



**Vaasan yliopisto**  
UNIVERSITY OF VAASA

Vinski Laaksonen

# **Active share and fund performance**

Evidence from Finland

School of Accounting and Finance  
Master's Thesis in Finance  
Finance

Vaasa 2023

---

**UNIVERSITY OF VAASA****School of Accounting and Finance**

**Author:** Vinski Laaksonen  
**Title of the Thesis:** Active share and fund performance: Evidence from Finland  
**Degree:** Master of Science in Economics and Business Administration  
**Programme:** Master's Degree Programme in Finance  
**Supervisor:** Stig Xenomorph  
**Year:** 2023 **Pages:** 64

---

**ABSTRACT:**

This thesis examines the level of active portfolio management within Finnish funds that mainly invest in Finnish equities. As passive funds have generally attracted more flows than their active counterparts, the discussion of the relevance of active investing is topical. In this thesis, the level of active management is observed using a measure called Active share. Active share measures the actual level of active portfolio management by comparing the overlap between portfolio holdings and portfolio benchmark index holdings. In contrast to traditional tracking error, Active share indicates the level of active management more precisely.

As there is no unanimous consensus on whether high Active share funds generate excess returns compared to the benchmark, there is room for further examination, especially in Finnish markets. In this thesis, the impacts of Active share are examined by studying pure data characteristics as well as utilizing a panel regression model. The relation between Active share and fund performance is tested using pure gross returns, risk-adjusted returns and benchmark-adjusted returns. The sample consists of 17 active Finnish mutual funds that mainly invest in Finnish equities during an examination period of 6/2016 – 6/2021. Additionally, the overall level of active management in Finland is studied as well as the presence of closet indexing.

The thesis provides evidence that Finnish funds mainly investing in Finnish equities are considerably passive and closet indexing is overrepresented within the sample. The share of truly active funds, measured by Active share, is minimal in Finland. As the funds are considerably passive, the sample does not give support for testing for the relation between high Active share funds and fund performance. Even though the used data does not provide significant evidence of Active share's ability to predict future returns, it provides a clear benefit of Active share being used in evaluating a suitable level of fees a fund charges for active management. The findings of this thesis suggest a fund investor to choose from the funds with the lowest fees as active management in Finland does not seem to generate reliably greater returns.

---

**KEYWORDS:** Active share, active portfolio management, passive portfolio management, mutual fund, closet indexing

---

**VAASAN YLIOPISTO****Laskentatoimen ja Rahoituksen yksikkö**

<b>Tekijä:</b>	Vinski Laaksonen
<b>Tutkielman nimi:</b>	Active share ja rahaston suorituskyky: Näyttö Suomesta
<b>Tutkinto:</b>	Kauppätieteiden maisteri
<b>Oppiaine:</b>	Rahoituksen maisteriohjelman
<b>Työn ohjaaja:</b>	Stig Xenomorph
<b>Valmistumisvuosi:</b>	2023 <b>Sivumäärä:</b> 64

---

**TIIVISTELMÄ:**

Tämä tutkielma tarkastelee aktiivista salkunhoitoa suomalaisten pääosin suomalaisiin osakkeisiin sijoittavien sijoitusrahastojen keskuudessa. Sijoitusvarallisuuden siirtyessä yhä passiivisempiin rahastovaihtoehtoihin, on aktiivisten sijoitusrahastojen olemassaolon relevanttius ajankohdasta. Tässä tutkielmassa aktiivista salkunhoitoa tutkitaan Active share -mittarilla. Active share mittaa sijoitusportfolion todellisen aktiivisuuden suhteuttamalla sijoitusten päällekkäisyyden portfolion sekä vertailuindeksin välillä. Perinteiseen Trackin erroriin verrattuna Active share indikoi portfolion aktiivisuutta täsmällisemmin.

Koska korkean Active sharen yhteyttä parempaan rahaston performanssiin ei ole yksimielisesti voitu todeta, on aihepiirissä tilaa uusille tutkimuksille, etenkin Suomen markkinoilla. Tässä tutkielmassa, Active sharen vaikutuksia tutkitaan tarkastelemalla aineiston ominaispiirteitä sekä regressiomallia hyödyntäen. Active sharen ja rahaston performanssin välistä yhteyttä testataan käyttäen puhdasta bruttotuottoa, riskikorjattua tuottoa sekä indeksikorjattua tuottoa. Aineisto koostuu 17 suomalaisesta rahastosta, jotka sijoittavat pääosin suomalaisiin osakkeisiin tarkastelujaksolla 6/2016 – 6/2021. Lisäksi tutkielmassa tarkastellaan yleistä salkunhoidon aktiivisuuden tasoa sekä kaappi-indeksoinnin esiintymistä tarkastelujoukossa.

Tutkielma osoittaa, että suomalaiset pääosin suomalaisiin osakkeisiin sijoittavat sijoitusrahastot ovat huomattavan passiivisia ja kaappi-indeksointia esiintyy runsaasti. Varsinaisten aktiivisten sijoitusrahastojen osuus on Suomessa minimaalinen Active sharella mitattuna. Aineiston rahastojen ollessa huomattavan passiivisia, aineisto ei tue korkean Active sharen ja rahaston tuoton välisen yhteyden testaamista pidemmälle. Vaikka tutkielman aineisto ei tarjoakaan merkittävää näyttöä Active sharen kyvystä ennustaa rahaston tulevia tuottoja, on Active sharen käytöstä selvää hyötyä arvioitaessa rahaston kulurakenteen sopivuutta. Tutkielman löydökset ohjaavat rahastosijoittajaa valitsemaan sijoituskohteen matalimman kulurakenteen rahastojen joukosta, sillä aktiivinen salkunhoito ei näytä Suomessa takaavan todistetusti parempia tuottoja.

---

**AVAINSANAT:** Active share, aktiivinen salkunhoito, passiivinen salkunhoito, sijoitusrahasto, kaappi-indeksointi

## Contents

1	Introduction	7
1.1	Purpose of the study	7
1.2	Structure of the study	9
2	Portfolio management	10
2.1	Modern portfolio theory	10
2.2	Efficient market hypothesis	11
2.3	Passive portfolio management	12
2.4	Active portfolio management and active portfolio management measures	14
2.4.1	Tracking error	15
2.4.2	Active share	17
2.4.3	Active weight	20
2.4.4	Other measures	21
3	Previous Active share studies	23
3.1	High Active share	23
3.2	High Active share and small fund size	27
3.3	High Active share and high fund competition	29
3.4	High Active share and long fund duration	31
3.5	Other Active share findings	34
4	Data and methodology	37
4.1	Data and limitations	37
4.2	Measures of portfolio performance	39
4.3	Methodology	40
5	Empirical findings	43
5.1	Sample characteristics	43
5.2	Active share	49
5.3	Fund performance	52
6	Conclusions	57
	References	60

Appendices	64
Appendix 1. List of sample funds	64

## Figures

Figure 1. Types of active and passive management (Cremers & Petajisto, 2009).....	24
Figure 2. Evolution of Active share (%). .....	44
Figure 3. Return (semiannual) and lagged Active share (%). .....	45
Figure 4. Sharpe ratio (semiannual) and lagged Active share (%). .....	45
Figure 5. Information ratio (semiannual) and lagged Active share (%). .....	46
Figure 6. Active share (%) and Tracking error (annualized). .....	47
Figure 7. Active share (%) and Total net asset value (M€). .....	48
Figure 8. Average Active share (%) and average Total expense ratio (%). .....	49

## Tables

Table 1. Calculation of Active share.	18
Table 2. Descriptive statistics of the sample variables.	43
Table 3. Determinants of Active share.	50
Table 4. Determinants of fund performance.	53

# 1 Introduction

Even though passive investing has become more popular in time with low-cost index funds and wide-ranging ETFs, active investing coexists on the side. The debate on the rationality of active investing has been going on for decades. The supporters of passive investing suggest that active, benchmark-deviant investing is a waste of resources. Instead, an investor or fund manager should replicate the index and settle on corresponding index returns. On the contrary, the supporters of active investing do not want to align themselves with index-like returns, they seek excess returns. To generate excess returns, the investment portfolio must diverge from the benchmark index by making active investment decisions. In this thesis, active investing and active portfolio management are considered as a divergence in holdings compared to the benchmark index, not necessarily as frequent trading.

In this thesis, I have chosen the perspective of active investing and to be more specific, I observe active investing through an active portfolio management measure called Active share. Active share is a rather new and coherent measure that reveals the true level of active management of a certain investment portfolio, typically an equity mutual fund. Active share measures the level of active management by comparing portfolio holdings to the benchmark index holdings and expresses the level of active management in percentages from 0 to 100.

## 1.1 Purpose of the study

I find multiple justifications for studying Active share deeper. Firstly, Active share is a rather new measure for active management and therefore not thoroughly studied. Secondly, most of the Active share studies are conducted with US data and there is no clear consensus in Finnish fund markets. The insufficient amount of Finnish evidence is most likely due to the absence of a collective database for Active shares. Therefore, Active shares need to be calculated manually and fund holdings data for the calculation is not

widely available. Overall, the purpose is to find evidence that supports active investing in contrast to passive index investing, even though the general sentiment is shifting away from active funds. Traditional active fund management measures such as Tracking error can not give a straightforward indication of the true level of active fund management and therefore, I find Active share as a more relevant measure for examining the performance of active funds.

Based on previous studies, I examine the effects of Active share within Finnish funds that mainly invest in Finnish equities. Active share is being examined further concentrating on the following three research questions:

*1. Are Finnish active funds truly active?*

An important viewpoint is to provide the characteristics of active Finnish funds. The sample time period is long enough to see potential shifts towards more passive or more active fund management. Reflecting on the historical Active shares in the US that have decreased significantly, since 1980, I expect a similar downward trend within Finnish funds. As the Finnish market is considerably smaller than the US equivalent, it may cause Active shares to be considerably low. The information of the true level of active fund management is valuable for an investor to avoid paying extra for a closet index fund. A closet index fund often claims to be active but in reality, overlaps largely with the benchmark index.

*2. Can Active share predict fund performance?*

The very first Active share studies suggest that Active share alone is useful for evaluating future fund performance for the funds with the highest Active shares. More recent studies however state that a high Active share alone is not enough to predict future fund performance but is useful jointly with another factor such as long fund duration or a highly competitive fund market. Although Active share is not comprehensively studied

in Finland, I stick to the original viewpoint and examine Active share mainly by itself in relation to fund performance.

### *3. Is it worth paying higher fees for active fund management in Finland?*

As much as a profit-seeking investor wants to maximize their returns, they similarly want to minimize their expenses. Generally, active funds charge higher fees from active portfolio management than their passive counterparts. However, if the active funds are not truly active, do they charge unjustified fees for semi-active management that generally yields index-like returns? Whether the fund fees are justified, I examine the relation between fund fees and fund returns as well as fund fees and the level of active management.

## **1.2 Structure of the study**

The thesis is divided into two main sections: theory in form of a literature review and empirical research. In the theory part, I present remarkable background theories behind portfolio management and investment markets such as modern portfolio theory and efficient market hypothesis. The theory proceeds to the differences between active and passive investing complemented by relevant measures of active portfolio management. The theory part ends with a summary of previous Active share studies and findings.

In the research part, I first introduce the sample and empirical methods being used. The methods and statistical models are followed by empirical findings. In the very last section of the thesis, the conclusion, I present the most important findings and try to provide answers to the research questions introduced earlier.

## 2 Portfolio management

In this chapter, I introduce some of the most significant previous studies and theories related to portfolio management. I begin with Harry Markowitz's modern portfolio theory which is considered the basis of portfolio management. I continue with Eugene Fama's efficient market hypothesis which is a primary framework for stock markets. Passive portfolio management is discussed after that as it is closely related to the assumptions of the efficient market hypothesis. In contrast to passive portfolio management, I continue with active portfolio management which is the principal theme of the thesis. Active portfolio management is followed by active management measures, such as Active share and tracking error.

### 2.1 Modern portfolio theory

Modern portfolio theory or mean-variance analysis is a Nobel-rewarded investment theory and a primary framework for assembling an optimal investment portfolio, originally developed by Harry Markowitz (1952). The theory aims to maximize the expected return at a certain level of risk. Modern portfolio theory assumes that investors chase low risk and high return and that the investors make rational investment decisions when they have all the existing information within their reach. According to the theory, the idea behind constructing a portfolio is either to try to maximize the expected portfolio returns for a given level of risk or to minimize the risk for a given level of expected return. Portfolios formulated following these two objectives are called efficient portfolios. Efficient portfolios construct an efficient frontier, a capital allocation line where expected returns are at their highest point with a certain level of risk. The risk in Markowitz's model is measured by the variance of the returns.

According to Markowitz's (1952) Modern portfolio theory, there is always a diversified portfolio that is preferable to any non-diversified portfolio. Markowitz's model assumes that a rational investor always chooses the least risky portfolio out of several portfolios

with identical expected returns. Modern portfolio theory suggests that assets of a portfolio should be selected as a well-diversified combination instead of concentrating on selecting individual assets by their individual characteristics. According to Modern portfolio theory, the investor needs to pay attention to the movements of the separate assets of the portfolio during different market scenarios and to consider the correlations between portfolio assets to maximize gains.

## **2.2 Efficient market hypothesis**

The Efficient market hypothesis assumes that the prices of securities in the security markets reflect all the available information (Fama, 1970). On the grounds of the Efficient market hypothesis, it is impossible to beat the market regularly as the prices of securities immediately adjust to newly released information. Under efficient market hypothesis assumptions, the investor cannot benefit from security mispricing as mispricing does not exist in the money markets. An investor can achieve greater returns than the market but exclusively by taking more risk. Risk-adjusted excess returns are unreachable under the efficient market hypothesis. Fama (1970) has divided market efficiency into three forms by the level of available information. The three forms of market efficiency are weak, semi-strong and strong forms which are explained in the following three paragraphs.

The weak form of market efficiency denotes that current security prices reflect all past information such as past trading volumes and past security prices. Under the weak form of market efficiency, predicting future security prices with past information is impossible. The condition above originates from Fama's theory which states that securities have no memory and therefore, future security prices follow a random walk. Technical analysis is not a relevant investment strategy under weak form circumstances as an investor cannot gain excess returns using past price information.

The semi-strong form of market efficiency denotes that current security prices reflect all the past and present information. Under the semi-strong form of market efficiency, all

the public information, such as financial statements and earnings forecasts, are already included in security prices and fundamental analysis is therefore needless.

The strong form of market efficiency denotes that current security prices reflect all the past, present and future (nonpublic) information. Under the strong form of market efficiency, in addition to the previous two forms, gaining excess returns with insider information is impossible as insider information is already included in the prices. The possibility of gaining excess returns under the strong form of market efficiency is fundamentally nonexistent.

The assumptions of Fama's (1970) efficient market hypothesis guide rational investors towards passive investing and support passive portfolio management. Fama's hypothesis denies the benefits of active fund management as active fund management cannot systematically lead to any excess returns under efficient market conditions. Therefore, active management cuts the returns by the number of portfolio costs and fees suggesting that active portfolio management is a waste of the portfolio managers' resources. Only passive portfolio management is rational under these circumstances. The efficient market hypothesis has received a lot of criticism because of the rather unrealistic assumptions of the model even though it has reached a considerable position in financial literature.

### **2.3 Passive portfolio management**

Passive portfolio management or passive investing is defined as a style of portfolio management where the portfolio holdings are similar to market index holdings (e.g. Garcia, Guijarro & Moya, 2013). In other words, the portfolio follows the movements of a certain benchmark index. Tracking a benchmark index is also called indexing or index investing. The aim of passive portfolio management is not to try to outperform the market but to yield index-like returns. The opposite of passive portfolio management is active portfolio management, which is considered in the following subchapter.

Passive portfolio management is closely related to Fama's (1970) efficient market hypothesis. As the theory assumes that the stock market is efficient and there is no mispricing of stocks, and thus individual stock picking is not profitable. Employing the Efficient market hypothesis, it is preferred to invest in a market index, by investing in an index fund, for instance.

The costs of passive portfolio management are typically lower than the costs of active portfolio management (e.g. Sharpe, 1991). Sharpe (1991) states that the costs of passive management are lower as the active fund manager trades more often and needs more research on their investment decisions. Simultaneously, the passive fund manager does not require any additional research or information as they invest the fund's capital according to the index and can thus charge less for fund management. The lower costs concerning passive investing are in line with Fama's (1970) efficient market hypothesis which suggests that active investing strategies lose to passive strategies by the difference in fees.

Regarding passive funds, index tracking can either be full or partial. When tracking is full, the fund holds equal stocks as the benchmark index and if the proportions of the stocks are completely identical to the benchmark index, the tracking is exact (e.g. Garcia et al., 2013). If the fund only includes a smaller subset of stocks belonging to the benchmark index, the tracking is partial (e.g. Garcia et al., 2013). There are some issues regarding full tracking compared to partial tracking which, for example, Garcia et al. (2013) point out in their research. They mention that full tracking becomes more expensive than partial tracking because of the higher trading costs as a full tracking index fund holds all the stocks belonging to the index, even the ones with very minimal weights in the index. At the same time, a partially tracking index fund holds a certain proportion of the index stocks and thus saves on trading fees. On the other hand, a partially tracking fund manager needs to pay more attention to stock-picking within the index as the fund does not include all the index stocks. Stock picking within an index, in turn, requires the utilization of extra resources compared to a full tracking fund (e.g. Ruiz-Torrubiano & Suarez, 2009).

## **2.4 Active portfolio management and active portfolio management measures**

Although Morningstar's (2021) U.S. fund flow report confirms the increasing popularity of passive investing in form of passive mutual funds, there is still a reason for the existence of active portfolio management. Passive investment products such as index funds offer market-like returns with relatively low costs and minimal effort. At the same time, passive funds that are built around a certain index prevent the investor from gaining any excess returns, alpha, that may be achieved through active fund management.

The fund managers of actively managed funds seek alpha by active investment decisions which refers to differing from the benchmark index. The fund manager may attempt to outperform the benchmark index in two general ways: stock selection or factor timing (or both) (Fama, 1972). The fund manager can deviate from the benchmark index by selecting stocks that are not included in the index or using factor timing, which is also known as tactical asset allocation. Factor timing refers to placing time-varying bets on any systematic risk related to the benchmark index, such as an entire industry or sectors of the economy.

As mentioned earlier, funds that actively try to outperform their indices generally have higher costs than passive funds as the active fund manager charges fund investors for active fund management. The challenge for the active fund manager is to create alpha that remains positive both before and after the fund fees.

French (2008) examines the costs of active investing. He states that before adding costs to the calculation, other investor's gain is another investor's loss, and therefore, after adding trading costs, active investors are playing a negative-sum game. On a larger scale, French does not consider active investing as a harmful matter as active investing tends to improve the efficiency of the stock market. French concludes that active investing is on average not as profitable as passive investing and the main reason for the above is the higher costs of active investing.

### 2.4.1 Tracking error

Tracking error is the traditional measure for active portfolio management and portfolio performance. It indicates the difference between portfolio returns and benchmark index returns. The fluctuations between portfolio returns and benchmark returns are in general measured by standard deviations. Typically, a high tracking error refers to an actively managed portfolio and a low equivalent refers to a passive, index-following portfolio. A high tracking error indicates wide divergence within portfolio holdings in contrast to benchmark holdings.

$$\text{Tracking error} = \sqrt{\frac{\sum_{i=1}^n (r_p - r_b)^2}{n-1}}, \quad (1)$$

where

$r_p$	is the return of the portfolio
$r_b$	is the return of the benchmark
$n$	is the number of return periods

As seen in the formula above, tracking error is the standard deviation of differences between portfolio and benchmark index returns during a certain period. For instance, a tracking error of 5% indicates that the portfolio returns have deviated 5% from the returns of the benchmark index. The portfolio has either outperformed or underperformed the benchmark by 5%.

In terms of tracking error, an active portfolio manager typically aims to create greater returns than the benchmark index and simultaneously keep the tracking error of their portfolio as low as possible. A high tracking error, high portfolio volatility in comparison to the benchmark index, exposes the portfolio to significantly underperform the benchmark index. A high tracking error is typically an indicator of higher risk that the fund manager has taken in order to outperform the benchmark.

As Roll (1992) states, it usually takes time for fund investors to notice the average performance of a fund manager and that is why many investors rely on tracking error when analyzing whether the fund manager can add excess value to the invested capital. Minimizing tracking error is an important matter for investors when selecting a fund and assessing the skill of the fund manager. A low tracking error suggests that the fund's performance is similar to the benchmark index and there should not be any substantially divergent result of the performance of the fund compared to the benchmark index.

Roll (1992) addresses that there are at least two main reasons why low tracking error is meaningful for investors as well as fund managers. The first reason he brings out is that in an ideal situation, the fund manager outperforms the benchmark index every month by a fixed amount of fund fees and expenses. That is a clear indicator of the fund manager adding value to the fund investors and requires zero tracking error. In that case, the investor would have lost the number of fees and expenses of an index fund, if they had decided to invest in such instead of the actively managed alternative. The second reason Roll (1992) brings out is that when reviewing fund manager performance, the observation period is usually too short. An overly short observation period does not give a reliable picture of the actual skill of the fund manager and a new manager could be chosen to replace the manager who underperformed in returns in the short term.

Cremers and Petajisto (2009) use an example of two mutual funds, where the first fund manager is a pure stock picker, who tries to create alpha by deliberately picking stocks within industries and at the same time aims to diversify the holdings versatily across industries. The second mutual fund is a fund with a sector rotation strategy, which means the fund chooses entire sectors and industries that the fund manager assumes to perform well in the medium term. The positions of the second fund are highly diversified within the sectors, so the positions inside the sectors can be considered passive, although the overall portfolio is rather active. The tracking error of the first fund is considerably lower than the tracking error of the second fund. On the grounds of the difference

between tracking errors, the first fund appears less active to alpha-seeking investors. By active management of the funds, based on tracking error, the alpha-seeking investor could decide to invest in the second fund. Nevertheless, tracking error is a slightly misleading indicator of active management as it is in this Cremers and Petajisto's (2009) example. The first fund's tracking error is lower because picking individual stocks allows greater sectoral diversification compared to putting overweight on entire sectors. Most likely the first fund differs from the benchmark holdings more than the second fund, even though tracking error refers to a completely different circumstance. Because of the weaknesses of tracking error, Cremers and Petajisto (2009) develop a new method, Active share, to measure the level of active fund management. Active share is discussed more comprehensively in the next subchapter.

#### 2.4.2 Active share

Active share is a rather new indicator for measuring the level of active management of a portfolio developed by Cremers and Petajisto (2009). Active share indicates the absolute difference between portfolio and benchmark index holdings, that tracking error cannot reveal perfectly. The measure represents the difference in holdings in percentage from 0 to 100, where 0% refers to completely identical portfolio investments with the benchmark index and 100% refers to completely diverse portfolio investments with the benchmark index. Theoretically, Active share is not limited to 100%; it may reach a higher value through short investment positions. Among mutual funds, the percentage limit is considered as 100 as mutual funds rarely take actual short positions.

$$Active\ share = \frac{1}{2} \sum_{i=1}^N |w_{p,i} - w_{i,i}|, \quad (2)$$

where

$N$  is the number of assets  
 $w_{p,i}$  is the weight of holding  $i$  in portfolio

$w_{i,i}$  is the weight of holding  $i$  in benchmark index

As seen above, the formula for calculating Active share is rather simple. The absolute difference of the weight of a certain asset in the portfolio and the weight of the equivalent asset in the benchmark index is divided by two. The absolute difference is divided by two to ensure that Active share takes a value between 0% and 100% (Cremers & Petajisto, 2009). The table below presents a simplified illustrative example of calculating Active share for a notional portfolio.

**Table 1.** Calculation of Active share.

	Weight in portfolio	Weight in index	Absolute difference	Active share
Asset 1	0%	30%	30%	15%
Asset 2	10%	20%	10%	5%
Asset 3	30%	10%	20%	10%
Asset 4	50%	0%	50%	25%
Asset 5	10%	40%	30%	15%
Total	100%	100%	140%	70%

The table above is completely figurative and not purposely based on an existing portfolio or index. The example consists of five assets and represents the weights of these five assets in a sample portfolio and in a sample index. The absolute differences are calculated on the second last column as in the Active share formula. The last column represents the contribution of each of the five assets in forming the eventual percentage number of Active share by dividing the absolute difference by two. The eventual Active share is seen on the bottom cell of the last column, 70%.

In the table above, both the portfolio and the index consist of four assets in total. The sample portfolio does not hold any of Asset 1 while the index holds 30% of the particular asset. In contrast, the portfolio holds a substantial share of Asset 4 with a 50% weight, while the index does not hold Asset 4 at all. For instance, one can easily observe from

the table above, that the theoretical portfolio manager has taken a considerably large active position with Asset 4, which also significantly increases the Active share of the portfolio.

Cremers and Petajisto (2009) suggest that Active share can predict the future performance of a mutual fund. They claim that the most active funds, measured by Active share, outperform their benchmark indices with a significant difference before and after fees and perform well constantly. On the contrary, they suggest that non-index funds with the lowest Active share, lose to their benchmark indices continuously. Index funds are not included in Cremers and Petajisto's (2009) theorem.

Cremers and Petajisto (2009) suggest that funds with an Active share of over 60% are considered truly active and funds with an Active share of less than 60% are not. Funds with an Active share of near-zero (index funds) are not included in the study as they do not principally try to outperform their benchmark indices with active fund management. I also consider the same limit of 60% Active share as a truly active fund limit even though the limit was self-declared by Cremers and Petajisto (2009). The rationale behind the 60% Active share limit is justified as a fund with an Active share of 50% is considered a hybrid between active and passive funds, so the limit needs to be higher than 50%. Applying the limit of 60%, the example portfolio in the table above passes as a truly active portfolio with an Active share of 70%.

One of the main benefits of Active share in contrast to tracking error is that Active share exposes closet indexing (Cremers & Petajisto, 2009; Petajisto, 2013). Closet indexing refers to a scenario when a fund claims to be active but in reality, has a significant overlap with the benchmark index. The true level of active management is important to know for an investor as actively managed funds typically charge more fees than passive equivalents. The investor should not pay any extra for "active" management for a closet index fund that mostly replicates the benchmark index. Cremers and Petajisto (2009) consider funds with an Active share of under 60% as closet indexers. The problem with tracking

error is that it does not necessarily expose closet indexing. For instance, in Cremers and Petajisto's (2009) example that is discussed earlier, the stock picker fund's tracking error is lower than the sector rotator fund's tracking error, even though the stock picker fund most certainly has a higher Active share.

### **2.4.3 Active weight**

Active weight is a similar measure to Active share, developed by Doshi, Elkamhi and Simutin (2015). Similarly to Active share, Active weight measures the difference between portfolio and benchmark index holdings. The main difference, however, is that Active weight does not compare portfolio holdings to an existing benchmark. Instead, Active weight method compares portfolio holdings to an imaginary benchmark that is constructed by value-weighting the existing portfolio holdings, making the Active weight calculation formula identical to the Active share formula.

Active weight has a deviating approach to closet indexing compared to Active share (Doshi et al., 2015). As most of the benchmark indices are value-weighted, Active weight method assumes that value-weighting fund components refers to closet indexing, even if the fund composition deviates from the benchmark composition. Based on these principles, Active weight evaluates the level of active fund management by purely concentrating on the assigned component weights, rather than also considering the selected components that may deviate from the benchmark, which Active share does. However, the calculation of Active weight is slightly more straightforward compared to Active share, only requiring portfolio component weights and component market capitalization values for a given time.

#### 2.4.4 Other measures

This subchapter includes two other measures for active portfolio management, R-squared value and information ratio. These two indicators are not as well-recognized as tracking error and Active share in measuring the level of active management.

R-squared, also known as  $R^2$  or coefficient of determination, statistically measures how changes in one variable explain changes in another variable (Bodie, Kane & Marcus, 2009, p. 268). R-squared is typically calculated from a statistical regression model. Regarding active portfolio management, R-squared may be used to measure how portfolio fluctuations can be explained by fluctuations in the benchmark index. R-squared takes a value between 0 and 1, and it is often represented as a percentage number between 0 and 100. A high number refers to a scenario, where the correlation between variables is high and most of the proxy variable movements can be explained with the movements of another variable. When considering active portfolio management, R-squared typically takes a high value with index funds, whose movements are similar to the benchmark movements. Active portfolios have, in principle, lower R-squared values than passive portfolios. In addition to active management, R-squared is a measure of risk as well. R-squared is not as clear measure for active management as Active share because an active portfolio can theoretically reach an R-squared value of 0 or 1 or anything in between.

Information ratio (IR) is an indicator measuring portfolio performance as well as active management. It represents the portfolio manager's ability to generate excess returns compared to benchmark index returns and considers whether the excess returns are due to the excess risk that has been taken or more due to managerial skill. The mathematical formula of IR includes tracking error as a risk indicator.

$$IR = \frac{r_p - r_b}{\text{Tracking error}}, \quad (3)$$

Information ratio can be defined more as a performance measure than an active management measure because tracking error is used as a risk factor. On the other hand,

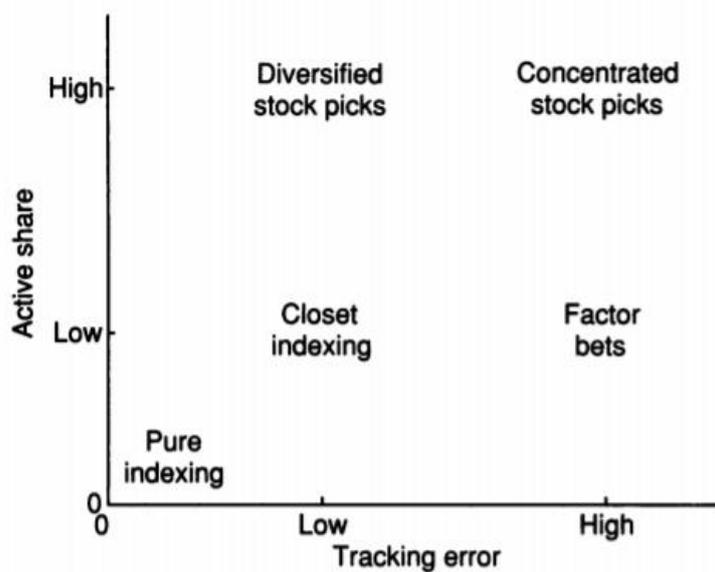
tracking error can be considered an active management measure in calculating IR in which case IR becomes more of an active management -based measure for excess returns. In any case, active management and excess returns go more or less hand in hand and should thus be considered together (Goodwin, 1998). In terms of interpreting the IR value, a higher value refers to successful active management as a high numerator value refers to higher portfolio returns in contrast to the benchmark, and a low denominator value refers to a similar portfolio risk level with the benchmark.

### **3 Previous Active share studies**

This chapter concerns Active share more precisely and consists of results of previous Active share -based studies. As there is no clear consensus on the relationship between Active share and fund performance, this chapter discusses the varying viewpoints on the reliability of Active share. The very first studies suggest Active share alone being capable of estimating future fund returns, while more recent studies suggest Active share may be used jointly with another fund factor in estimating future performance. Some studies, on the other hand, do not find significant evidence of the relation between Active share and fund performance.

#### **3.1 High Active share**

Cremers and Petajisto develop Active share in their 2009 published paper in which they examine whether high Active share funds can accomplish benchmark-adjusted excess returns on a statistically significant level among US mutual funds. They calculate Active shares for domestic US equity mutual funds for the examination period of 1980-2003 and examine the evolution of fund-specific Active shares in addition to examining whether high Active share can yield alpha to fund investors. Cremers and Petajisto (2009) use a benchmark index that has the highest percentage of overlap with a given fund in such a way that every fund's Active share is at its possible lowest at each given time during the examination period.



**Figure 1.** Types of active and passive management (Cremers & Petajisto, 2009).

Cremers and Petajisto (2009) share the funds into groups based on the style of management as seen in the figure above. They use both Active share and tracking error in a matrix that is roughly divided into five categories by the level of active management. The most passive category according to both Active share and tracking error is pure indexing which refers to an index fund.

When both Active share and tracking error are low and not relatively close to zero, the management type is called closet indexing which typically means that the fund is marketed as an active fund, but the fund mostly mimics the benchmark index with only a few active, benchmark-differing positions.

When tracking error is high and Active share is low, the fund is placed in factor bets category, where low Active share is often caused by over- or underweighting certain industries or economic sectors. This kind of investment policy can lead to a lot of deviation (tracking error) in fund returns compared to the benchmark index but not necessarily leads to high Active share.

When Active share is high and tracking error low, the type of management is diversified stock picks. Diversified stock picker fund often has very active stock positions within industries and because of the large diversification, tracking error is not necessarily high.

Concentrated stock picks -category is a combination of the two lastly introduced with both indicators, Active share and tracking error, being high. A concentrated stock picker can either bet on systematic factors such as industries, pick active stock positions within industries or both. Categorizing funds is an important part of the study as actively managed funds are typically considered as one group instead of four separate groups, divided by the type of active management.

Firstly, Cremers and Petajisto (2009) use the categories above to find out other characteristics that are related to each category, for example, fund size, fund fees, money flows and prior returns. Secondly, they study the period from 1980 to 2003 to see the evolution of management types and changes in these types. Thirdly, they examine whether the most active funds' managers have the ability to create excess returns for the fund investors even after fees and expenses.

Cremers and Petajisto (2009) find that over the examination period, the number of pure index funds grew significantly from 1% to 15% of mutual fund assets. They also find that the number of closet index funds grew even more substantially than index funds, from nearly 0% to about 30%. A fund is defined as a closet index fund in the study if Active share is between 20% and 60%. The limit of a truly active fund is 60% as defined by Cremers and Petajisto. The increase in closet indexing in the examination period brings down the average Active share within non-index large-cap mutual funds from about 80% to 60%.

Cremers and Petajisto (2009) find that funds with the highest Active shares outperform their benchmark indices by 1.51-2.40% yearly before fees and expenses and by 1.13-1.15% yearly after fees and expenses. Funds with the lowest Active share underperform

their benchmark indices by 1.42-1.83% yearly after fees and expenses. Cremers and Petajisto do not find any significant relation between fund returns and Tracking error. They suggest that not all four active management types can perform better than the market but those with high Active share can. Their results suggest that the most active stock-picking managers can beat their benchmarks with managerial skill, even after fees and transaction costs. Factor bet fund managers seem to not have the potential to beat their benchmarks and closet indexers evidently underperform after expenses.

Cremers and Petajisto (2009) find a significant continuity in the highest Active share quintile when measured by a benchmark-adjusted Carhart's (1997) four-factor model. Funds in the highest quintile in both Active share and prior-year returns have outperformed their benchmarks by 5.10% per year. The results measured with Carhart's model without benchmark adjustment do not have any significant relation with Active share. The reason behind the differences between benchmark-adjusted and non-benchmark-adjusted returns is that the benchmark indices of the highest Active share funds tend to have a great negative alpha and similarly, the benchmark indices of the lowest Active share funds tend to have a great positive alpha.

In conclusion, Cremers and Petajisto (2009) suggest that Active share is a more relevant measure for active management than traditional tracking error. They conclude that the highest Active share funds outperform their benchmarks before and after fees and high Active share also predicts future performance relative to the benchmark index. The most appealing funds for an investor according to findings of Cremers and Petajisto (2009) are funds with the highest Active share, smallest amount of assets under management and best one-year performance. These types of funds outperform their benchmark indices by 6.5% after fees and expenses.

Petajisto (2013) continues his and Cremers' earlier study on Active share with more recent data and a few other adjustments, such as using the same benchmark index as the fund mentions in their prospectus, instead of using the benchmark that generates the

lowest Active share. Petajisto (2013) extends the sample period by six years so that the new sample period is 1980-2009. Otherwise, Petajisto (2013) uses the same Active share and tracking error -based active management matrix and management types.

The findings of Petajisto (2013) are similar to Cremers and Petajisto (2009). In his study, an average actively managed fund does not beat its benchmark, but the most active stock picker funds can beat their benchmarks by about 1.26% net of fees. Factor bet funds are not able to create any alpha after fees and expenses and closet index funds evidently follow their benchmark indices. The returns of closet index funds sink significantly below zero after fees and expenses. Petajisto (2013) suggests that it is relevant for an investor to pick from the most active (measured by Active share) stock picker funds or the most inexpensive index funds. Only the extremes in the scale of Active share are reasonable investments when seeking the best-performing funds, according to Petajisto's (2013) results. He states that high tracking error is not necessarily a positive sign as it refers to factor bets that have not performed well. Petajisto (2013) concludes that closet index funds are nearly guaranteed to underperform and as mentioned in the earlier study, Active share is very valuable in exposing closet index funds.

Cremers, Fulkerson and Riley (2022a) find similar results to Cremers and Petajisto (2009) and Petajisto (2013) within US separate accounts. Although the characteristics of separate accounts slightly vary from mutual funds, Active share is a relevant measure for them as well. Cremers et al. (2022a) find positive performance persistence among high Active share separate accounts. In their study, a portfolio formed by high Active share funds with robust past performance generates a net annual alpha of 1.38%.

### **3.2 High Active share and small fund size**

Cremers and Petajisto's (2009) research is based on Active share and the performance of active funds. In addition to Active share, they also investigate whether fund size has any relation to fund performance and get results based on their data sample. However,

the size-performance-relation part is rather brief in Cremers and Petajisto's (2009) study. Chen, Hong, Ming and Kubik (2004), for instance, conduct a broader study on fund size and fund performance relation.

Chen et al. (2004) investigate whether big fund size erodes fund returns as a large-scale downside. They assume that a fund that has a significantly large market share, does not benefit from large-scale endlessly, even though some practitioners justify scale advantages with benefits such as lower expense ratios and greater budget for research expenses. The other point of view for large-scale funds is that the performance of significantly large funds decreases as the fund buys and sells securities in large proportions while the market for such proportions is not liquid enough. Another large-scale downside that is being pointed out is the fact that large funds sometimes must take bigger positions than optimal because of the substantial amount of capital the large fund manager or managers needs to allocate. By taking undesirably large positions in certain securities, the fund exposes itself to a greater asset-specific risk than planned.

Chen et al. (2004) use US equity mutual fund data from the period of 1962-1999. They exclude bond, international and specialized sector funds from their sample as well as small funds with less than \$15 million in assets under management. By executing a regression model with monthly gross returns, they find that fund performance is inversely correlated with lagged fund size. By closer inspection, they find that size is correlated more with funds that identify as small-cap funds. Within other fund types, size does not have a significant effect on fund performance. These findings lead Chen et al. (2004) to continue their liquidity hypothesis which assumes that the illiquidity of small-cap stocks has a substantial role in small-cap funds and erodes fund performance when fund size is bigger. They also point out that small funds are often managed by a single manager whereas large funds have various managers and thus have their own difficulties in decision-making policies, for instance.

Chen et al. (2004) are one of the first researchers studying size-performance-relation and come up with a conclusion stating that fund size indeed erodes performance. As mentioned earlier, they suggest that the negative size-performance-correlation is most pronounced for small-cap funds as illiquidity has a strong presence within small-cap funds. They claim liquidity to be one of the explanatory factors for why size erodes performance. Moreover, Chen et al. (2004) consider the number of managers per fund also as an explanatory variable as to why size erodes performance.

The findings of Chen et al. (2004) are much in line with Cremers and Petajisto's (2009) findings even though the latter authors have a little different point of view in their study. Cremers and Petajisto (2009) end up to a similar conclusion by investigating Active share, where large funds (measured by market capitalization) tend to be less active and thus perform poorer. Active share can, therefore, be added as one of the explanatory variables explaining why fund size erodes performance.

### **3.3 High Active share and high fund competition**

Cremers, Ferreira, Matos & Starks (2016) investigate the competition in the mutual fund industry worldwide. Typically, the fund industry is divided into two parts: actively and passively managed funds. Based on previous literature, Cremers et al. (2016) assume that active and passive funds are substitutes for each other and that the increased competition from index funds in the fund market leads active funds to compete with passive equivalents by lowering their management fees or being more active (diverging more from the benchmark index). Their alternative hypothesis is that passive and active fund management do not have any relation and enhanced competition from passive funds does not affect active funds' behavior.

The research of Cremers et al. (2016) is based on fund and ETF data from 32 different countries with a sample period of 2002-2010. Active share has a great role in their research as Active share is used as an active management measure to expose closet

indexing. They find many relations regarding fund competition, Active share and fund expenses that vary in different countries. The research is conducted using various types of fund data such as name, domicile, sponsor, benchmark, monthly returns, total net assets, fees and expenses.

The findings of Cremers et al. (2016) are in line with their hypothesis that competition in the fund industry is advisable for actively managed funds in a way that it either increases the Active share of actively managed funds or decreases the fees of active management. Thus, they accept their assumption that passive and active funds are substitutes for each other. They find evidence that in markets where the pressure from index funds or ETFs is tough, actively managed funds have more divergence from the benchmark index and charge lower fees. On the contrary, in markets where the level of fund competition is low, a larger amount of actively managed funds charge more and are less active. Closet indexing appears more widely in less competitive markets. The main implication of Cremers et al. (2016) is the fact that investing in funds in more competitive markets is more efficient for the investor and they do not necessarily pay unreasonable premiums for active management. Based on the results, the investor can also rely on that closet indexing is not as widespread in competitive fund markets as in less competitive fund markets.

In addition to the results above, Cremers et al. (2016) find that active funds have higher Active shares in more competitive markets and average alphas generated by active management are higher as well. Vice versa, the alphas are lower in markets with less competition. Cremers et al. (2016) find many Active share -related relations such as Active shares are higher for funds with higher tracking error, higher total shareholder cost, younger fund age and affiliations with smaller fund families. Their results are in line with Cremers and Petajisto's (2009) and Petajisto's (2013) earlier studies in the US fund market that suggest that Active share can be used to predict future fund performance. The findings of Cremers et al. (2016) suggesting that Active share predicts future fund performance are statistically and economically significant, across the world. They also

replicate the same results as Cremers and Petajisto (2009), the most active stock picker funds outperform, and factor bet funds as well as closet index funds, underperform. Another finding worth noticing from the study of Cremers et al. (2016) is that poorly performed active managers tend to stick more to the benchmark index to prevent future losses and by that far, increase the level of closet indexing. Well-performed fund managers, on the other hand, tend to get spurred by the increased level of fund competition and become more motivated to increase the level of active management and thus, create more alpha for the fund investors.

The results of Cremers et al. (2016) support the assumption that a high Active share is related to achieving excess returns. Their study adds one more perspective to high Active share investing which is to invest in high Active share funds in such markets where the level of competition is high.

### **3.4 High Active share and long fund duration**

Cremers and Pareek (2016) investigate whether fund duration (trading frequency) has any relation with fund performance within actively managed funds. They measure active management by Active share and find differing results on Active share's future performance predicting ability, compared to the earlier Active share studies of Cremers and Petajisto (2009) and Petajisto (2013). Cremers and Pareek (2016) study the relation between active management and fund duration within active US funds with a sample period of 1990-2013. Fund duration is measured by the quarter-end holdings method, which measures the weighted average holding length of each security the fund holds during the quarter. Fund performance is being compared to a fund's self-declared benchmark instead the one causing the lowest Active share, similar to Petajisto's (2013) study.

Cremers and Pareek's (2016) study is based on the assumption that active management exists for a reason and that excess returns can be achieved only by differing from the

benchmark. They suggest that the holding time of assets needs to be long enough in addition to active management so the fund can yield excess returns. To justify their suggestion, they point out the following conflict that is related to holding assets for a rather long time. When a fund manager has made a long-term bet with a singular stock and believes it to go upwards in the future, they have to wait for the price-increasing event to take place. However, during the waiting time, the stock may go down in the short-term. In such a scenario, the fund manager may be pressured by superiors to close their position as unprofitable to keep their job. An event described above is one of the reasons why long-term bets are not firstly favorable for fund managers.

Cremers and Pareek (2016) measure the fund performance with a specific seven-factor model, based on factors introduced in the financial literature. The model consists of the following factors: market, size, value, momentum, systematic liquidity, low versus high beta and earnings quality. The penultimate factor, low versus high beta, is a factor developed by Frazzini and Pedersen (2014). This so-called Betting against beta (BAB) factor is built by being long in low beta stocks and being short in high beta stocks. The last factor, earnings quality or quality minus junk (QMJ) factor, is built by being long on profitable, growing, less uncertain and higher payout stocks and shorting the opposite.

Cremers and Pareek (2016) find that there is an existing relation between Active share and fund duration. Active share must be high and duration must be long enough to have better possibilities to outperform the benchmark. The best-yielding strategy in Cremers and Pareek's (2016) study is called the "active and patient" strategy as the level of active management is high and trading frequency, on the other hand, is rather patient. Funds that trade frequently, generally underperform the benchmark, no matter whether their Active share is high or low. After noticing that frequent trading does not necessarily lead to excess returns, the second important matter is to eliminate closet index funds from the long-duration (low trading frequency) funds' category (Cremers & Pareek, 2016).

Based on the results of Cremers and Pareek (2016), the best managers with high Active shares and low trading frequencies beat the benchmark by over 2% on average. The out-performance of active and patient managers is mostly explained by Betting against beta and Quality minus junk -factors. The active and patient strategy is rather rare in Cremers and Pareek's (2016) sample, as most of the patient funds (low duration) tend to have low Active shares. The difficulty of succeeding in long-term bets needs to be considered and because of the high difficulty, long-term bets are relatively more rewarding, when successful (Cremers & Pareek, 2016). Long-term bets also require patience from the investor to give time for the manager's bets to succeed. For that reason, the funds that invest according to the active and patient strategy, are not relevant for the most impatient investors. However, an important note to point out is that the results are prone to index manipulation as funds' self-declared indices are used as benchmarks.

Cremers and Pareek's (2016) results suggest that Active share is not alone valid for predicting future fund performance, the trading frequency also needs to be low. Therefore, the results are not completely in line with earlier studies suggesting Active share alone can predict future excess returns. Cremers and Pareek (2016) suggest that managers with active and patient strategies have the most managerial skill: the ability to pick either low beta, value or high-quality stocks, that do not necessarily belong to the benchmark index and hold them for a relatively long period. The finding is later supported by Cremers (2017) who states that Active high Active share funds do not outperform as a group but by including another factor, such as long duration, the fund can perform better. Collecting all the findings of Active share highlighted in the thesis so far, an alpha-seeking investor should search high Active share funds that have a relatively long duration, operate in competitive fund markets and have a relatively small amount of capital under management.

### 3.5 Other Active share findings

Frazzini, Friedman and Pomorski (2016) investigate Active share and get rather differing results compared to earlier Active share studies that broadly suggest that Active share has future performance predicting ability. Frazzini et al. (2016) use the same sample of actively managed US mutual funds as Cremers and Petajisto (2009) and Petajisto (2013) with a sample period of 1990-2009. Frazzini et al. (2016) study whether high Active share funds yield any excess returns.

Frazzini et al. (2016) find no significant statistical evidence of the correlation between Active share and fund returns using their sample data. However, they find three main results. Firstly, high Active share funds typically have a small-cap benchmark and low Active share funds typically have a large-cap benchmark. Therefore, they suggest that sorting funds based on Active shares is the same as sorting funds by benchmark type. Secondly, Frazzini et al. (2016) find no statistical difference between high and low Active share fund returns. And thirdly, no statistical evidence is found that high Active share funds perform better than low Active share funds compared to a given benchmark. Frazzini et al. (2016) suggest that Cremers and Petajisto's (2009) and Petajisto's (2013) results were due to misinterpretation caused by benchmark-adjusting returns. Frazzini et al. (2016) do not suggest using Active share as a significant criterion when searching for the best fund.

Even though Frazzini et al. (2016) do not consider Active share as a reliable fund performance indicator, they highlight other abilities that make Active share useful. According to Frazzini et al. (2016), Active share is considered useful in exposing closet indexing and thereby estimating fund fees to make sure that fund fees are in line with the level of active risk the fund is taking. If an active fund delivers index-like returns, it is unjustified to pay more fees than an index fund charges.

De Rossi and Brar (2018) do not primarily doubt Active share's predicting ability, but they assume that there are certain ranges for a suitable Active share. They create an

optimization tool that can be used to determine whether to increase or decrease a given fund's Active share. However, they do not argue about Active share being a successful performance-predicting measure of an active fund.

Andreu, Forner and Sarto (2021) study the differences in Active share values during publicly reported months compared to the surrounding months within Spanish funds. They find that, on average, funds with high Active shares and low Tracking error, tend to increase their Active shares before the holdings are publicly reported. Also, funds with these characteristics perform the worst while attracting the highest money flows. The findings of Andreu et al. (2021) suggest critically considering Active shares as the values may be prone to window-dressing before public holdings reporting.

Cremers, Fulkerson and Riley (2022b) point out a deviating method for using Active share in fund evaluation. Besides exposing closet indexing, Active share exposes benchmark discrepancies referring to a situation where a fund self-declares a benchmark index that does not match the riskiness of the fund. A fund manager may have the incentive to choose a less risky benchmark compared to fund holdings and thus allow a greater potential to outperform the benchmark. In such a scenario, the outperformance is not necessarily due to fund superiority but due to greater risk taken. Active share may be calculated in contrast to any potential benchmark index and by finding the benchmark causing the lowest Active share, the true risk-adjusted performance of a fund may be evaluated. This is similar to what Cremers and Petajisto (2009) do in their study; instead of choosing the benchmark declared in the prospectus, they use a benchmark that has the greatest overlap in holdings.

Chue and Mian (2022) examine the fluctuations in Active shares in relation to investor sentiment. They find that Active shares tend to decrease when investor sentiment increases. Their findings are in line with previous studies (Baker & Wurgler, 2006, 2007; Stambaugh, Yu & Yuan, 2012, 2014; Antoniou, Doukas & Subrahmanyam, 2013, 2016) claiming that high investor sentiment increases cross-sectional mispricing. The finding

implies that fund managers tend to become more passive during times of high investor sentiment, even though greater performance is achievable due to greater mispricing.

Despite the differing research results on Active share, one factor is generally claimed to be explanatory of fund performance. That factor is managerial skill. Most of the papers regarding Active share refer to the managerial skill being the driver for an outperformance, with the only difference of how managerial skill is measured. Some authors claim that Active share can indicate managerial skill through excess returns and others claim that other indicators are better for measuring the skill of a manager. Managerial skill is an important matter to mention in this thesis regarding active management and excess returns, but it is not discussed any further in this thesis due to the complexity in measuring.

## 4 Data and methodology

In this chapter, I introduce the data and methods being used in the following main chapter. Data limitations such as data unavailability and potential biases are also concerned in this chapter. Both non-statistical and statistical methods are explained as well as the regression models used in the following chapter.

### 4.1 Data and limitations

The data for the thesis is collected from multiple sources. Firstly, the fund sample has been selected from equity funds that mainly invest in Finnish stocks. The existing funds during the examination period of 06/2016 – 06/2021 have been selected from Mutual Fund Report Archive reports (Investment Research Finland, 2021). The reports only include funds that are domiciled in Finland meaning that foreign funds investing in Finnish equities are therefore excluded. The sample is free of survivorship bias as merged and terminated funds are included. The data of the merged funds is merged accordingly so that the observations of the acquired fund end to the last date being independent, and the observations of the acquiring fund continue including the acquired fund from the first observation date after the merger. Out of funds with several fund share classes outstanding, I have chosen growth (reinvested dividends) share classes with the lowest minimum investment amount to harmonize the sample. In cases where a growth share class is unavailable for a fund, the dividends have been taken into account using total return index (includes dividends) instead of normal return index. Small-cap funds and funds with only institutional share classes available have been excluded from the sample due to different benchmark indices and deviating expense structures, respectively. By doing the exclusions, the sample becomes more comparable.

Active share values are mainly unobtainable directly from a database or fund reports, so I have calculated semiannual Active share values for each observation date for each fund. Fund composition data is from Bloomberg and benchmark index OMX Helsinki Cap

Growth Index (OMXHCAPGI) composition data is privately received from Nasdaq. The benchmark applies to all sample funds either based on their self-declared benchmark or investment scope mentioned in the fund prospectus, even though one sample fund uses a dual index benchmark structure. According to the fund prospectus, OMXHCAPGI is the main benchmark for the fund with a weight of 60% and with a secondary benchmark weight of 40%. Therefore, worth noticing is that the use of a dual index benchmark causes upward biased Active share values for this particular fund as Active shares are calculated using OMXHCAPGI holdings exclusively. The possibility of biased results is taken into account when reviewing the findings. As Active share is the core of this thesis, funds without historical composition data available are excluded from the sample as Active share values are unobtainable. After the exclusions, the final sample consists of 17 Finnish equity funds that mainly invest in Finnish equities. The final sample contains only active funds as well as funds that have been merged or terminated during the sample time period. Regarding Active shares, cash positions are considered active investment decisions.

Other fund variables such as total returns, total net asset values and total expense ratios data is mainly collected from Refinitiv Datastream and partly complemented with Bloomberg data regarding missing values. Benchmark index OMXHCAPGI return data is collected from Refinitiv Datastream. As the risk-free rate for Sharpe ratio calculations, I use 10 Year German Government Bond yield values, collected from Bloomberg.

Overall, the sample is rather small with 17 funds and a five-year time period with semi-annual observations. However, historical fund holdings are not publicly available and Active share values need to be calculated separately from obtained holdings data. Based on the difficulties in obtaining Active shares, I find the sample size sufficient for the examination. Also, worth mentioning is that none of the funds have used leverage during the semiannual observation dates limiting the hypothetical Active share to 100, which is in line with the assumptions of Cremers and Petajisto (2009). Potential derivatives

positions are not considered separately nor converted into hypothetical stock positions as equity mutual funds rarely take derivative positions.

## 4.2 Measures of portfolio performance

To measure fund performance, three different measures are being used. As the observations are semiannual, the performance measures are similarly semiannual. The first measure, normal semiannual return, is mainly used for descriptive performance as it does not take fund risk into account. More sophisticatedly and alike previous studies, the more important performance measures are risk-adjusted. For this thesis, I use Sharpe ratio and Information ratio due to their suitability and straightforward characteristics.

Sharpe ratio is a risk-adjusted performance measure, developed by William Sharpe, in 1966. Sharpe ratio is a model for expressing the amount of excess return per unit of risk taken. It utilizes standard error as a risk factor and gives a risk-adjusted return number. Sharpe ratio can be calculated either for an individual asset or a whole portfolio, as is in this thesis.

$$\text{Sharpe ratio} = \frac{r_i - r_f}{\sigma_i}, \quad (4)$$

where

$\sigma_i$  is the standard deviation of security  $i$  returns

The formula of the Sharpe ratio is rather simple. The risk-free rate is subtracted from the return of security  $i$  and the difference is divided by security  $i$ 's standard deviation. In case of calculating the Sharpe ratio for a portfolio instead of an individual asset, security return and security standard error are replaced with portfolio equivalents. Investors seek

high Sharpe ratios to possibly achieve excess returns in the future as well. The ratio takes a high value when either the security return is relatively high, or the security standard error is relatively low. Sharpe ratio is not relevant to be compared with normal percentual returns as normal returns are uncontrolled for risk. To compare potential investments based on their Sharpe ratios, an investor requires corresponding ratios for each investment separately.

Sharpe ratio is very similar to Treynor ratio (Treynor, 1965), the only difference is in the denominator. Treynor ratio uses security or portfolio beta coefficient instead of standard deviation. The ratios are very similar to each other, and Sharpe ratio is chosen to be used in this thesis as it is developed later on the grounds of Treynor ratio.

Sharpe ratio uses a risk-free rate as a benchmark, which may not be the best suited for a fund that attempts to overperform a certain benchmark index return, rather than a risk-free return. Therefore, in addition to normal return and Sharpe ratio, I use Information ratio as the third performance measure as it adjusts portfolio returns to benchmark returns. Information ratio as introduced earlier is calculated by dividing the portfolio excess return over the benchmark index with Tracking error. The use of these two risk-adjusted measures is justified based on the deviating viewpoints.

### **4.3 Methodology**

For the examination, I apply similar methodology as Cremers and Petajisto (2009) and Doshi et al. (2015). Firstly, the sample is examined based on its raw characteristics. The funds are divided into five categories depending on their level of active management, measured by both Active share and Tracking error. The five groups are as mentioned earlier: Pure indexing (Active share and Tracking error close to zero), Closet indexing (low Active share and low Tracking error), Factor bets (low Active share and high Tracking error), Diversified stock picks (high Active share and low Tracking error) and Concentrated stock picks (high Active share and high Tracking error). Both dimensions are used

to understand the difference between factor timing and individual stock selection. In addition to comparing Active share with annualized Tracking error, Active share is viewed in relation to other variables such as semiannual returns, total net assets and total expense ratio.

Secondly, the sample is examined using panel regressions to find if data characteristics are statistically significant. The main assumption for the examination is that high Active share is related to higher fund performance, similar to e.g., Cremers and Petajisto (2009) and Cremers (2013). The assumption is slightly differing from more recent studies of e.g. Cremers et al. (2016) who state that in addition to high Active share, the fund is required to operate in a high competition environment or Cremers and Pareek (2016) who suggest that in addition to high Active share, fund duration needs to be low enough (i.e. trading frequency) or Frazzini et al. (2016) who do not find a clear relation between fund performance and Active share but highlight the usefulness of Active share in evaluating the level of fund fees. The reason I still mainly focus on Active share alone is that Active share is not widely studied within Finnish fund markets due to difficulties in gathering fund holdings data.

Panel regression models are used as the sample data is cross-sectional. For the used panel regression model, I have chosen between pooled ordinary least squared (OLS), fixed effects and random effects models. A pooled OLS is ignored due to potential omitted variable bias. The choice between fixed and random effects is made based on the Hausman test. After choosing the applicable model, the results are controlled for potential heteroskedasticity before presenting them.

First, Active share is used as a dependent variable to see whether the other variables can explain Active share. Explanatory variables for the regression are selected using the same logic as Cremers and Petajisto (2009): variables that are under the fund manager's control such as the number of stocks and total expense ratio and factors that are out of

the fund manager's control such as prior returns and fund size. Thus, the regression is as follows:

$$\begin{aligned} \text{Active share} = & \alpha_i + \beta_{1,i}\text{Prior fund return} + \beta_{2,i}\text{Prior index return} + \quad (5) \\ & \beta_{3,i}\log\text{Fund size} + \beta_{4,i}\text{Total expense ratio} + \\ & \beta_{5,i}\text{Tracking error} + \beta_{6,i}\log\text{Number of shares} + \varepsilon_i, \end{aligned}$$

where all the observations are semiannual.

After figuring out the potential relations between Active share and other explanatory variables, the fund performance is used as the dependent variable using a regression model chosen based on the methodology explained above. For the fund performance examination, I use independent variables that are commonly considered in previous studies and expected to affect fund performance and have data available. Thereby, the panel regression model is as follows:

$$\begin{aligned} \text{Fund performance}_t = & \alpha_i + \beta_{1,i}\text{Active share}_{t-1} + \beta_{2,i}\text{Fund return}_{t-1} + \quad (6) \\ & \beta_{3,i}\text{Index return}_t + \beta_{4,i}\text{Index return}_{t-1} + \\ & \beta_{5,i}\log\text{Total net assets}_{t-1} + \\ & \beta_{6,i}\text{Total expense ratio}_{t-1} + \\ & \beta_{7,i}\text{Tracking error}_{t-1} + \beta_{8,i}\log\text{Number of shares}_{t-1} + \\ & \varepsilon_i, \end{aligned}$$

where all the observations are semiannual and  $i$  refers to a certain fund and  $t$  to a certain time period. Note that most of the variables are lagged, such as Active share at time  $t$  which affects the fund performance from  $t$  to  $t+1$ . As mentioned,  $\text{Fund performance}_t$  is evaluated by pure gross returns, Sharpe ratio and IR. The regression results are presented in the next chapter.

## 5 Empirical findings

In this chapter, I present the findings of the study. First, the sample is examined purely based on distinguishable characteristics from the perspectives of sample descriptive statistics and factor-specific relations. Continuing with more sophisticated statistical methods, I first build a panel regression model to notice whether other fund variables are able to explain Active share. Finally, similar panel regression models are built trying to explain fund performance using multiple performance measures.

### 5.1 Sample characteristics

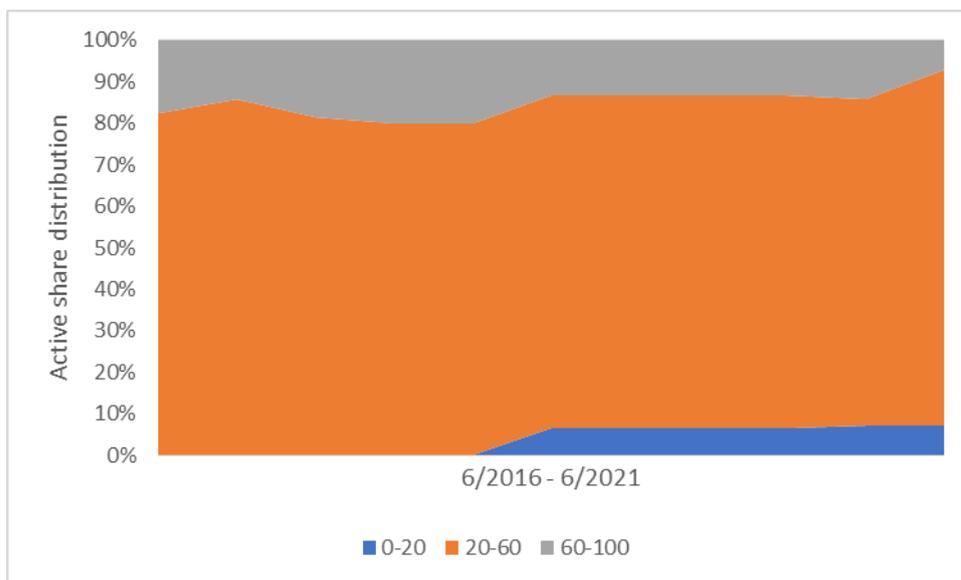
As mentioned in the previous chapter, the sample consists of 17 Finnish funds that mainly invest in Finnish equities for a time period of 6/2016 – 6/2021 with semiannual observations. Nine base variables are used in this examination. The descriptive statistics of the used variables are presented in the table below. Worth noticing is, that Sharpe ratio and Information ratio are semiannual values. Tracking error values are semiannual as well but annualized to make numbers indicatively comparable with Cremers and Petajisto's (2009) tracking error numbers. The fund return values are gross returns.

**Table 2.** Descriptive statistics of the sample variables.

	Mean	Median	Standard Error	Standard Deviation	Kurtosis	Skewness	Minimum	Maximum	Count
Fund return	0.052	0.064	0.009	0.110	-0.579	0.142	-0.174	0.309	153
Index return	0.055	0.067	0.007	0.089	-0.901	-0.447	-0.117	0.172	153
Sharpe ratio	0.653	0.766	0.093	1.151	-0.752	-0.006	-1.947	3.192	153
Information ratio	-0.089	-0.115	0.056	0.691	0.949	0.622	-1.809	1.999	153
Active share (%)	45.2	42.5	1.3	16.4	1.604	0.987	4.5	89.9	151
Tracking error	0.070	0.050	0.004	0.049	4.288	1.856	0.018	0.274	153
Total expense ratio (%)	1.59	1.60	0.03	0.35	0.887	-1.150	0.70	2.01	146
Total net assets (M€)	186	107	17	212	5.917	2.403	14	1023	153
Number of shares	36.4	34.0	1.0	11.8	2.853	1.770	17.0	72.0	151

Looking at the descriptive statistics above, an average active Finnish fund investing in Finnish equities yields 5,2% semiannually before fees and expenses during the examination period while the Finnish benchmark index OMXHCAPGI yields 5,5%. The sample

funds are relatively passive with an average Active share of 45% while only one of the 17 funds is considered passive during the latter half of the period. An average fund charges 1,59% annually, manages assets worth 186 million euros and holds 36 different shares (including possible cash positions), on average.

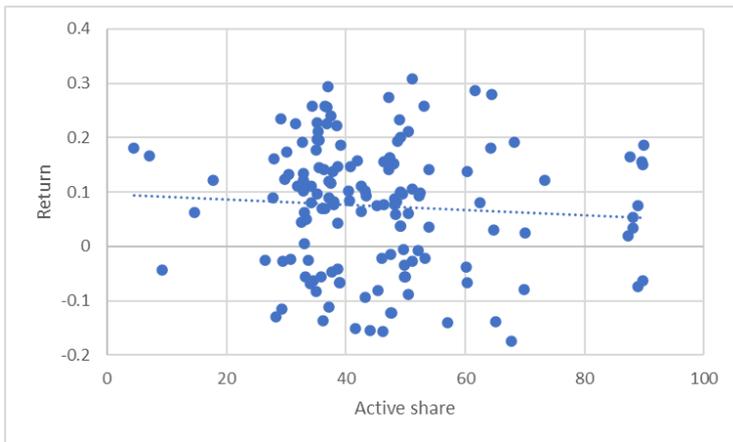


**Figure 2.** Evolution of Active share (%).

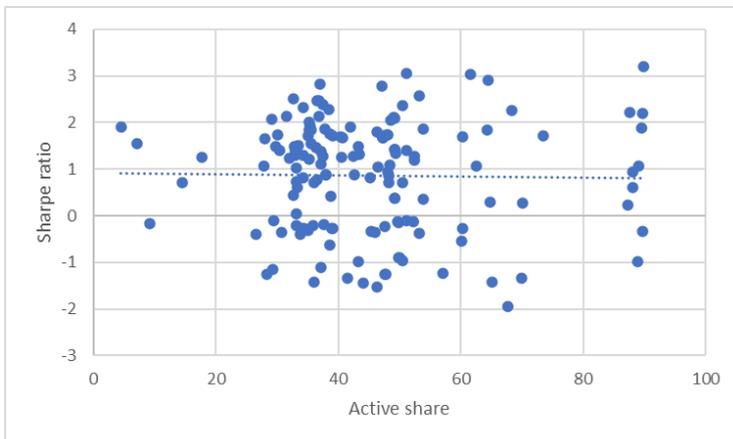
The figure above represents the evolution of the sample Active share distribution over the sample period from 2016 to 2021. The share of truly active funds (Active share over 60%) has decreased from 18% to 7% while the share of passive funds (Active share below 20%) has increased from 0% to 7% over time. The share of closet index funds (Active share between 20% and 60%) has remained mostly the same at over 80% of the distribution over time. The trend of decreasing Active shares shown in the figure is in line with the findings of Cremers and Petajisto (2009) and Doshi et al. (2015), who find a similar downward trend in actively managed funds during their examination periods of 1980-2003 and 1980-2013, respectively.

Next, I examine the potential relations between Active share and other variables. The relation I am most interested in is the relation between Active share and fund performance. In the three figures below, Active share is plotted with three different

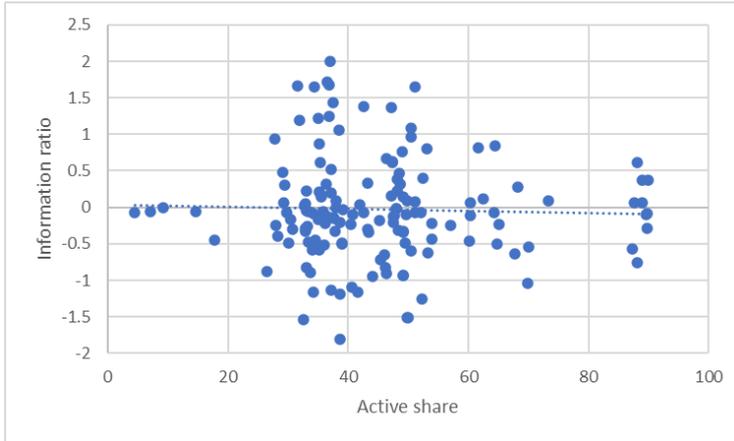
performance measures, presenting all Active share-Return observation pairs that exist in the sample. Again, it is important to notice that an upper outlier of Active share values is mostly due to a certain fund using a dual benchmark, while Active share is calculated only using the main benchmark.



**Figure 3.** Return (semiannual) and lagged Active share (%).



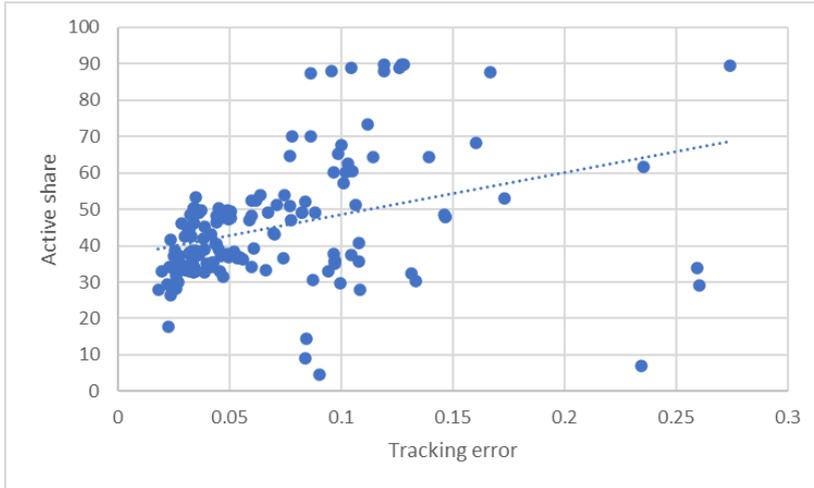
**Figure 4.** Sharpe ratio (semiannual) and lagged Active share (%).



**Figure 5.** Information ratio (semiannual) and lagged Active share (%).

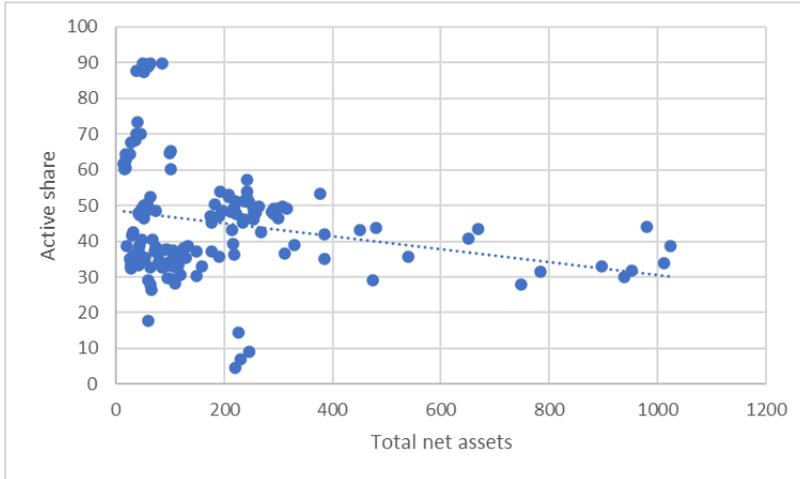
Observing the figures above, there seems not to be any notable relation between Active share and either normal gross returns or adjusted returns. The finding is not in line with my assumptions based on Cremers and Petajisto's (2009) results of the existing relation between high Active share and higher fund performance. However, the significance of the relation is examined further from a more statistical point of view later in this thesis. Other than a clear performance-enhancing effect from Active share, higher Active share seems to increase the volatility of Sharpe ratio values. IR, on the other hand, well visualizes the border of index funds and active funds at slightly above 20% Active share in the figure. The most passive funds yield a rather constant IR of near zero, while the IR values start to deviate more when moving toward more active funds. This finding justifies the use of the 20% limit as the minimum Active share for active funds in larger samples, similar to Cremers and Petajisto's (2009) first proposal.

Continuing with the characteristics of Active share, I try to understand the relations between Active share and the other control variables Tracking error, Total net asset value and Total expense ratio. The observations are plotted below.



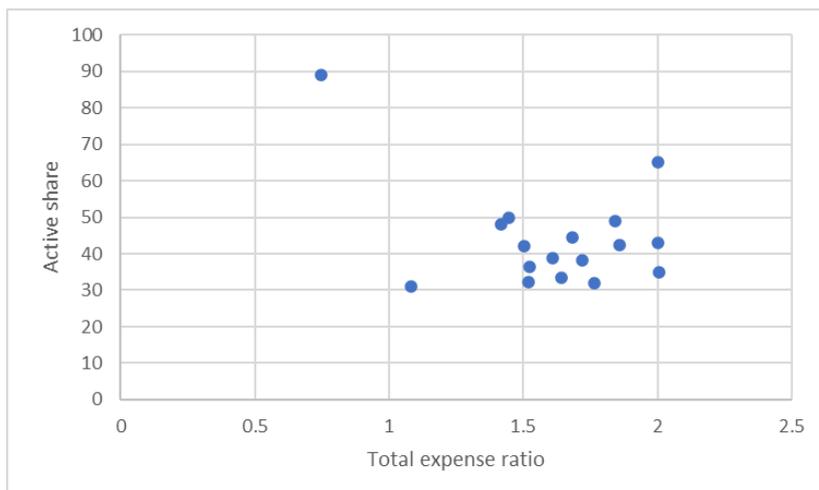
**Figure 6.** Active share (%) and Tracking error (annualized).

Similar to Cremers and Petajisto (2009), I use the same Active share categories in describing the true level of active management. Active share values below 20% refer to index funds and Active share values above 60% refer to truly active funds. Active share values between 20% and 60% refer to closet index funds. As seen in the figure above, most of the fund semiannual Active share observations rank in the closet index fund category. One reason for relatively low overall Active shares is likely the small size of the Finnish equity universe compared to the US equivalent, for example. A smaller selection of stocks to choose from makes it more difficult for the fund manager to deviate from the benchmark index. Regarding truly active funds (Active share over 60%), most of the active fund observations rank in the diversified stock picker category (Cremers & Petajisto, 2009) combining relatively high Active share with relatively low Tracking error.



**Figure 7.** Active share (%) and Total net asset value (M€).

The figure above indicates the relation between Active share and fund size within all the observations. Similar to what Cremers and Petajisto (2009) find, the plot suggests a slight inverse relationship between fund size and Active share. The finding is in line with previous studies of Cremers and Petajisto (2009) who suggest a larger amount of capital under management complicates active asset allocation. When plotting for fund size and fund returns, on the other hand, there is no remarkable outperformance when comparing small funds and large funds. The findings of e.g. Chen et al. (2004) suggest small funds tend to outperform large funds due to easier allocation and insignificant liquidity challenges. However, the lack of performance-related findings from my sample is likely due to the matter that none of the sample funds clearly outperform the benchmark during the sample period.



**Figure 8.** Average Active share (%) and average Total expense ratio (%).

As seen in the figure above, Total expense ratio tends to increase on average when the average Active share increases. The finding is in line with Cremers and Petajisto (2009) who find a similar tendency. The relation is justified as a higher Active share enables returns greater than the benchmark. Now that the observations on the figure are fund-specific averages instead of date-specific individual fund observations, the figure reveals that on average, only two of the sample funds are truly active with an average Active share over 60%. Again, worth noticing is that one of the two truly active funds gives an upward-biased Active share due to the use of a dual benchmark while having a low TER. Simultaneously, none of the funds appear as an index fund, on average. Thus, 15 of the 17 sample funds are on average closet index funds. The great amount of closet index funds does not give out an attractive image of active Finnish funds and assumingly provides worse performance results when examining fund performance from a more statistical point of view.

## 5.2 Active share

Firstly, the characteristics of Active share are examined further using the regression model presented in the methodology section. The purpose of this is to see whether Active share is relevant to use in a model explaining fund performance or whether Active

share is explained by other independent variables and adds no additional value to the model. I use the same explanatory variables as Cremers and Petajisto (2009) as far as the data is available. Due to data limitations, the final independent variables trying to explain Active share are prior fund return, prior index return, log total net assets, total expense ratio, tracking error and log number of shares.

As the sample is a panel dataset, the use of an ordinary least squares (OLS) regression model is ignored due to the inability to capture fixed and/or random effects. Therefore, I examine the possibility to use a fixed effects model or random effects model instead. After building both fixed and random effects models separately, I choose the preferred model out of these two using a Hausman test. The regression output of a random effects model, suggested by the Hausman test, is seen in the table below. The results are controlled for heteroskedasticity and autocorrelation.

**Table 3.** Determinants of Active share.

	Active share
Intercept	125.95 (3.621)
Prior fund return	-26.14 (-2.531)
Prior index return	35.69 (2.600)
Log(TNA)	-1.76 (-1.038)
TER	5.41 (0.573)
Tracking error	-1.78 (-0.104)
Log(Shares)	-23.15 (-2.247)
N	144
Adj. R-squared	0.233

The table above shows the regression coefficients of the independent variables with the corresponding test statistic values. On a 5% significance level, only prior fund return, prior index return and log number of shares give out statistically significant values. Out of these variables, prior index return seems to have the most explanatory power on Active share, suggesting a positive relation between these two. However, prior index nor prior fund return do not have recognizable economic significance as semiannual return typically takes a value lower than 0.1. In short, the fund managers lower their active exposure when the prior fund return increases and increase their active exposure when the prior index return increases. This indicates that the fund managers become more risk-averse when they have had a good performance during prior semi-annual trading periods and possibly defend their gains. In case of relatively good index performance, fund managers tend to increase their fund's active exposure accordingly to possibly try to keep up with the benchmark gains. The finding is contrary to Cremers and Petajisto's (2009) original study, where they find a positive and significant relation between the prior fund return and Active share suggesting that managers who have been successful in the past, choose to increase their level of active management.

Log number of shares, however, has a greater economic impact on Active share as log number of shares values tend to settle between 3 and 4. The inverse relation between Active share and the log number of shares is logical and intuitive as increasing the number of shares most probably increases the fund's overlap with the benchmark index. Overall, funds with fewer stocks under management are more active measured by Active share.

The most interesting finding regarding Tracking error is not in line with Cremers and Petajisto's (2009) result. They find Tracking error being the most significant explanatory variable in explaining Active share while my sample suggests a minimal and insignificant impact on Active share. The sample finding does not deviate from the scatter plot of Active share and Tracking error, which does not provide a clear relation either. Even though the finding deviates from Cremers and Petajisto (2009), it suggests that Active

share and Tracking error have varying approaches to measuring the level of portfolio active management and can thereby complement each other if used jointly.

The overall conclusion from the regression is that Active share cannot be explained comprehensively by other independent variables and is thus a relevant addition to use in the upcoming fund performance regressions. The main purpose of the upcoming regressions is to prove whether Active share has any potential predicting power over future fund returns.

### **5.3 Fund performance**

Using a similar approach as in the previous Active share regression, I build three regression models trying to explain fund performance. The regression model is presented in the earlier chapter. Performance is the dependent variable, measured by fund gross return, Sharpe ratio and Information ratio. As independent variables, I have chosen a similar selection of variables as in the models of Cremers and Petajisto (2009) and Doshi et al. (2015), making exclusions due to data limitations. The final independent variables for the performance regressions in addition to Active share are prior fund return, prior index return, index return, log total net assets, total expense ratio, tracking error and log number of shares.

The main interest regarding the regressions is the potential relation between Active share and fund performance. Based on the characteristics of the sample Active share, the expectation is that the sample cannot provide significant evidence of high Active share yielding higher returns. One of the main drivers for the expected results is generally low sample Active shares. It is unlikely that a sample overrepresented by closet funds would yield significant Active share findings. However, it is still relevant to compare Active share and Tracking error as they serve a very similar purpose.

As Active share may not yield favorable results within the sample, it is relevant to observe other fund characteristics and whether they affect fund performance. Small fund size generally correlates with better fund performance (e.g. Chen et al., 2004) and it will be the center of focus in addition to Active share. Besides Active share and fund size, the fund expense ratio is another factor that is observed more thoroughly as it is important to know whether a fund investor gets value for their money.

The methods are the same as in the previous Active share regression, the models only have more variables. Building regression models separately for each observed dependent variable, I test for fixed and random effects to choose a correct model. Using a Hausman test to decide between fixed and random effects models, the test suggests using a random effects model for each of the three different regressions. The output of the regression models is presented in the table below, where return, Sharpe ratio and IR measures performance, respectively. The numbers are controlled for heteroskedasticity and autocorrelation.

**Table 4.** Determinants of fund performance.

	Return	Sharpe	IR
Intercept	0.0264 (1.189)	0.3007 (0.960)	0.4440 (0.854)
Active share <sub>t-1</sub>	-0.0002 (-1.168)	0.0011 (0.519)	-0.0028 (-1.498)
Fund return <sub>t-1</sub>	0.1129 (1.011)	0.9838 (1.165)	3.0780 (1.274)
Index return <sub>t</sub>	1.1046 (35.071)	11.3396 (47.411)	1.6595 (2.634)
Index return <sub>t-1</sub>	-0.2098 (-1.589)	-1.8475 (-1.813)	-4.6354 (-1.62)
Log(TNA) <sub>t-1</sub>	-0.0019 (-1.042)	-0.0370 (-1.629)	-0.0343 (-1.64)
TER <sub>t-1</sub>	0.0070 (1.056)	-0.0456 (-0.617)	0.0849 (0.715)
Tracking error <sub>t</sub>	0.0798 (1.493)	0.6002 (0.835)	0.6309 (0.972)
Log(Shares) <sub>t-1</sub>	-0.0071 (-1.007)	-0.0150 (-0.173)	-0.1034 (-0.737)
N	141	141	141
Adj. R-squared	0.897	0.874	0.101

In the table above, we see the coefficients of the regression model with three different dependent variables: Return (1), Sharpe (2) and IR (3) with the corresponding test statistics. Index return<sub>t</sub> is the most significant variable in all three regressions and significant at a 5% significance level. As the sample is overweighted by closet index funds, yielding index-like returns is virtually unavoidable. Having a relatively large overlap in portfolio components increases the possibility of generating similar performance numbers. Interestingly, the Index return<sub>t</sub> coefficient in regression (1) is greater than 1, indicating a semiannual fund outperformance of 10.4% compared to the index. The finding, therefore, is also economically significant. Descriptive statistics, on the other hand, suggest that an average fund loses by 3% (0.2 percentage points) semiannually. However, the regression output is corrected for cross-sectional (fund-specific) effects increasing its relevance. Even though the finding suggests a great fund outperformance, it is important to look at the whole regression model output instead of only one factor. The Index return<sub>t</sub> factor is economically significant in the other models (2) and (3) as well. Prior return factors, prior fund nor prior index return seem to have no impact at all suggesting that there is no semiannual momentum effect within Finnish funds.

Examining the main factor, Active share, the results are not as promising as Cremers and Petajisto (2009) provide using their US sample. Active share is not statistically significant on any of the regression models indicating a nonexistent relation between Active share and fund performance. However, the sign of the coefficient is contrary to the expected. According to previous studies, the relation is positive, indicating an increased performance when the level of active management is higher. The contrary relation is most probably due to the homogenous Active share values of the sample funds. As the sample Active shares are generally low, the results are slightly biased. A possible interpretation is that an increase in Active share for a closet index fund (Active share between 20% and 60%) causes a decrease in performance. As the sample only has roughly two truly active funds, it does not support examining the relation between high Active share (over 60% Active share) funds and fund performance. Due to the sample characteristics, the

relation between high Active share funds and their performance remains unexamined in this thesis.

Even though I do not find similar results of Active share alone being capable of predicting future fund performance, my results are more in line with more recent studies. For instance, Cremers and Pareek (2016) suggest that in addition to high Active share, fund trading frequency needs to be low enough. Frazzini et al. (2016), on the other hand, do not find significance in the relation between Active share and fund performance using a similar dataset to Cremers and Petajisto (2009) and Petajisto (2013).

Another viewpoint to take into account is the differences between Active share and Tracking error. In terms of both of these measures, I do not find any evidence on their usefulness in evaluating fund performance. However as seen in the scatter plot in the first section of this chapter, there seems to be a positive correlation of some sort between these two measures. While Active share does not reveal the type of fund active management, Tracking error is better for examining whether the fund allocation rests on sector overweight or diversified stock picks, for instance. As Cremers and Petajisto (2009) state, Tracking is useful besides Active share.

A measure that remains out of Active share and Tracking error comparison is Active weight. The reason Active weight is kept out of the examination is its slightly deviating base logic of benchmark selection. According to the principles of Active weight, a fund's benchmark should be a handmade benchmark constructed by value-weighting the fund components. The calculation formula of Active weight is identical to the Active share formula and therefore it would be interesting to see whether Active weight can predict future fund performance better than Active share. However, Active weight is not examined further in this thesis. In any case, as the Active shares are generally low and Active weights are by nature lower than Active shares, the sample Active weights would presumably be particularly low.

Fund size, measured by log Total net assets, does not provide any statistically significant results in any of the regression models, even at a 10% significance level. The result is not in line with previous studies finding an inverse relation between fund size and performance (e.g. Chen et al., 2004). This may be due to sample distribution, where relatively small funds are overrepresented, not allowing the effect to be realized. When the fund size is small enough, the difficulties of taking large positions and making the transactions do not principally erode fund performance. Even though the fund size coefficients lack significance, the sign is similar to expected suggesting an inverse relation between fund size and performance.

Finally, from the individual regression factors, I examine fund fees, TER, further. The log number of shares is purely a control variable and therefore not discussed more closely. TER is statistically significant on a 5% significance level in regression (1) but remains insignificant in the other regressions. However, TER is not statistically significant as a 0.10 percentage point increase in TER would indicate a 0.076 percentage point increase in semiannual return after transforming both variables into comparable percentage form. Referring to earlier Active share regression in Table 3, TER does not have any significant explanatory power over Active share either. To conclude the finding, it seems that the expenses of Finnish funds are generally not in line with either fund performance or the level of active management.

The explanatory power of the models is rather volatile and some additional factors would most probably add value to each model but are excluded due to data limitations. Additional control variables that have been used in similar previous studies are for instance fund age, fund manager experience, fund manager age and fund turnover/duration. Fund duration is especially interesting as Cremers and Pareek (2016) state that funds with high duration (low trading frequency) and high Active share outperform their benchmark indices. Although as mentioned earlier, the Active shares are relatively very low within the sample, and therefore dividing funds based on Active share and duration would most likely cause unreliable results.

## 6 Conclusions

This thesis studies the effects of Active share within Finnish equity funds mainly investing in Finnish equities. Examining the group of 17 mutual funds for a period starting in 6/2016 and ending in 06/2021, I find somewhat similar results as in previous Active share studies. I conclude the results by reflecting on my findings on the research questions presented in the introduction chapter.

### 1. *Are Finnish funds truly active?*

As observable from the results, Finnish funds tend to be rather passive. Out of the sample of 17 funds, on average none of the funds is an index fund, two funds are truly active (Active share over 60%) and the rest are closet index funds (Active share between 40% and 60%). Worth mentioning is also that one of the truly active funds uses a dual benchmark, while Active shares are calculated only considering the main benchmark index, which causes upward-biased Active shares for that one particular fund. Looking at the sample characteristics, an average Finnish fund investing in Finnish equities has an average Active share of 45.2%. This means that an average sample fund has roughly a 55% overlap with the benchmark index.

As the sample lacks truly active funds, I do not find it relevant to try to find the relation between high Active share funds and fund performance as the number of observations for the highest Active share group would be relatively low. The most likely reason for considerably low Active shares is the small size of the Finnish stock market. The constricted selection of Finnish equities makes it difficult for a fund manager to deviate significantly from the benchmark. This is indirectly in line with a previous finding indicating that funds are relatively more active when the competition is tougher (Cremers et al., 2016). Another finding worth mentioning is the general trend of decreasing Active shares within Finnish funds. Similarly to the US, the share of truly active funds has decreased significantly and the share of more passive alternatives has increased significantly.

## *2. Can Active share predict future fund performance?*

Based on my results, Active share cannot predict future fund performance as the panel regressions do not provide statistical significance on the Active share factor. The finding is understandable as Active share tends to be most reliable within the extremes of Active share, within most active and most passive funds (Petajisto, 2013). As my sample mostly consists of closet index funds, the potential relation between either high Active share and performance or low Active share and performance does not have statistical support to appear. The extreme Active shares having a stronger impact on fund performance suggests the relation is non-linear.

The Active share results from the panel regression models suggest an inverse relation between Active share and fund performance. The relation is expected to be positive, meaning that a higher Active share leads to better fund performance. Now that the sample is overrepresented by closet index funds, the relation mainly reflects the impact of Active share within closet index funds. Regardless of the inverse relation suggested by the regressions, the finding lacks significance.

## *3. Is it worth paying higher fees for active fund management in Finland?*

Generally, higher fees are justified if the fund can outperform its benchmark and the only way to do so is by holding divergent positions from the benchmark. Mathematically, the probability of delivering index-divergent returns increases when the overlap in holdings decreases. After deducting fund fees and expenses from the return, it is difficult to outperform the benchmark without a high level of Active share. Therefore, I suggest similarly to previous studies that an investor should not pay higher fees for closet index funds, which the sample largely represents. The average TER of the sample is 1.59% while the minimum TER is 0.70%. This implies that an investor would pay almost one percentage point extra for similar returns they can achieve by paying 0.70% in expenses.

It is easy to agree with the consensus of Active share being a great indicator of whether the fund charges are relatively too high compared to its return potential.

Overall, my findings do not provide reliable insight for a fund investor regarding achieving excess returns. However, the results provide evidence that Active shares within Finnish funds mainly investing in Finnish equities are considerably low. The fund fees, on the other hand, are not correspondingly lower. Therefore, I see no clear motivation for a fund investor to seek excess returns by choosing active funds in the Finnish fund market. Instead, the results suggest that a low-cost index fund is a relevant option for a profit-maximizing fund investor. Based on my results, Active share is unreliable in evaluating expected future fund performance but instead, a more valid measure for assessing whether the fund fees are too high.

Worth keeping in mind when interpreting the results is the fact that the sample used in this thesis includes 17 funds and a time period of five years. The relevance of the findings would be stronger if the sample was wider and the time period longer. The observations used are semiannual and the overall picture would be more accurate with more frequent observations, monthly or quarterly, for instance.

The findings of the thesis leave room for further Active share examination within Finnish equity mutual funds. Due to data limitations, the sample does not include all possible funds that could fit into the research framework. Also, the examination period could be longer than five years. A wider study would likely require a more collective method for collecting Active share data. Due to identification challenges, I framed the fund domicile only to Finland but increasing the sample is possible by also including funds domiciled outside Finland. If the potential sample includes a sufficient amount of truly active funds, it is relevant to test for different factor combinations in order to increase fund performance, such as high Active share and small fund size. A potential expansion to my study is expanding the geographical horizon to Nordic countries as the market characteristics are similar to Finnish equity markets.

## References

- Andreu, L., Forner, C., & Sarto, J. L. (2021). Window dressing in the Active Share scores in publicly reported portfolios. *Business research quarterly*, 234094442110246. <https://doi.org/10.1177/2340944421102465>
- Antoniou, C., Doukas, J., & Subrahmanyam, A. (2013). Cognitive Dissonance, Sentiment, and Momentum. *Journal of financial and quantitative analysis*, 48(1), 245-275. <https://doi.org/10.1017/S0022109012000592>
- Antoniou, C., Doukas, J., & Subrahmanyam, A. (2016). Investor Sentiment, Beta, and the Cost of Equity Capital. *Management science*, 62(2), 347-367. <https://doi.org/10.1287/mnsc.2014.2101>
- Baker, M., & Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of finance (New York)*, 61(4), 1645-1680. <https://doi.org/10.1111/j.1540-6261.2006.00885.x>
- Baker, M., & Wurgler, J. (2007). Investor Sentiment in the Stock Market. *The Journal of economic perspectives*, 21(2), 129-151. <https://doi.org/10.1257/jep.21.2.129>
- Bodie, Z., Kane, A. & Marcus, A. (2009). *Investments* (8. ed.). McGraw-Hill.
- Carhart, M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance*, 52(1), 57-82. <https://doi.org/10.1111/j.1540-6261.1997.tb03808.x>
- Chen, J., Hong, H., Huang, M. & Kubik, J. (2004). Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization. *American Economic Review*, 94(5), 1276-1302. <https://doi.org/10.1257/0002828043052277>
- Chue, T., & Mian, M. (2022). Investor sentiment and mutual fund stock picking. *Applied economics letters*, 29(17), 1620-1625. <https://doi.org/10.1080/13504851.2021.1951440>
- Cremers, M. (2017). Active share and the three pillars of active management: Skill, conviction, and opportunity. *Financial analysts journal*, 73(2), 61-79. <https://doi.org/10.2469/faj.v73.n2.4>
- Cremers, M., Ferreira, M., Matos, P. & Starks, L. (2016). Indexing and active fund management: International evidence. *Journal of Financial Economics*, 120(3), 539-560. <https://doi.org/10.1016/j.jfineco.2016.02.008>

- Cremers, M., Fulkerson, J., & Riley, T. (2022a). Active Share and the Predictability of the Performance of Separate Accounts. *Financial analysts journal*, 78(1), 39-57. <https://doi.org/10.1080/0015198X.2021.1984826>
- Cremers, M., Fulkerson, J., & Riley, T. (2022b). Benchmark Discrepancies and Mutual Fund Performance Evaluation. *Journal of financial and quantitative analysis*, 57(2), 543-571. <https://doi.org/10.1017/S0022109021000119>
- Cremers, M. & Pareek, A. (2016). Patient capital outperformance: The investment skill of high active share managers who trade infrequently. *Journal of Financial Economics*, 122(2), 288-306. <https://doi.org/10.1016/j.jfineco.2016.08.003>
- Cremers, M. & Petajisto A. (2009). How Active Is Your Fund Manager? A New Measure That Predicts Performance. *The Review of Financial Studies*, 22(9), 3329-3365. <https://doi.org/10.1093/rfs/hhp057>
- De Rossi, G. & Brar, G. (2018). Is high active share always good? *Journal of Asset Management*, 19(7), 460-471. <https://doi.org/10.1057/s41260-018-0096-5>
- Doshi, H., Elkamhi, R., & Simutin, M. (2015). Managerial activeness and mutual fund performance. *Review of asset pricing studies*, 5(2), 156-184. <https://doi.org/10.1093/rapstu/rav005>
- Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25(2), 383-417. <https://doi.org/10.1111/j.1540-6261.1970.tb00518.x>
- Fama, E. (1972). Components of Investment Performance. *Journal of Finance*, 27(3), 551-568. <https://doi.org/10.1111/j.1540-6261.1972.tb00984.x>
- Frazzini, A., Friedman, J. & Pomorski, L. (2016). Deactivating active share. *Financial analysts' journal*, 72(2), 14-21. <https://doi.org/10.2469/faj.v72.n2.2>
- Frazzini, A. & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, 111(1), 1-25. <https://doi.org/10.1016/j.jfineco.2013.10.005>
- French, K. (2008). Presidential Address: The Cost of Active Investing. *Journal of Finance*, 63(4), 1537-1573. <https://doi.org/10.1111/j.1540-6261.2008.01368.x>
- García, F., Guijarro, F. & Moya, I. (2013). A multiobjective model for passive portfolio management: An application on the S&P 100 index. *Journal of Business*

- Economics and Management*, 14(4), 758-775.  
<https://doi.org/10.3846/16111699.2012.668859>
- Goodwin, T. (1998). The information ratio. *Financial Analysts Journal*, 54(4), 34-43.  
<https://doi.org/10.2469/faj.v54.n4.2196>
- Investment Research Finland. (2021, December). Mutual Fund Report Archive. Investment Research Finland. <https://www.sijoitustutkimus.fi/en/solutions/institutional-investors/mutual-fund-report/mutual-fund-report-archive/>
- Jegadeesh, N. & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance*, 48(1), 65-91.  
<https://doi.org/10.1111/j.1540-6261.1993.tb04702.x>
- Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, 7(1), 77-91.  
<https://doi.org/10.1111/j.1540-6261.1952.tb01525.x>
- Morningstar. (2021, December). *Morningstar U.S. Fund Flows: Records smashed in 2021*. Morningstar research. <https://www.morningstar.com/lp/fund-flows-direct> [Restricted access]
- Petajisto, A. (2013). Active Share and Mutual Fund Performance. *Financial Analysts Journal*, 69(4), 73-93. <https://doi.org/10.2469/faj.v69.n4.7>
- Roll, R. (1992). A Mean/Variance Analysis of Tracking Error. *Journal of Portfolio Management*, 18(4), 13-22. <https://doi.org/10.3905/jpm.1992.701922>
- Ruiz-Torrubiano, R. & Suarez, A. (2009). A hybrid optimization approach to index tracking. *Annals of Operations Research*, 166(1), 57-71.  
<https://doi.org/10.1007/s10479-008-0404-4>
- Sharpe, W. F. (1966). Mutual Fund Performance. *The Journal of Business*, 39(1), 119-138. <https://doi.org/10.1086/294846>
- Sharpe, W. (1991). The Arithmetic of Active Management. *Financial Analysts Journal*, 47(1), 7-9. <https://doi.org/10.2469/faj.v47.n1.7>
- Stambaugh, R., Yu, J., & Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of financial economics*, 104(2), 288-302.  
<https://doi.org/10.1016/j.jfineco.2011.12.001>

Stambaugh, R., Yu, J., & Yuan, Y. (2014). The long of it: Odds that investor sentiment spuriously predicts anomaly returns. *Journal of financial economics*, *114*(3), 613-619. <https://doi.org/10.1016/j.jfineco.2014.07.008>

Treynor, J. L. (1965). How to rate management of investment funds. *Harvard business review*, *43*(1), 63-75. <https://doi.org/10.1002/9781119196679.ch10>

## Appendices

### Appendix 1. List of sample funds

Aktia Capital

Elite Alfred Berg Finland Focus

Danske Invest Finland Dividend Plus

Danske Invest Finland High Dividend

Danske Invest Finnish Equity

EQ Finland 1 K

Evli Finland Select B

FIM Fenno

Fondita Equity Spice

LähiTapiola ESG Dividend Finland

Nordea Finnish Stars K

OP Delta

OP Finland

POP Finland

Säästöpankki Kotimaa

SEB Finlandia B

Seligson Phoebus