Fund Activity and performance: A closer look at the swedish mutual fund industry

Henrik Ivarsson and Hugo Olofsson

May 24, 2016

UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Bachelor of Science in Financial Economics

Bachelor Thesis Spring 2016

Supervisor: Ted Lindblom

Abstract

This paper investigates the level of fund activity and its effect on fund performance on the Swedish market. By analyzing fund characteristics between 2005 to 2015, we find that among the active fund alternatives, the most active funds generates the highest level of alpha. These funds however, can still not beat index funds and are not able to predict positive returns compared to their benchmark. Further, we can conclude that an increasing fraction of explicit indexing increases the overall level of activity among active funds. Finally, we find that the passively managed funds that pretend to be active, *i.e.* closet indexers, do not seem to perform any worse compared to other fund categories in exception of the category active stock pickers. A finding contradictory to earlier studies and tells us that these funds have lowered their fees to a large extent. Investors should rather avoid investing in the categories factor bets and moderately active funds than closet indexers, since they have a high fee in proportion to their level of activity.

Keywords: Fund Activity, Fund Performance, Mutual Funds, Index Funds, Active Share, Tracking Error, Fees

JEL Classifications: G15, G18, G23

Contents

1 Introduction **1**

2 Theoretical Framework and Empirical Evidence **4**
2.1 Previous Research on the Level of Activity 4
2.2 Previous Research on Fund Performance 7
2.3 Previous Research on the Swedish Market 8

3 Data and Methodology **9**
3.1 Databases 9
3.2 Sample Selection 10
3.3 Definition of Activity Measures 11
3.3.1 Tracking Error 12
3.3.2 Active Share 13
3.3.3 Definition of Fund Categorization 15
3.4 Other Variables 17
3.4.1 Flow 17
3.4.2 Fund Industry Size 18
3.4.3 Explicit Indexing 18
3.4.4 Age 18
3.4.5 Carhart Model 19
3.5 Analysis model 19

4 Results **20**
4.1 Descriptive Statistics 20
4.1.1 Active Share 20
4.1.2 Explicit Indexing 23
4.1.3 Summary Statistics 24
4.1.4 Fund Performance for Categories & Index Funds 26
4.2 Regression Results 27
4.2.1 Results on the level of Activity 27
4.2.2 Results on Performance 28

5 Potential Biases and Robustness Checks **31**

6 Conclusion **34**

1 Introduction

When an investor decides to use mutual funds as an investment vehicle, there are a wide range of aspects that needs to be considered: what geographical focus is appropriate? Should the fund only keep long positions in stocks or also be able to short the market? More importantly, should the investor choose passive investment management such as an index fund or rely on the expertise of a professional fund manager that claims to have stock picking abilities? If the latter alternative seems the most appropriate for the investor, what level of return can this actively managed fund realize and is it even possible for an active fund manager to beat its benchmark? According to various researches, the answer is often no.[1]. However if having invested in an actively managed fund, investors want to be given the chance of beating the benchmark and by definition this can only be achieved by deviating from a benchmark portfolio in terms of holdings and thereby taking on more risk according to traditional theories, Markowitz (1952). In theory, there are several ways of determining the level of activity of a fund, but in practice it can be rather difficult and time demanding due to the lack of transparency between investors and fund managers.

In research conducted by Cremers and Petajisto (2009), a measure of fund activity called active share is introduced. The measurement is used by the authors to distinguish between different levels of activity between funds. By using the active share measure, they discover that there seems to be many funds that have a low level of activity but still charges a fee that represents the opposite. The funds with these characteristics are called closet indexers. A closet indexer is an actively managed fund that never will give itself the opportunity to beat its benchmark, simply due to the fact that the holdings are too similar to the benchmark. Hence, a closet indexer is like an expensive version of an explicit index fund where investors pay for active management but receive passive management.

In further research conducted by Cremers et al. (2015), the authors use the active share measure in the same way as the paper by Cremers and Petajisto (2009). They find that the average proportion of closet indexers of the domestic funds, on a worldwide basis, is 21 percent.[2] For the Swedish fund industry the same figure is 73 percent, which is one of

[1] See amongst Jensen (1968), Gruber (1996) and Wermers (2000)

[2] Funds investing in their country of domicile are for instance a Swedish mutual fund investing in the Swedish market, *i.e.* domestic funds

the higher figures in the sample. Closet indexing is something that seems to be commonly occurring on the Swedish fund market where the managers are not as active as they claim they are.

The debate about the closet indexing problem has in the last couple of years intensified in Sweden due to the large proportion of actively managed funds in the Swedish fund industry that in fact are passively managed. One particular Swedish fund family (Swedbank Robur) has been in the full glare of publicity the last two years after the Swedish association *Aktiespararna* sued them in 2013 for this wolf in sheep's clothing phenomenon called closet indexing. Two of their largest funds claimed to be active but in fact tended to act more as index funds. This leads to the results that after deducting fees from benchmark-adjusted returns, investing in a closet indexer receives a lower net return than when investing in a passive index fund. The investor is worse off because of the managers inability to act in the interest of the investors in terms of maximizing profits.

A fund manager is supposed to act in the interest of the investors when making strategic investment decisions. Investors trust the managers' stock picking abilities and in return, the actively managed funds should at least give their investors an opportunity to beat the benchmark. This is something that according to Carhart (1997) has been proven to be a non-existing anomaly. Even if the fund managers deviate a lot from the funds' benchmark indices, the deviation is only relevant if the level of activity has a clear connection to the level of performance of the fund. This is a relationship which will be investigated in this paper. If not deviating from the fund's benchmark, investors might be paying large fees for the same beta exposure as they get from an explicit index fund, *i.e.* the correlation for an closet indexer with the market portfolio will be the same as for explicit index funds and ultimately, the return will also be in line with explicit index funds.

Before the active share measured was introduced in 2009, activity was traditionally determined by the tracking error of a fund. The tracking error measures the volatility in the returns of a fund versus the returns in the comparable index itself. An important difference between the two measures is that active share analyzes the actual portfolio holdings of the fund, whereas the tracking error measures the risk connected to the activity, also called active risk. Hence, there are different interpretations of these two activity measures. Peta-

jisto (2013) conducts a study where the measurements are developed further by combining both the tracking error and the active share measurement. By doing so, Petajisto is able to categorize the funds into various types of active management. In the study, he finds evidence that funds with the most active stock pickers tend to outperform their benchmark indices even after fees, whereas closet indexers tend to underperform on average.

Having this background and previous research in consideration, we will use a quantitative approach in order to answer the following research questions:

- Can the activity measurement active share alone predict fund performance or should it be combined with tracking error in order to distinguish between fund strategies and their effect on fund performance?
- Does closet indexers underperform in terms of benchmark and- risk-adjusted returns, compared to other fund categories?
- Can the fraction of explicit indexing funds serve as a proxy for increased competition in the Swedish mutual fund industry and how will the competition affect the level of activity and fees among the other fund categories?

The purpose of this paper is to explain what factors, in terms of fund characteristics of Swedish mutual funds, that can predict the level of active share. In more detail, we will as in the paper by Cremers (2015), test whether the fraction of explicit indexing can serve as a proxy for increased competition between passively managed funds and actively managed funds. To our knowledge, this relationship has never before been tested on the Swedish market. Since explicit index funds are a growing product segment in the country according to Flam & Vestman (2014), our findings may be of economic significance for the fund industry as well of importance to further research.

Further, this paper aims at distinguishing between different levels of active management, measured both in terms of active share and tracking error, and how these measurements can predict future returns. By doing this, we aim to see what fund strategy that has the largest opportunity of creating wealth for investors. This is in line with the paper by Petajisto (2013), however our research focuses on Swedish equity mutual funds investing solely on the Swedish market.

Research of activity and performance in the Swedish mutual fund industry is of great economic significance since the investment industry in Sweden is well-developed. The Swedish market covers 3.2 per cent of total fund assets managed in Europe, Efama (2016), which can be compared to Germany that currently manages 3.8 per cent of total assets despite having a population more than five times as large as the population of Sweden.[3]

This paper proceeds as follows: in section 2 we will present a theoretical framework and empirical evidence on the level of activity and fund performance as well as previous studies on the Swedish market. Further we will present data and methodology in section 3, which then leads to section 4 consisting of results that include descriptive statistics, regressions and analysis of the results. The paper will end with a potential bias and robustness check in section 5 and finally a conclusion in section 6.

2 Theoretical Framework and Empirical Evidence

This section is divided into three parts. The section starts with previous research on the level of activity, continues with previous research on performance and finishes of with previous research on the Swedish market.

2.1 Previous Research on the Level of Activity

Tracking error has been used for a long time as a measurement of activity of a mutual fund. Tracking error measures how much the fund is deviating from its benchmark or, expressed in another way, how well it is following the benchmark, Grinold & Kahn (1999). Tracking error is often measured in standard deviations but could also be measured in terms of returns. A high tracking error of a fund indicates that the fund is deviating from its benchmark and a low tracking error means that the fund is following its benchmark rather tightly. One should however remember, by deviating more from the benchmark the fund managers take on more risk, hence the measurement is often also referred to as active risk. A bigger deviation indicates an excess risk in comparison to the benchmark portfolio.

[3]Since Sweden and Germany are similar in terms of economic and social development, we find this a relevant peer for this comparison

This excessive risk is something that an investor should be rewarded for in terms of returns compared to the benchmark in accordance with traditional theories such as the efficient frontier, Markowitz (1952).

A fund manager can only outperform the market either by investing in stocks different from the benchmark, either by the strategy of stock selection or by the strategy of factor timing according to Cremers & Petajisto (2009). When implementing these strategies the fund incurs a higher degree of deviation and it is only then, by definition, a fund has the potential of beating the index or benchmark. A stock selection strategy involves picking stocks that the manager think has the chance to outperform comparable firms. When using a factor timing strategy, the manager invests solely to a specific industry or sector.

Since the tracking error is just measuring activity through standard deviations of the returns, Cremers & Petajisto (2009) introduced a new measurement called active share. Active share analyzes a funds actual holdings and not just its volatility. A fund that completely engages in stock picking, and in the same time diversifies across different industries to minimize risk, will generate a low tracking error *i.e.* low risk. A fund that engages in a factor bets strategy and invests completely in one sector to generate abnormal returns, will not diversify its risk and inevitably generate a high tracking error. If you then only measure a manager's activity by a funds tracking error, this would indicate that the stock picker that diversifies is not as active as the fund engaging in factor bets. This would be an inadequate conclusion due to the fact that the fund´s actual holdings are not analyzed and hence the active share measure is not taken into account, which would show a different results where both strategies generates a high active share.

Due to the issue described above with the two different strategies and their effect on the different activity measures, Petajisto (2013) develops a categorization method of mutual funds, a categorization which is based on the funds' level of active share and tracking error. Here, he uses active share as a measure of stock selection abilities and tracking error as a measure of exposure to systematic risk *i.e.* market risk. By doing so, he comes up with five different categories for actively managed funds: *active stock pickers, factor bets, concentrated funds, moderately active funds and closet indexers.*

Active stock pickers invests in stocks they believe have the ability to outperform the market and applies this strategy to various industries and sectors to generate a diversified portfolio, hence they incur a lower tracking error but a high active share. Closet indexers uses neither a stock picking strategy nor a factor bet strategy and generates a low tracking error and low active share. According to Petajisto (2013), a factor bet fund isolates their investments to one specific industry or sector *i.e.* only implementing a factor bet strategy and no stock selection strategy, hence incurring large tracking error compared to the index. Concentrated funds combine active stock picking strategy and factor bet strategy, hence picks stocks in a specific industry or sectors, which leads to a very high systematic risk and active share. A moderately active fund is a fund without a well-defined strategy. The research by Petajisto (2013) was based on US data and the author found that closet indexing has increased in popularity since 2007. He also finds evidence for that closet indexers, who has holdings very similar to their benchmark after deducting fees, will underperform compared to its benchmark. The one category that has provided value to their investors, according to the author, is active stock pickers. The active stock pickers have, net of fees and benchmark returns, generated a 1.26% return per year compared to their comparable benchmark.

In the study by Cremers (2015), the authors use the active share measure as in the research by Cremers & Petajisto (2009), but expand their research from US data solely to a cross-country analysis including 32 countries. In this paper, the authors also hypothesis about the presence of explicit index funds on the market and its effect on how active managers are. They find significant evidence that an increasing presence of explicit indexing funds on the market tend to make the actively managed funds more active. The reason for this phenomenon, according to the authors, is that closet indexers must further differentiate their funds to keep up with the increased level of competition in the market. They can, according to the authors, do this either by lowering their fee or by creating bigger excess returns in comparison to their benchmark, hence becoming more active since the only way they can create excess return is by investing in stocks outside the benchmark. Their findings, that explicit indexing has a positive effect on active share, has the reverse relationship in countries where investors have limited options of explicit index funds. Here, active managers tend to engage more in closet indexing.

2.2 Previous Research on Fund Performance

The pioneer researcher that covers the topic whether a fund manager have the ability to beat the market is Michael C. Jensen. He develops a risk-adjusted return-measure called Jensen's alpha in order to see if fund managers have some predictive abilities in evaluating future mutual fund performance, Jensen (1968). By combining measurements developed by *e.g.* Sharpe, Lindtner and Treynor, he analyzes the risk-adjusted return of 115 mutual funds in the years 1945-1964. He finds that on average, mutual funds are not able to predict security prices well enough to outperform the market.

The model nowadays most widely used to adjust for risk when evaluating fund performance, is developed by Carhart (1997), and is referred to generally as Carhart's four-factor model. The Carhart model adds one additional factor, the momentum factor, to the Fama-French three factor model. Carhart shows that these four factors almost exclusively can explain the performance persistence among mutual funds.

In a paper by Chen et al. (2004), the authors investigate the effect of fund size on performance. There are obvious reasons one might argue that fund size should have a positive relationship with fund performance, *i.e.* a larger fund can spread the fixed costs over a larger set of assets. However, the authors find that fund returns decline with fund size, one explanation for this is that large funds may have difficulties investing in small and illiquid stocks.

In line with Chen (2004), a more recent study conducted by Ferreira et al. (2011) goes even further in predicting what factors that might explain fund performance by performing a cross-country study with 27 countries included in their data, more than any previous study. Their findings of size effect on mutual fund performance concludes that, outside of the United States, size of fund affects returns positively while the negative relationship holds when looking at U.S. funds. In the same paper, they also look at fund flow and its impact on fund performance, they find that more investment tend to flow to funds with high future returns, *i.e.* fund flow is a positive and significant explanatory variable to fund performance, this holds only in the case of non-U.S. funds.

Fund flow and its effects on fund performance were earlier analyzed more detailed in a paper by Sapp & Tiwari (2004). They find that on the U.S market, fund flow can almost exclusively be predicted by the momentum effect, *i.e.* when including the momentum factor in the regression trying to predict fund returns, flow is no longer a significant variable. Earlier studies such as Gruber (1996) and Zheng (1999), did find a positive effect in funds where cash-inflow was prominent and the opposite where cash-outflow was prominent. When controlling for the fourth factor momentum, Sapp & Tiwari (2004) could later dismiss this effect.

2.3 Previous Research on the Swedish Market

Research conducted solely on the Swedish market is rather scarce. The research that has been made, has contributed to the debate of fund's abilities of beating the market rather than the focus of this paper, how the level of activity of a fund can be attributable to the performance of a fund.

Research conducted by Dahlquist et al. (2000), focuses on performance, fund attributes on the Swedish market and performance of funds. The authors found that good performance occurs among small equity funds, low fee funds and funds with a high trading activity. These findings, especially the activity finding, are similar to the findings made by Petajisto (2013) despite that they did not use the active share and tracking error measure. Instead, they measure activity by turnover and by commission in relation to fund size. Another finding by Dahlquist (2000) is that large equity funds tend to perform worse than small equity funds. One explanation for this could be that large equity funds are actually very large in proportion to the market, another explanation they present is that the large equity funds may be too large for aggressive trading.

Other research isolated to the Swedish market which is more up to date is one conducted by Flam & Vestman (2014). This study, similar to Dahlquist (2000), focuses on mutual funds' performance and persistence, they also add a new perspective to the research on the Swedish fund market by investigating excess return of index funds. In the beginning of their report they say the following: *"We also estimate index funds, a product segment that is still small but growing in Sweden"* Flam & Vestman, ((2014), p. 2). These facts

of explicit index funds being a small product segment in Sweden, aligns and provides evidence for the hypothesis of Cremers (2015), where more index funds on the market makes the actively managed funds more active. Further, the paper by Flam & Vestman (2014) finds that index funds in Sweden, on average deliver returns equal to the benchmark index they are following. When calculating for net returns *i.e.* deducting fees, the index funds substantially underperform their benchmarks, there is however a large dispersion between the index funds investigated and the median index fund has performed worse than the median actively managed fund.

3 Data and Methodology

This section starts with a description of how the data was collected and how the sample was constructed, we continue by explaining the variables used in the regressions as well as how the fund categories are created. Finally, we end with a description of the Carhart model and our interpretation of it as well as a description of how we analyze our data.

3.1 Databases

In order to generate active share we needed to obtain holdings data of each fund over time. Since all Swedish funds by law are bounded to report their stock holdings on a quarterly basis, We were able to receive this from the Swedish Financial Supervising Authority (*Finansinspektionen*). From these filings, we were able to withdraw institutional number, fund name, fund market size, fund size. The difference between fund market size and fund size is that fund market size is net of cash. We also extracted each funds respective stock-holdings and its current market value in the fund portfolio at each report date.

From the database Lipper, a fund screening tool from Thomson Reuters, we obtained a list of Swedish mutual equity funds. This list was used as our reference list for funds to be included in our sample, and was sorted on country registered for sale, country of domicile, asset class focus, asset universe, large-cap funds and mid/small-cap funds. The fund list consisted of 178 funds, where active, merged and liquidated funds were included in order to free our sample from a survivorship bias. This list was then complemented with ISIN-codes collected from Bloomberg, where we also retrieved expense ratios of the funds.

From Morningstar Direct we extracted both quarterly and monthly returns when sorting the funds by domicile, region of sale, large cap funds and mid/small cap funds. We also obtained each funds inception date, primary prospectus benchmark and ISIN-codes. We extracted the ISIN codes from this database for matching purposes with the previously collected ISIN-codes from Bloomberg.

From the database AQR Frazzini & Pedersen, (2016), we were able to obtain factor loadings data in order to compute risk adjusted Carhart alphas. The four-factor loadings are: monthly market return, measured as value-weighted portfolio of all available stocks in Sweden in excess of the risk free rate (MKT), small minus big (SMB), High Minus Low (HML) using the Book to Market definition of Fama French and the Up Minus Down (UMD) factor, which adjust for momentum effects. By accumulating the monthly returns, we got the loadings per quarter needed for us to run the regressions to obtain our alpha. The database of AQR is based on the methodology of Asness and Frazzini (2013) and Asness, Frazzini & Pedersen (2013).

3.2 Sample Selection

The reports from *Finansinspektionen* were available from the third quarter of 2000. However, there were several quarters missing in 2001, 2004 and 2005. To yield a consistency for our study, free from interruptions, the inception date of the sample period was chosen as the third quarter of 2005, since there were no missing quarters from that date and forward. This results in a sample period of 42 quarters starting from 2005/09/30 and ends 2015/12/31.

By using the list extracted from Lipper as a primary sortation of the funds of interest, we continued by merging the ISIN-codes from this dataset with the dataset from Finansinspektionen. We also kept the institutional number for each fund; this number is unique for every fund in the sense that it follows one particular fund throughout the whole sample period, even if the fund is acquired. For instance, this means that when Banco Etisk Sverige was bought by Swedbank Robur in 2009 and changed its name to Swedbank Robur Ethica Sverige, it still kept its unique institutional number. By this method, we could easily track

a fund despite the fact that it was acquired by another company or changed its name. This was another method for assuring that our sample was free from a survivorship bias.

After the match from the Lipper-list with the data from *Finansinspektionen* by its unique institutional number, we had a total of 129 funds. However, since this paper only will investigate long only equity mutual funds investing solely on the Swedish market, there were still a lot of funds in our dataset that had to be dropped. We therefore manually read through each fund's prospectus and adjusted the list by dropping those funds that did not fulfill our criteria, *e.g* funds that had the mandate to allocate a fraction of the portfolio to foreign markets and funds that had the ability to short the market.

After manually adjusting for funds of no interest for this paper, we ended up with a sample of 58 mutual funds consisting of 1957 observations and 10 index funds consisting of 291 observations.The index funds were kept in the sample in order to make a relevant comparison between actively managed funds and passive funds. Hence, they were not included in the sample when running regressions on active share and fund performance. We are well aware of that there are more funds available on the Swedish market, but due to the limitations in the databases used, this was the largest sample available and we believe the sample used is representative for the whole market. For a complete list of the funds investigated we refer to Table 13 in the appendix.

The sample is unbalanced due to the fact that new funds are opened, some funds are liquidated or merged and hence the number of observations vary in composition. Further, there are a total of 12 observations that have data on active share but not on quarterly returns, this depends on missing values from Morningstar. It can also depend on the reason that the fund recently launched and therefore quarterly returns cannot yet be computed. The missing values of quarterly returns also makes it impossible to compute tracking error for those quarters, which also incurs a missing value of fund category.

3.3 Definition of Activity Measures

In this paper we use two measures of activity: tracking error and active share, which both are defined below. However, a general definition of passive and active management

would be in place. Following the definition used by Cremers & Petajisto (2009), passive management of a portfolio is simply a management strategy that replicates the return on a target index by holding the same stocks and weightings that constitutes the same index. For instance, a passive managed portfolio could track OMXS30 by holding the index members *i.e.* the 30 largest companies on the Swedish stock market in the same proportions as the official index. Active management however is defined as any deviation from passive management and hence, the degree of management activity is defined as the "degree of deviation" from passive management according to Cremers & Petajisto (2009).

3.3.1 Tracking Error

Higher tracking error means higher active risk. In this paper we use the definition below, which expresses the standard deviation of the difference between the returns of the fund and the index it is trying to beat. Because the measure is computed in standard deviations the measurement is also called active risk.

$$TrackingError = Stdev(R_{fund} - R_{index}) \tag{1}$$

Where R_{fund} is the return of the fund and R_{index} is the return of the benchmark index. When computing the tracking error, we used monthly returns of our funds and indices and then calculated the tracking error on a quarterly basis *i.e.* calculating the differences for three months and then computing the standard deviation of those three differences. We computed tracking error on a quarterly basis in order to have the same time horizon on both of our activity measure to be able to compute the different categories of active management following the methodology by Petajisto (2013), which is explained in section 3.3.3.

The indices used when calculating tracking error were SIXRX for all large cap funds throughout the whole sample period and for all small/mid cap funds approximately halfway throughout the sample period. The other half of the sample period for the small/mid cap funds, Carnegie Small Cap Return GR (CSRXSE) was used for calculation of tracking error. The particular indices were chosen because the majority of our funds has these two indices as their primary prospectus benchmark, hence it will give the best reflection of the funds tracking error as it is the self-chosen benchmark they are trying to beat. CSRXSE

was used for the small/mid cap funds after its inception date because it gives smaller discrepancies between the funds returns and the index returns, hence a smaller and more accurate tracking error. CSRXSE has an inception date of March 2011 and was first used in our calculations at the third quarter of 2011, since this was the first whole consecutive quarter after its inception and hence the first quarter were the TE could be computed.

3.3.2 Active Share

The second measure we are using in this paper in order to distinguish the activity of a fund is active share, which is defined as the sum of the absolute differences between the weightings in the fund portfolio and index portfolio divided by two:

$$ActiveShare = \frac{1}{2}\sum_{i=1}^{N} |w_{fund_i} - w_{index_i}| \tag{2}$$

Where w_{fund_i} is the weight of asset i in the fund and w_{index_i} is the weight of asset i in the benchmark used for comparison. Active share is calculated separately for each fund and each quarter. The active share can take a value between zero and one for all funds that can take long positions only. Funds that can take short positions can have an active share above one, however as mentioned above, these types of funds are not included in our sample since Cremers & Petajisto (2009) sort out these kind of funds. A high active share indicates a high activity whereas a lower active share indicates a lower level of activity of the fund Cremers & Petajisto (2009).

The active share measure can increase in three different ways:

1. Differences in weightings between the fund and the index, where the active share of the fund increases with the absolute value divided by two.
2. If the fund has a holding which the index does not have. The active share of the fund increases by the weight that the stock has in the fund portfolio divided by two.
3. If the index has a holding that the fund does not have, this is an active choice by the manager to not invest in the stock and hence the active share of the fund increases with the weighting that the stock has in the index divided by two.

Table I: Active Share an Example

Security	Portfolio Weight % (x)	Index Weight % (y)	Active Weight %	Absolute Value %
1	25	15	10	10
2	75	0	75	75
3	0	85	-85	85
Sum	100	100	0	170
			Active Share	**85**

The Table is based on the three ways the active share measure can change, where security 1 is based on example 1 and so on.

When deriving the active share measure we used the data obtained from *Finansinspektionen*, then we divided each holding's market value by the total fund size, which includes cash, in order to get the weights of the fund portfolio. We chose to calculate the weights against fund size, since holding cash in a portfolio is an active choice instead of attributing the investors' assets to a certain stock. Hence, this method gives a more accurate measure of the level of active share than calculating with total assets net of cash. By using the active share measure, a fund can still be regarded as active when the weighting in the index changes but the weightings in the fund remains the same. Compared to *e.g.* a turnover measure, active share interprets any differences in holdings compared to index as active decisions, when the decisions themselves are based on passiveness. Hence, the word active cannot be literally interpreted at all times.

As for the index part of the active share calculations we chose the SIXRX index since this is the benchmark index the majority of our fund sample is trying to beat. However we could not obtain historical weightings of the index from any database. In order to solve this problem, we chose an index fund as a proxy of our benchmark. This index fund is called *Handelsbanken Sverigefond Index* and tracks the SIXRX index, the same as used in the tracking error calculation. Since an index fund´s only task is to track its benchmark, this methodology will yield approximately the same results as if the official benchmark would be used. The methodology of using an index fund as a direct proxy of an index, is also a method used by Morningstar, Lindmark (2013). In fact, they use the particular index fund we have chosen to use in their report. The same index fund was used for all funds in our sample, unlike the methodology used for the computation of the tracking error

where CSRXSE was used when available. The reason for not using CSRXSE is that we could not get hold of neither historical weightings of the index itself nor an index fund that tracks the CSRXSE, hence the small/mid cap funds might have a potential bias, resulting in an active share bigger for than what it actually may be. Further discussion of this will follow in section 5.

3.3.3 Definition of Fund Categorization

This paper will use the methodology presented by Petajisto (2013) where the two different activity measures are combined in order to sort funds into five different investment strategies. To illustrate the problem with only using one activity measure when determining how active a fund is, let's say we have two portfolios consisting of 50 shares each and both have a high level of active share. One of the portfolios is investing solely in tech-industry shares and thereby not only taking active bets in stock picking but also in an industry. The other portfolio picks the best stocks of 50 different industries and thereby diversifying systematic risk. The difference between the two portfolios is that portfolio 1 (concentrated stock picks) has a high tracking error, *i.e.* high risk, as well as a high active share, whereas portfolio 2 (diversified stock pickers) has a high active share and low tracking error *i.e.* low risk. The portfolios described in the example are the alternatives in the top row in Figure 1.

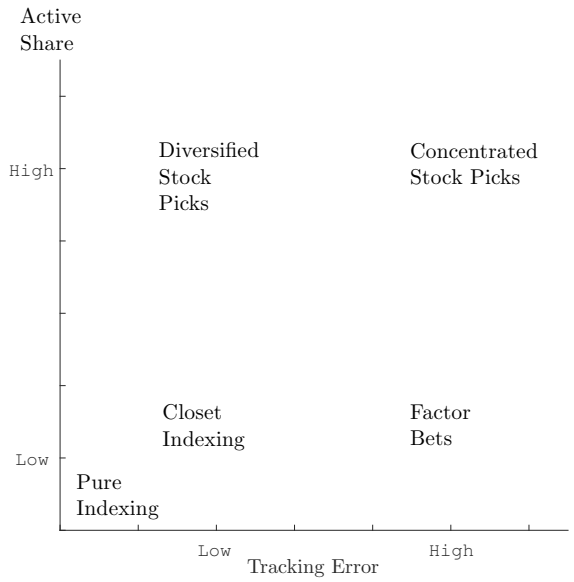

Figure 1: The figure shows different types of active management used in this paper, which are based on a funds level of tracking error and active share.

Because of the relationship between active share and tracking error, we use the methodology presented by Petajisto (2013) and divide our fund-sample in to five categories in order to incorporate the different strategies' effect on the degree of activity. We do this by sorting our sample by quintiles on both active share and tracking error.

Table II: Categorization Matrix

	Tracking-Error Quintile						
Active Share Quintiles	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*

Table 2 is a matrix which returns a group value depending on what quintile of the both measures the fund is placed in. This matrix is used when dividing our funds in to our categories.

Closet indexers place themselves in the lowest active share quintile and in quintile 1-4 in regards to tracking error and return a group value of 1.

Factor bets however place themselves in the third active share quintile but in the fifth tracking error quintile returning a value of 3. Factor bet funds have small active positions that are exposed to major systematic risk *i.e.* betting on a particular sector or industry and incurring a high firm specific risk.

Concentrated stock picks place themselves in the fifth quintile of active share and fifth quintile of tracking error, returning a group value of 4. Concentrated stock picks exposes their large active positions to large systematic risks. A strategy, which combines a high level of stock selection in one, or few, particular industries or sectors incurring a high volatility.

The active stock pickers place themselves in the highest active share quintile and percentiles 1-4 in regards of tracking error, returning a group value of 5. Active stock pickers engage in a strategy similar to concentrated funds. They implement a strategy with a high level of stock selection but unlike concentrated funds they diversify across industries.

The rest of the funds that does not have a combination of quintiles touched upon above, are called moderately active funds and returns a group value of 2. These funds do not have a well-defined strategy, *e.g.* neither a high/low active share nor a high/low tracking error. To illustrate, imagine a fund placing them in the active share quintile 3 and tracking error quintile 3. These funds are in-between an active stock picking strategy and betting on particular industries or sectors strategy. However one should be aware of that a moderately active fund is still active enough to exclude themselves from the closet indexing group. There are also funds that do not have a value for tracking error as mentioned previously; these funds are excluded from the categorization.

3.4 Other Variables

3.4.1 Flow

Flow is defined as percentage growth of assets under management during a time period, in our case quarterly. By multiplying total net assets during the previous period with the return during current period, we can net out the effect of the internal growth of already

managed assets. The remaing difference is defined as new money coming into the fund *i.e.* fund flow and can be calculated as follows:

$$Flow_{i,t} = \frac{(TNA_{i,t} - TNA_{i,t-1}\,(1 + R_{i,t})\,)}{TNA_{i,t-1}} \tag{3}$$

Where $TNA_{i,t}$ is the total net assets of a fund i at time t, $R_{i,t}$ is fund i return between $t-1$ and t and $Flow_{i,t}$ the flow of fund i at time t.

3.4.2 Fund Industry Size

Fund industry size is the total assets for all funds included in our sample summed up at each time period, including the index funds. We are well aware that the total Swedish fund industry is bigger than this, but we use this variable to show how the industry for our sample has grown during the sample period and too investigate an effect on active share. The variable is calculated on a quarterly basis.

3.4.3 Explicit Indexing

The variable explicit indexing is the index funds fraction of the total fund industry size calculated for every quarter. Explicit indexing is calculated as follows:

$$ExplicitIndexing = \sum \frac{TNA_{indexfund_i}}{TNA_{fundindustry}} \tag{4}$$

Where $TNA_{indexfund}$ is the total amount of assets for index fund i and $TNA_{fundindustry}$ is the sum of total assets of the fund included in our sample.

3.4.4 Age

Fund age is measured as number of years since inception date of the fund until the start of year 2016. The age of the fund is static, *i.e.* it does not change for fund i from one year to another. Since we already capture all time-varying effects in the panel regression, it does not matter if the age variable is varying with time or not. The cross-sectional effects between funds are still captured this way, thus we can determine whether the age of a fund can explain the level of active share. To be able to determine the effect, we need to develop average measures of the other variables included in the regression, *i.e.* we calculate the

average active share, average fund size, average fund flow and average tracking error for every fund and run a cross-sectional regression to see if age is an explanatory variable for active share. The same methodology and reasoning is also used for the variable expense ratio.

3.4.5 Carhart Model

We use ordinary least square regressions (OLS), in order to estimate the alpha of the funds. The alphas calculated are based on the Carhart four-factor model below, where R_i is the return of fund i net of fees, R_f the risk free rate, $R_{MKT} - R_f$ the market return minus the risk free rate *i.e.* market premium, SMB the size factor, HLM the value factor and MOM the momentum factor.Alpha α is the alpha of the fund, and the betas are the fund specific factor loadings. The four-factor model is given by the regression:

$$R_i = \alpha + R_f + \beta_{MKT}\left(R_{MKT} - R_f\right) + \beta_{SMB}SMB + \beta_{HML}HML + \beta_{MOM}MOM \tag{5}$$

The MKT, SMB and HML factors are exactly the same as in the Fama French three factor model. The MOM factor is the fourth factor and it is the monthly average return on the two high prior-return portfolios minus the monthly average return on the two low-prior return portfolios:

$$MOM = \frac{SmallHigh + BigHigh - SmallLow - BigLow}{2} \tag{6}$$

The alphas are calculated on the quarterly returns of each fund. However, for the funds that have less than five consecutive quarters, an alpha cannot be computed.

3.5 Analysis model

In order to be able to find relations between the variables used and provide answers to the research questions discussed in the introduction, we will preform multivariate panel data regressions controlling for fixed effects. In addition to this we will compute descriptive and sample statistics in combination with illustrative figures to explain important relations of main variables. The analysis can be found in the result section since we are analyzing our findings were they are presented in order to more easily interpret the effects. Further,

potential econometric problems, such as multicollinearity, will be analyzed and discussed in more detail in section five.

4 Results

We begin the result section by presenting descriptive and sample statistics of various variables and fund categories. We will then present our regression results and analyze their effects.

4.1 Descriptive Statistics

4.1.1 Active Share

As shown in Figure 2, the level of activity among Swedish mutual funds investing in the Swedish large-cap market have had an upward-sloping trend since the start date of our sample 2005/09/30. The level of activity dropped during the peak of the financial crisis in the end of 2008, and has since then almost exclusively been rising. One could question the pattern of activity around 2008. If the figure would show performance instead of activity, this finding would not be surprising since the market dropped substantially, but that the activity drops substantially cannot be as easily interpreted. We cannot find any explanation to this in previous studies but we believe one reason for such a behavior could be that in a downward trend, the fund managers are more careful in how to allocate the portfolio, thereby mimicking the index portfolio to a larger extent. They would do this in a financial chock to minimize their overall portfolio risk, and more importantly not to generate losses larger than the benchmark portfolio, which would be the case if the market sensitivity of the fund portfolio was higher than the benchmarks. Another explanation for the drop in activity could be that during the crisis 2008, there were a lot of cutbacks in the management team at various banks and institutions. Such a cutback could result in a lack of stock coverage of a fund and hence a risk of potentially missing out on exploiting opportunities to increase the fund´s active share.

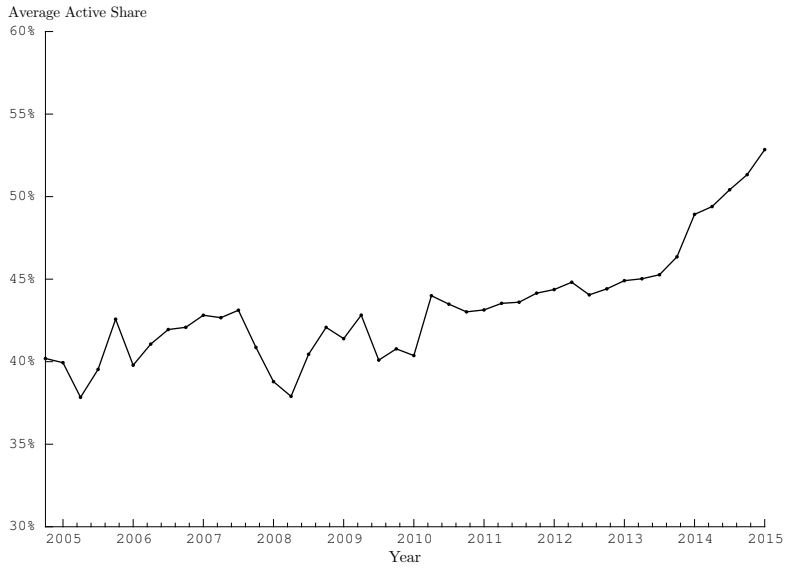

Figure 2: The figure shows the average active share for large-cap funds during the sample period.The amount of observations at every quarter can vary due to missing values and that new fund enter the sample.

However, small/mid-cap funds have had an active share that has been higher on average than the large-cap funds. This finding however, may be a result of a potential bias due to the lack of fit of the index used in the computation of active share. For a more detailed explanation, please refer to section 5.

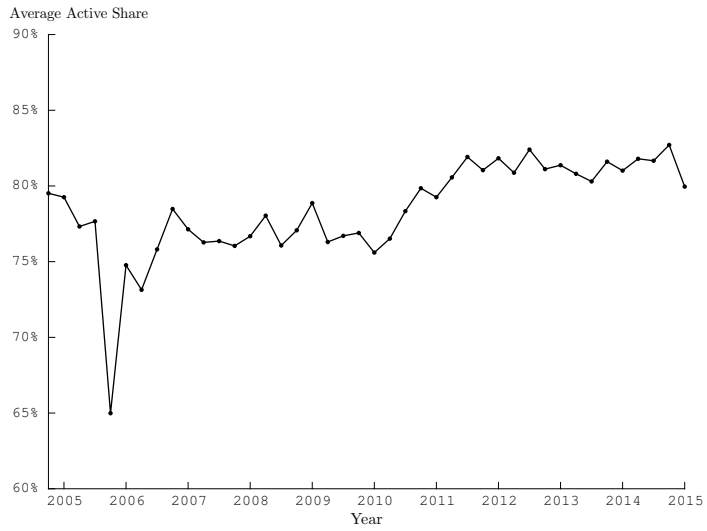

Figure 3: The figure shows the average active share for small and mid cap funds during the sample period. The amount of observations at every quarter can vary due to missing values and that new fund enter the sample.

Further, as can be seen in Figure 3, active share for the small/mid-cap funds is behaving rather peculiar during 2006 with a rapid decrease which is not due to the financial crisis. Even more peculiar is that active share during the next quarter increased to almost the same level. An explanatory reason for this behavior, could be that the index and the funds' holdings for some reason matched particularly well for that quarter. It could be that a certain stock that small/mid-cap funds tend to invest in, got large enough in terms of market size and therefore was included in the SIXRX index. This would mean that, when included in the index, the active share would decrease for the funds that held the stock beforehand.

To illustrate this phenomenon we refer back to table 1. Imagine that up to the time before the stock was included in the index we had an active share calculation for that particular stock as the second security example in table 1. However, when the stock was included in the benchmark we have an active share calculation as in the security 1 example, and hence a lower ex post active share. The reason the average active share then returns to almost the same level again the next quarter, is that when the stock is included in the

index, the small/mid-cap funds will decrease their holdings compared to earlier since the stock now no longer is considered as a small/mid-cap size stock, which increases the active share again.

4.1.2 Explicit Indexing

The index funds TNA as a fraction of total TNA of the fund industry market, *i.e.* explicit indexing is described in figure 4. Here we can clearly see an increase of the level of explicit indexing throughout the sample period.

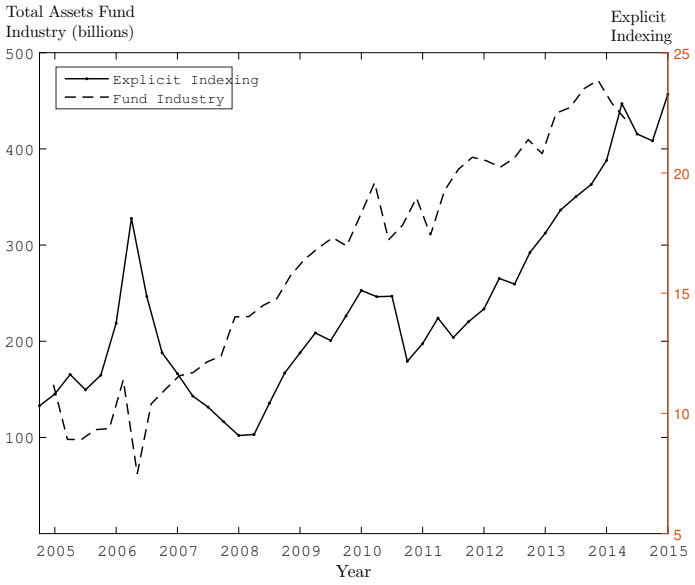

Figure 4: The figure displays the development of explicit indexing and fund industry size during our sample period.

At the beginning of the sample period, the fraction of explicit index funds in relationship to total net assets of the total fund industry in our sample was around 10%, which could be compared to a level of around 20% at the end of the sample period. At the same time, TNA has increased from low 150 billion SEK to around 450 billion SEK at the end. According to Flam & Vestman (2014), index funds are a small but growing fund segment on the Swedish market. Consistent with the results from Cremers (2015) where explicit index funds have steadily increased as a fraction of the total fund industry, our sample also shows that this segment continues to grow. One explanation for this might be the increasing transparency and access of information of fund characteristics. Another explanation might be the media

attention index fund has indirectly received by the damaged reputation surrounding the actively managed funds.

4.1.3 Summary Statistics

As shown in Table 3 we can see that on average, concentrated funds have the highest average active share and highest tracking error and that closet indexers are on the other end of the spectrum, with the lowest active share and lowest tracking error of the categories. This result is rather expected, since the categories are created by sorting active share and tracking error into quintiles. However, our finding that concentrated funds have the highest level of active share is not obvious, since both categories stock pickers and concentrated funds are based on the upper quintile of active share, the results are in line with the findings in Petajisto (2013).

Table III: Mean Values

Group	Label	No. of obs.	TNA (millions)	Active Share (%)	Tracking Error	Expense Ratio (%)	Flow (%)	Fund Age
5	Active Stock Pickers	249	3404	86.70	0.80	1.50	3.82	14
4	Concentrated Funds	143	2776	87.11	2.40	1.42	23.57	16
3	Factor Bets	274	2407	62.74	2.26	1.43	7.52	16
2	Moderately Active	938	3545	50.64	0.71	1.40	1.93	19
1	Closet Indexers	380	6621	22.27	2.51	1.13	0.07	25
Total		**1957**	**3924**	**53.91**	**1.00**	**1.36**	**4.12**	**19**

Table IV: Standard Deviations

Group	Label	No. of obs.	TNA (millions)	Active Share (%)	Tracking Error	Expense Ratio (%)	Flow (%)	Fund Age
5	Active Stock Pickers		3455	3.29	0.37	0.34	33	7
4	Concentrated Funds		2486	3.24	0.89	0.32	179	6
3	Factor Bets		3603	15.33	0.86	0.24	31	7
2	Moderately Active		5258	14.94	0.36	0.31	36	8
1	Closet Indexers		6137	4.52	0.28	0.62	7	9
Total			**5119**	**23.86**	**0.83**	**0.40**	**57**	**9**

The tables display the sample statistics for the categories shown in Table 2 that further will be used in following regressions.

Another finding in the sample statistics is that closet indexers hold the largest amount of TNA among the categories which means that on average, the least active funds tend to be the largest funds. Despite this, closet indexers are the group that has the lowest flow, *i.e.* there is a low fraction of new money flowing into closet indexing funds. Fund flow is high in the group Concentrated Funds, we believe that a flow bias ,where lack of adjusting for merger effects, contributes to this result. Further explanation can be found in section 5.

In terms of average expense ratio, the active stock pickers are the most expensive ones, closely followed by the rest of the truly active funds. However, somewhat surprisingly, is that closet indexers has such a large difference in fee compared to the active funds at a level of 1,13%. Our expectation was that their fee should be somewhat similar as the fees of the other actively managed funds. However, it is substantially lower and it seems like the managers of the funds are aware of that they have a low activity, hence lowering the fee as compensation to the investors. The findings are somewhat in line with Cremers (2015) and their discussion around increased competition between explicit index funds and active funds. They discuss around an actively managed funds ability to diversify its product by either lowering its fee or increasing its activity as a diversification strategy. By looking at the descriptive statistics of our sample, it seems like the closet indexers are using the former diversification strategy rather than the latter, *i.e.* lowering their fees, which compensate their investors for the funds low net returns. Instead, it seems like the moderately active funds and factor bets have not yet lowered their fees to such extent that would be appropriate with regards to the large difference in active share compared to the two most active fund categories.

This finding seems positive for the investors, but the average index fund in our sample has an average fee of 0,43%, almost a third of the fees of the closet indexers. Thus, the discrepancy between these two similar fund types gives explicit index funds a major advantage compared to closet indexers when adjusting returns for fees.

When it comes to standard deviations of variables included in our regressions, we can tell that the level of activity has the largest spread among factor bets and moderately active funds. Another interesting finding is that closet indexers deviate the most when looking at expense ratio, this might infer that the group closet indexers either have low fees, somewhat similar to index funds, or tend to charge fees that are similar to actively managed funds. This relative large deviation shows that some of the closet indexers have gone further in adapting their fees as a competitive decision while others still charge a fee not representative to the level of activity.

4.1.4 Fund Performance for Categories & Index Funds

In our sample, the most active funds, *i.e.* stock pickers and concentrated funds, are the ones that have the highest average benchmark-adjusted returns, both gross and net of fees. Closet indexers are the funds that have the lowest gross and net benchmark-adjusted returns. This shows that there is a relation between the level of activity and performance.

When adjusting the returns for risk in terms of Carhart four-factor alphas, active stock pickers generates the highest gross and net alphas and closet indexers the lowest. The results that active stock pickers have the highest gross alpha are in accordance with Petajisto (2013) and are somewhat expected. Since active stock pickers implement a stock selection strategy and in the same time diversifies across industries, they lower the exposure to systematic risk. This lower exposure generates a higher risk-adjusted alpha than concentrated funds who are exposed to a high systematic risk due to the lack of diversification across industries. These findings of risk-adjusted alphas are in accordance with the findings in Petajisto (2013).

Our alphas are only given as one per fund and we are aware of that the results may be limited due to the lack of observations. Since a few funds have not been running for more than five consecutive quarters, the number of alphas are not that many. Due to this, the alphas should be interpreted with some caution.

Table V: Fund Performance, September 2005 - December 2015, (%)

		Gross Return		Net Return	
Group	**Label**	**Benchmark Adjusted**	**Four-Factor Alpha**	**Benchmark Adjusted**	**Four-Factor-Alpha**
5	*Active Stock Pickers*	3.57	1.31	3.32	0.93
4	*Concentrated Funds*	5.96	1.07	5.61	0.71
3	*Factor Bets*	0.68	0.11	0.73	0.76
2	*Moderately Active*	0.20	0.93	0.10	0.58
1	*Closet Indexers*	-0.26	0.85	-0.55	0.56
	Index Funds	-0.07	1.48	-0.18	1.38

In Table 5 we display the average performance index funds and fund categories in our sample. Performance is measured in terms of Carhart four-factor alphas and benchmark-adjusted returns, both gross and net of fees, of which the benchmark-adjusted return is calculated as the return of the fund minus the return of SIXRX if it is a large fund, SIXRX or CSRXSE if the fund is a small/mid-cap fund, depending on year of observation.

4.2 Regression Results

4.2.1 Results on the level of Activity

In table 6 we present the results of the regression on active share. Similar to our expectations, we find that the fraction of explicit indexing has a significant and positive effect on active share. We refer to the same explanation for this as in Cremers (2015), that explicit indexing as percentage of TNA can serve as a proxy for competition between passive and actively managed funds. Increased presence of explicit index funds increases the activity of other funds on the market. This effect can also be seen from the figures presented in the descriptive statistics section where we can see that the average active share for both our large and small/mid-cap funds in figure 2 & 3 have increased in the same time as explicit indexing in Figure 4 has increased throughout our sample period .

Table VI: Regression on Active Share

	Active Share	**Average Active Share**
TNAlog	0.0061	0.0141
	(1.67)	(1.01)
Tracking Error	0.0145	0.3024
	(6.44)**	(4.18)**
Explicit Indexing	0.4791	-0.9083
	(10.59)**	(-0.92)
Flow	-0.00004	0.3219
	(-0.01)	(2.07)**
Expense Ratio		0.1544
		(2.10)**
Fund Age		-0.0085
		-(2.57)**
Constant	0.3145	0.0348
	(4.24)**	(0.09)
N	1945	57
R^2	0.9201	0.6505

The table shows a regression where active share is used as the dependent variable. TNAlog is the log of the total net assets of the funds, tracking error is the tracking error of the funds, explicit indexing is the fraction of index fund of total TNA on the market, flow is the fund's flow lagged with one quarter, expense ratio is the funds annual fee and fund age is the age of the fund. The significance level used is 5% and each significant variable is marked by **. We only use one significance level of 5% since it is a predetermined level and if changed afterwards we would engage in data mining. The regression are run with robust standard errors.

Tracking error has a positive effect on active share as expected, however the effect is rather small. We also expect a significant coefficient on tracking error since the two activity measures correlates to some extent, as can be shown in the Table 9 in the appendix. The correlation between the two measures (0.425) is of neither problem since it is low, nor of any surprise since the two measurements both capture the level of activity to some extent.

In Table 6 we also investigate the effect of expense ratio on active share. Our findings shows that expense ratio has a positive and significant effect on active share, this relationship should be considered with some limitation due to a possible reverse causality problem. A high level of activity demands a higher fee in order to pay for increasing transaction costs as well as more time spent on analyzing upcoming investments.

Consistent with the findings by Cremers & Petajisto (2009), older funds tend to be less active. Fund age have a positive and significant relationship with the level of activity in our sample. This shows that older funds tend to be less active in terms of active share.

Despite the fact that TNA(Log) and fund flow are not significant variables in predicting active share, our R^2 is 0.9201. This is very surprising and much higher compared to previous studies, *e.g.* Cremers & Petajisto (2009). The big difference from their study is that they computed the active share and all the explanatory variables on a yearly basis which may cause a lower level of precision and in turn, their regression did not capture variation within the year. Based on our regression results, we can almost exclusively define what variables that affect the level of activity of a fund. However, the relevance of active share and what variables that might be explanatory occurs when the relationship to fund performance is investigated.

4.2.2 Results on Performance

Active share as an explanatory variable cannot explain benchmark adjusted performance since there is no significance in any of the regressions. This is different compared to previous studies and not according to our primary beliefs. In contrast to Cremers (2015), where 88 indices where used to compute active share, the lack of fit between funds in our dataset compared to the index used may be a contributing factor that leads to the insignificance.

The insignificance may also, most probably, be attributable to the small sample size used in this paper. Earlier studies had a much larger sample and as an example, the paper by Grinblatt & Titman (1993) were criticized for having a sample too small when including at most 274 mutual funds in a time period over 10 years. This paper has 58 funds on a ten year basis and hence, a limited amount of observations might affect the models ability to predict active share.

In contrast to our active share calculations, we used different indices when computing tracking error; overall this made the measure more representative. In our regression, the tracking error has a positive and significant effect on performance. This indicates that a larger active risk generates higher performance. However, one should remember that the regression does not predict tracking errors effect on a risk-adjusted performance measure. If that would have been the case, the result would probably be of another distinction. An example of such a distinction could be seen by observing the descriptive statistics in Table. 5, where the category taking the highest active risk *i.e.* concentrated stock pickers, incurs a lower alpha than both active Stock pickers and factor bets. Because of this, interpreting that a higher tracking error alone results in higher performance, would be and inadequate result since there are more dimensions that matter, as also pointed out by (Petajisto, 2013).

Our regression findings on performance suggest that the most active stock pickers tend to perform best compared to the other categories: closet indexers, stock pickers and factor bets, when measured by their benchmark-adjusted net-returns. The results are significant and in line with the findings in Petajisto (2013). It is worth mentioning that none of the categories have the ability to outperform the benchmark, hence with regards to our sample, an investor cannot on average beat the market by investing in a certain fund category based on active share and tracking error. Somewhat surprisingly is that closet indexers seem to perform better than factor bets and moderately active funds despite that they are much less active. We believe this relates to the fact that closet indexers seem to have lowered their fees to a higher extent than the rest of the active categories and hence this effect occurs since the dependent performance measures used in the regression is net of fees.

Table VII: Regression on fund performance

Benchmark-adjusted return			
	(1)	(2)	(3)
Active share	-0.0173	-0.0138	
	(-1.03)	(-0.14)	
Active*large		-0.0040	
		(-0.04)	
Active*small/mid		-0.0019	
		(-0.02)	
Tracking error	0.0058	0.0058	
	(2.01)**	(2.01)**	
TNAlog	0.0142	0.0142	0.0142
	(4.60)**	(4.58)**	(4.50)**
Flow	-0.0008	-0.0008	-0.0008
	(-0.27)	(-0.28)	(-0.41)
Stock pickers			-0.0302
			(-2.48)**
Factor bets			-0.0356
			(-2.57)**
Moderately active			-0.0355
			(-2.81)**
Closet indexers			-0.0335
			(-2.65)**
Constant	-0.2886	-0.2889	-0.2523
	(-4.50)**	(-4.46)**	(-3.86)**
N	1945	1945	1945
R^2	0.1222	0.1222	0.1304

The table displays a multivariate panel regression with benchmark-adjusted return as the dependent variable. The first column includes a regression with explanatory variables of active share, tracking error, flow and total net assets logged. The second column includes interaction terms with active share and dummies of large or- mid/small-cap, by doing this we can distinguish whether fund performance may be affected differently by the level of activity in large-cap compared to small/mid-cap funds. The third column represents a regression where we include the fund categories. We exclude one of the groups due to avoid a dummy variable trap. Because the categories are generated from active share and tracking error we excludes these variables since they are jointly incorporated in the categories themselves. The significance level used is 5% and each significant variable is marked by **. We only use one significance level of 5% since it is a predetermined level and if changed afterwards we would engage in data mining. The regression are run with robust standard errors.

When it comes to fund size, our findings show that there seems to be a positive scale effect among the Swedish mutual funds. Similar to explanations by Chen (2004), this

positive size effect might occur due to that cost of operations can be distributed among a larger asset base. Our findings do not entail a liquidity constraint effect as mentioned by Dahlquist (2000), *i.e.* Swedish funds are not too big for the Swedish market in order find liquid investment opportunities that can generate positive returns.

Different from the results in Ferreira (2011), flow has no significant effect on fund performance. Their results showed positive relationship between fund flow and performance, although this effect could almost exclusively be explained by momentum effect, which previously also been proved in Sapp & Tiwari (2004). The insignificance of the variable could also be due to a potential bias further discussed in in chapter 5.

The R^2 in our tests has a maximum of 0.1304 which indicates that our model cannot fully predict fund performance. This shows that there are a lot of research to be made in order to achieve a better understanding of how, and if, fund performance can be predicted. Even if our model would have been better at predicting performance, it is not sure that the interpretation would have given us a better understanding of what funds to pick to achieve the highest level of return.

5 Potential Biases and Robustness Checks

The regression results described in Section 4 can to some extent be affected by potential biases in our variables, which will be described in more detail in this section.

We perform various robustness checks for our model, however we are not shortening our sample period, since we believe this would be harmful for the model overall because the number of observations would decrease. A more appropriate robustness check, in terms of sample period, would be to increase the number of years observed. Since there is no possibility of getting hold of consistent holdings data before our start date, this check cannot be made.

As discussed briefly in the methodology section, it is impossible to find an index that fits ultimately for all of our funds when calculating active share. We address this first problem by choosing a broad and frequently used index for our calculations. Despite this choice, the small/mid-cap funds may have a larger active share than what is actually true.

However, the use of SIXRX as the benchmark index for our small/mid-cap funds is the best approximation available since we cannot get hold of historical weightings of CSRXSE. In order to address this potential bias, and exclude the risk of an inadequate effect of the small-cap funds, we perform a robustness check were the small/mid-cap funds were excluded from the sample. By doing this, we could investigate whether the results on active share and fund performance are approximately the same as in Table 6 and Table 7. The results are presented in Table. 8 which can be found in the appendix.

In the first two columns of the same table, the benchmark-adjusted return net of fees is used as the dependent variable and active share as one of the independent variables. In this robustness check regression, none of the variables are significant. We believe this depends on the small size of the sample, hence a hard time fitting and interpret any relation between the variables investigated. Because of this problem, the regression table is omitted and our original regression should be interpreted as the most reliable effects.

In column three of Table 8, where active share is used as a dependent variable, the explanatory power of the independent variables are lower than in the original regression. This shows that this adjusted model, without the small/mid-cap funds, is not as good in predicting active share as when using the original data-set. When looking at the result of each explanatory variable in the regression with our adjusted data-set, it gives us the same interpretation as before, except for fund flow that now becomes a significant variable of determining active share. One explanatory reason behind flow, and its negative effect now being significant, might be that large funds are slower in turning new assets to investments, and instead holds cash during a longer period compared to small/mid-cap funds. There seems to be a problem for the Swedish large-cap funds in placing their money in the market fast, presumably because of their inflexibility. However, in Table. 7 when running the regression of fund performance and not removing the small/mid-cap funds, we see that fund size has a positive and significant effect on fund performance. While flow might have a negative effect on the level of activity, the advantages of being a large fund is dominant, especially since active share alone cannot predict fund performance.

Compared to the study by Frazzini and Owen (2008), we cannot access data that describes when a fund is merged or acquired by another fund family and because of this, we

cannot adjust our flow variable for this merger effect. In the case of a merger of two funds, the total net assets will, from one quarter to another, increase by an unusual amount and thereby affect the fund flow. Taking this limitation of the data into consideration, we perform a robustness check where all variables of flow above 100% are removed, *i.e.* removing outliers that might make the interpretation of the flow variable unfair. As shown in Table 11, we can see that the variable remain insignificant in the regression trying to predict active share. Because the variable remained insignificant, we do not risk excluding true extraordinary high flow variables from the regression that might occur when a new fund is launched and a large amount of capital flows into the fund.

Since there are no historical data on fees for mutual funds available, we made the assumption that the yearly fees reported in the latest of the funds' prospectus, are also maintained on the same level historically. Because of this assumption, our performance measures used throughout the paper might be biased on a historical basis and hence, the interpretation in the regressions where performance net of fees is used as a dependent variable, might be wrongful. However, we are aware that fees most definitely have changed over time but believe that the current fee is somewhat persistent, hence a better estimation than excluding the variable in total. In addition, we believe that the differences in fees among funds, have also remained persistent to some degree. In addition, since a fund's net returns is of absolute importance, not adjusting for fees would make a relevant comparison between fund types impossible to make.

Swedish mutual funds are regulated in terms of maximum allowance in allocation of the fund portfolio, which can be a potential explanation for a low level of active share for the funds in our sample. For instance, funds are only allowed to invest 10% of the total assets of a fund in one single security in accordance to chapter four §6 in law (2004:46) by Munck (2016). To illustrate, imagine a stock picker seeing potential in a stock that lies outside the funds comparable benchmark and would lead to an increase in active share. Due the above mentioned regulations, the stock picker is not allowed to increase his position in the stock to more than 10%, which leads to an active share lower than intended by the fund manager. However, it is no conclusive argument for a low active share since there are alternatives for an increase in active share. In addition, achieving a high level of active share seems possible since we observe funds with a high level of active share in our sample.

As mentioned in Ferreira (2011), explicit indexing is expected to correlate with fund industry size. Because of this we run additional regressions on active share where the variable fund industry size is included, the results can be found in Table 12 in the appendix. In order to determine the combined explanatory power of the other independent variables on explicit indexing, we calculate the variance inflation factor, more commonly known as the VIF. VIF is a measure of investigating to what extent the other independent variables explains the increase in variation of our explanatory variable of interest, Wooldridge (2008). In our case we are interested in $Var\beta_{explicitindexing}$ and $Var\beta_{fundindustrysize}$. There is not a particular limit of what level of VIF can be determined too high, VIF is calculated as follows:

$$VIF_j = \frac{1}{1 - R_j^2} \tag{7}$$

Where VIF_j is the VIF of Variable j and R_j^2 is the explanatory power of the other independent variables on variable j. Our calculation indicates a VIF of 2.44 on explicit indexing and 3.25 on fund industry size and as we can see in Table 10 in the appendix, these two variables have a correlation of more than 0.7.

When removing this factor, there might be a positive bias in the explicit indexing factor, i.e. the effect of explicit indexing on active share might be too big, according to Wooldridge (2008) one might be hesitant to drop explanatory variables in order to infer causality by the variable of interest that correlates with the dropped variable, in this case explicit indexing that correlates with fund industry size. Since explicit indexing is the variable of most interest of the two and the VIF of fund industry size is the larger, we decide to remove fund industry size from the regression on active share.

6 Conclusion

From our findings, we can conclude that the large fraction of closet indexing funds is of less worry than previously expected. When using mutual funds as an investment vehicle on the Swedish market one might think that, due to the large proportion of closet indexing as mentioned by Cremers (2015), there is high probability of paying fees not being in parity to the level of activity. From our results, we can conclude that the presence of index

funds indeed tend to affect the way that active funds diversify their investment products. On the Swedish market, the diversification in terms of lowering fees is mostly occurring, and therefore the investors in a closet indexer gets compensated for the low level of activity.

Instead of worrying about closet indexers, which previous research emphasizes, an investor on the Swedish market should worry about the groups moderately active and factor bets. These funds, despite having a higher level of activity in terms of active share and tracking error, will generate substantially lower returns both adjusting for benchmark and adjusting for risk. If investing in any category of the active funds, it should be in the category active stock pickers since they tend to perform least worst compared to index and has the highest alpha of the categories observed. Active stock pickers are followed by closet indexers that are predicted to perform slightly worse than active stock pickers.

Further, we find a strong relationship between the fraction of explicit indexing and active share. A dynamic that is consistent with Cremers (2015), and also a finding of great importance for the fund industry and its future development, due to the increased level of competition between active and passively managed funds. The predictive model of active share shows a high explanatory power when predicting the level of active share, compared to previous studies performed on the U.S market and also compared to the cross-country performed in Cremers (2015). This finding is of use if an investor wants to choose the fund that has the greatest theoretical opportunity of generating high returns by deviating from its benchmark. However our paper can neither, as many other papers has succeed to do, provide theoretical evidence nor, draw a relationship between the active share and increased performance and hence one can question the relevance of the measure. This inability of drawing a relation between the variables is though, probably due to a small sample and a bias in the active share variable.

Due to this fact, that active share on its own cannot give any prediction of future fund performance, an investor should jointly combine the activity measures presented in this paper, in order to incorporate a fund's strategy and not just its level of activity. This paper shows that a fund's strategy is a key determinant for predicting fund returns, since all of our categories investigated are significant. However, none of the categories perform well enough in terms of benchmark-adjusted returns, we can therefore conclude, as previously

shown on the Swedish market by Flam & Vestman (2014) as well as by Dahlquist (2000), that an investor is better off when investing in an explicit index fund. We can also conclude that no fund manager have the ability of generating abnormal returns over time.

If historical fees of funds were available, it would be interesting to test whether explicit indexing has any relationship with fees charged by active funds, this because one might expect that increased competition would lead to lower prices on the investment products that are affected by the new increased level of competition. Since we cannot get hold of this data we can only hope to see this in future research. If possible in the future, when a bigger data set and portfolio holdings of other indices can be obtained, it would be appropriate to challenge the actual relevance of the variable active share since extensive criticism of the variable is rather scarce. Combined with the fact that the majority of the research to date have been conducted by a few researchers in various constellations, this might also be a signal that entail evidence of irrelevance in terms of predicting fund performance. Whether active share alone is a relevant measure or not is still unanswered, as shown in this paper, and overall active share calculations are very sensitive to the choice of benchmark. Due to this sensitivity, the use of the measure needs to be further analyzed.

7 References

Asness and Frazzini, C. A. (2013, September). *The Devil in HML´s details.* Journal of Portfolio Management.

Asness Frazzini and Pedersen, C. S. (2013). *Low-Risk Investing without Industry Bets.* Financial Analysts Journal(4).

Carhart, M. M. (1997, March q). *On Persistence in Mutual Fund Performance.* The Journal of Finance, 52(1), 57-82.

Chen J, Hong H, Huang M, Kubik J.D, (2004, December). *Does Fund Size Erode Mutual Fund Performance?* The Role of Liquidity and Organization . The American Economy Review, 94(5), 1276-1302.

Cremers and Petajisto. (2009). *How active is your fund manager? A new measure that predicts performance.* Review of Financial Studies 22, 22(3329-3365).

Cremers M, Ferreira M.A, Matos P, Starks L, (2015, February). *Indexing and Active Fund Management: International Evidence.* Journal of Financial Economics.

Dahlquist M, Engström S, Söderlind P, (2000, September). *Performance and Characteristics of Swedish Mutual Funds.* Journal of Financial and Quantitative Analysis, 35(3), pp. 409-423.

Daniel K, Grinblatt M, Titman S, Wermers R, (1997, July). *Measuring Mutual Fund Performance with Characteristic-Based Benchmarks.* The Journal of Finance, 52(3), pp. 1035-1058.

Efama. (2016). *Trends in the European Investment Fund Industry in the Fourth Quarter of 2015 & Results for the Full Year of 2015.* Quarterly statistical release. Efama, European Fund and Asset Management Association.

Ferreira M.A, Keswani A, Miguel A.F, Ramos S.B, (2011, July 27). *The Determinants of Mutual Fund Performance: A Cross-Country Study*.* Swiss Finance Institute Research Paper, 31.

Flam and Vestman, H. F. (2014, February). *Swedish Equity Mutual Funds: Performance, Persistence and Presence of Skill.* Swedish House of Finance Research Paper, 14(4).

Frazzini and Owen, A. F. (2008, February 23). *Dumb money: Mutual fund flows and the cross section of stock returns.* Journal of financial economics(88).

Frazzini and Pedersen. (2016, February 29).*Betting Against Beta: Equity Factors, Monthly. Retrieved from AQR:* https://www.aqr.com/library/data-sets/betting-against-beta-equity-factors-monthly

Grinblatt and Titman, M. G. (1993). *Performance Measurement without Benchmarks: An Examination of Mutual Fund Returns.* Journal of Business, 47-68.

Grinold, R. C., & Kahn, R. N. (1999). *Active Portfolio Management: A quantitaive approach for providing superior returns and controlling risk* (2 ed., Vol. 2). New York, USA: McGraw-Hill.

Gruber, M. J. (1996, July). *Another Puzzle: The Growth in Actively Managed Mutual Funds.* Journal of Finance, 51(3), 783-810.

Jensen, M. C. (1968, May). *The Performance of Mutual Funds in the Period 1945-1964.* Journal of Finance, 23(2), 389-416.

Lindmark, J. (2013, May 23). *Tabell Active Share Sverigefonder.* Retrieved from Morningstar: http://www.morningstar.se/Articles/Analysis.aspx?title=tabell-active-share-sverigefonder

Markowitz, H. (1952, March). *Portfolio Selection.* The Journal of Finance, 7, 77-91.

Munck, J. (2016). *Sveriges Rikes Lag.* Wolters Kluwer.

Petajisto, A. (2013). *Active Share and Mutual Fund Performance.* Financial Analysts Journal, 69(4), 73-93.

Sapp and Tiwari, T. S. (2004). *Does Stock Return Momentum Explain the "Smart Money" Effect?* Journal of Finance, 2605-2622.

Wermers, R. (2000, August). *Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transaction costs, and Expenses.* Journal of Finance, 55(4), 1655-1703.

Wooldridge, J. M. (2008). *Introductory Econometrics:* A Modern Approach. South-Western College.

Zheng, L. (1999). *Is Money Smart? A Study of Mutual Fund Investors' Fund Selection Ability.* Journal of Finance, 901-933.

8 Appendix

Table VIII: Regression on Activity and Performance, Adjusted for Bias

	Benchmark-adjusted net return	**Active share**
Active share	0.0031	
	(0.37)	
Tracking error	0.0018	0.0217
	(1.06)	(5.11)**
TNAlog	0.0004	0.0096
	(0.24)	(1.65)
Flow	-0.0012	-0.0091
	(-0.63)	(-1.97)**
Explicit indexing		0.4356
		(7.16)**
Constant	-0.0153	0.1264
	(-0.48)	(1.06)
N	1331	1331
R^2	0.0372	0.8389

The table shows a multivariate panel regression where fund performance and active share are used as dependent variables. TNAlog is the log of the total net assets of the funds, tracking error is the tracking error of the funds, flow is the fund's flow lagged with one quarter and explicit indexing is total net assets as percentage of fund industry TNA. The significance level used is 5% and each significant variable is marked by **. We only use one significance level of 5% since it is a predetermined level and if changed afterwards we would engage in data mining. The regression are run with robust standard errors..

Table IX: Correlation Matrix

	Active Share	**Tracking Error**	**Explicit Indexing**	**TNAlog**	**Expense Ratio**	**Fund Age**	**Flow**
Active Share	1.0000						
Tracking Error	0.4250	1.0000					
Explicit Indexing	0.1603	-0.0915	1.0000				
TNAlog	-0.2016	-0.1432	0.0375	1.0000			
Expense Ratio	0.2418	0.0996	0.0214	-0.3201	1.0000		
Fund Age	-0.4898	-0.2134	-0.1542	0.4347	0.1100	1.0000	
Flow	0.0514	0.0305	0.0154	-0.0500	-0.0158	-0.0803	1.0000

Table X: Correlation Matrix 2

	Active share	Explicit indexing	Fund age	Tracking error	Benchmark-adjusted return	Flow	Expense ratio	TNAlog	Fund industry size(log)
Active share	1.0000								
Explicit indexing	0 1603	1.0000							
Fund age	-0.4898	-0.1543	1.0000						
Tracking error	0.4252	-0.0916	-0.2142	1.0000					
Benchmark-adjusted return	0.2487	0.0751	-0.1363	0.1361	1.0000				
Flow	0.0515	0.0154	-0.0804	0.0305	0.0130	1.0000			
Expense ratio	0.2420	0.0214	0.1092	0.0991	0.0312	-0.0159	1.0000		
TNAlog	-0.2015	0.0376	0.4337	-0.1445	0.0357	-0.0502	-0.3221	1.0000	
Fund industry size(log)	0.1412	0.7523	-0.1232	-0.2427	0.1454	-0.0121	0.0145	0.1128	1.0000

The matrix includes the variables benchmark adjusted return and fund industry size(log).

Table XI: Regression with Adjusted Flow

	Benchmark adjusted net return			**Active share**
	(1)	(2)	(3)	
Active share	-0.0164	-0.0118		
	(-0.97)	(-0.12)		
Active*large		-0.0045		
		(-0.05)		
Active*small/mid		-0.0054		
		(-0.04)		
Tracking error	0.0058	0.0058		0.0146
	(2.00)**	(1.99)**		(6.47)**
Stock pickers			-0.0299	
			(-2.47)**	
Factor bets			-0.0356	
			(-2.57)**	
Moderately active			-0.0353	
			(-2.80)**	
Closet indexers			-0.0337	
			(-2.66)**	
TNAlog	0.0141	0.0141	0.0137	0.0063
	(4.57)**	(4.55)**	(4.47)**	-1.73
Flow	0.0124	0.0124	0.0127	-0.0198
	(1.11)	(1.11)	-1.15	(-1.41)
Explicit indexing				0.4765
				(10.52)**
Constant	-0.2864	-0.2864	-0.2496	0.3099
	(-4.48)**	(-4.44)**	(-3.82)**	(4.18)**
N	1945	1945	1945	1945
R^2	0.1226	0.1226	0.1308	0.9202

The table shows four multivariate panel regression where fund performance and active share are used as dependent variables. Fund flow represent the adjusted dataset where the outliers are removed. Active share*large is the level of active share multiplied with a dummy if the fund uses a large-cap benchmark as comparison, and multiplied with the small/mid dummy if the fund compares to a small/mid benchmark. TNAlog is the log of the total net assets of the funds, tracking error is the tracking error of the funds, flow is the fund's flow lagged with one quarter and explicit indexing is total net assets as percentage of fund industry TNA. The significance level used is 5% and each significant variable is marked by **. We only use one significance level of 5% since it is a predetermined level and if changed afterwards we would engage in data mining. The regression are run with robust standard errors.

Table XII: Regression and VIF of Explicit Indexing and Fund Industry Size

	Explicit indexing	**Fund industry size (log)**
Explicit indexing		4.3874
		(21.75)**
Fund industry size (log)	0.0834	
	(26.57)**	
Benchmark adjusted return	-0.0403	0.4908
	(-4.53)**	(6.05)**
Active share	0.0462	0.5890
	(3.74)**	(7.26)**
TNAlog	0.0063	0.2003
	(3.64)**	(12.85)**
Tracking error	0.0061	-0.0964
	(6.66)**	(-13.09)**
Flow	0.0005	-0-0335
	(0.69)	(-2.64)**
Constant	-2.1742	20.9310
	(-38.72)**	(66.93)**
N	1945	1945
R^2	0.5982	0.6926
VIF	2.4888	3.2531

The table shows two regressions where explicit indexing and fund industry size are used as dependent variables. The VIF is calculated as $1/(1 - R^2)$. TNAlog is the log of the total net assets of the funds, tracking error is the tracking error of the funds and flow is the fund's flow lagged with one quarter. The significance level used is 5% and each significant variable is marked by **. We only use one significance level of 5% since it is a predetermined level and if changed afterwards we would engage in data mining. The regression are run with robust standard errors.

Table XIII: Funds Used in Sample

Active Funds		Index Funds
Swedbank Robur Ethica Sverige	SEB Sverigefond	Catella Sverige Index
Handelsbanken Bosparfond	Swedbank Robur Ethica Sverige MEGA	Handelsbanken Sverigefond Index
Aktie-Ansvar Sverige	Spiltan Aktiefond Sverige	Aktiespararna Topp Sverige
Carnegie Sverigefond	Spiltan Aktiefond Stabil	Handelsbanken Sverige Index Criteria
Human Sverigefond	Swedbank Robur Sverigefond	Avanza Zero
Didner & Gerge Aktiefond	Gustavia Sverige	Nordea Indexfond Sverige
Handelsbanken Svenska Smabolagsfond	AMF Aktiefond Smabolag	SEB Sverige Indexfond
Swedbank Robur Sverigefond MEGA	Nordea Olympiafond	Lansforsakringar Sverige Index
SEB Sverigefond Smabolag Chans/Risk	Skandia Cancerfonden	Handelsbanken Sverige OMXSB Index
Lansforsakringar Smabolag Sverige	Skandia Varldsnaturfonden	Human Index Sverige
SEB Stiftelsefond Sverige	Swedbank Robur Humanfond	
Catella Reavinstfond	Danske Invest Sverige Fokus	
Catella Smabolagsfond	Nordic Equities Sweden	
Danske Invest Sverige	Spiltan Aktiefond Dalarna	
AMF Aktiefond Sverige	Cicero Focus	
Skandia Smabolag Sverige	Spiltan Aktiefond Smaland	
Folksam LO Sverige	Lannebo Sverige Plus	
Folksam LO Vastfonden	Gustavia Smabolag	
Nordea Swedish Stars	Simplicity Sverige	
Handelsbanken Sverigefond	Solidar Fonder Sverige	
Swedbank Robur Exportfond	Inside Sweden	
Swedbank Robur Smabolagsfond Sverige	Nordea Smabolagsfond Sverige	
Enter Sverige	Granit Smabolag	
SEB Sverigefond Smabolag	Carnegie Smabolagsfond	
Alfred Berg Sverige Plus	PriorNilsson Sverige Aktiv	
Lannebo Sverige	Handelsbanken Sverige Selektiv	
Lannebo Smabolag	Cicero Focus SRI	
Lansforsakringar Sverige Aktiv	Enter Smabolagsfond	
Skandia Sverige	Swedbank Robur Access Sverige	

The table shows the funds included in our sample, both active managed and index funds. The funds are ordered after institutional number which refers to inception date, as used in the dataset from *Finansinspektionen*