first nine and final three months of each fund calendar year for young and small fraction fee and incentive fee funds in Panels A and B, respectively. Young funds are funds with prior existence of fewer than seven fund-years. Small funds have a lower average TNA than USD 100 million. Funds are ranked on their absolute differential return to their Technical Indicator in the first nine months. The rightmost column shows the average absolute change in tracking error. The fourth row of each sub-panel displays the difference between the top and quintiles.

	First subperiod		Second subperiod		
	$R_i - R_m(\%)$	$TE(\%)$	$R_i - R_m(\%)$	$TE(\%)$	$\Delta TE(\%)$
Panel A: Fraction fee.					
Top 20%	10.31	2.91	1.81	3.35	0.44
All	-1.4	2.33	-0.43	2.6	0.27
Bottom 20%	-12.99	3.53	-2.06	3.07	-0.46
Difference	-23.3***	0.62	-3.87***	-0.28	-0.9
Panel B: Incentive fee.					
Top 20%	9.88	2.62	2.47	1.81	-0.81
All	-1.6	1.94	-0.77	1.91	-0.02
Bottom 20%	-10.23	2.57	-2.98	2.49	-0.08
Difference	-20.11**	-0.05	-5.45**	0.69	0.73

Note: *p<0.1; **p<0.05; ***p<0.01

D Mapping funds on risk features

Table D.1: Descriptive statistics for the clusters obtained from the 8x8 SOM.

The tables show cluster statistics for each of the clusters obtained. The features used as input for the SOM are to the left of the dotted vertical line. $M = 64$, $\eta_0 = 0.01$, $\sigma_0 = 4$, $I = 50\,000$ and $k = 7$.

	$t_{\beta_{MKT}}$	$t_{\beta_{SMB}}$	$t_{\beta_{HML}}$	$t_{\beta_{MOM}}$	R^2	σ_r	γ	TE_{TI}	σ_{TE}	$\bar{\alpha}(\%)$	TNA	TE_{FM}	Flow	β_{MKT}	β_{SMB}	β_{HML}	β_{MOM}	N
Panel A: K-means clustering.																		
C0	-0.07	-0.17	0.69	-0.45	0.89	0.05	0.27	1.93	0.59	-2.47	433.18	1.88	-0.27	1.01	-0.0	0.15	-0.07	4/75
C1	-5.08	0.75	-0.19	-0.34	0.59	0.05	0.26	3.78	0.96	-2.54	34.66	4.17	-0.15	0.61	0.19	0.0	-0.07	0/21
C2	0.36	2.33	-0.84	0.24	0.88	0.06	0.39	2.3	0.58	-1.84	251.35	3.06	-0.32	1.04	0.48	-0.17	0.05	3/71
C3	-0.2	-0.14	0.33	-0.28	0.71	0.07	0.54	4.33	3.02	-3.94	421.87	3.41	-0.49	0.97	-0.08	0.22	-0.13	0/29
C4	0.17	-0.41	-0.6	0.55	0.92	0.05	0.26	1.54	0.42	-2.04	212.97	1.79	0.45	1.02	-0.05	-0.08	0.09	0/105
C5	-1.38	0.23	0.42	0.36	0.69	0.04	0.21	2.6	0.78	-4.09	62.43	2.8	-0.3	0.77	0.09	0.13	0.09	2/65
C6	0.29	0.25	-0.17	-0.07	0.75	0.07	0.56	2.25	0.77	-1.83	334.75	2.23	-1.03	1.05	0.1	-0.06	-0.03	4/56
Panel B: Hierarchical clustering.																		
D0	-0.13	0.32	0.68	-0.78	0.88	0.04	0.23	2.33	0.77	-2.74	215.68	2.13	-0.63	0.99	0.11	0.12	-0.1	2/45
D1	-2.96	0.41	0.21	-0.03	0.63	0.04	0.21	3.37	0.96	-3.94	63.48	3.66	-0.24	0.68	0.12	0.08	-0.01	0/58
D2	0.78	2.95	-1.43	0.56	0.9	0.06	0.37	2.11	0.53	-1.8	293.99	2.58	-0.26	1.08	0.61	-0.31	0.11	2/33
D3	-0.01	0.08	0.09	-0.21	0.69	0.08	0.64	3.21	1.95	-2.84	345.43	2.75	-0.54	1.0	0.04	0.1	-0.11	2/52
D4	-0.14	-0.1	0.24	0.72	0.88	0.04	0.21	1.72	0.44	-2.41	131.43	1.85	0.1	0.98	-0.0	0.06	0.14	1/81
D5	-0.09	-0.6	-1.79	0.22	0.93	0.04	0.25	1.56	0.4	-1.55	349.88	1.96	0.74	1.0	-0.09	-0.27	0.02	0/36
D6	0.16	0.35	-0.02	0.07	0.85	0.06	0.4	2.03	0.62	-2.2	340.94	2.37	-0.43	1.01	0.09	0.02	0.01	6/117

Table D.2: Descriptive statistics for the clusters obtained from the 12x12 SOM.

The tables show cluster statistics for each of the clusters obtained. The features used as input for the SOM are to the left of the dotted vertical line. $M = 144$, $\eta_0 = 0.03$, $\sigma_0 = 6$, $I = 50\,000$ and $k = 7$.

	$t_{\beta_{MKT}}$	$t_{\beta_{SMB}}$	$t_{\beta_{HML}}$	$t_{\beta_{MOM}}$	R^2	σ_r	γ	TE_{TI}	σ_{TE}	$\bar{\alpha}(\%)$	TNA	TE_{FM}	Flow	β_{MKT}	β_{SMB}	β_{HML}	β_{MOM}	N
Panel A: K-means clustering.																		
E0	-0.19	0.15	0.81	-0.64	0.87	0.05	0.26	2.3	0.65	-2.37	229.74	2.25	-0.39	0.98	0.06	0.15	-0.09	3/63
E1	-3.33	0.55	0.25	0.06	0.61	0.04	0.2	3.0	0.85	-3.7	51.08	3.54	-0.26	0.64	0.16	0.09	0.03	0/50
E2	0.43	2.41	-1.05	0.38	0.88	0.06	0.39	2.32	0.57	-2.0	250.52	2.99	-0.55	1.05	0.5	-0.21	0.07	3/57
E3	-0.47	-0.04	0.24	-0.24	0.69	0.07	0.55	4.41	2.83	-4.6	372.79	3.65	-0.48	0.94	-0.02	0.21	-0.12	0/33
E4	0.02	-0.61	-1.06	0.16	0.92	0.05	0.26	1.61	0.51	-2.13	256.75	1.95	0.54	1.0	-0.08	-0.17	0.02	0/76
E5	-0.1	0.15	0.35	0.98	0.86	0.04	0.22	1.71	0.4	-2.54	136.58	1.75	0.26	0.98	0.03	0.09	0.19	1/59
E6	0.31	0.19	0.01	-0.01	0.78	0.07	0.5	2.07	0.71	-1.84	416.32	2.06	-0.76	1.04	0.07	0.01	-0.02	6/84
Panel B: Hierarchical clustering.																		
F0	-0.13	0.46	0.45	-0.63	0.88	0.05	0.26	2.28	0.63	-2.31	208.66	2.37	-0.57	0.99	0.13	0.07	-0.08	4/71
F1	-2.89	0.54	0.24	0.18	0.64	0.04	0.2	2.91	0.83	-3.48	48.05	3.42	-0.2	0.68	0.16	0.1	0.06	0/61
F2	0.97	3.35	-1.56	0.7	0.9	0.06	0.4	2.35	0.64	-2.19	234.99	3.01	-0.62	1.1	0.68	-0.32	0.13	2/24
F3	-0.54	-0.19	0.3	-0.3	0.7	0.07	0.54	4.51	3.12	-4.24	424.41	3.6	-0.63	0.92	-0.09	0.25	-0.15	0/28
F4	-0.05	-0.14	-0.6	0.61	0.9	0.04	0.23	1.67	0.41	-2.11	226.47	1.88	0.32	1.0	-0.0	-0.08	0.1	1/105
F5	0.35	-0.28	0.23	0.03	0.9	0.06	0.37	1.66	0.56	-2.45	449.12	1.96	0.27	1.04	-0.04	0.08	-0.01	2/59
F6	0.19	0.54	-0.21	0.07	0.75	0.07	0.53	2.36	0.77	-2.06	268.08	2.44	-0.79	1.02	0.16	-0.06	-0.0	4/74

Bibliography

Admati, A. and Pfleiderer, P. (1997). Does it all add up? Benchmarks and the compensation of active portfolio managers. *The Journal of Business*, 70(3):323–350.

Agarwal, V. and Naïk, N. Y. (2000). Performance evaluation of hedge funds with option-based and buy-and-hold strategies. LBS working paper.

Baghai-Wadji, R., El-Berry, R., Klocker, S., and Schwaiger, M. (2006). Changing investment styles: Style creep and style gaming in the hedge fund industry. *Intelligent systems in accounting, finance and management*, 25(1):96–143.

Basak, S., Pavlova, A., and Sharpiro, A. (2007). Optimal asset allocation and risk shifting in money management. *The Review of Financial Studies*, 20(5):1583–1621.

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

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

Blake, E., Elton, M., and Gruber, C. (1999). Common factors in active and passive portfolios. *European Finance Review*, 3(1):53–78.

Brown, S. J. (1992). Survivorship bias in performance studies. *Review of financial studies*, 5(4):553–80.

Brown, S. J. and Goetzmann, W. N. (1997). Mutual fund styles. *Journal of Financial Economics*, 43(3):373–399.

Børsheim, E. and Eilertsen, B. (2016). Mutual fund performance in Norway and its effect on investor capital allocation. Master's thesis, University of Stavanger.

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

Carpenter, J. (1989). Does option compensation increase managerial risk apetite? *The Journal of Finance*, 55(5):2311–2331.

Chen, H. L. and Pennachi, G. G. (2005). Does prior performance affect mutual fund choice of risk? Theory and further empirical evidence. *The Journal of Financial and Quantitative Analysis*, 44(4):745–775.

Chen, J., Su, M., Cao, R., Hsu, S., and Lu, J. (2017). A self organizing map optimization based image recognition and processing model for bridge crack inspection. *Automation in Construction*, 73(1):58–66.

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

Cuoco, D. and Kaniel, R. (2010). Equilibrium prices in the presence of delegated portfolio management. *The Journal of Financial Economics*, 101(2):264–296.

Das, S. R. and Sundaram, R. K. (2002). Fee speech: Signaling, risk-sharing, and the impact of fee structures on investor welfare. *The Review of Financial Studies*, 15(5):1465–1497.

Davies, D. and Bouldin, D. (1979). A cluster separation measure. *IEEE Transactions on*

Pattern Analysis and Machine Intelligence, 31(1):224–227.

Deboeck, G. (1998). Financial applications of self-organizing maps. *Neural Network World*, 8(2):213–241.

Deetz, M., Poddig, T., and Varmaz, A. (2009). Klassifizierung von hedge-fonds durch das k-means clustering von self-organizing maps: Eine renditebasierte analyse zur selbsteinstufungsgüte und stiländerungsproblematik. *x*, x(x):x.

Elton, E. J., Gruber, M. J., and Blake, C. R. (2003). Incentive fees and mutual funds. *The Journal of Finance*, 58(2):779–804.

Ervik, S. N. and Qvale, L. I. (2017). Resultatavhengig forvaltningsgodtgjørelse. Master's thesis, Norwegian School of Economics.

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

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

Fama, E. F. and French, K. R. (2012). Size, value, and momentum in international stock returns. *The Journal of Financial Economics*, 105(3):457–472.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *The Journal of Banking and Finance*, 55(3):117–129.

French, K. R. (2020). *Homepage Kenneth R. French*. `https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/`. Accessed at 27/01/2020.

Fung, W. and Hsieh, D. A. (2001). The risk in hedge fund strategies: Theory and evidence from trend followers. *The Review of Financial Studies*, 12(2):313–341.

Gallefoss, K., Hansen, H., Haukaas, E., and Molnár, P. (2015). What daily data can tell us about mutual funds: Evidence from Norway. *The Journal of Banking and Finance*, 55(3):117–129.

Golec, J. and Starks, L. (2004). Performance fee contract contract change and mutual fund risk. *The Journal of Financial Economics*, 73(1):93–118.

Grimeland, S. J. (2018). Anomalies and the five-factor model in the Norwegian stock market. Master's thesis, Norwegian University of Science and Technology.

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

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

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

Gruber, M. J. (2001). Identifying the risk structure of mutual fund returns. *European Financial Management*, 7(2):147–159.

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

Huang, F., Yin, K., Huang, J., Gui, L., and Wang, P. (2017). Landslide susceptibility mapping based on the self-organizing map network and extreme learning machine. *Engineering Geology*, 223(1):11–22.

Huang, J., Sialm, C., and Zhang, H. (2011). Risk shifting and mutual fund performance. *The Review of Financial Studies*, 24(8):2575–2616.

Ibert, M. (2018). What do mutual fund managers' private portfolios tell us about their skills? Paris December 2018 Finance Meeting EUROFIDAI - AFFI.

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.

Jensen, M. C. and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4):305–360.

Kaufman, L. and Roussew, P. J. (1990). *Finding groups in data - an introduction to cluster analysis*. John Wiley Sons.

Kempf, A. and Ruenzi, S. (2007). Tournaments in mutual-fund families. *The Review of Financial Studies*, 21(2):1013–1036.

Kempf, A., Ruenzi, S., and Thiele, T. (2009). Employment risk, compensation incentives, and managerial risk taking: Evidence from the mutual fund industry. *The Journal of Financial Economics*, 91(1):92–108.

Kiang, M. Y. (2001). Extending the kohonen self-organizing map networks for clustering analysis. *Computational Statistics and Data Analysis*, 38(2):161–180.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480.

Kohonen, T. (2014). *MATLAB Implementations and Applications of the Self-Organizing Map*. Unigrafia Bookstore Helsinki, Unigrafia Oy, Helsinki, Finland.

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.

Li, W. and Tiwari, A. (2009). Incentive contracts in delegated portfolio management. *The Review of Financial Studies*, 22(11):4681–4714.

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

Lisi, F. and Otranto, E. (2010). Clustering mutual funds by return and risk levels. *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, 3(1):183–191.

Lokesh, S., Priyan, M., and Ramya, D. (2019). An automatic tamil speech recognition system by using bidirectional recurrent neural network with self-organizing map. *Neural Computing and Applications*, 31(5):1521–1531.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, pages 281–297, Berkeley, Calif. University of California Press.

Maillet, B. and Rousset, P. (2003). Classifying hedge funds with kohonen maps: A first attempt. *Connectionist Approaches in Economics and Management Sciences*, 6(2):233–259.

Mantegna, R. N. (1998). Hierarchical structure in financial markets. *The European Physical*

Journal B - Condensed Matter and Complex Systems, 11(1):193–197.

Marate, A. and Shawky, H. A. (1999). Categorizing mutual funds using clusters. *Advances in Quantitative Analysis of Finance and Accounting*, 7(1):199–204.

Massa, M. and Patgiri, R. (2009). Incentives and mutual fund performance: Higher performance or just higher risk taking? *The Review of Financial Studies*, 22(5):1777–1815.

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

Ou-Yang, H. (2003). Optimal contracts in a continuous-time delegated portfolio management problem. *The Review of Financial Studies*, 16(1):173–208.

Palomino, F. and Prat, A. (2003). Risk taking and optimal contracts for money managers. *The RAND Journal of Economics*, 34(1):113–137.

Pattarin, F., Paterlini, S., and Minerva, T. (2004). Clustering financial time series: an application to mutual funds style analysis. *Computational Statistics Data Analysis*, 47(2):353–372.

Pölzlbauer, G. (2004). Survey and comparison of quality measures for self-organizing maps. *Proceedings of the Fifth Workshop on Data Analysis (WDA04)*.

Rajanen (Marghescu), D. (2006). Evaluating the effectiveness of projection techniques in visual data mining.

Schmidt, C. R., Rey, S. J., and Skupin, A. (2011). Effects of irregular topology in spherical self-organizing maps. *International Regional Science Reivew*, 34(2):215–229.

Sensoy, B. A. (2009). Performance evaluation and self-designated benchmark indexes in the mutual fund industry. *Journal of Financial Economics*, 92(1):25–39.

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

Sharpe, W. F. (1992). Asset allocation: Management style and performance measurement. *The Journal of Portfolio Management*, 25(2):7–19.

Solidoro, C., Bandelj, V., and Barbieri, P. (2007). Understanding dynamic of biogeochemical properties in the northern adriatic sea by using self-organizing maps and k-means clustering. *Journal of Geophysical Research*, 112(7):197–215.

Starks, L. T. (1987). Performance incentive fees: An agency theoretic approach. *The Journal of Financial and Quantitative Analysis*, 22(1):17–32.

Stoughton, N. M. (1993). Moral hazard and the portfolio management problem. *The Journal of Finance*, 48(5):2009–2028.

Sun, Z., Wang, A., and Zheng, L. (2012). The road less traveled: Strategy distinctiveness and hedge fund performance. *The Review of Financial Studies*, 14(1):157–177.

Sørensen, L. Q. (2009). Mutual fund performance on the Oslo Stock Exchange.

Vatanen, T. (2015). Self-organization and missing values in som and gtm. *Neurocomputing*, 147(1):60–70.

Vesanto, J. and Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions of Neural Networks*, 11(3):586–600.

Wu, s. and Chow, T. (2004). Clustering of the self-organizing map using a clustering validity index based on inter-cluster and intra-cluster density. *Pattern Recognition*, 37(2):175–188.

Zhu, D., Liu, Y., and Sun, B. (2017). Biologically inspired self-organizing map applied to task assignment and path planning of an auv system. *Transactions on Cognitive and Developmental Systems*, 10(2):304–313.

Ødegaard, B. A. (2020). *Homepage of Bernt Arne Ødegaard*. `http://finance.bi.no/~bernt/`. Accessed at 27/01/2020.

NTNU

Norwegian University of
Science and Technology