Making Reasonable Mutual Fund Investment Decisions in the Recovering German Equity Market

---

Gaining Insight into Active Share

Adam Eliasson
ID #: 900713

Jannik Haskamp
ID #: 890722

Supervisor: Naciye Sekerci
Date: June 19, 2015

Abstract

This study investigates influences of certain fund characteristics as independent variables on a funds *active share*, as introduced by Cremers & Petajisto in 2009. Furthermore, along with *active share*, the potential use of these variables to predict future mutual fund returns is tested. The study is conducted for the recovering German equity mutual fund market between 2010 and 2015 as a recovery period was assumed to show significantly different results compared to long-term studies. To investigate influences, a cross-sectional study was performed with partially different results than initially expected. Compared to empirical foundations, the findings for *active share* are less statistically and economically significant during economic recovery in Germany. Although *active share* can somehow be explained through certain variables, the study disproves pervious findings as *active share* cannot predict fund performance. Funds with high active share even tend to underperform their benchmarks during recovery. On this basis, a fund investment strategy for the recovering market is proposed, finding exchange traded funds as the more suitable investment vehicle. This mainly derives from the significantly lower fees charged by ETFs.

Keywords: Active share, exchange traded funds, mutual fund, tracking error, investments

Acknowledgements

We would like to thank our supervisor Naciye Sekerci for her guidance and professional comments during the process of this study.

Furthermore, we also want to thank the professors of the "Corporate and Financial Management" master's program at Lund University for providing solid ground which enabled us to conduct professional research in the field of corporate finance and investments.

Additionally, we would like to our fellow students for their valuable comments which positively improved the end result of this paper.

Definitions and Abbreviations

Accumulating distribution policy – All capital gains are reinvested in the fund

Active share – How does each weighted holding in a fund deviated from each weighted holding in the respective benchmark

Alpha – Performance compared to a benchmark

Benchmark – Related objective index against a fund measures it's performance

DAX – Deutscher Aktien Index

ETF – Exchange traded fund, passively managed fund following a certain index

Fees – In this article, fees are referred to as annual management fees

Institutional/restricted share class – Share classes where only certain investors have access and a minimum investment might be required

Investment policy – Binding document between portfolio manager and investor, outlining general rules for management

Investment strategy – An investor strategy in order to make reasonable investment decisions on risk tolerance, asset allocation and time horizon

Mutual funds – Refers to funds that have the possibility to invest outside benchmark, actively

OLS – Ordinary Least Squares Method

Recovering market –Post financial crisis period (2010 – 2015)

Share class – A distinction of different security types within one fund

Sharpe ratio – A measure for calculating risk-adjusted return by William F. Sharpe

Total Expense Ratio – Costs incurred compared to the fund value

Tracking error – Divergence between price behavior of a fund and the respective benchmark

Umbrella Fund – Fund of fund, investing in other funds only

Table of Contents

Equations

Graphs

Tables

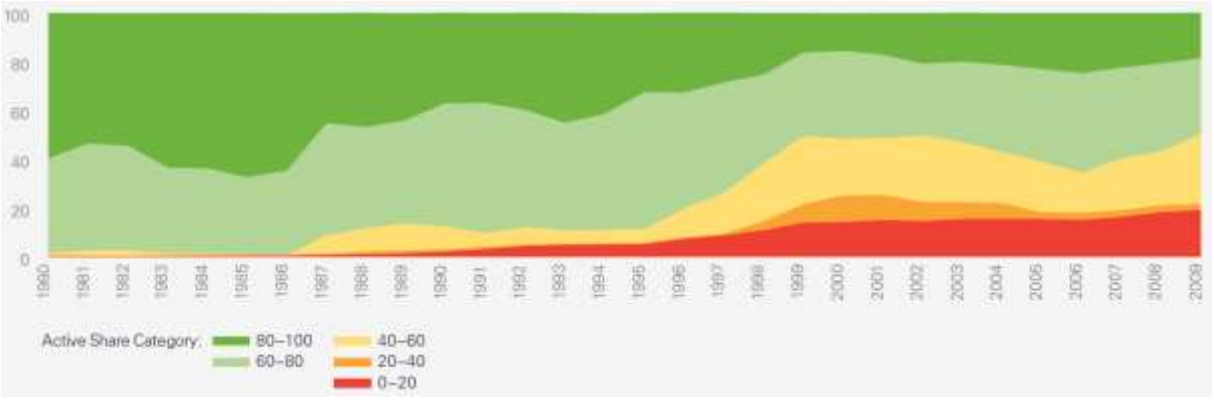
1. Introduction

This chapter provides an overview of the research. The background shows the necessity and urgency for such study in this setting related to active share. Based on such, objectives are presented and research questions are formulated. Finally, research limitations are introduced.

1.1. Background

Why do investors invest in actively managed mutual funds (hereafter mutual funds) as costs involved are significantly higher compared to similar investments such as exchange traded funds (hereafter ETFs)? The only logical explanation would be that mutual funds can outperform ETFs through actively picking investments outside a benchmark, which in turn justifies higher costs (Stalter, 2014). However, highly active mutual funds experienced a sharp decline in appearance while less active funds experienced a rise in appearance in recent years. This development can be seen in the graph below:

Graph 1 - Rise in Closet Indexing



Source: Petajisto, 2009; Khusainova & Meir, 2014

Through active share, a new measurement invented by Cremers & Petajisto (2009), fund (managers) activeness can be tracked, providing the possibility to select best performing mutual funds, making this measure crucial for investment decisions.

As a reaction to this research, European regulators are currently discussing whether active share disclosure should be required in order to provide more transparency. (Evans, 2014) This would provide investors with the possibility to see whether a fund manager is actively engaged in investment decisions or only closely tracking an index, similar to ETFs. Inevitably, close index

tracking mutual funds provide returns similar to index returns but also have high fees which would be unjustified. In an interview published by Citywire, Iain Richards (2014), head of governance and responsible investment at Threadneedle, commented:

" [...]. Investors are paying active fees but getting [a fund manager] who hugs the index. Looking at the active share [level] breaks down the stocks and gives us an insight into the [active] investments."

Richards further stated that investors should have the right to know the active share of a fund as it would help them make a difference between active and passive investments and maybe more importantly whether their investment strategy potentially adds value. (McGagh, 2014)

This recent development however shows a crucial research gap since studies on active share were conducted solely on US based equity funds (Khusainova & Mier, 2014). Furthermore, as this topic is broadly discussed in international equity markets, this research at hand may confirm or refute previous studies as it will be conducted on German equity funds during a recovering market situation from 2010 to 2015. This provides a unique research base as no other study has yet been conducted on Europe's biggest economy. Moreover, no research has yet been conducted on a specific market situation solely. This setting provides the possibility to judge on the findings of Jones & Wermers (2011) who argue that active share largely depends on market situations. Additionally, different variables are commented on which have not yet been tested in a similar setting.

1.2. Purpose & Aims

This research has two main purposes which should fill the research gap explained previously. The first is split up in two parts; evaluating variables that influence active share and, as a result, putting these variables in context with performance to judge on their practicability to predict future mutual fund returns. Secondly, returns of mutual funds and ETFs are compared to adjusted benchmarks in order to conclude if actively managed mutual funds tend to create positive alpha.

The next step aims to make judgments on the suitability of management fee structures, where gross and net returns are being compared for ETFs and mutual funds. This provides insight into whether mutual funds, despite their higher fees, still qualify as suitable investments and outperformance is not significantly diminished through management fees. Generally, a difference between mutual fund and ETF returns is expected, as it is every active manager's goal to outperform, which can be seen as the cause of return differences. Reasons for this assumption are based on the findings of Cremers & Petajisto (2009) who argue that highly active mutual funds tend to outperform their benchmarks significantly.

Taking this as a basis, the aim of this research is to advise a fund investment strategy based on the concept of active share, involving ETFs and mutual funds, for economic recovery periods in German markets.

1.3. Objective & Research Questions

As this study aims at finding variables describing active share and drawing conclusions on its usability to predict returns on which basis an investment strategy is conducted, the research questions will help to guide this research. This guarantees that all information needed is gathered and analyzed in this paper to answer the necessary questions. Please find the main question, which is split in two parts, and related sub-questions below:

Main research question:

1. How are tested variables influencing active share and is active share along with other variables tested useful to predict future fund returns?

Sub-questions:

2. How related are fees to the activeness of mutual funds?
3. How can active share be used for a fund investment strategy and for predicting returns?
4. How significant is the difference in returns between mutual funds and ETFs?
5. How do mutual funds and ETFs perform compared to their respective benchmarks?

1.4. Research Limitations

As with other research, this study has a limited scope due to different reasons. Limitations are given primarily due to the short time frame (two months) in which this study had to be conducted. Additionally, a lack of historical data availability exposed severe difficulties. This drove the decision to focus on the German equity market for mutual funds and ETFs solely. As Thomson Reuters Eikon (hereafter Eikon) only displays historical returns for funds dating five years back, this was chosen as the recovery period following the 2008 financial crisis. Consequently, it is difficult to generalize the behavior of active share and returns to other markets and economic environments. The overall data this research is based on had therefore to be taken for the same time frame. Hence, the recovery period will partly be influenced by the Euro crisis where the German economy was only impacted marginally compared to other economies Broyer et al. (2012). This qualifies researching Germany over other European economies during recovery after the 2008 financial crisis.

Due to necessary restrictions such as evaluating only on accumulating share classes, explained in detail in chapter 4.2., another limitation presented itself to be the small sample size, existing of only 41 mutual funds and six ETFs. This also limits transferability. However, the most crucial limitation is the impossibility to make judgements on changes in active share over time, as data for active share could only be extracted for the final date of the studied time frame. Hence, a cross-sectional analysis was conducted instead of a panel data analysis. Despite of this drawback, the validity of the results is regarded not to be influenced, as Cremers & Petajisto (2009) found that active share tends to show consistency over time.

All in all, the given limitations make it difficult for this study to compare the results to Cremers & Petajisto (2009). As their paper presents the only existing benchmark for this research, comparisons were conducted with tremendous care as different market conditions, time frames, and regression settings were studied. Additionally, it is difficult to review historically proven theoretical frameworks, as active share is such a new phenomenon.

2. Theoretical Background

This section gives a brief introduction of mutual funds and the novelty of active share as well as ETFs. Attention is drawn on the features both investment vehicles carry as well as active share, management and fee structures involved.

2.1. Mutual Funds & Active Share

Investors deciding to invest in mutual funds instead of ETFs believe in mispriced securities. Hence, fund mutual fund managers are expected to identify and invest in such to outperform the stock market (Stalter, 2014). This leads to the conclusion that the investment policy, tacking error, and active share are highly important facts and measures. They give indications to investors how flexible a fund can invest compared to their benchmark to primarily generate positive alpha by investing in mispriced securities. Furthermore, restrictions on how often and how many assets can be reallocated differ significantly between mutual funds. Another difference of mutual funds compared to ETFs is that mutual funds are priced only once a day after market close which makes them tradable only once a day (Ferri, 2008). This means that the investor will have to wait until day-end for the net asset value (hereafter NAV) before they will know the purchase or sales price. Furthermore, because of the fact that a mutual fund's manager is actively involved in investing, the fees to compensate for active management are usually higher compared to ETFs. As a consequence of illiquidity and fees related to mutual funds, a long-term investment strategy is best when investing in mutual funds. (Stalter, 2014)

Additionally, mutual funds usually have different share classes. These do not correspond to voting rights as they do for stocks, but to differences in fees charged and/or initial required investment subscription as well as dividends being distributed or retained etc. Moreover, the share class system enables fund managers to attract a wider range of investors with different preferences. (Morgan Stanley, 2015) Especially the dividend feature needs to be taken into consideration when identifying the relevant share classes of each fund taken into the sample.

All mutual funds are managed according to their unique investment policy that is required to be disclosed in Key Information Documents for Investment. This policy states how much managers may be allowed to reallocate assets. Furthermore, mutual funds must state an

objective[1] benchmark to compare performance and returns. (European Commission, 2012) As policies differ and benchmarks must be objective, they vary across funds. Unfortunately, the investment policy alone does not provide deep insight into active vs passive investments.

Here, active share comes into play, presenting activeness through a simple percentage number. As already explained, active share was invented by Cremers & Petajisto in 2009. This measurement shows the share of portfolio holdings that differ from the benchmark at a certain point in time. In other words, it shows the investment flexibility of a mutual fund. Per definition, the only possible way for a fund manager to outperform the benchmark is by investing in securities outside the benchmark or by reallocating investment weights. This differentiation is captured when calculating active share. In general, the higher the active share the more potential a fund managers sees in actively picking stocks. (Cremers & Petajisto, 2009) A detailed introduction into active share can be found in chapter 3.

2.2. Exchange-Traded Funds

ETFs are passive fund investment solutions traded on stock exchanges. Those count as the most popular investment type in the exchange-traded product (ETP) group. Typically, ETFs are benchmarked against indices or any basket of securities or goods. They can hold different assets such as commodities, bonds or equities. According to Richard A. Ferri, (2008) ETFs build upon the advantages of traditional mutual funds but provide the investor with higher trading flexibility, carry lower operating fees for the investor, offer more transparency, and especially tax advantages. However, there are drawbacks as well including, but not limited to, the learning complexities of products as well as the trading costs involved. In the following, these benefits and drawbacks are being described shortly on the basis of Richard A Ferri's "The ETF Book". (Ferri, 2008)

The first big advantage of ETF investments lies within the possibility of flexible trading. This means, similar to stocks and bonds, ETFs can be traded intraday where prices are updated constantly. This adds attractiveness to investors who are interested in liquidity. Liquidity might be crucial for umbrella funds as it gives the managers the possibility to quickly implement updated portfolio investment policies. The second advantage is related to risk concerned

[1] Objective in this case refers to a reasonable benchmark, e.g. a German equity fund should not use the S&P500 as benchmark.

investors who quickly want to change exposure to a certain market or industry. This is possible with long or short ETF investments due to their great investment diversification. However, it is also interesting to an investor that might be restricted to investments in a certain area where ETFs can provide him with the solution. The third advantage important to mention are ETF related tax advantages. This feature, however, is disregarded herein as it is not the purpose of this research.

The fourth and probably most important advantage when deciding upon ETF investments are significantly lower fees as compared to traditional mutual funds. Costs can occur in different forms such as management fees, administrative expenses, distribution fees etc. Ferri (2008, p. 61) comments on the returns related to fees as; "the lower the costs of investing in a fund, the higher the expected return on that fund". The lower costs for ETFs are due to lower occurring expenses as these funds neither provide and nor require extensive client services and staff related expenses. Moreover, they are not required to provide investors with regular information. Lastly, ETFs usually do not carry any redemption fees.

Despite the advantages, ETFs also have disadvantages. Here, especially brokerage commissions play an important role that negatively affects investments return. Thus, it would make more sense to make a large one time investment compared to several small investments. Additionally, ETFs would never positively contribute to an investor's alpha in a certain market, as market movement is exactly reflected. Moreover, investors should keep bid and ask spreads in mind and should carefully evaluate the expense ratios as not all ETFs are low cost investments.

Tracking error is potentially the most well regarded measure for ETFs. It accounts for the differences between fund and the benchmark performance. However, indices do not hold cash while ETFs do, due to liquidity reasons. Moreover, dividend matching is difficult especially when the fund is reinvesting dividends. All of which automatically creates tracking error. (Ferri, 2008, p. 70)

Another problem for investors might be the settlement period. Hence, it can take time for an investor until cash is actually invested or until the investor receives back his investment when selling. According to Richard A. Ferri (2008), this can take up to three days depending on their investment policy, whereas mutual funds usually settle the next day.

2.3. Agency Theory

Another interesting theory to address in this paper is the agency theory developed by Jensen & Meckling (1976). In their article the authors regard a firm as a set of contracts and shareholders as the principal of a company, hiring agents, the managers, to manage the firm. However, each individual is pursuing their own goals. Such might differentiate dramatically due to the fact that both parties have different incentives. Thus, problems stemming from asymmetric information might occur, which corporate governance is applied to resolve.

Many follow up articles used agency theory as their basis for research in the field of corporate governance and the principal agent conflict. This is potentially interesting for funds as well, where research is rarely conducted in this context. However, it is interesting to see that smaller firms, due to less monitoring, tend to have higher management flexibility (Garcia-Teruel & Martinez-Solano, 2008). A relationship between size and fund flexibility in terms of active share was already briefly introduced by Cremers & Petajisto (2009), but not related to agency theory yet. If this relationship holds for funds as well, it makes it highly interesting to look at when pursuing the goal of implementing an investment strategy based on active share, as smaller funds can under these terms assumed to be more flexible.

2.4. The Utility of Wealth

"The Utility of Wealth" is a theory developed by Friedman & Savage (1948) and was revised by Harry Markowitz in 1952. The theory describes three different investor types, the risk-averse, risk-neutral, and risk-seeking investor. Each investor has a different view on risk. A risk-averse investor has more utility when entering lower risk investments, with the potential for lower gains, while a risk-seeking investor enters high risk investments with the potential for higher gains, as more utility arises (Markowitz H. M., 1952). Although these findings cannot describe every investment behavior for recent market developments (Rabin & Thaler, 2001), the theory helps to give insights into investors behaviors driving their decisions in investing in mutual funds or ETFs. Therefore, it is of use when developing a mutual fund and ETF investment strategies for the private investor, related to active share, in the end of this paper.

3. Literature Review

The literature review section introduces the deeper relevance of studying active share. This provides reference for a clear understanding of the importance of further research on this topic. In addition to the introduction of the active share, literature, and empirical findings, previous research papers and evaluating articles are being elaborated on in order to provide insight on the developments regarding active share. Based on the theoretical and empirical findings evaluated, relevant hypotheses are developed. Such are furthermore set into context to the theory presented in chapter 2.

3.1. Active Share

The active share measurement was first introduced by Cremers & Petajisto in their article "*How Active Is Your Fund Manager? A New Measure That Predicts Performance"*, published on March 31, 2009. The article focusses on a new measure for mutual fund activity, as they argue that tracking error alone does not provide a sound basis for fund investment decisions. As the study was presented in 2009, active share can be regarded as relatively new but increasingly relevant for mutual funds.

Their study is especially concerned with two key factors of how fund managers, through differentiation from their benchmark, try to outperform. Such factors are defined as:

- Stock selection – picking individual stocks managers expected to outperform
- Factor timing – involves time-varying bets on systematic risk factors such as entire industries, sectors, generally any systemic risk related to benchmark

To explain these two methods of benchmark differentiation, tracking error as a stand-alone measure, capturing only time-series standard deviation of return differences, does not suffice as it only provides insights into factor timing bets, not stock selection. Tracking error is calculated through excess returns, in accordance to Cremers & Petajisto (2009):

Equation 1 – Tracking Error

$$Tracking\ Error = Stdev\left[\varepsilon_{fund,t}\right]$$

$$R_{fund,t} - R_{f,t} = \alpha_{fund} + \beta_{fund}\left(R_{index,t} - R_{f,t}\right) + \varepsilon_{fund,t}$$

Given these thoughts, using tracking error alone does not disclose much about active investments as it only regards return differences. Hence, funds and related benchmarks must be compared to display a measure for activity in terms of stock selection along with factor timing. This can be achieved by comparing their single line weightings which is defined as[2]:

Equation 2 – Active Share

$$Active\ Share = \frac{1}{2}\sum_{i=1}^{N}\left|w_{fund,i} - w_{index,i}\right|$$

Active share, as presented in Equation 2, is divided by two as it captures both, the overlaps with the benchmark and the stock weightings. If not divided by two, the active share could be at 200%. Hence, a fund with zero overlap to the benchmark gets a 100% active share.

As sample, the authors used all-equity funds over a time period of 100 days to compute tracking error. The benchmark indices were selected from their stated benchmarks but more close benchmarks were selected to compare against where active share showed a lower result. This cannot be confused with selection biased as it only helps to identify the best fitting benchmark of a fund and is taken upon for the regressions in this study. Please find more information on the benchmark adjustments in chapter 4.2.

As a first measure, Cremers & Petajisto (2009) compared active share and tracking error, showing their positive correlation. Moreover, influences of the following variables on active share were tested:

- Endogenous variables, controllable by fund managers:
    - tracking error (closely related)
    - turnover ratio (neither statistically nor economically significant)
    - expense ratio (significant, economically weak significance)
    - number of stocks (only statistically significant)

[2] $W_{fund,i}$ and $w_{index,i}$ are portfolios weights of asset i in the fund and index respective

- Exogenous variables, beyond managers control:
    - fund size (related to active share, nonlinear and weak)
    - fund age (longer age, lower activeness)
    - manager tenure (longer tenure, higher activeness)
    - prior inflows (does not really matter)
    - prior benchmark returns (statistically significant)
    - prior benchmark-adjusted returns (related to active share, successful managers chose more active share)

This regression yields the outcome that active share is difficult to explain through these variables but was found that active share is steady over time.

To make comments on another hypothesis being higher activeness yielding higher gross returns, Cremers & Petajisto (2009) conducted another regression to evaluate insights of active share on gross returns. The results show that funds with low active share and high tracking error tend to perform worst and the best performer are concentrated stock pickers, followed by diversified stock pickers. This regression is seized by this study and used to draw precise conclusions when establishing the fund investment strategy, helping investors to identify the potentially best performing fund.

The final outcome of Cremers' & Petajisto's (2009) study suggests that investors should pick funds based on 3 measures: Active share (min. 80%), fund size, and prior one year return.

Compared to their study, this paper examines German equity funds between 2010 and 2015. Thus, the developments during the economic recovery after the 2008 crisis are examined, while Cremers & Petajisto (2009) focused on earlier developments between 1992 and 2003.

In a second article published in July/August 2013, named *"Active Share and Mutual Fund Performance"*, Petajisto elaborates on this first paper about active share. He argues that paying for active fund management does on average not turn out to yield higher returns, as many mutual funds underperform ETFs net of fees. The performance patterns found previously also apply during the 2008/2009 financial crisis and within market-cap styles. In addition to his first paper, Petajisto (2013) evaluates active of over time for the famous Fidelity Magellan fund with its

well-known manager Peter Lynch. He finds that under Lynch's management, the fund started very active. However, when the fund managers were replaced, active share levels changed. This finding is also taken on in this study to find results for different management tenure related to active share and performance.

Besides this practical application of active share on the Magellan fund, trends in closet indexing[3] are discussed. Petajisto (2013) sees the reasons for an increase in closet indexing funds between 1981 and 2009 by the SEC regulations introduced in 1989 that mutual funds are required to disclose a benchmark. This is regarded as a favorable development but might make managers afraid of yielding returns below benchmark, decreasing their investment activity in mutual funds and thus decoying the private investor who beliefs being investing in an active fund while pays higher fees for activity.

Based on his new findings, a 5x5 grid was developed, labelling mutual funds according to their active share and tracking-error quantile. From these five categories, stock pickers tend to have the highest average active share ratio, followed by concentrated funds with the second highest. Still, when Petajisto (2009) examined return consistence, he found concentrated funds to yield the most stable returns. This investment grid is used as an example for establishing an easier to understand investment strategy for a private investor, based on the findings of this study. Please find the overview of the grid in Table 1:

Table 1 - Different Types of Active Management

	Tracking Error Quintile						
Active Share Quintile	1 (low)	2	3	4	5 (high)	Group	Label
5 (high)	5	5	5	5	4	5	Stock Pickers
4	2	2	2	2	3	4	Concentrated
3	2	2	2	2	3	3	Factor Bets
2	2	2	2	2	3	2	Moderately Active
1 (low)	1	1	1	1	3	1	Closet Indexers

Source: Petajisto, 2013

Concluding his paper, the same results were found in his previous paper as inefficiencies in the markets were found which can be exploited by active stock selection. However, on average

[3] Closet indexing refers to funds that are more or less invested in the index benchmark, thus not active.

mutual funds underperform passive investments net of fees and only funds with the highest levels of active share outperform. This still requires confirmation in different market settings.

3.2. Mutual Funds and their Performance

In the study "*The Difficulty of Selecting Superior Mutual Fund Performance"* conducted by McGuigan in 2006, the author tries to quantify the relative performance between domestic stock actively managed mutual funds and passive funds. The study was performed over a 20 year period on funds that hold 85% or more in North American equities. For the 171 funds data was available back to 1983 where the Vanguard 500 index met the requirements to work as benchmark. The study was conducted for 5, 10, 15 and 20 year timeframes from 1983-2003. McGuigans finds that, the longer the timeframe, the harder it was for active managers to outperform the benchmark index. Concluding, as the probability of finding a superior fund is low, it is rather expensive to make investments into actively managed funds. Throughout the entire period only a few funds outperformed the benchmark at all times. Anticipating which fund shows this behavior is very difficult.

Thus, McGuigan proposes that the majority of a portfolio should be invested in passively managed funds with both, mid- and large-cap focus, as he follows the theory that only a few funds outperform the market. This argument is investigated further in this study and compared to the findings for active share.

Another study, named *"Another Puzzle: The Growth in Actively Managed Mutual Funds"*, published in 1996 by Gruber, the importance of mutual funds is examined and the reasons for holding mutual funds are analyzed. Gruber examines the performance of actively managed vs passively managed funds against an appropriate index between 1985 and 1994. The performance is measured applying three different factors (Gruber, 1996, p. 785):

1. A measure of return relative to the market
2. The excess return from a single index model
3. The excess return from a four index model

Although Gruber (1996) uses all three measurements, he puts most emphasis on the third as: "[...], failure to include indices that span the major types of securities held by the fund during study can lead to incorrect conclusions about performance. The indices selected for the four

index model span the major types of securities held by nonspecialized domestic stock fund, the type of funds examined in this study." (Gruber, 1996, p. 787)

His findings, using a single index model, are that the risk adjusted return, on average, is estimated to be – 1.56% compared to its index taken as benchmark. However, using the four index model instead, an underperformance of negative 65 basis points is estimated. Gruber (1996) furthermore estimates that an actively managed mutual fund on average underperforms the market by 1.94%. This underperformance happened as mutual funds did not relocate their assets during this time, staying investing in underperforming corporations. Thus, his proposed adjusted index model can catch mutual funds alphas better in terms of being comparable. This confirms the usage of an adjusted benchmark as described by (Cremers & Petajisto, 2009) in order to draw conclusions on active share and is also used for this study.

In closing, Gruber (1996) too states that investor should invest in ETFs as, despite the higher exposure to systematic risk, mutual funds systematically underperformed all benchmark models. As the four factor model is difficult to apply for the individual investor, this paper tries to find a more simplistic benchmark model for the evaluation of active share.

3.3. Evaluating Active Share

In the paper *"Active Management in Mostly Efficient Markets"*, Jones & Wermers (2011) try to answer the following three questions regarding active management:

1. Does active management add value?
2. Can we identify superior active managers ex ante?
3. How much active risk should investors include in their portfolios?

To answer the questions, the research refers to Cremers' & Petajisto's (2009) active share study. Jones & Wermers (2011) find that their results are not entirely consistent and that conclusions often depend on the methodologies' studied period in terms of economic development. Moreover, the authors find that mutual funds net of fees historically do not capture positive alpha.

The study furthermore finds that during the 2008 financial crisis, active equity management has "[...] been a bit more favorable [...]". (Jones & Wermers, 2011, p. 31). Through their study they

develop the theory of the superior active managers (SAMs) performs better than the inferior active managers (IAMs). This was already found by Ding & Wermers (2009) who found that experienced managers outperform less experienced by 92 basis points per annum. SAMs can, according to their definition, display any kind of fund manager that, according to certain measures, provides superior fund returns over other fund managers. Hence, an SAM is not solely defined as being more experienced or older or someone who is more tenured.

Despite of these findings, Jones & Wermers (2011) argue that during periods where SAMs outperformed the market, they are likely to underperform in the next period, following the mean reversion pattern in line with Bogley (2009). This concludes that when managers had weak performance during one period, they are more likely to be replaced or change their strategy, resulting in better future performance.

As a concluding statement, active management and thus active share can be related to outperformance in times financial crises, but does not necessarily need to hold during market recovery. Identifying higher activeness and SAMs improve the overall Sharpe ratio significantly according to the authors. Hence, the ratio is therefore added in the regression of this study as an independent variable.

Another evaluation on active share was conducted by Khusainova & Mier in their 2014 article "*Taking a closer look at active share"*. The authors find active share as providing additional enhancement over traditional tracking error, being in line with related literature. They find that funds applying stock selection tend to invest in a smaller amount of individual stocks related and larger amount of individual stocks being unrelated to the benchmark. Fund managers using factor timing base the weighting on their view on the systematic economic risk. Here, they mention the distinct advantage of active share over tracking error, represented through ability of displaying the effects of stock selection.

Again, Khusainova's & Mier's (2014) results also show that actively managed funds tend to underperform ETFs net of fees. This phenomenon is referred to as "closet indexers" again, similar to Petajisto (2013). Thus, these funds act similar to ETFs while charging higher fees. Active share is thus a good instrument to identify such funds. The cutoff of an active fund is set at 60% in their study, stating that an active share below this level has the potential to be a "closet indexer". Moreover, the authors find that there is no relationship between fund fees and active

share, contrary to Cremers & Petajisto (2009). The "closet indexers" however show fees below the average, but are still more expensive compared to investments in ETFs to their study.

In closing the authors further stress that, since Cremers' & Petajisto's study is based on active share for US based mutual funds, solid conclusions for other markets are difficult, underlining the importance of this study. Although results could be different as also anticipated by Jones & Wermers (2011), Khusainova's & Mier's (2014) argue for the use of active share as it provides valuable information for investors.

3.4. Summary & Contribution

In this section, the examined theory is being assessed and further related to this research in order to understand the differences and need for further research, thus, underlining the contribution of this paper.

As this research is closely related to active share, which is a rather new measurement, Cremers' & Petajisto's (2009) introduced research is of crucial importance to this research as it represents the basis for active share. This research therefore builds upon their thoughts and findings but under a different economic impact, namely economic recovery as suggested by Jones & Wermers (2011). The recovery period is tested solely for the Germany equity market in this study, also providing the first information of mutual fund active share in a European setting. Additionally, not all variables are treated similar because of two reasons as other variables are assumed to also influence active share and/or mutual fund returns. Thus, this research introduces the Sharpe ratio to previous research, picking up again on the research of Jones & Wermers (2011). Moreover, the cross-sectional study conducted in this paper is valid, as Cremers & Petajisto (2009) find that active share is highly consistent over time. Thus, active share is only taken at a specific point in time.

Also Petajisto's (2013) second research on active share focusses solely on the US mutual fund market. Hence, this research can be seen as a validation or rejection of his findings. Based on his 5x5 investment grid, this paper uses his findings and the new findings found for the German recovery market to conduct a new investment criteria table valid for this setting.

Additionally to the fact that Jones & Wermers (2011) basically stress to test active share under different economic influences, they also argue that active management is favorable during

economic distress. Hence, this paper provides information if active management is also favorable during economic recovery.

Additionally, this research takes upon the conclusions that SAMs for German equity funds can also be indicated through holdings and past performance. Contrary to McGaugans (2006) argument that it is difficult to find good actively managed mutual funds based on available data, this research tries to find present suitable variables to choose above average mutual funds. This would provide the possibility to increase the chances of picking a mutual fund that is likely to outperform the market.

Compared to Gruber's (1996) research, this paper takes on his thoughts of benchmarks being unreliable but does not use his 4-index model. Instead, benchmarks are selected similar to Cremers & Petajisto (2009) who propose that the best benchmark to compare performance to, is the one that yields the lowest active share as explained in chapter 3.2.

As Khusainova & Mier (2014) argue that due to the positive correlation of tracking error and active share, together they present a good measure for investment decisions, also found by Cremers & Petajisto (2009). This relationship will be tested for the German market and taken into thought for the proposed investment strategy in a recovering market setting.

3.5. Hypotheses Development

Based on the presented literature and the empirical findings introduced in this chapter, hypotheses were developed to be tested through statistical regressions. Firstly, this gives the possibility to make judgments on whether historical empirical findings are actually in line with the new findings. Secondly, it gives the possibility to make judgments on the likelihood of relationships between variables (Robson, 2002). The hypotheses are developed on the common model of H_0 and H_1. H_0, the null hypothesis, assumes that the variables tested are not significantly different from zero. H_1, the alternative hypothesis, assumes that there is a significant difference from zero. The significant difference is found by evaluating the p-value of the regression outcomes which is based on the t-statistics, indicating a two tailed testing for each conducted regression. However, not only the statistical significance is important, but also the economic significance or impact (Ziliak & McCloskey, 2003) and such are therefore also captured in the analysis. The economic significance is represented through the magnitude of the coefficient, introduced in chapter 5.

In the following, the hypotheses this paper is evaluating are presented in the explained form:

Hypothesis 1:

H_0: The activeness of a mutual fund is random and future activeness cannot be predicted through another variable.

$$H_0 = \beta_{independent\ variable} = 0$$

H_1: The activeness of a mutual fund can be predicted through other variables and is therefore significantly different from 0.

$$H_1 = \beta_{independent\ variable} \neq 0$$

Hypothesis 2:

H_0: Fees, measured as total expense ratio (TER), are not related to the funds active share.

$$H_0 = \beta_{TER} = 0$$

H_1: Fees, measured as total expense ratio (TER), are related to the funds active share.

$$H_1 = \beta_{TER} \neq 0$$

Hypothesis 3:

H_0: During recovery, there is no significant different between active fund and market returns.

$$H_0 = \beta_{adjusted\ alpha} = 0$$

H_1: During recovery, there is a significant different between active fund and market returns.

$$H_1 = \beta_{adjusted\ alpha} \neq 0$$

Hypothesis 4:

H_0: On average, EFTs and mutual funds perform the same gross of fees during recovery.

$$H_0 = \beta_{ETF\ return\ (gross)} = 0$$

H_1: On average, ETFs perform significantly different to mutual funds gross of fees during recovery.

$$H_1 = \beta_{ETF\ return\ (gross)} \neq 0$$

Hypothesis 5:

H_0: On average, EFTs and mutual funds perform the same net of fees during recovery.

$$H_0 = \beta_{ETF\ return\ (net)} = 0$$

H_1: On average, ETFs perform significantly different to mutual funds net of fees during recovery.

$$H_1 = \beta_{ETF\ return\ (net)} \neq 0$$

4. Methodology

The methodological framework uses Saunders' et al. (2009) *"research onion", to establish the research philosophy and approach, as well as to present strategy, time horizon, and techniques used. Additionally, the data collection method which includes data delimitation, sample selection, choices of variables and regression models is introduced. In the data analysis section, reliability, validity, and robustness measures are explained and approaches presented by Brooks* (2008) *were used to guarantee the unbiasedness and efficiency of the regression model.*

4.1. Research Design: Philosophy, Approach, and Strategy

As this research makes use of hypothesis testing, a positivist research philosophy is adopted. The hypotheses tested were established based on theory and empirical findings presented in chapter 3. It is of particular interest to see whether the existing empirical findings can be confirmed or rejected following this methodology. According to Saunders et al. (2009), a positivist philosophy is a value-free way, meaning that the researcher has no influence on the sample selected. This value-free philosophy can also be referred to as axiology where "the researcher is independent of and neither affects nor is affected by the subject of the research" (Remenyi et al., 1998, p. 33). Moreover, as the study rather focusses on facts and observations than on the researchers' opinions, a positivist philosophy is pursued.

The research approach is of deductive nature as theoretical frameworks and empirical findings were used to establish conclusions on the selected sample. Thus, focus lies on drawing conclusions towards the latter, based on facts that can be measured quantitatively and are therefore, according to Saunders et al. (2009, p. 125) "operationalized". However, as this research does not strictly copy an existing paper and the regression variables were adjusted based on the researchers gained experiences, the inductive approach plays a subordinated role. According to Saunders et al. (2009) it is perfectly reasonable to combine these two approaches and often this proofed to be beneficial.

As this study examines the "casual" links between variables, the strategy is of experimental nature. This was also the basis for selecting an appropriate sample from the known population of all German equity mutual funds and ETFs. A mono method quantitative data collection technique is applied on this sample as mainly Eikon or the respective funds annual reports and prospectus' were used to gather data. This data was analyzed through a linear multiple

regression model. Despite a chosen time horizon of 5 years (2010 – 2015) for comparing returns of all mutual funds, ETFs, and indices, Saunders et al. (2009) refers to this as a cross-sectional study as the focus lies on capturing a specific point in time for many of the variables. Thus, five year averages, fitting the time horizon, were taken for return calculations and other variables were used in the regression. This shows a longitudinal study approach.

4.2. Data Collection Method

4.2.1. Data Set

The data set used can be referred to as "probability sampling" (Saunders et al., 2009, p. 214). This means that the sample is selected in a way that all possible research questions can be answered and objectives can be achieved as the sample represents the population perfectly. Thus, statistical results should yield the same for the whole population.

Due to the scope of this research, the sample selection was limited to German equity mutual funds and ETFs. The study period chosen represents the financial recovery period after the 2008/2009 financial crisis. The exact time frame studied is 2010 – 2015. The reason for choosing this time frame is to study only recovering market phenomenon's as different results are expected compare do Cremers & Petajisto (2009), based on the findings of Jones & Wermers (2011). 41 mutual equity funds and 6 ETFs build the sample, which might be considered a small data set. However, due to necessary restrictions discussed later, the sample reflects the whole population for this study, putting it in line with "probability sampling". Because Eikon does not provide historical data older than five years, this study is limited to this period. As a result of this, it is arguable that part of 2009 can be considered recovery period as well and that the period is influenced by the European debt crisis which was already described in chapter 1.4.

In order to select fitting asset classes for mutual funds, restrictions had to be made in Eikon. The fund screener function in the Eikon provided the possibility to receive the whole population for this research for both mutual funds and ETFs. For both, a similar approach was used and is explained in the following:

Mutual Funds

The restrictions in Eikon were set to Germany as the geographic range. As security type, German equity funds were selected and income distribution was set retained. The retained income distribution is crucial in this case, as all funds need to reflect performance the same way, including dividends. This is guaranteed only when taking accumulating share classes. Moreover, only currently active mutual fund share classes were selected to make sure that actual investments would be possible to investors. These restrictions showed 86 funds. From this, all leveraged and short funds were removed in order to investigate pure long investing mutual funds only, and to guarantee comparability. This states the whole population for the research. Moreover, all restricted share classes such as institutional share classes with high subscription limit were removed based on the input found on the fundinfo website (2015). This leaves one share class per fund in the mutual fund sample, which exists of 41 funds in total.

ETF's

The same criteria selection was conducted when approaching the German ETFs. The only difference in the criteria entered into Eikon was to change the selection mutual funds to ETFs. This provided 14 ETFs in total. Again, all short and leveraged ETFs were deleted to receive a valid population, which is comparable to the mutual fund population. Moreover, only DAX30 tracking ETFs were taken into the final sample in case multiple management companies were displayed. Applying this provides the total sample of 6 ETFs. Although the sample size of 6 ETFs might seem small, all showed to yield approximately the same results in terms of returns as they all track the DAX with slightly differentiating tracking errors. Hence, as only average returns are taking into consideration and are being compared, the small sample size would be of no consequence.

4.2.2. Data Collection

All data collected in this study is secondary data retrieved from reliable sources. The process of locating and collecting the data is, according to Saunders et al. (2009), primarily used to establish and define the type of data needed. Secondly, the precise data needed was located.

The first data collected is the net asset value (hereafter NAV) for all mutual funds and ETFs, retrieved through Eikon. The NAV is used to calculate the monthly and annual returns of the funds. The returns display ln returns. This is advantageous as ln returns empirically and statistically are more convenient as they are more likely to be normally distributed which

increases the comparability. (Brooks, 2008). Furthermore, all benchmark data is retrieved through Eikon. This data is needed to calculate relative fund performance (alpha) and the active share to their best fitting benchmark, decreasing the active share (Cremers & Petajisto, 2009). Other independent variables were also collected using Eikon. However, if data was not available through Eikon, the data was calculated manually from the annual reports or Morningstar according to the applicable formulas. The independent variables for which data is used are active share, tracking error, turnover ratio (hereafter TR), total expense ratio (hereafter TER), number of invested securities, latest fund value (hereafter TNA), Sharpe ratio, management tenure and performances gross/net as well as compared to adjusted benchmark (alphas). Different performance measures are used as dependent and independent variables in this research.

In the following, these variables are being explained in more detailed and if variables needed manual calculation, the calculation steps are elaborated. Furthermore, please note that these variables are divided up into two groups, endogenous and exogenous. Endogenous variables can be affected by the fund manager, while exogenous variables cannot. However, these terms of endogenous and exogenous variables are not related to OLS violations and are therefore not causing endogeneity problems.

Active Share

Active share was calculated according to Cremers & Petajisto (2009) using their active share formula presented in equation 2 in chapter 3.1.

As fund companies usually do not disclose active share, active share was calculated manually. To calculate active share manually, the required data was retrieved from the fund's annual reports where all invested securities with ISINs and holdings weightings are listed. Following, these were matched against both, the displayed fund benchmark taken from the fact sheet and the CDAX. The benchmark holding's ISINs and weightings were retrieved from Deutsche Börse Group (2015) for the respective DAX indices. The data for the MSCI indices was collected through the MSCI support which was contacted at clientservice@msci.com. Using the active share formula as explained in chapter 3.1., each fund's active share was calculated. In case the active share displays a lower result for the CDAX, this was taken as the more appropriate benchmark, following Cremers' & Petajisto's (2009) approach of finding the better suited benchmark and being in line with Gruber (1996). This is due to the nature of active share

measurement showing that the fund is closer invested to the CDAX. Additionally, active share was only taken at the most recent point in time as only the last annual report was available. This is appropriate as Cremers & Petajisto (2009) find active share to be stable over time, validates this approach.

Active share dummy

In order to explain active share further, a dummy variable for active share is also included in the regressions in model 2. This dummy variable is based on the median TNA, and is assigned a zero if below median and vice versa, following the same pattern as with the previous active share variable. This way a distinction can be drawn between active share influences for above and below median size total net assets funds, increasing the interpretation possibilities related to agency theory and fund size.

TE - Tracking Error

The five year historical average tracking error data was retrieved through Eikon, matching the research's time horizon. Tracking error in general is calculated as presented in chapter 3.1. in equation 1.

PTR

Portfolio turnover ratio (TR) is defined as the percentage a fund's holdings that have been "turned over" in a given year. In other words, the assets a fund replaces and invests in compared to the total assets during a given year, expressed in percentage. The data for turnover ratio was retrieved through the funds fact sheets where only the most recent annual data was used. As it is not required to disclose this measurement, the turnover ratio had to be calculated manually for several funds. This was done according to the PTR formula[4] which includes the redemptions and new investments data taken from multiple annual reports:

Equation 3 - PTR

$$PTR = \frac{\min(X, Y)}{M}$$

[4] min = minimum element
X = Value of purchases
Y = Value of redemptions
M = Average fund value

TER – Total Expense Ratio

The total expense ratio represents the costs an investor faces when investing in a fund. It was calculated in percent affecting the return and is needed in order to receive the net return of a fund. For example, if a fund shows a TER of 4% and total return of 7%, the net return would roughly be 3%. The formula for the TER is shown below:

Equation 4 - TER

$$TER = \frac{Total\ Fund\ Costs}{Total\ Fund\ Assets}$$

The data to calculate each fund's TER was found through Eikon.

Securities

Number of securities represents each single line a fund is currently invested in. The data for the total number of securities was taken from Eikon.

TNA – Total Net Assets

The total net asset value of a fund is basically the total market value of all securities the fund is invested in. This data was retrieved through Eikon.

Sharpe Ratio

The Sharpe ratio is a widely used ratio to proxy for the risk-adjusted return. Thus, the ratio basically provides the average return per unit of risk, measured through volatility, over the risk-free rate. There are drawbacks to the Sharpe ratio as it can be inaccurate for portfolios which do not show a normal distribution (Getfaoui, 2010). This might be the case for financial portfolios and thus for the selected funds because of the existence of fat tails which was found in empirical studies (Haas & Pigorsch, 2007). However, this will be disregarded and the assumptions of Markowitz (1952), that portfolios show normally distributed returns, in this paper are assumed to hold during the recovery period where volatility is expected to stay low.

Under these assumptions, it is valid to use the Sharpe ratio which is calculated according to the following equation[5] developed by William F. Sharpe (1994):

Equation 5 – Sharpe ratio

$$Sharpe\ ratio = \frac{\bar{r_p} - r_f}{\sigma_p}$$

The ratio was retrieved through Eikon. However, the Sharpe ratio was not available for the Allianz Thesaurus fund and was therefore taken from Morningstar.

Tenure

Management Tenure describes the time a fund manager has been managing a fund in total, expressed in years, months, or days. The data was retrieved manually by researching each manager separately in Eikon or Morningstar and a dummy variable was constructed.

Performances

As performance variables, different measures are used as independent and dependent variables throughout the regressions as explained in the beginning of this section.

The first variable is *adjusted alpha*, representing the gross out/underperformance to the adjusted benchmark. This variable describes the performance of a fund compared to its benchmark and is calculated the following way:

Equation 6 – Adjusted Alpha

$$Adjusted\ Alpha = \ Fund\ perf. - adjusted\ benchmark\ perf.$$

The performance towards the adjusted benchmark is taken in accordance with Cremers & Petajisto (2009) as the variable is used to explain active share.

As a second performance variable, solely acting as dependent variable, the *funds average gross LN return* is used. The monthly return data was collected through Eikon and the ln's were used in order to make the outcome more reliable (Brooks, 2008).

---

[5] r_p = Expected portfolio return
r_f = Risk-free rate
σ_p = Portfolio standard deviation

As the third performance variable, again solely acting as a dependent variable, the *adjusted alpha (net)* was used, which is the net performance compared to the adjusted benchmark. This can be calculated the same way as explained with Equation 6 net of fees. Thus, fees were deducted from the fund performance.

4.3. Regression

The regressions conducted in this study are divided up into two parts, each representing a different "model". The first model is evaluating which of the chosen variables have an impact on active share. The second model looks further into which variables represent good predictors for fund returns. Model 2 tests three different return types; the gross ln returns of the funds, adjusted alphas, as well as the adjusted alphas (net). The rationale behind the choice of regressions and the variables tested is further explained below.

Cross-sectional data is used as input to run the regressions. Thus, the underlying assumptions and the reliability of the data were tested in order to make sure the results are correct and reliable. According to Brooks (2008), cross-sectional data needs to be tested for the following violations of Ordinary Least Square (hereafter OLS) assumptions: Assumptions 2 and 4, potentially causing heteroscedasticity, and assumption 5, potentially causing non-normality issues. Furthermore, Brooks (2008) stresses the importance to test for non-linearity and multicollinearity. These results are being presented in Chapter 4.4.

4.3.1. Model 1 – Active Share

With model 1, the objective is to determine what influences active share. Active share defines the fund managers' activeness in this model and therefore represents the dependent variable in this regression. Factors influencing the active share are stated in detail below and are divided into two factors the fund manager can (endogenous) and cannot (exogenous) affect. Consequently, the regression formula for active share is the following:

Equation 7 – Regression Model 1

$$Active\ share_i = \alpha + \sum_{i=1}^{T} \beta\, A_i + \sum_{i=1}^{T} \gamma\ B_i + \varepsilon_i$$

Where *active share*$_i$ is presented in percentage of funds differences in investments compared to a benchmark, i denotes the regressions cross-section. A_i represents the influential or endogenous

variables tested for the same cross-section and country and B_i represents the non-influential or exogenous variables. ε_i denotes the regressions' error term. Please refer to Appendix I, Table A for the regression output for the active share regression.

In the following, the different endogenous and exogenous variables tested are presented.

Dependent Variable

Active share, the measure for fund managers' activeness, is evaluated in this regression in order to find out which variables have an influence and are reliable to predict active share. As active share measures both, stock selection and factor timing, it will be interesting to find out if other variables can also make predictions on these two approaches for investment differentiation compared to a benchmark.

Independent Variables

The first independent variable chosen is *tracking error*, which is a precursor to the active share and therefore potentially captures some of the effects of the latter. As argued in the literature review, active share does estimate the effect of stock selection better which is why it is interesting to see the impact of tracking error, not capturing this effect, on active share (Khusainova & Mier, 2014). To test this, tracking error is included in order to evaluate the relationship to active share and to see if the positive relationship found in literature is also present in this sample and during recovery. Furthermore, it is intended to see if, along with active share, the measurements are useful for investment strategies.

Total expense ratio was selected to provide a possibility to show investors if higher fees are verified and the manager is actively managing the fund or if fees are not verified as the manager is simply closely tracking the benchmark. Please refer to equation 4 for more information.

The *portfolio turnover ratio*, as presented previously, is another interesting variable and is included in order to make this regression more comparable to similar empirical studies. However, this was not the only reason for regressing turnover against active share and returns. Additionally, it was interesting for the researchers to find out if higher turnover shows higher/lower fund activeness or higher/lower returns as more transactions automatically increase the total transaction costs for a fund. Thus, funds might be reluctant to adjust their portfolios as their returns might be diminished or the fees need to be adjusted upwards to compensate for additional expenses.

The *number of securities* that the fund is currently invested in was included in order to see whether diversification of the fund assets can have an effect on the activeness of the fund manager. Following Cremers & Petajisto (2009), this variable was made use of in the active share regression in order to investigate possible transferability of their findings to the German market and especially to times of economic recovery.

The *gross return* variable represents the overall fund return before fees and is included in order to evaluate active share independent of a funds benchmark.

The *adjusted alpha* represents another independent variable. The reason is that investors might want to invest in mutual funds with higher fees as they expect these, through their activity, to capture more upside potential. Thus, it is interesting to make judgements on if active share is actually related to gross performance. Furthermore, this variable was also tested by Cremers & Petajisto (2009) and found to be significantly positively related to active share.

The *Sharpe ratio* measures performance above risk-free investments. Thus, it is a good measure to see if the additional risk engaged in with fund mutual investments yields appropriate returns. Furthermore, it would be interesting to know if the above risk-free return is correlated to active share.

A funds active share is definitely influenced by the fund manager, as he has to make investment decisions. Thus, *manager tenure*, also researched previously (Cremers & Petajisto, 2009), might be a good independent variable. This variable is the only variable presented as a dummy variable for regression model 1 and shows the following conditions:

0 – Fund only has one fund manager over the entire time horizon investigated

1 – Fund has more than one fund manager over the entire time horizon investigated

In the end it will be interesting to see if there is a relationship between active share and management tenure as found by Cremers & Petajisto (2009) where longer tenure is associated with higher fund active share.

4.3.2. Model 2 – Returns

Model 2 tries to establish the relationship between three different return measurements acting as dependent variables in separate regressions against a number of independent variables. The return variables are in resemblance to Model 1 also tested against a set of exogenous and endogenous variables. Consequently, the regression formula for the returns is the following:

Equation 8 – Regression Model 2

$$Returns_i = \alpha + \sum_{i=1}^{T} \beta\, C_i + \sum_{i=1}^{T} \gamma\ D_i + \varepsilon_i$$

Where $Returns_i$ is presented in percentage of a funds gross return, adjusted alpha, or adjusted alpha (net), and i denotes the regressions cross-section unit. C_i represents the influential or endogenous variables tested for the same cross-section and country and D_i represents the non-influential or exogenous variables. ε_i denotes the error term of the regression. Please refer to Appendix II, Table F, of the regression output for the returns regression.

In the following, the different endogenous and exogenous variables tested in this model are presented.

Dependent Variables

The *gross returns*, *adjusted alphas*, as well as the *adjusted alphas (net)* are in this model tested against a set of independent variables. The ln gross returns are important to look at in order to evaluate the returns, independent of the fund benchmark and, to see how the fund performed overall. This is because it provides a possibility, independent of out- and underperformance measures, to estimate how the selected independent variables affect returns. However, both, gross and net alphas of the adjusted benchmark were additionally regressed against the same independent variables to evaluate their impact on performance relative to the benchmark. The reason for including both is to see whether the size of the fees affects the significance or the impact of the independent variable. All these variables are also tested similarly in the paper by Cremers & Petajisto (2009). However, this paper also drops and includes certain variables as explained for Model 1 in chapter 4.3.1.

Independent Variables

For an explanation of the choice of independent variables, please refer to the independent variables presented in the previous chapter 4.3.1. These are the same in this model, except for the inclusion of *active share* as additional independent variable. Furthermore, this regression excludes the return performance variable, now used as a dependent variable. Taking active share as an independent variable is of great importance for this regression, as the aim is to explain the impact of the latter and find out if superior returns are a result of active fund management. However, it is also used to draw conclusions on whether the alphas of the adjusted benchmarks are related to active share. This regression was also run by Cremers & Petajisto (2009) which underlines the importance to include active share as an independent variable. The inclusion of the active share dummy variable relates to their findings as well, where a negative relationship between the size of the fund, TNA, and the active share is found. Moreover, it enables judgement on the assumption derived from agency theory that smaller firms tend to have more control over their operations (Garcia-Teruel & Martinez-Solano, 2008). Thus, they can act more independently which might equivalently be the case for smaller asset sized funds.

For all other variables and in contrast to model one, the dummy variables are included in this regression in order to explain returns. Based on this, it is possible to establish an investment strategy for funds investing in the German equity market which is also based on return predictability.

4.4. Data Analysis

This section addresses potential inconsistencies and violations of general OLS assumptions. Thus, it is presented how data is filtered in order to enable unbiased interpretation of results.

4.4.1. Validity and Reliability

As this paper is using secondary data from reliable sources such as Eikon, Morningstar, annual reports etc., the reliability of these is presumed. Thus, most of the data is what Saunders et al. (2009) refer to as "raw data". In other words, it represents data where none or only little processing is needed. This raw data has been translated into applicable measures following explicit formulas and theories as developed by the presented literature. With regards to the limited time of this study, the usage of secondary data has one big advantage, namely; saving

resources such as time and money. The data is also collected in line with the overall objectives and goals of this study and is therefore following the criteria set by Saunders et al. (2009).

As briefly introduced earlier, violations of commonly known OLS assumptions along with other issues have been resolved and controlled for. This provides the possibility to present unbiased results in the end and to draw correct conclusions. In Table 2, the tests and results for each of the potential violations are presented along with possible solutions. Cross-sectional data is exposed to heteroscedasticity, non-normality, multicollinearity and non-linearity (Brooks, 2008). The existence of the named potential issues has been tested and the outcome is presented in the following:

Table 2 - OLS Assumptions & Other Issues

OLS Assumptions	**Tests**		
1: E (ut) = 0	*No test*		
2: Var (ut) = σ2< ∞	White test		
3: Cov (ui,uj) = 0	*Not required*		
4: Cov (ui, xi) = 0	*Not required*		
5: ut – N (0,σ2)	Jarque-Bera test		
Other issues			
Multicollinearity	Correlation matrix		
Non-linearity	Ramsey RESET test		
Model one*		**Score**	**Result**
Heteroscedasticity**	White test	0.7428	Not rejected
Non-normality	Jarque-Bera test	Appendix I Table D	Some variables adjusted
Multicollinearity	Correlation matrix	X<0,8	No multicollinearity
Non-linearity***	Ramsey RESET test	2.2149	Not rejected
Model two*			
Heteroscedasticity** (Gross Returns (LN))	White test	1.5092	Not rejected
Heteroscedasticity** (Adjusted Alpha)	White test	0.5739	Not rejected
Heteroscedasticity** (Adjusted Alpha (net))	White test	0.5526	Not rejected
Non-normality (Gross Returns (LN))	Jarque-Bera test	Appendix II Table G	Some variables adjusted
Non-normality (Adjusted Alpha)	Jarque-Bera test	Appendix II Table H	Some variables adjusted
Non-normality (Adjusted Alpha (net))	Jarque-Bera test	Appendix II Table I	Some variables adjusted
Multicollinearity (All)	Correlation matrix	X<0,8	No multicollinearity
Non-linearity*** (Gross Returns (LN))	Ramsey RESET test	0.1259	Not rejected
Non-linearity*** (Adjusted Alpha)	Ramsey RESET test	1.4932	Not rejected
Non-linearity*** (Adjusted Alpha (net))	Ramsey RESET test	2.6009	Not rejected

Source: Brooks, 2008

** For further information on the outcome of the specific tests, please refer to Appendix I & II.*
*** Find White test outputs in Tables E, J, K, and L in the same order in Appendix I & II.*
**** Find Ramsey RESET test in Tables B & M in Appendix I & II.*

Any presence of *heteroscedasticity* would lead to a violation of assumption (2) and possibly (4). Homoscedasticity is an important OLS assumption as it assumes that variance of the errors

to be constant. Heteroscedasticity, which violates this assumption, would in a scatter plot show significant patterns. For this study, the presence of heteroscedasticity is tested with the White test, but no OLS violation was found. Furthermore, a violation of assumption 4 could cause endogeneity problems. However, all variables are assumed to be exogenous, following the approaches of related literature disregarding this (Cremers & Petajisto, 2009) (Petajisto, 2013).

Non-Normality violates OLS assumption (5) and existence is tested with the Jarque-Bera test. Under OLS assumption (5) a data series is determined normal if the residuals follow a normal distribution and skewness and kurtosis are not present. If residual distribution is right or left skewed there is non-normality. In order to control for non-normality, the natural logarithm was taken of the residual output where applicable. The results are shown in Table 2. However, dummy variables were disregarded as this is not possible for such variables. (Brooks, 2008)

Multicollinearity issues arise when independent variables are correlated with each other. There can be two types of multicollinearity, perfect or near multicollinearity. Perfect multicollinearity is when independent variables are perfectly correlated to another. Near multicollinearity is, according to a rule of thumb, when two independent variables have a correlation higher than 0.8. In our dataset this is tested with a correlation matrix and it is found that no independent variable displays a correlation higher than 0.8 except for TNA and active share dummy. This was already expected as the active share dummy is based on the funds TNA and can therefore be disregarded. As a result, multicollinearity was found to not be an issue in this study and the results can be found in Appendix I, Table C. (Brooks, 2008)

Non-linearity is another important issue to address which could be present in a cross sectional regression. Non-linearity is when the scattered relationship between the dependent and independent variables is following a non-linear pattern instead of the expected linear relationship. For the regressions tested in this paper, non-linearity was not detected through the Ramsey RESET test. As linearity is given, there will be no biasedness in the results.

4.4.2. Robustness and Model Accuracy

In order to guarantee a robust and accurate model, a stepwise regression procedure has been followed. The approach follows a step-by-step selection of variables, starting with only two variables in the regression for active share (Brooks, 2008). The main reason for using this procedure is to guarantee a good measure of fit in terms of R^2 (Mark & Goldberg, 2001) as an increasing R^2 along with added variables justifies the inclusion of the variables. Along with the

so called “forward selection”, the accuracy was further tested through backward elimination. The reason for using backward elimination is to find the variables that have an impact on the model by deleting the statistically and/or economically insignificant variables. This is seen to be a less biased approach compared to other selection approaches according to Dunkler et al. (2014).

4.5. Summary

Through the methodological approach adapted from Saunders et al (2009), this research retrieved the used data according to probability sampling where the selected sample reflects the movement of the population. All data was collected through reliable secondary sources such as Eikon, annual reports, Morningstar etc. Additionally, each variable is explained in detail along with the respective formula for calculating it in the case it seemed necessary to understand the concept behind the variable in more detail. The regressions are divided into two different models to explain active share as well as returns through several exogenous and endogenous variables. Lastly, the data analysis part serves as a part to check the data for potential inconsistencies and potential violations of OLS assumptions were tested. Furthermore, other issues such as multicollinearity and non-linearity were tested. All issues found were resolved according to Brooks’ (2008) most applicable solution. Moreover, the stepwise selection process as well as the backward elimination is introduced, underlining the validity and reliability of the model. This also shows the robustness and accuracy of the model, as results are relatively robust to changes. Please find more information on the coefficient changes in chapter 5.1.

5. Empirical Findings, Analysis & Discussion

Firstly, the findings, analysis, and discussion part depicts the regression findings. Secondly, these are analyzed and finally a discussion on their impact with respect to the research hypotheses is conducted. Additionally, an investment strategy is proposed based on historical ETF and mutual fund returns.

5.1. Regression Results

As described in chapter 4, regression results are presented in the identical layout as two different regressions were conducted. Furthermore, the findings part and the analysis & discussion part are presented separately in order to make it easier to understand.

5.1.1. Regression 1 – Active Share

Regression 1 analyzes the influences of certain independent variables on active share as the dependent variable. Seven regressions, including different variables, were conducted, following a stepwise regression procedure and finally backward elimination.

Findings

The variables including coefficients and t-statistics are presented in Appendix 1. The results are being presented by first looking at the endogenous variables that yield a significant result followed by the ones that were found to be insignificant. The same approach is then followed for the exogenous variables.

The first significant endogenous variable is *tracking error* which is significant on the 1% level on every step tested. TE shows a positive correlation to active share of 64.5% as can be found in the correlation matrix in Appendix I, Table C. A high correlation was expected, however this does not cause any problems as described by Cremers & Petajisto (2009). The correlation is based on the fact that the two variables measure very similar effects although approaches differ. The same accounts for the economic significance, which was found for TE in any of the regression steps as TE positively influences active share. When taking solely regression (1), a 100% increase in TE lets active share increase by about 40%. For regression (7), the backwards elimination model, the influence is weaker at about 19%. This was expected since the variable was included in order to see how well TE can explain the active share and it was also to see if Cremers' & Petajisto's (2009) study results hold for the recovery period in the German equity

market. Their study yielded the same results for TE. *Total expense ratio* also shows statistically significant results in all regressions. However, the significance level is strongest for the backward elimination regression, stating that there is a positive relationship between active share and fees. This is also supported by the correlation matrix, presenting a positive correlation of 45% between the variables. This concludes that, if investors investing in a mutual fund with higher fees are likely to invest in a fund with a fund manager actively engaging in stock picking and/or factor timing. This also makes sense through the eyes of a rational investor, who would potentially be willing to pay more to a manager who enters into separate investments, outside the benchmark index to receive superior returns. The last significant endogenous variable is *securities*. This variable is significant in regressions (6) and (7) and furthermore almost significant for the overall model in regression (5). This states that there is a negative relationship between active share and the number of securities in the fund. Concluding this means that the less different securities the fund invests in, the more likely the fund is to be highly active. This is also supported through presence of economic significance which however is weak. *Turnover ratio* is found insignificant in all steps tested. This is supported by the correlation matrix which indicates almost no correlation between the two variables.

The first significant exogenous variable is *Sharp ratio*. It is found to have both a statistically and an economically significant impact on active share. This means that the higher the activeness of the fund manager, the higher the return of the fund compared to a risk-free alternative and implies that fund managers are able to increase returns relative to risk. *Gross returns* are almost statistically significant on the 10% level in the backwards elimination step and furthermore, economically significant on all tested steps. The outcome suggests that funds with lower 5 year historical returns tend to show higher active share. *Total net assets* as a measurement of fund size, has a statistically insignificant impact on active share and has a coefficient being close to zero for all steps tested, stating that there is neither statistic nor economic significance. *Adjusted alpha* is also statistically insignificant but shows economic significance for the sixth step, excluding tracking error. *Management tenure* is the last variable tested and was found to be both, statistically and economically insignificant, with a negative coefficient.

Analysis & Discussion

The backward elimination step regression (7) presents the best fit for explaining active share and shows a measure of goodness through an R^2 of 76.5%. Even though regression (5), including all variables, yields a higher R^2 of 77.17%, the backward elimination step is the preferred option to estimate active share. This is because step (5) includes more variables which were found to be insignificant, but only adds 0.65% to R^2 compared to (7). Thus, for this slight increase in R^2 a lot of data is need which would be inefficient. Therefore the backward elimination regression is seen to be the best fitting model to make predictions on active share.

As an effect of the R^2 of 76.5%, H_0 for hypothesis 1 can be rejected. This means that, with a likelihood of 76.5%, the variables used in regression (7) can predict active share. This rejects H_0 for hypothesis 1 (chapter 3.5.), as an investor willing to invest in a fund that is actively managed should look at the latter.

Overall, the regression results found in this study show a couple of differences compared to Cremers' & Petajisto's (2009) outcomes. First of all their R^2 is lower. This potentially stems from the fact that they used panel data, while this study uses cross-sectional data which presents a more simplistic model, increasing the goodness of fit. *Tracking error* is found to be significant on a statistical as well as economic level. As this study reveals no econimc significance, the result can be interpreted as tracking error having less impact on the active share during a recovery period on the German equity market or that funds try to follow the market in times the market rises anyway. The same goes for *total expense ratio* where this study only finds statistical signifiance while others found economic significance too (Cremers & Petajisto, 2009). This draws the conclusion that mutual funds with higher fees are also carrying a higher active share. The economic impact however is minimal. This leads to a rejection of H_0 for hypothesis 2 (chapter 3.5.) due to the statistically significant relationship proving that fees influence the fund managers activeness measured as active share. An investor, whishing to invest in an active mutual fund should, according to these findings choose among the funds with the highest TER to have the highest possibility of picking a fund with a high active share. This is contradicting to Khusainova's & Mier's (2014) study whos paper does not display any significance for managers charging higher fees to be more active. Regarding both, *turnover ratio* and the *number of securities*, Cremers & Petajistos (2009) study found similar results. This supports the findings of this paper, stating that TR and securities can be disregarded as explainatory variables for active share.

As for the exogenous variables, *total net asset value* shows no significant influence on active share in the German equity market. Contrary to this finding, Cremers' & Petajisto's (2009) study found a significant positive relationship. This can be interperated as fund size has a big effect during other periods but economic recovery period and moreover especially on the US equity market. For Germany and the recovering markets, fund size does not play a big role when it comes to active share as active share varies within larger and smaller asset funds.

The *Sharpe ratio*, representing the only exogenous variable in regression (7), shows stastistically significant results. This was not tested by Cremers & Petajisto (2009) and is therefore particularly interesting for this study. As stated previously, it is both statistically and economically significant. This means that there is a strong relatinship between the Sharp ratio and active share for this study. The new variable could possibly increase the results from previous studies and further support the possibility to foresee the future activeness of the fund manager. Furthermore, the Sharpe ratio's significance doubles for the backward elimination step. In this step, the active share would as an effect of a one unit increase in Sharpe ratio increase by 287%, which states strong economic significance. As a conclusion, if funds show higher returns compared to risk-free returns, they also tend to show higher active share. The next return measure, *gross return*, was found to be economically significant but displayed no statistical significance. However, the backward elimination step shows an almost significant result on the 10% level. For the backward elimination step, the impact of a postivie 100% change in Gross returns would yield a reduction of -18.16% in the active share. Gross returns are not tested in Cremers & Petajistos (2009) study but they have other return measurements where the results are similar. This leads to the conclusion that during economic recovery, a fund with a high active share tends to have lower gross returns due to the observed negative relationship. *Adjusted alpha*, the next return measure, refers to the gross return over the adjusted benchmark. This variable was found to be economically significant in regression (6) where TE is removed. This means, a 100% change in adjusted alpha would change the active share by -7.7%. The reasoning behind this impact is explained further in the next section where regression model two is analyzed. However, it can be interpreted in a way that a high active share decreases the possibility of a high adjusted alpha. Thus, outperformacne decreases during econimc recovery. When including TE (7), the possibility of having a high adjusted alpha is lower in the model, thereby decreasing the economic impact of the adjusted alpha variable.

The last variable that is tested in this regression is *management tenure* which shows to be insignificant in regression (5) with a minor negative relationship to active share in terms of its coefficient. In contrary to this study, Cremers & Petajisto (2009) find a statistic significance with a positive correlation of management tenure effecting active share. Their finding is in other words stating that the longer the tenure of the manager the higher will the activeness be. However, this is done in a different way in their study since they look at different time periods through the usage of panel data, which is as a consequence of unavailable data not feasible for this study. The results of this study show that management tenure does not have any influence on active share. This might also be an effect created by the time period studied for economic recovery only where all managers potentially try not to underperform the already positive market returns through more or less following the index.

5.1.2. Regression 2 – Returns

Regression 2 tests the independent variables influences on the three stated return variables being gross returns, adjusted alpha and adjusted alpha (net). Each is regressed two times as the active share dummy variable is only included as an independent variable for every second regression. Hence, another approach is used compared to model 1 which applies a stepwise regression approach.

Findings

When looking at Appendix II, Table F, one can find the overall regression outputs, stating the significance of the variables. Surprisingly, almost all of the endogenous variables which are controllable by the fund management do not display statistical significance. However, the exogenous variables, except for management tenure, are statistically significant and thus are influencing returns on different levels. The findings for the endogenous variables are introduced first, followed by the findings for the exogenous variables.

In terms of significance on the endogenous level, only the *total expense ratio* was found to be significantly influencing dependent variables. However, gross returns and adjusted alpha were not found to be influenced by this ratio. Hence, the adjusted alpha net of fees is influenced by the TER negatively, as the coefficient was found to be negative. Moreover, the significance level was different for the two regressions with the dependent variable adjusted alpha (net). Regression (5) shows significance on the 5% level while regression (6), including the active share dummy, only shows significance on the 10% level. All other independent variables

exhibit a statistically insignificant result in terms of p-values. The relationships however differ. *Active share* and of course the *active share dummy*, as well as the *tracking error* show a negative relationship to all three dependent return variables. The same accounts for the insignificant results for TER for the gross returns and adjusted alpha as dependent variables. Surprisingly, the *number of securities* invested, also showing no statistically significance, has neither a positive nor a negative influence.

As for the exogenous variables, two out of three variables were found to be statistically significant. The fund size, measured in *TNA*, was found to be significantly influencing all three different return measures, with a positive coefficient. However, the significance level differs as it is significant on a 1% level only for regression (1) where it is regressed without active share dummy against gross returns as the dependent variable. For all other regression, the variable was found to significantly influence the dependent variables on a 5% level. Despite the statistical significance, the economic significance is close to zero. The second statistically significant independent variable is the *Sharpe ratio*. The latter was found to be significant against all dependent return variables on different significance levels. The variable is overall positively correlated through the coefficient, without economic impact. The only insignificant variable of all exogenous variables is *management tenure*. According to the regression results, this variable has neither statistical nor economic significance, but shows a negative coefficient. For gross returns, regressions (1) and (2), the t-values are higher compared to the other regressions against the two different alpha measures.

Analysis & Discussion

This part again serves to analyze the previously stated findings and to discuss their impact.

The attentive reader can immediately capture from Appendix II, Table F, that the measure of fit (R^2) is always higher for the regressions including the adjusted active share dummy variable. Thus, including this variable also increases the measure of fit of the regression model to draw conclusions on the independent variables impact on the different return variables. The best fitting regression model in this case is the gross return regression (2) with an R^2 of 63.37%. Hence, the independent variables best describe returns in terms of gross fund returns, excluding any performance compared to benchmarks gross and net of fees. Still, only a few variables were found to be significantly influencing. This displays significantly different results compared to Cremers & Petajisto (2009) as well as Petajisto (2013). This difference however, confirms the

findings of McGuigan (2006), who found that it is very difficult to find funds that outperform throughout a longer time period.

Regarding the endogenous variables, *active share*, the actually most interesting variable to test for influences on performance, was found to be statistically insignificant. Thus, the findings of Cremers & Petajisto (2009) are rejected for the German equity market in times of economic recovery. In their paper, they provide evidence for a positive relationship between the variables in their model, indicating that active managers are likely to earn superior returns to the market returns. Regression model 2 however, showing negative coefficients for active share, shows that in times of economic recovery in Germany, higher active share yields below market returns. For regressions (5) and (6) with the dependent variable being adjusted alpha (net), this negative relationship and impact was found to be marginally lower, indicating that if fees are deducted from performance, mutual funds with high active share perform even worse compared to the market and thus ETF investments. This of course is a logical consequence but confirms the validity of the variable included in the regression. This also confirms the findings of Petajisto (2013), who found that on average mutual fund underperform indices net of fees. However, as all coefficients are close to zero and there is no statistical significance, the impact of active share is economically as well as statistically nonexistent. As a result, "H_0: During recovery, there is no significant difference between active fund and market returns" of hypothesis 3 has failed to be rejected as there is no significant difference between the returns. Please find all hypotheses in chapter 3.5.

The next endogenous variable, the *active share dummy* variable, was also found not to be statistically significant. This consequently proves again the opposite results compared to Cremers' & Petajisto's (2009) findings. However, the relationship between active share for below medium asset size funds and returns was found to be stronger, as t-statistics yield values closer to significant values. Thus, it is arguable that for below median sized asset funds, a higher active share can yield better returns, due to a coefficient closer to zero. The finding for active share dummy can lead to the conclusion that funds' active share is somehow tied to the funds size. Thus, similar to the assumptions of agency theory, where smaller firms are assumed to have more freedom (Garcia-Teruel & Martinez-Solano, 2008), the same may be valid for mutual funds as smaller funds might restrict their fund managers less in terms of investment policy. This enables them to increase their activities in stock picking or factor timing, simultaneously increasing active share.

The findings for the two active share variables can be related to the findings of Jones & Wermers (2011). Their analysis found that funds with high active share usually perform better during times of financial distress, e.g. the financial crisis in 2008, which cannot be rejected by the outcome of this analysis. However, it can be assumed that mutual fund investments pay out during times of financial distress, while they tend to not pay out during times of financial recovery as indicated by the findings in regression model 2. Hence, during financial recovery, it might be worthwhile investing in ETFs. This is being evaluated thoroughly in part 5.2.

The next endogenous variable in the second regression model is *tracking error*. When referring to the literature and findings of Cremers & Petajisto (2009), the finding is surprising as TE, similarly to active share, is also negatively correlated to returns. However, in their findings, a high active share along with a low TE was found to yield the best returns. Thus, their regression coefficients are positive, while the TE coefficients are negative. Please refer to Table 3 to receive a clear picture of all possible combinations and their impacts on returns according to Cremers' & Petajisto's (2009) findings:

Table 3 - Active Share vs. Tracking Error

	High Tracking Error	**Low Tracking Error**
High Active Share	Highest returns	Medium high returns
Low Active Share	Lowest returns	Medium low returns

Source: Cremers & Petajisto (2009)

This table is also being drawn for the outcome of this study. However, due to the fact that both coefficients, for active share as well as for TE, are negative, the outcome is the opposite:

Table 4 - Active Share vs. Tracking Error

	High Tracking Error	**Low Tracking Error**
High Active Share	Lowest returns	Medium low returns
Low Active Share	Medium high returns	Highest returns

Table 4 clearly indicates that the regression coefficient output in Appendix II, Table F, favors investments in low TE and low active share funds during economic recovery. This seems reasonable as also the active share variable has a negative coefficient, indicating a below index performance. Thus, the lower the tracking error, the closer the fund is invested in the index which was already found to yield superior returns. Moreover, the findings are different to the

findings in Petajisto's (2013) follow- up research where he developed the 5x5 grid, presenting the different fund manager strategies depending on tracking error and active share. Please refer to Table 1 for more details.

The *Turnover Ratio* again shows no statistically significant impact. However, compared to the other presented variables, the coefficient is positive, indicating a positive relationship between TR and returns. However, there is basically no economic impact which shows that turnover has no impact on returns at all in times of economic recovery. This finding underlines the more dated findings by Ippolito & Turner (1987) who found that there is no indication that portfolio performance is related to total portfolio turnover. Furthermore, the findings are in line with the evidence provided by Cremers & Petajisto (2009), who also found no relationship between TR and returns. Nonetheless, their relationship is negative compared to the positive relationship of this research. This discrepancy is likely to be due to the different economic circumstances the researches were conducted in.

The *number of securities* the funds are invested in was found to have no significant influence on the performance. Thus, there is basically no difference between the performances of funds that are diversified largely compared to funds that are diversified to a smaller extent. The coefficient, being basically zero, leads to no economic impact on the model. Hence, it can be concluded that, during recovery in Germany, reducing risk through dispersed investments yields the same result as a more focused strategy. This relationship was however not found by Cremers & Petajisto (2009). In their research, a positive and statistically significant relationship was found. The divergence can derive from the different market settings researched. During recoving markets, lower volaility is expected (Markowitz H. , 1952). Hence, it would not make a difference if a portfolio is largely or just slightly diversified in terms of return characteristics.

The last endogenous variable tested in model 2 is the *total expense ratio*. TER is the only endogenous variable that shows statiastical significance against the adjusted alpha (net) for regressions (5) and (6). Hence, it can be concluded that the fees of mutual funds have a significant negative influence on final returns earned by the investor. As the significance level decreases from 5% in regression (5), to 10% in regression (6) where the active share dummy is included, it is arguable to assume that funds with smaller asset size can be expected to carry lower fees. This can however derive from the generally better returns found for smaller TNA funds. These findings are in line with the findings by Cremers & Petajisto (2009), Jones & Wermers (2011) as well as Khusainova & Mier (2014), who consistently found a negative

relationship between fees and returns presented as alphas. Despite the statistically significant findings, the economic impact is alsmost nonexistent for regressons (5) & (6).

Following the endogenous variables, the exogenous variables are examined. Here, it is found that *management tenure* has no statistically sigificant influence on returns. As a result, in times of economic recovery, funds of fund managers who are new or old to the office show the same returns which can be seen as reasonable since fund managers probably tend to follow the positive market developments. This also confirms the findings of Cremers & Petajisto (2009), as they only find a relationship to the characteristics by which fund managers pick stocks. Due to time constraints, the characteristic variable had to be disregarded for this research.

The last two exogenous variables in the model are displaying the only two variables significant for all 6 regressions. The *total net assets*, is found to significantly influence returns positively. However, the economic impact is close to zero as a 100% change in TNA would only result in a 0.05% - 0.09% change in returns. This again is contrary to the findings by Cremers & Petajisto (2009), who found that TNA is negatively correlated to returns. The positive relationship found in this study confirms the findings that smaller funds tend to have a higher active share, resulting in worse performance.

The *Sharpe ratio* was also found to positively influence returns. This is due to the fact that first of all both, the explanatory variable as well as the explained variable are performance measures. Thus, the positive coefficient was expected. This signals that during recovery periods, mutual funds outperform risk-free investment instruments. The economic impact however is low as a 100% change in Sharpe ratio only changes return measures between 1.6% and 2.8%.

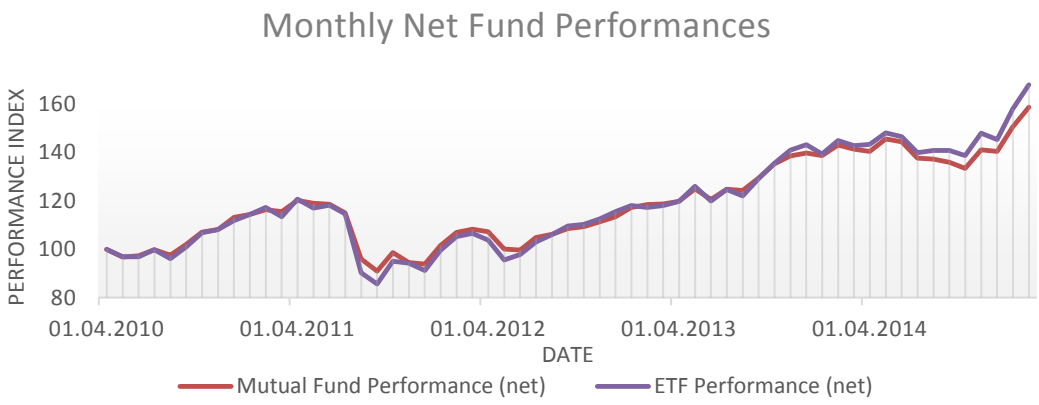
5.2. Proposed Fund Investment Strategy

In this part, the performance of the selected ETFs is compared against the performance of the selected mutual funds for the German equity market. Please find a detailed overview of the performances in Appendix III, Tables N & O.

In Table 5, the annual returns and fees are summarized for both, mutual funds and ETFs. The table shows that on average, both before and after fees, ETFs outperform mutual funds. The net per month performance can be seen in Graph 2, which shows a net of fee outperformance of ETFs compared to mutual funds, especially during the recent strong economic upswing. The

numbers used are however averages and an investor, by using the previous results, might argue that splitting investments into mutual funds and ETFs is valid. Most interesting in this context is the gross/net out- or underperformance against the benchmark, labeled adjusted alphas. In Appendix II, Table F, the regression results can be found. The adjusted alpha regression, with all variables included, shows a R^2 just below 50%, which can be interpreted as the variables can explain the gross out- or underperformance of the fund benchmark with 50% predictability. This presents a fragile explanation and an investor cannot increase the chances of picking an outperforming fund in the German equity market during a period of financial recovery. This is in line with the findings of McGuigan (2006). With a slightly better R^2 of 56.95%, the adjusted alpha (net) can provide a better explanation. But the slight improvement still leaves a big gap unexplained. A conclusion would be that during periods of financial distress on the German equity market, one cannot easily find mutual funds with potential to outperform based on the variables used in the regressions. An investor is therefore better off investing in a passively managed ETF during financial recovery.

Graph 2 - Monthly Net Fund Performances



Source: Eikon, 2015

Using this result, hypothesis 4 and 5 can be answered. Hypothesis 4 fails to be rejected since gross of fees, as presented in Table 5, there is no significant difference between the annual returns of mutual funds and ETFs. However, net of fees for hypothesis 5, H_0 can be rejected since the difference in annual returns becomes significant. As a result for this rejection, the above conclusion further holds and an investor is better off investing in ETFs during financial recovery in the German equity market. This is also confirmed by Gruber (1996) who arrived at the same conclusion. His paper states that mutual funds on average offer a negative risk adjusted return, showing that investments in index funds are more favorable to investors. Further

analysis finds that investing into mutual funds that charge higher annual fees does not pay out as on average their performance stays in line with the benchmark gross of fees and even underperforms net of fees.

Table 5 - Average Annual Returns & Fees

	Annual Gross Return	Annual Fees
Mutual Funds Average	12.22%	1.65%
ETFs Average	12.38%	0.15%

Source: Eikon, 2015

If however, an investor has the feeling a mutual fund is likely to outperform the index, a risk-neutral and risk-averse investor would still choose this fund to increase the chances of receiving superior returns. Markowitz (1952) introduced the revised aspect of different utility curves depending on the level of risk involved. For a risk-averse investor the utility curve is concave, meaning that at a certain point, more return by engaging into more risk does not generate higher utility. This theory says that a risk-neutral investor shows a linear utility curve leading to increased utility for increased risk. For a risk-seeking investor the utility curve is convex and shows an exponential increase in utility for higher risk. Both risk-neutral and risk-seeking investors may therefore still choose to invest in an actively managed mutual fund based solely on the increased utility received through higher returns which can be generated by engaging into riskier investments. For these investors, the results from model 1 could be used, since an active manager is what investors should strive for if they want their investment to outperform an index and being riskier in terms of return predictions compared to ETFs. The riskiness basically derives from additional systematic risk exposure; however the added risk usually does not provide higher compensation in forms of better returns (Gruber, 1996).

5.3. Summary

In the first part of this chapter, regression model 1, explaining active share is presented. The outcome indicates that for endogenous variables, tracking error is highly correlated to active share at the 1% level. Regarding the hypothesis related to this model, TER is found to be significant, explaining that higher active share funds tend to also charge higher fees. This rejects H_0 of hypothesis 2 that fees do indicate higher active share. Through the backward elimination, TE, TER, Securities, Gross returns, and Sharpe ratio are found to best describe active share in

this model with an R^2 of 76.5%. Consequently, H_0 of hypothesis 1 can be rejected, as active share is not random and can somehow be described through other variables.

In the second part, regression model 2 for returns is presented. The regression shows the overall impact of different variable on three different return variables being the funds gross return, the adjusted alpha, and the adjusted alpha (net). Overall, the analysis of the regression fails to reject H_0 for hypothesis 3. Hence, active funds were found to yield insignificantly different returns towards their index, rejecting the findings of related literature for the Germany recovering equity market.

In the third part, the findings are converted into a fund investment strategy, including the theories for ETFs introduced in chapter 2. By looking at Table 5, in times of economic recovery, ETFs outperform mutual funds slightly gross of fees. However, as mutual funds carry higher fees, ETFs net of fees meaningfully outperform. As a result, the indifferent investor, who bases investment decisions on statistics, should invest primarily in ETFs during times of recovering markets. The proposed strategy was based on Markowitz (1952) *"The Utility of Wealth"* theory and is therefore transferrable to all three different types of investors: Risk-averse, risk-neutral, and risk-seeking.

6. Conclusion

The overall purpose of this paper was to achieve an understanding of which drivers influence fund activeness, measured as active share. The best predicting drivers found for active share in the recovering German equity market between 2010 and 2015 were estimated through the statistical backwards elimination. In this model, Tracking Error, Total Expense Ratio, and Sharpe ratio show a significant positive influence on Active Share, while the number of securities variable shows a significant negative influence. Moreover, three variables, namely Tracking Error, Gross returns, and Sharpe ratio have a significant economic impact on the model. This directly answers the first part of the overall research question of this study.

With this finding, an investor deciding to invest in an actively managed fund should look at these variables. Please refer to Table 6 below, to receive an overview of the discussed results, including the influences on the hypotheses based on the literature review.

Table 6 - Results Model 1

			Significance		Economic Impact	
Variables		Hypothesis	Regression model	Backward elimination model	Regression model	Backward elimination model
	Endogenous					
	Tracking Error	H1: 1	***	***	positive	positive
	Turnover Ratio	n/a	-	n/a	no	n/a
	Total Expense Ratio	H1: 1, H2:1	***	***	no	no
	Securities	H1: 1	-	*	no	no
	Exogenous					
	Total Net Assets	n/a	-	n/a	no	n/a
	Gross returns	H1: 1	-	-	negative	negative
	Adjusted Alpha	n/a	-	n/a	no	n/a
	Sharp Ratio	H1: 1	***	***	positive	positive
	Tenure	n/a	-	n/a	no	n/a

Regression model 1 provides the possibility to answer the first part of research questions 1 as well as research question 2. The model provides a good explanation of active share through influencing independent variables and finds that active share and fees are positively related. This would justify the fact that the manager is actually receiving a higher payment for an active investment strategy in terms of fees paid by the investor.

As all other independent variables show no significant results, the backward elimination step reveals the most useful variables. A detailed analysis of this regression model can be found in chapter 5.1.1.

The findings of this study differ from the findings of Cremers & Petajisto (2009) in a way that during times of economic recovery, other variables influence active share. Overall, the study shows that between 2010 and 2015 less variables are found to significantly predict active share. This is in line with the study of Jones & Wermers (2011) who also conclude that active share largely depends on the level of economic distress and thus, the overall economic situation. Hence, a fund's active share level is and its impact is, according to the econimc situation, different over time.

The main purpose of regression model 2 was to conclude whether active share has an influence on mutual fund returns during the economic recovery period from 2010 – 2015 in Germany. This again is closely related to the studies of Cremers & Petajisto (2009) and Petajisto (2013), but the findings significantly differ. The findings of this paper show that active share has no significant influence on any return measurement taken as dependent variable. Hence, this paper concludes that previous empirical findings on active share do not hold during times of economic recovery. However, the time period covered in related papers by Cremers & Petajisto (2009) and Petajisto (2013) is longer and broader, as the sample includes over 8000 US based equity funds. This result clearly answers the second part of research question 1 as well as research question 3. It shows that active share is not a good measurement for predicting mutual fund performance during times of economic recovery. However, as the coefficient is negative, it is assumed that funds with higher active share tend to perform worse during recovery, which can however not be statistically proven.

Contrary to active share, Total Expense Ratio, Total Net Assets, and Sharpe ratio are found to significantly influence returns on different levels. TER infleunce on adjusted alpha (net) shows that fees significantly influence the investor's overall retuns negatively compared to pure market returns. Potentially ETFs, carrying smaller fees, do not show this pattern. TNA and Sharpe ratio on the other hand, show a weak positive coefficient. The active share dummy variable together with TNA leads to the conclusion that smaller sized funds show higher active share and performed worse during 2010 – 2015 compared to larger sized funds. This can be related to agency theory where Garcia-Teruel & Martinez-Solano (2008) found that smaller firms can act more freely. The findings can lead to the conclusion that smaller asset size funds

tend to be less influenced by investment policies. Thus, their fund managers have more freedom to make use of factor timing and stock selection, thereby increasing active share. The outcome for the Sharpe ratio reveals that overall mutual funds with active share outperform risk-free securities. All other variables show neither statistically nor economically significant results. Please refer to Table 7 below for more detail:

Table 7 - Results Model 2

		Hypothesis	Significance	Economic Impact
Variables	*Endogenous variables*			
	Active Share	H3: 0	-	no
	Active share dummy	H3: 0	-	no
	Tracking Error	n/a	-	no
	Turnover Ratio	n/a	-	no
	Total Expense Ratio	n/a	* (Only for net returns)	no
	Securities	n/a	-	no
	Exogenous variables			
	Total Net Assets	n/a	**	no
	Sharp Ratio	n/a	* to *** (dependent on model)	no
	Tenure	n/a	-	no

As mentioned earlier, the model exhibits a difference to the models of Cremers & Petajisto (2009) and Petajisto (2013) who display more influencing variables. More interestingly, they conclude that active share influences returns positively. The conclusion of this study however would advise to disregard active share as a measure for investors attempting to find outperforming funds during recovery.

To compose the investment strategy for the recovering German equity market, the analyses of the two regressions is used. However, the optimal strategy may vary for each investor as risk tolerances differ. As mentioned earlier, Markowitz (1952) introduced the concept of three different investor categories with different risk tolerances: The risk-averse, risk-neutral, and risk-seeking investor. The investment strategy has thus been established depending on these

categories. Each type should, when making fund investment decisions, base them on different measurements. This is contrary to the findings of Gruber (1996), who in his study stated that investors continuing to invest in mutual funds are either unsophisticated, represented by another investor group, or been holding them for a longer period and do not sell because of tax advantages. According to him, the rational investor should solely invest in ETFs. Please refer to Table 8 to receive a detailed overview of the variables influencing the decisions:

Table 8 – Variables Influencing Investment Decisions

	Investor Types		
	Risk-averse	Risk-Neutral	Risk-Seeking
Variables: Active share		X	X
Tracking Error	X	X	X
Total Expense Ratio	X	X	X
Securities		X	X
Gross Return		X	
Sharp Ratio	X	X	X
Invest in:	**ETFs**	**Mutual funds and/or ETFs**	**Mutual funds**

This table shows that the risk-averse investor should, during times of recovery, invest in ETFs and base investment decisions on TE, TER, and the Sharpe ratio. On the contrary, the risk-seeking investor should base investment decisions on active share and number of securities additionally. This is based on the fact that the risk-seeking investor has a higher utility for higher risk potentially generating superior returns. An investor is more likely to find funds offering returns different to market returns by looking at the active share combined with the other proposed measurements. Hence, investments in mutual funds should be made as they do not closely follow the market movement. The risk-neutral investor, being indifferent in making investment decisions based on risk, should still use all measurements in order to prevent the possibility to always end up on one side only. This means, by chance a risk-neutral investor could always make risk-averse or risk-seeking investments. However, the empirical findings of this study indicate the opposite of what the theory proposes, as on average mutual funds with higher risk yield lower returns. This makes the risk-neutral behavior inefficient. Although the return difference between mutual funds and ETFs is not significant gross of fees, answering

research question 4, there is a significant underperformance of mutual funds net of fees. As described in chapter 5.2., this underperformance compared to benchmark should influence the rational investor towards deciding to only invest in ETFs during economic recovery. This finding finally answers the last research question 5.

Despite the fact that this research provided the possibility to answer all crucial research questions, it is difficult to generalize. This mainly derives from the small sample compared to other, greater samples tested by Cremers & Petajisto in 2009 as well as in Petajisto's follow up study in 2013. Therefore, future research should apply the study on other markets, attempting to find if these differences exist due to the different market or due to the economic setting tested. It might well be that both together are the reason for the discrepancies. Additionally, due to time and resource constrains, future research should also have a look at the behavior of active share in distressed markets. Therefore, another panel data study is suggested, including the possibility to differentiate between different levels of economic distress related to time. In other words, crises and recovery periods should be researched and evaluated independently. This would then also follow up on the research of Jones & Wermers (2011), enabling a further evaluation of their findings.

References

Bogley, J. C. (2009). *Common Sense on Mutual Funds.* Hoboken: John Wiley & Sons, Inc.

Brooks, C. (2008). *Introductory Econometrics for Finance.* Cambridge: Cambridge University Press.

Broyer, C., Petersen, A.-K., & Schneider, R. (2012). *Impact of the euro crisis on the German economy.* Frankfurt am Main: Allianz.

Cremers, M., & Petajisto, A. (2009). How Active Is Your Fund Manager? A New Measure That Predicts Performance. *Oxford University Press*, 1-56.

Deutsche Börse Group. (2015, April 03). *Dax-Indices*. Retrieved April 03, 2015, from Dax-Indices: http://goo.gl/VKRRHT

Ding, B., & Wermers, R. (2009). Mutual Fund Performance and Governance Structure: The Role of Portfolio Managers and Boards of Directors. *Working paper, SUNY at Albany and University of Maryland*.

Dunkler, D., Plischke, M., Leffondre, K., & Heinzer, G. (2014). Augmented Backward Elimination: A Pragmatic and Purposeful Way to Develop Statistical Models. *PLOS One*, 1-19.

European Commission. (2012, July 03). Proposal for a Regulation of the European Parliament and of the Council on key information documents for investment products. Strassbourg, Alsace, France.

Evans, J. (2014, November 26). *Financial Times*. Retrieved from Financial Times: http://goo.gl/eBC8TL

Ferri, R. A. (2008). *The ETF Book - All You Need To Know About Exchange-Traded Funds.* Hoboken: John Wiley & Sons, Inc.

Friedman, M., & Savage, L. J. (1948). The Utility Analysis of Chices Involving Risk. *The Journal of Political Economy*, 289-304.

fundinfo. (2015). *fundinfo*. Retrieved 04 03, 2015, from fundinfo: http://www.fundinfo.com

Garcia-Teruel, P. J., & Martinez-Solano, P. (2008). On the Determinants of SME Cash Holdings: Evidence from Spain. *Journal of Business Finance & Accounting*, 127-149.

Getfaoui, H. (2010). Deviation from Normaility and Share Ratio Behavior: A Brief Simulation Study. *Investment Management and Financial Innovations*, 106-118.

Gruber, M. J. (1996). Another Puzzle: The Growth in Actively Managed Mutual Funds. *Journal of Finance*, 783-810.

Haas, M., & Pigorsch, C. (2007). *Financial Economics, Fat-tailed Distributions.* Bonn: German Research Foundation.

Ippolito, R. A., & Turner, J. A. (1987). Turnover, Fees and Pension Plan Performance. *Financial Analysts Journal*, 16-26.

Jensen, M. C., & Meckling, W. H. (1976). Theorz of the FIrm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics*, 305-360.

Jones, R. C., & Wermers, R. (2011). Active Management in Mostly Efficient Markets. *Financial Analysts Journal*, 29-45.

Khusainova, E., & Mier, J. (2014). Taking a Closer Look at Active Share. *Investment Research*.

Mark, J., & Goldberg, M. A. (2001). Multiple Regression Analysis and Mass Assessment: A Review of the Issues. *The Appraisal Journal*, 89-109.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 77-91.

Markowitz, H. M. (1952). The Utility of Wealth. *Journal of Political Economy*, 151-158.

McGagh, M. (2014, October 30). *Citywire*. Retrieved April 20, 2015, from Citywire: http://goo.gl/dQxfDx

McGuigan, T. P. (2006). The Difficulty of Selecting Superior Mutual Fund Performance. *Journal of Financial Planning*, 50-56.

Moore, C. (2014). Active Management: Passive skeletons in the active closet. *Investment & Pensions Europe*.

Morgan Stanley. (2015). *Mutual Fund Share Classes and Compensation.* Morgan Stanley Smith Barney LLC.

Petajisto, A. (2013). Active Share and Mutual Fund Performance. *Financial Analysts Journal*, 73-93.

Rabin, M., & Thaler, R. H. (2001). Anomalies - Risk Aversion. *Journal of Economic Perspectives*, 219-232.

Remenyi et al. (1998). *Doing Research in Business and Management.* London: Sage Publications.

Richards, I. (2014, October 30). Is your fund manager active enough? (M. McGagh, Interviewer) Retrieved April 20, 2015, from Citywire: http://goo.gl/dQxfDx

Robson, C. (2002). *Real World Research: A Resource for Social Scientists and Practitioner-Researchers.* Hoboken: John Wiley & Sons, Inc.

Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research Methods for Business Students.* Essex: Pearson Education Limited.

Sharpe, W. F. (1994). The Sharpe Ratio. *The Journal of Portfolio Management*, 49-58.

Stalter, K. (2014, November 10). *Money*. Retrieved April 28, 2015, from Money: http://goo.gl/sS52UC

Ziliak, S. T., & McCloskey, D. N. (2003). Size matters: the standard error of regressions in the American Economic Review. *The Journal of Socio-Economics*, 527–546.

Appendices

Appendix I – Active Share Regression Tables

Table A - Determinants of Active Share for German all-equity mutual funds 2010-2015

The dependent variable is active share; all other variables are computed as previously described. Based on standard error, the t-statistics can be found (in parentheses), the respective coefficient above each t-statistics. ***shows significance on a 1% level, ** on a 5% level and * on a 10% level in terms of p-values.

	Active Share						
Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Independent Variables:							
Endogenous Variables:							
TE 5 years (ln)	0.3968***	0.3908***	0.3453***	0.3433***	0.1950***		0.1936***
	(5.27)	(5.19)	(4.48)	(4.44)	(2.90)		(3.13)
TR (ln)			-0.0113	-0.0103	-0.0139	-0.0077	
			(-0.50)	(-0.45)	(-0.82)	(-0.41)	
TER (ln)			0.2155**	0.2049*	0.2044**	0.2414***	0.2094***
			(2.13)	(2.00)	(2.55)	(2.75)	(2.84)
Securities					-0.0031	-0.0045**	-0.0030*
					(-1.62)	(-2.20)	(-1.79)
Exogenous Variables:							
TNA (ln)		0.0000		0.0000	-0.0068	-0.0145	
		(-1.09)		(-0.84)	(-0.41)	(-0.79)	
Gross return (ln)					-14.0783	-18.0823	-18.1556
					(-0.78)	(-0.90)	(-1.63)
Adjusted Alpha					-0.7944	-7.7014	
					(-0.05)	(-0.48)	
Sharpe ratio 5 years					2.8323***	3.4089***	2.8703***
					(5.59)	(6.59)	(6.08)
Tenure					-0.0143	-0.0063	
					(-0.25)	(-0.10)	
Constant	0.3641***	0.3781***	0.2469***	0.2639***	0.0164	0.1678	-0.0568
	(8.22)	(8.22)	(3.37)	(3.46)	(0.05)	(0.49)	(-0.47)
N	41	41	41	41	41	41	41
R^2	0.4160	0.4338	0.4837	0.4937	0.7717	0.7096	0.7651

Definitions:

Endogenous Variables refer to variables controllable by management. **TE** refers to Tracking Error. **TR** refers to turnover ratio. **TER** is the Total Expense Ratio and **Securities** the number of stocks invested.

Exogenous Variables refer to variables not controllable for management. **TNA** refers to Total Net Assets. **Gross Return** refers to the fund average gross return and **Adjusted Alpha** refers to the gross return over the adjusted benchmark. **Sharpe Ratio** represents the average 5 years Sharpe Ratio while **Tenure** stands for the management tenure (dummy). **Constant** equals to the intercept of the regression. **N** stands for the sample size tested and **R^2** the coefficient of determination functioning as a measure of fit.

Table B – Ramsey RESET Test for Non-Linearity

The RESET Test tests for the existence of non-linearity which is existent when rejecting H_0 from the f-statistics.

Ramsey RESET Test for regression model 1

Omitted Variables: Squares of fitted values

	Value	df	Probability
t-statistic	1.4882	30	0.1471
F-statistic	**2.2149**	**(1, 30)**	**0.1471**
Likelihood ratio	2.9205	1	0.0875

F-test summary:

	Sum of Sq.	df	Mean Squares
Test SSR	0.0514	1	0.0514
Restricted SSR	0.7480	31	0.0241
Unrestricted SSR	0.6965	30	0.0232

LR test summary:

	Value	df
Restricted LogL	23.903	31
Unrestricted LogL	25.364	30

Table C – Correlation Matrix (Active Share Regression)

The correlation matrix presents the simple correlation between each variable tested for the active share regression. The correlation matrix is valid for all regressions.

Sample: 1 41
Included observations:
41

Correlation	Active Share	Active Share Dummy	TE 5 years (ln)	TR (ln)	TER (ln)	Securities	TNA (ln)	Sharpe ratio 5 years	Tenure	Gross return (ln)	Adjusted Alpha	Adjusted Alpha (net)
Active Share	1.0000											
Active Share Dummy	-0.1919	1.0000										
TE 5 years (ln)	0.6450	-0.3393	1.0000									
TR (ln)	-0.0288	0.2018	0.0581	1.0000								
TER (ln)	0.4503	-0.1420	0.3267	-0.0098	1.0000							
Securities	-0.0235	0.1082	-0.1400	0.1160	0.0291	1.0000						
TNA (ln)	-0.2704	0.8096	-0.3742	0.0422	-0.2692	0.0803	1.0000					
Sharpe ratio 5 years	0.6308	0.1750	0.2515	0.1025	0.0506	0.3297	0.0878	1.0000				
Tenure	-0.0059	-0.2302	0.0993	-0.1016	-0.1398	-0.2893	-0.1076	-0.1061	1.0000			
Gross return (ln)	-0.0684	0.4927	-0.2849	0.1497	-0.3231	0.2730	0.5637	0.4867	-0.2254	1.0000		
Adjusted Alpha	-0.2561	0.4115	-0.3862	0.1355	-0.3988	0.0168	0.5487	0.1769	-0.0932	0.7673	1.0000	
Adjusted Alpha (net)	-0.2991	0.3996	-0.4198	0.1190	-0.5225	0.0161	0.5498	0.1737	-0.0517	0.7682	0.9841	1.0000

Definitions:

Active Share refers to the current active share while **Active Share Dummy** shows the below median size active share dummy. **TE** refers to Tracking Error. **TR** refers to turnover ratio. **TER** is the Total Expense Ratio and **Securities** the number of stocks invested. **TNA** refers to Total Net Assets. **Sharpe ratio** represents the average 5 years Sharpe ratio while **Tenure** stands for the management tenure (dummy). **Gross Return** refers to the fund average gross return and **Adjusted Alpha** refers to the gross return over the adjusted benchmark. **Adjusted Alpha (net)** refers to the alpha less annual fees.

Table D – Jarque-Bera Test for Non-Normality (Active Share Regression)

The table shows the Jarque-Bera test for non-normality. The **p-values** show that H0 for non-normality is rejected which leaves all variables normally distributed.

	Active Share	TE 5 years (ln)	TR (ln)	TER (ln)	Securities	TNA (ln)	Gross return (ln)	Adjusted Alpha	Sharpe ratio 5 years	Tenure
Skewness	0.2153	0.2277	-0.3964	-0.0847	0.9039	0.1825	-0.1320	-0.1363	0.6213	0.9113
Kurtosis	1.9503	1.9790	2.5003	4.1970	3.3369	2.9363	2.4192	2.6925	3.7536	1.8305
Jarque-Bera	2.1989	2.1354	1.5002	2.4966	5.7767	0.2346	0.6954	0.2884	3.6081	8.0115
Probability	**0.3331**	**0.3438**	**0.4723**	**0.2870**	**0.0557**	**0.8893**	**0.7063**	**0.8657**	**0.1646**	**0.0182**
Observations	41	41	41	41	41	41	41	41	41	41

Table E – White Test for Heteroscedasticity (Active Share Regression)

The White Test is dependent on the **F-Statistic**. The test equation is using the squared residuals as a dependent variable to test for heteroscedasticity. All independent variables are expressed as squared variables accordingly.

Sample: 1 41

Included observations: 41

F-statistic	**0.7428**	Prob. F(9,31)		0.6675
Obs*R-squared	7.2730	Prob. Chi-Square(9)		0.6087
Scaled explained SS	4.5309	Prob. Chi-Square(9)		0.8731

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	0.0622	0.0310	2.0032	0.0540
TE 5 years (ln)^2	-0.0033	0.0157	-0.2082	0.8365
TR (ln) ^{2}	0.0000	0.0007	0.0602	0.9524
TER (ln) ^{2}	-0.0085	0.0121	-0.7035	0.4870
Securities^2	0.0000	0.0000	-1.9279	0.0631
TNA (ln) ^{2}	-0.0001	0.0001	-1.1456	0.2607
Gross return (ln) ^{2}	27.6849	144.0153	0.1922	0.8488
Adjusted Alpha^2	205.5665	517.1696	0.3975	0.6937
Sharpe ratio 5 years^2	-0.0069	0.1879	-0.0368	0.9709
Tenure^2	0.0018	0.0111	0.1583	0.8752

R-squared	0.1774	Mean dependent var	0.0182
Adjusted R-squared	-0.0614	S.D. dependent var	0.0273
S.E. of regression	0.0281	Akaike info criterion	-4.0984
Sum squared resid	0.0245	Schwarz criterion	-3.6804
Log likelihood	94.0164	Hannan-Quinn criter.	-3.9462
F-statistic	0.7428	Durbin-Watson stat	1.7977
Prob(F-statistic)	0.6675		

Definitions:
TE refers to Tracking Error. **TR** refers to turnover ratio. **TER** is the Total Expense Ratio and **Securities** the number of stocks invested.
TNA refers to Total Net Assets. **Gross Return** refers to the fund average gross return and **Adjusted Alpha** refers to the gross return over the adjusted benchmark. **Sharpe ratio** represents the average 5 years Sharpe ratio while **Tenure** stands for the management tenure (dummy).

Appendix II – Returns Regression Tables

Table F – Predictive Fund Performance Regression 2010 – 2015

The dependent variable is active share; all other variables are computed as previously described. Based on standard error, the t-statistics can be found (in parentheses), the respective coefficient above each t-statistics. ***shows significance on a 1% level, **on a 5% level and *on a 10% level in terms of p-values.

	Gross Returns (ln)		Adjusted Alpha		Adjusted Alpha (net)	
Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)
Independent Variables:						
Endogenous Variables:						
Active Share	-0.0024	-0.0026	-0.0020	-0.0023	-0.0021	-0.0024
	(-1.09)	(-1.15)	(-0.74)	(-0.85)	(-0.73)	(-0.85)
Active Share		-0.0009		-0.0017		-0.0018
(Dummy)*		(-0.76)		(-1.16)		(-1.23)
TE5 years (ln)	-0.0007	-0.0007	-0.0012	-0.0014	-0.0015	-0.0016
	(-0.69)	(-0.77)	(-1.06)	(-1.19)	(-1.22)	(-1.37)
TR (ln)	0.0001	0.0002	0.0002	0.0003	0.0002	0.0003
	(0.52)	(0.70)	(0.77)	(1.05)	(0.71)	(1.03)
TER (ln)	-0.0011	-0.0009	-0.0015	-0.0012	-0.0029**	-0.0026*
	(-1.00)	(-0.85)	(-1.10)	(-0.89)	(-2.10)	(-1.88)
Securities	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(-0.19)	(-0.28)	(-1.26)	(-1.40)	(-1.27)	(-1.43)
Exogenous Variables:						
TNA (ln)	0.0005***	0.0007**	0.0006**	0.0009**	0.0006**	0.0009**
	(2.86)	(2.37)	(2.36)	(2.39)	(2.32)	(2.42)
Sharpe ratio 5 years	0.0266***	0.0279***	0.0164*	0.0187*	0.0180*	0.0205**
	(3.46)	(3.52)	(1.72)	(1.93)	(1.85)	(2.08)
Tenure	-0.0008	-0.0009	-0.0004	-0.0006	-0.0003	-0.0005
	(-1.14)	(-1.26)	(-0.49)	(-0.72)	(-0.35)	(-0.60)
Constant	-0.0032	-0.0060	-0.0102**	-0.0154**	-0.0112**	-0.0168**
	(-0.87)	(-1.15)	(-2.24)	(-2.42)	(-2.41)	(-2.59)
N	41	41	41	41	41	41
R^2	0.6268	0.6337	0.4681	0.4902	0.5484	0.5695

(Dummy)*: 0 = below median, 1 = above median

Definitions:

Endogenous Variables refer to variables controllable by management. **Active Share** refers to the current active share while **Active Share Dummy** shows the below median size active share dummy. **TE** refers to Tracking Error. **TR** refers to turnover ratio. **TER** is the Total Expense Ratio and **Securities** the number of stocks invested.

Exogenous Variables refer to variables not controllable for management. **TNA** refers to Total Net Assets. **Sharpe ratio** represents the average 5 years Sharpe ratio while **Tenure** stands for the management tenure (dummy).

Constant equals to the intercept of the regression. **N** stands for the sample size tested and **R^2** the coefficient of determination functioning as a measure of fit.

Table G – Jarque-Bera Test for Non-Normality (Gross Returns (ln))

The table shows the Jarque-Bera test for non-normality. The **p-values** show that H_0 for non-normality is rejected which leaves all variables normally distributed.

	Gross Returns (ln)	Active Share	Active Share Dummy	TE5 years (ln)	TR (ln)	TER (ln)	Securities	TNA (ln)	Sharpe ratio 5 years	Tenure
Skewness	-0.1320	0.2153	-0.0488	0.2277	-0.3964	-0.0847	0.9039	0.1825	0.6213	0.9113
Kurtosis	2.4192	1.9503	1.0024	1.9790	2.5003	4.1970	3.3369	2.9363	3.7536	1.8305
Jarque-Bera	0.6954	2.1989	6.8333	2.1354	1.5002	2.4966	5.7767	0.2346	3.6081	8.0115
Probability	**0.7063**	**0.3331**	**0.0328**	**0.3438**	**0.4723**	**0.2870**	**0.0557**	**0.8893**	**0.1646**	**0.0182**
Observations	41	41	41	41	41	41	41	41	41	41

Table H – Jarque-Bera Test for Non-Normality (Adjusted Alpha Regression)

The table shows the Jarque-Bera test for non-normality. The **p-values** show that H_0 for non-normality is rejected which leaves all variables normally distributed.

	Adjusted Alpha	Active Share	Active Share Dummy	TE5 years (ln)	TR (ln)	TER (ln)	Securities	TNA (ln)	Sharpe ratio 5 years	Tenure
Skewness	-0.1363	0.2153	-0.0488	0.2277	-0.3964	-0.0847	0.9039	0.1825	0.6213	0.9113
Kurtosis	2.6925	1.9503	1.0024	1.9790	2.5003	4.1970	3.3369	2.9363	3.7536	1.8305
Jarque-Bera	0.2884	2.1989	6.8333	2.1354	1.5002	2.4966	5.7767	0.2346	3.6081	8.0115
Probability	**0.8657**	**0.3331**	**0.0328**	**0.3438**	**0.4723**	**0.2870**	**0.0557**	**0.8893**	**0.1646**	**0.0182**
Observations	41	41	41	41	41	41	41	41	41	41

Table I – Jarque-Bera Test for Non-Normality (Adjusted Alpha (net))

The table shows the Jarque-Bera test for non-normality. The **p-values** show that H_0 for non-normality is rejected which leaves all variables normally distributed.

	Adjusted Alpha (net)	Active Share	Active Share Dummy	TE5 years (ln)	TR (ln)	TER (ln)	Securities	TNA (ln)	Sharpe ratio 5 years	Tenure
Skewness	-0.3513	0.2153	-0.0488	0.2277	-0.3964	-0.0847	0.9039	0.1825	0.6213	0.9113
Kurtosis	2.9156	1.9503	1.0024	1.9790	2.5003	4.1970	3.3369	2.9363	3.7536	1.8305
Jarque-Bera	0.8557	2.1989	6.8333	2.1354	1.5002	2.4966	5.7767	0.2346	3.6081	8.0115
Probability	**0.6519**	**0.3331**	**0.0328**	**0.3438**	**0.4723**	**0.2870**	**0.0557**	**0.8893**	**0.1646**	**0.0182**
Observations	41	41	41	41	41	41	41	41	41	41

Table J – White Test for Heteroscedasticity (Gross Returns (ln))

The White Test is dependent on the **F-Statistic**. The test equation is using the squared residuals as a dependent variable to test for heteroscedasticity. All independent variables are expressed as squared variables accordingly.

Sample: 1 41
Included observations: 41

F-statistic	**1.5092**	Prob. F(9,31)	0.1884
Obs*R-squared	12.4910	Prob. Chi-Square(9)	0.1870
Scaled explained SS	11.2709	Prob. Chi-Square(9)	0.2576

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	0.0000	0.0000	-1.2856	0.2081
Active Share^2	0.0000	0.0000	0.7463	0.4611
Active Share Dummy^2	0.0000	0.0000	-1.9337	0.0623
TE 5 years (ln)^2	0.0000	0.0000	1.6253	0.1142
TR (ln) ^{2}	0.0000	0.0000	-0.9332	0.3579
TER (ln) ^{2}	0.0000	0.0000	-0.5405	0.5927
Securities^2	0.0000	0.0000	0.5450	0.5897
TNA (ln) ^{2}	0.0000	0.0000	1.9636	0.0586
Sharpe ratio 5 years^2	0.0000	0.0000	-0.4181	0.6788
Tenure^2	0.0000	0.0000	-0.4116	0.6834

R-squared	0.3047	Mean dependent var	0.0000
Adjusted R-squared	0.1028	S.D. dependent var	0.0000
S.E. of regression	0.0000	Akaike info criterion	-21.3576
Sum squared resid	0.0000	Schwarz criterion	-20.9397
Log likelihood	447.8314	Hannan-Quinn criter.	-21.2054
F-statistic	1.5092	Durbin-Watson stat	2.2367
Prob(F-statistic)	0.1884		

Definitions:
Active Share refers to the current active share while **Active Share Dummy** shows the below median size active share dummy. **TE** refers to Tracking Error. **TR** refers to turnover ratio. **TER** is the Total Expense Ratio and **Securities** the number of stocks invested.
TNA refers to Total Net Assets. **Sharpe ratio** represents the average 5 years Sharpe ratio while **Tenure** stands for the management tenure (dummy).

Table K – White Test for Heteroscedasticity (Adjusted Alpha Regression)

The White Test is dependent on the **F-Statistic**. The test equation is using the squared residuals as a dependent variable to test for heteroscedasticity. All independent variables are expressed as squared variables accordingly.

Sample: 1 41

Included observations: 41

F-statistic	**0.5739**	Prob. F(9,31)	0.8078
Obs*R-squared	5.8558	Prob. Chi-Square(9)	0.7543
Scaled explained SS	4.2858	Prob. Chi-Square(9)	0.8916

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	0.0000	0.0000	-0.5125	0.6119
Active Share^2	0.0000	0.0000	-1.0272	0.3123
Active Share Dummy^2	0.0000	0.0000	-1.9752	0.0572
TE 5 years (ln)^2	0.0000	0.0000	0.6822	0.5002
TR (ln) ^{2}	0.0000	0.0000	-0.3806	0.7061
TER (ln) ^{2}	0.0000	0.0000	0.0212	0.9832
Securities^2	0.0000	0.0000	-0.3373	0.7382
TNA (ln) ^{2}	0.0000	0.0000	1.2800	0.2100
Sharpe ratio 5 years^2	0.0001	0.0001	1.1380	0.2638
Tenure^2	0.0000	0.0000	-0.1146	0.9095

R-squared	0.1428	Mean dependent var	4.42E-06
Adjusted R-squared	-0.1060	S.D. dependent var	7.17E-06
S.E. of regression	7.54E-06	Akaike info criterion	-20.54534
Sum squared resid	1.76E-09	Schwarz criterion	-20.1273
Log likelihood	431.1794	Hannan-Quinn criter.	-20.3931
F-statistic	0.5739	Durbin-Watson stat	1.65494
Prob(F-statistic)	0.8078		

Definitions:
Active Share refers to the current active share while **Active Share Dummy** shows the below median size active share dummy. **TE** refers to Tracking Error. **TR** refers to turnover ratio. **TER** is the Total Expense Ratio and **Securities** the number of stocks invested.
TNA refers to Total Net Assets. **Sharpe ratio** represents the average 5 years Sharpe ratio while **Tenure** stands for the management tenure (dummy).

Table L – White Test for Heteroscedasticity (Adjusted Alpha Regression)

The White Test is dependent on the **F-Statistic**. The test equation is using the squared residuals as a dependent variable to test for heteroscedasticity. All independent variables are expressed as squared variables accordingly.

Sample: 1 41

Included observations: 41

F-statistic	**0.5526**	Prob. F(9,31)		0.8243
Obs*R-squared	5.6684	Prob. Chi-Square(9)		0.7726
Scaled explained SS	4.1800	Prob. Chi-Square(9)		0.8992

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	0.0000	0.0000	-0.4701	0.6416
Active Share^2	0.0000	0.0000	-0.8834	0.3838
Active Share Dummy^2	0.0000	0.0000	-2.1482	0.0396
TE 5 years (ln)^2	0.0000	0.0000	0.2273	0.8217
TR (ln) ^{2}	0.0000	0.0000	-0.2922	0.7721
TER (ln) ^{2}	0.0000	0.0000	0.2668	0.7914
Securities^2	0.0000	0.0000	-0.6920	0.4941
TNA (ln) ^{2}	0.0000	0.0000	1.3621	0.1830
Sharpe ratio 5 years^2	0.0001	0.0001	1.0896	0.2843
Tenure^2	0.0000	0.0000	-0.3883	0.7004

R-squared	0.1383	Mean dependent var	0.0000
Adjusted R-squared	-0.1119	S.D. dependent var	0.0000
S.E. of regression	0.0000	Akaike info criterion	-20.4671
Sum squared resid	0.0000	Schwarz criterion	-20.0492
Log likelihood	429.5758	Hannan-Quinn criter.	-20.3149
F-statistic	0.5526	Durbin-Watson stat	1.7234
Prob(F-statistic)	0.8243		

Definitions:
Active Share refers to the current active share while **Active Share Dummy** shows the below median size active share dummy. **TE** refers to Tracking Error. **TR** refers to turnover ratio. **TER** is the Total Expense Ratio and **Securities** the number of stocks invested.
TNA refers to Total Net Assets. **Sharpe ratio** represents the average 5 years Sharpe ratio while **Tenure** stands for the management tenure (dummy).

Table M – Ramsey RESET Test for Non-Linearity

The RESET Test tests for the existence of non-linearity which is existent when rejecting H_0 from the f-statistics.

Ramsey RESET Test for regression model 2 – (1)

Omitted Variables: Squares of fitted values

	Value	df	Probability
t-statistic	0.3548	31	0.7251
F-statistic	**0.1259**	**(1, 31)**	**0.7251**
Likelihood ratio	0.1662	1	0.6835

F-test summary:

	Sum of Sq.	df	Mean Squares
Test SSR	4.98E-07	1	4.98E-07
Restricted SSR	0.0001	32	3.85E-06
Unrestricted SSR	0.0001	31	3.95E-06

LR test summary:

	Value	df
Restricted LogL	202.5055	32
Unrestricted LogL	202.5887	31

Ramsey RESET Test for regression model 2 – (3)

Omitted Variables: Squares of fitted values

	Value	df	Probability
t-statistic	1.2219	31	0.2309
F-statistic	**1.4931**	**(1, 31)**	**0.2309**
Likelihood ratio	1.9287	1	0.1649

F-test summary:

	Sum of Sq.	df	Mean Squares
Test SSR	8.70E-06	1	8.70E-06
Restricted SSR	0.0001	32	5.91E-06
Unrestricted SSR	0.0001	31	5.82E-06

LR test summary:

	Value	df
Restricted LogL	193.6905	32
Unrestricted LogL	194.6549	31

Ramsey RESET Test for regression model 2 – (5)

Omitted Variables: Squares of fitted values

	Value	df	Probability
t-statistic	1.6127	31	0.1169
F-statistic	**2.6008**	**(1, 31)**	**0.1169**
Likelihood ratio	3.3031	1	0.0691

F-test summary:

	Sum of Sq.	df	Mean Squares
Test SSR	1.52E-05	1	1.52E-05
Restricted SSR	0.0001	32	6.14E-06
Unrestricted SSR	0.0001	31	5.85E-06

LR test summary:

	Value	df
Restricted LogL	192.9064	32
Unrestricted LogL	194.5579	31

Appendix III – Return Tables

Table N – ETF Average Returns 2010 – 2015

ETF Funds	Monthly Returns	Annual Returns	Annual Fees
Amundi ETF MSCI Germany UCITS ETF	1,06%	12,71%	0,25%
ComStage DAX TR UCITS ETF	1,06%	12,67%	0,08%
db x-trackers DAX UCITS ETF (DR)	1,05%	12,57%	0,09%
Deka DAX® UCITS ETF	1,00%	11,96%	0,15%
iShares DAX® (DE)	0,99%	11,87%	0,15%
Lyxor UCITS ETF DAX	1,04%	12,49%	0,15%
Average:	1,03%	12,38%	0,15%

Source: Eikon, 2015

Table O – Mutual Fund Average Returns 2010 – 2015

Mutual Funds	Monthly Returns	Annual Returns	Annual Fees
Actis Champs SEL – Actis Aktien Dtl	1,20%	14,35%	2,53%
Aktienfonds Deutschland Spezial R VA	0,46%	5,53%	1,65%
Allianz German Equity - AT - EUR	1,22%	14,68%	1,80%
Allianz Thesaurus - AT - EUR	1,00%	11,97%	1,80%
Candriam Equities L Germany C Cap	0,95%	11,41%	1,50%
CS (Lux) Small and Mid Cap Germany	1,37%	16,46%	1,92%
D&R TA Deutsche Aktien P	0,50%	5,99%	1,80%
DB Platinum III Platow R1C	1,58%	18,92%	1,00%
DKO-Lux-Aktien Deutschland	0,75%	9,02%	1,94%
DWS Aktien Strategie Deutschland	1,49%	17,88%	1,45%
DWS Deutschland	1,28%	15,38%	1,40%
DWS German Equities Typ O	1,29%	15,51%	1,45%
DWS German Small/Mid Cap	1,47%	17,65%	1,40%
FPM Funds Stockpicker Germany All C	0,98%	11,73%	0,90%
FPM Funds Stockpicker Ger Small/Mid Cap	1,20%	14,38%	1,25%
Generali IS German Equities D Cap EUR	0,91%	10,88%	1,60%
HAIG MB Max Value	0,78%	9,34%	1,95%
HSBC Trinkaus German Equity	0,94%	11,32%	1,00%
ICM Portfolio - Select	0,79%	9,49%	1,59%
JB EF German Value-EUR B	1,15%	13,77%	1,20%
Jyske Invest German Equities	1,07%	12,86%	1,00%
KR FONDS Deutsche Aktien Spezial P	0,52%	6,21%	1,30%
LBBW Aktien Small & MidCaps Detuschl.	1,01%	12,17%	1,50%
LBBW Exportstrategie Deutschland	0,89%	10,72%	1,50%
Lupus alpha Smaller German Champions A	1,44%	17,25%	1,50%
Lux-Euro-Stocks TecDAX	1,08%	12,94%	1,80%
MainFirst - Germany Fund A	1,38%	16,60%	1,50%
MAS Value - Select	0,83%	10,01%	2,90%
MAV Invest - Aktienfonds	0,52%	6,28%	4,91%
Mediolanum Challenge Germany Equity	0,71%	8,54%	1,72%
NDACinvest - Aktienfonds	0,55%	6,58%	3,83%
PARVEST Equity Germany C C EUR	0,96%	11,47%	2,02%
Pioneer Investments Aktien Deutschland	0,96%	11,54%	0,50%
Pioneer Investments German Equity	1,06%	12,73%	1,50%
SSgA Germany Index Equity Fund P	1,04%	12,53%	0,77%
Swiss & Global AM Deutsche Aktien	1,15%	13,80%	1,50%
UBAM Dr Ehrhardt German Equity	0,81%	9,77%	1,36%
UBS (D) Equity Fund - Mid Caps Ger	1,21%	14,58%	1,80%
UniDeutschland	0,98%	11,75%	0,90%
UniDeutschland XS	1,32%	15,88%	1,55%
Value - Holdings Capital Partners Fund	0,95%	11,35%	1,25%
Average:	1,02%	12,22%	1,65%

Source: Eikon, 2015