UNIVERSITY OF TURKU

ABSTRACT

	Bachelor's thesis
X	Master's thesis
	Licentiate's thesis
	Doctoral dissertation

Subject	Accounting and Finance	Date	28.11.2024
Author(s)	Jere Jokiniemi	Number of pages	82+appendices
Title	Performance of passive versus active fund management – Empirical analysis on the characteristics of passive and active exchange-traded funds on US markets		
Supervisor(s)	D.Sc. Abu Chowdhury, Prof. Mika Vaihekoski		

Abstract

The rise in popularity of passive investing and exchange-traded funds has spawned a swath of studies with the goal of finding out whether the active fund management has a future in the age of index funds and ever lowering expense ratios. This warrants the question whether the active funds can perform better ceteris paribus, and whether they have something unique that can explain their better performance compared to passive funds.

In this paper, we empirically analyze how different fund characteristics such as expense ratio, size and age affect its performance, and that whether active fund management can be said to perform better. In line with prior studies, in this study the expense ratio is shown to be the most significant factor in the performance puzzle, with the relation being consistently negative across the board but more sensitive to changes among passive funds. Diseconomies of scale are observed among passive funds, with the results being insignificant for active funds. It is also found that older funds tend to perform better, possibly displaying a learn-on-the-job effect. Finally, the results suggest that the difference in abnormal returns is insignificant between an active and a passive fund when they have otherwise similar characteristics and accounted for interaction terms. This suggests that the fund style is unlikely to have predictive power on future performance that is not explainable via other characteristics.

Key words	Fund management, exchange-traded funds, ETFs, fund characteristics, comparative performance

TURUN YLIOPISTO

TIIVISTELMÄ

	Kandidaatintutkielma
X	Pro gradu -tutkielma
	Lisensiaatintutkielma
	Väitöskirja

Oppiaine	Laskentatoimi ja rahoitus	Päivämäärä	28.11.2024
Tekijä(t)	Jere Jokiniemi	Sivumäärä	82+liitteet
Otsikko	Passiivinen ja aktiivinen rahastonhoito vastakkain – Empiirinen analyysi passiivisten ja aktiivisten pörssinoteerattujen rahastojen ominaisuuksista Yhdysvaltain markkinoilla		
Ohjaaja(t)	KTT Abu Chowdhury, Prof. Mika Vaihekoski		

Tiivistelmä

Passiivisen sijoittamisen ja pörssinoteerattujen rahastojen suosion kasvu on johtanut lukuisiin tutkimuksiin tavoitteinaan selvittää onko aktiivisella rahastonhoidolla tulevaisuutta näin indeksirahastojen ja alati laskevien juoksevien kulujen aikakaudella. Tästä johtuen on aihetta kysyä voivatko aktiiviset rahastot suoriutua passiivisia paremmin ceteris paribus, ja että onko niillä jotakin ominaista joka selittäisi niiden paremman suorituskyvyn verrattaessa passiivisiin rahastoihin.

Tässä tutkimuksessa tutkitaan empiirisesti miten erilaiset rahaston ominaisuudet kuten juoksevat kulut, koko tai ikä vaikuttavat sen suorituskykyyn, ja sitä voidaanko aktiivisen rahastonhoidon sanoa suoriutuvan paremmin. Linjassa aikaisempien tutkimusten kanssa, tässä tutkimuksessa rahaston juoksevat kulut osoittautuivat yhdeksi merkittävimmistä tekijöistä selittämään sen suorituskykyä. Suhde juoksevien kulujen ja rahaston suorituskyvyn välillä oli kauttaaltaan negatiivinen, ja se negatiivinen vaikutus oli voimakkaampi passiivisten rahastojen kohdalla. Passiivisten rahastojen kohdalla havaittiin mittakaavahaittoja, tulosten ollessa epämerkitseviä aktiivisten rahastojen kohdalla. Lisäksi tulokset myös osoittavat vanhempien rahastojen pärjäävän yleensä paremmin, viitaten siihen että rahastojen ja niiden hoitajien taidot paranevat ajan kuluessa. Lopuksi, tulosten mukaan ero poikkeavissa tuotoissa muilta ominaisuuksiltaan samanlaisten aktiivisten ja passiivisten rahastojen välillä on merkityksetön kun muuttujien vuorovaikutukset otetaan huomioon. Tämän perusteella rahaston tyyli ei vaikuttaisi olevan selittävä tekijä sen tulevaisuuden tuottoihin, vaan ne olisi selitettävissä muiden ominaisuuksien perusteella.

Avainsanat	Rahastonhoito, pörssinoteeratut rahastot, ETF, rahaston ominaisuudet, suorituskykyvertailu

UNIVERSITY
OF TURKU
Turku School of
Economics

PERFORMANCE OF PASSIVE VERSUS ACTIVE FUND MANAGEMENT

Empirical analysis on the characteristics of passive and active exchange-traded funds on US markets

Master's Thesis
in Finance

Author(s):
Jere Jokiniemi

Supervisor(s):
D.Sc. Abu Chowdhury
Prof. Mika Vaihekoski

28.11.2024
Turku

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Background and motivation

The ease of accessibility, lower costs and simplicity of index tracking funds has increased their popularity significantly in the past decades. Investing in these index tracking funds is more commonly called "passive investing" or "index investing". From 1995 to 2020 the portion of passively managed investment vehicles, open-end mutual funds and exchange-traded funds (ETFs) increased from around 3% in 1995 to 41% in 2020. In terms of assets under management (AUM) this means an increase from around 1 trillion dollars to 20 trillion dollars over 25 years. (Anadu et al. 2020.) The trend has been clear and even with the index funds having gained massive popularity, ETFs have become the most accessible way to invest especially for retail investors. As the ETFs can be bought as is from the markets, they have the benefit of great liquidity, broad exposure and diversification through one instrument, low costs and transparency of holdings. From the perspective of the ETF, what makes their form superior to index funds is the lack of requirement in holding comfortable cash balances, being able to be fully invested and not having to worry about cash redemptions which in turn aids with tax efficiency as they do not need to incur capital gains by liquidating their positions to pay the investors. Additionally ETFs deals with fund inflows and outflows often in kind, meaning the investors can trade in assets that match the fund's portfolio and receive the cash equivalent amount of shares, and investors redeeming their investments can receive securities equivalent of their shares, and these transactions are not taxable exchanges. This also ensures that the ETF's price often stays very close to their holdings' total market price thanks to arbitrageurs.

Morningstar (2024) has been tracking the performance of active managers against passive counterparts in US markets over the years in their semiannual "Active/Passive Barometer" spanning over 8000 unique funds. In their latest mid-year review from July 2023 to June 2024, they found that only about 50% of the actively managed strategies in the analysis managed to outperform their average passive peers. When the time span is extended to a decade, this percentage falls to as low as 20% for large-cap funds with the chances being slightly higher for mid-cap and small-cap. From the evidence over the decades, it has become apparent that most active strategies will lose to the passive competition, which has driven majority of the fund flows towards passive investing over the years.

With the abudance of ETFs to choose from, it is important from the investor's perspective to know whether indicators exists that are observable from the outside and which are predictive of its future performance. Fund characteristics such as expense ratio, fund and family size, age, portfolio turnover and securities lending ratio have been widely used in literature to explain excess returns, and the relations between these characteristics tend to be different between the two management styles. Arguably the largest drag on the fund's returns is the expense ratio, which almost always is way higher for active funds, even by an order of magnitude. This has put into question the function of active research and trading if the benefits gained do not surpass the spent resources. Regardless of the mixed results on whether active funds overperform before fees, in practice it is often irrelevant from the perspective of the investor as they are only interested in what they end up pocketing after the management has taken their cut. A negative relation between expense ratio and performance among active mutual funds has been shown in e.g. Sheng et al. (2023), and among passive mutual funds and ETFs in Elton et al. (2019a) and Paudel and Naka (2023). The negative effect appears to be stronger among passive funds and they are more sensitive to changes in expense ratio. This could be attributed to the fact that these funds are not very differentiated from each other and have very limited ways of gaining an edge over the competition. Despite of the expense ratio being emphasized so much, it does not necessarily always give a complete picture of the fund's efficiency as it is missing factors such as transaction costs which might differ significantly between funds.

The size of the fund has been shown to impact its performance, where the larger it is, the more it is able to benefit from the economies of scale and leverage its size for example when dealing with brokers. On the other hand, as an active fund it is possible for the fund's size to exceed the optimal allocation, where it may become more difficult to find as profitable strategies as it used to be while smaller, leading to diseconomies of scale (e.g. Pástor et al. 2015). Theoretically, there must be limits for economies of scale, as otherwise the optimal fund would assume all of the wealth and become the whole market. Similarly, a constant returns to scale would also be unrealistic by implying that the fund's strategy is infinitely scalable. (Zhu 2018.) Taking this into consideration, if this relation were to be found, it should be only temporary as there is optimal size each of the fund types. The family that the fund belongs to has been shown to have an impact on its performance. This is thanks to the large available resources and access to more efficient trading desks, aiding them at minimizing transaction costs greatly (Cici et al. 2018). Among active funds being a part of a prestigious fund family increases the fund inflows and makes investors less sensitive to poor performance and more hesitant at switching their positions elsewhere.

Another characteristic that has been widely used to predict the fund's performance is its age, but with mixed results. The reasoning for seeing positive relation between performance and age is often attributed to the learning-on-the-job effect in addition with increase in name recognition. There are also proponents that suggest that younger funds can outperform the older funds thanks to their unique new strategies, and being more skilled than the incumbents (Pástor et al. 2015).

Passive funds are very limited in their options on how they can affect their costs and returns. One way for them to generate additional profits is the act of securities lending, in which they temporarily lend shares to counterparties (usually shortsellers) in return for a fee. The securities lending ratio has been shown to be one of the most significant predictors of performance among passive index funds and ETFs (Elton et al. 2019a). Minimizing transaction costs is extremely important when attempting to compete against low cost passive ETFs. These transaction costs are often proxied by portfolio turnover ratio in models, which represents essentially how many times over the fund's holdings have rotated over the year. In general, passive funds aim to trade as little as possible, whereas for active funds it is not as cut-and-dried. For active funds the relation between trading and performance largely depends on how successful its trading strategy is and whether the additional profits gained through trading can cover the increased transaction costs.

All of these characteristics have been widely used to estimate the predictive power they have on the fund's performance, and how their impact differs between the two different fund management styles. But there is little research that attempts to isolate the performance difference that can be attributed to the management style specifically. If such gap exists based on the model, it is of interest which factors can account for such overperformance for active funds. One factor that has been shown to advantage the active funds is the effect of index premium, which essentially is the premium that passive funds have to pay due to their rigid trading requirements around index additions and deletions. The theories on whether it is caused by price pressure, downward-sloping demand, certification, information cost or liquidity, all have implications on its size and how it affects the active versus passive performance puzzle. Generally it is agreed that active funds lose to the overall market in the longterm after-fees, but what is the actual performance of the average active trading strategy compared to the passive one if the other fund characteristics were not a factor or kept constant?

This paper contributes to the discussion of the rise in passive investing, ETFs and the active versus passive fund management question by researching the previous literature on the topic and determining how the overall changes have likely affected the market landscape. One of the main focuses in this paper is the difference between these two styles and how their difference in performance can be attributed to their properties and strategies, whether it is due to superior trading, cost efficiency or lack of restrictions. This difference is measured using an assigned dummy variable on a combined sample which includes funds of both styles. Using the base regression which accounts for the expense ratio, fund size and age, we find superior performance for the active style dummy variable which is statistically significant and around 65 bp a month ceteris paribus. This finding is based on the assumption of the other independent variables in the model behaving similarly between the two groups. If interaction terms are introduced to test for robustness, allowing the variables in the model to decouple between the two groups causes the active effect to become insignificant, but still remaining positive. This suggests that the observed superior performance for active funds can be explained by the fund characteristics, and the initial result could be due to omitted variable bias.

Another contribution of the paper is to assess how different fund characteristics among both active and passive funds affect their respective performances. The results show a positive but insignificant relation between AUM and abnormal returns among active funds, but a negative relation among passive funds with the result being statistically significant on one period. The importance of expense ratio is reinforced and consistent with prior literature, as it is highly negatively correlated with the after-fees performance. Our findings suggest that the passive funds are more sensitive to an increase in expense ratio, showing a statistically significant 135–143 bp decrease in monthly abnormal returns per percentage point increase in expense ratio. Among the active funds the decrease was only between 39–55 bp, but statistically insignificant. These results are robust even when alternative scales are used for variables such as AUM. The relation between fund age and the performance was inconclusive among active funds showing mixed results, but statistically significant and positive among passive funds. This suggests that older passive funds have a tendency to perform better, while the fund's age does not have such predictive power among active funds. The adjusted R-squared is expected to be low when attempting to explain abnormal returns, but the very low values especially among the active sample suggests that there might be omitted variables or that the amount of sample data is too small.

More broadly the result suggest that there is no significant difference in abnormal returns between two funds with similar characteristics which only differ in their fund management style, when interactions between the groups is accounted for. This finding aids the investor in making an informed choice, and encourages them to pay close attention to the fund's characteristics, especially expense ratio, as these are the main determinants on how well it is expected to perform in the future.

1.2 Research question

Passive versus active fund management and how different fund characteristics affect the fund's performance have long been highly debated topics among investors and within the academic field. In this thesis I will review the latest academic research available on the topic of fund characteristics especially among exchange-traded funds and what is the impact that the fund's management style has on its performance. In this thesis I will examine the previous research on the subject of fund characteristics and management styles, and seek to find answers to the following research questions:

- What are the characteristics that affect a fund's performance and how?
- How does expense ratio, fund size, lifetime, and management style explain the abnormal returns for actively and passively managed exchange-traded funds?
- What is the performance difference between actively and passively managed exchange-traded funds and what could it be attributed to, e.g. index price effects?
- How has the popularization of passive investing and exchange-traded funds affected the financial landscape and market efficiency?

1.3 Limitations and structure of the thesis

The choice of using the largest funds by size for respective management types limits our sample to quite large funds, which might ignore some effects if they only affect very small funds. Similarly, as funds tend to generally grow in size over time, younger funds might be underrepresented in the sample. The selection might also be biased towards better performing funds due to survivorship bias, however this effect should affect both groups alike if we do not make assumptions about which of these groups is more likely to disband for example due to poor performance. Limitations on data availability for variables such as portfolio turnover ratio and securities lending ratio have limited the factors which were selected for use in the models.

This thesis proceeds as follows. Section 2 reviews the theoretical background on the market efficiency, rise in passive investing, the benefits of active management, the index premium as the hidden cost of passive funds, how fund characteristics affect the fund's performance and how the possible gap between active and passive managed funds can be explained according to the previous historical research on this topic. Section 3 presents the gathered data, where it is from, how it has been filtered and then processed for this thesis. The section also explains the research methods used in the analysis that follows next. Section 4 starts with examining the whole dataset and presents the descriptive statistics for the sample. Various tests follow which are ran on the sample data to determine whether it conforms to the model requirements. Section continues with the main multi-factor regression model, its characteristics and how the variables are defined. This is followed by the results and analysis, with the results compared to existing literature. This section also includes alternative theories to explain the results with tests and robustness checks. Section 5 concludes this thesis.

2 THEORETICAL BACKGROUND

2.1 Efficient market hypothesis

Efficient market hypothesis is a fundamental theory in finance which proposes that stock prices in the global market already encompass all possible known information, whether it be history, news, forecasts or expectations. Fama (1965) was one of the major proponents who introduced an early version of this theory, and argued that what we nowadays call technical analysis, or as he called it "chart reading" is very little use to the average investor as the stock market is ruled by the random-walk. This meant that the sequence of prices is independent of previous sequences in history, though he stated that perfect independence would be difficult to find in practice. Another aspect that brings independence to the prices is the presence of sophisticated traders that have the ability to ingest and evaluate newly released information, to buy or sell accordingly making sure that the current market prices reflect the available information.

As Malkiel (2003) restates, if efficient market hypothesis held true, even fundamental analysis would not provide any greater returns along with technical analysis when compared to someone just holding up any diversified portfolio with similar risk. Malkiel however counters some of this perfect world thinking by reminding that investors, even if professional ones, are still humans and they make mistakes. This means that the markets can still for short periods have inefficiencies, which in turn enables the professional traders to make their profits in it. If this wasn't the case and there was no money to be made in the market, this would in itself deter traders and further increase inefficiency.

Fama (1970) expanded on his original hypothesis, introducing three different forms of efficiency based on well the prices reflect the information. The first one, weak form implies that the prices reflect all previous price and returns history. If this holds true, the second is semi-strong form which also takes into account how quickly and well other publicly available information is adjusted into the market price. This can include announcements, stock splits, financial statements and reports, security concerns, leaks or any other information that is available to public and can affect the price of the stock or how it is perceived. Lastly, in the strong form market efficiency, in addition to all public, also all private information is included already in the valuation of the stock price. This would mean that even someone with insider information would not be able to take advantage of the stock's current pricing. Even very early it was apparent that it was very unlikely for the strong form market efficiency to hold up in the real world, and there was already empirical evidence of certain groups such as one with monopolistic information on NYSE order fullfillment, being able to use their private information generate profits. Fama's empirical results through weak and semi-strong form tests were largely in support of efficient market hypothesis and he found it as a quite good representation of the real world. If we were to assume that semi-strong market efficiency held true in the real world, it would cast doubt on the role of active stock management as they shouldn’t be able to generate any higher profits than passively managed funds using buy-and-hold strategy on stocks with similar risk characteristics, quite the contrary when accounting for higher fees.

2.1.1 Alternatives to the efficient market hypothesis

As not every trader on the market can be completely rational, and to be aware of all the facts and handle complex valuation models, there inevitably will be transactions which move the value of a stock away from its “true” value. This allows certain “smart money” traders to take advantage of these short-term mispricings and anomalies, which partly explains how the participants who do their part in keeping the market efficient, can still gain profits. Eventhough efficient market hypothesis is still held as a cornerstone of the classical finance theory, behavioural finance has gained some popularity in attempting to explain some of the observed deviations from the rational behaviour. One of these alternatives to efficient market hypothesis is a theory called noise trader approach, introduced by Shleifer and Summers (1990). This theory suggests that investors with certain sentiment on a stock create excessive demand on it, which is not fully counteracted by the arbitrageurs and thus can actually affect the price of the stock. The rational traders who do not share the same sentiment consider this risky and will only have a limited impact on correcting the price discrepancy caused by the other group. Other investors in the market should take these noise traders into account when making their own decisions, and there has been empirical evidence of the investor sentiment being able to affect the markets. (Nyakurukwa & Seetharam 2023.)

Some other alternatives presented in the previous literature include fractal market hypothesis, which supposes that the variations in the markets are caused by the different time horizons the investors target when making their decisions. It purports that the stability in the market is achieved when both short and long horizon investors both participate in the market, and deviations in balance occur when one of these groups for some reason reduces their participation. The reason can be for example economical turmoil, in which the long-term traders which under normal circumstances are the ones providing liquidity to the short-term investors, can abstain from the market causing instability. The heterogenous market hypothesis expanded on this, but did away with the assumption of investor homogeneity. Rather than all the investors with a similar time horizon acting as a group, they would instead be able to interpret the news differently within the group and not act as a unit. Further alternatives include adaptive market hypothesis which supposes that investors act rationally while the markets are stable, but when the markets encounter instability and disruptions, the investors will act irrationally either by fear of missing out on possible opportunities, or in fear of losing value on their investments. (Nyakurukwa & Seetharam 2023.)

While not exhaustive, the few previously mentioned theories and hypotheses provide some possible alternatives that attempt to supplement some of the weaknesses that can be argued to exist in the original efficient market hypothesis from the assumption of investor homogenuity, instant incorporation of information into prices, to rational investor behaviour. Thus a more holistic understanding of the price behaviour can be gained by viewing the continuous price development as a multifaceted problem that encompasses also human aspects that can be researched within other disciplinaries such as psychology and sociology.

2.2 Value of active management

In Sharpe's (1991) famous "The Arithmetic of Active Management" he states that within the parameters that "active" and "passive" management styles are defined appropriately, it must follow that "(1) before costs, the return of the actively managed managed dollar will equal the return on the average passively managed dollar" and that "(2) after costs, the average actively managed dollar will be less than the return on the average passively managed dollar". According to him, this will hold for any time period. Clear distinction must be made on how active and passive investors are classified. He differentiates these two based on how much they trade, with the passive investors attempting to hold every security in the market, what you would call "market portfolio" which is a portfolio weighted based on market capitalization. The active investor is someone that does not follow the aforementioned strategy, rather they tend to base their trades on perceived mispricings of the market and they also tend to trade more frequently as the name suggests.

The first point on returns being equal between active and passive returns before costs can be shown with simple arithmetics: as passively managed portfolios evidently provide market returns before costs, and as market return needs to be the weighted average of passive and active managed segments' returns, it must follow that the average actively managed portfolio also obtains market returns before costs. With the assumption of actively managed funds having higher fees due to security analysis, brokers, traders, specialists and other market-makers having their cost, it is clear that with the returns before costs being equal with passively managed funds, the returns after costs must be lower. In addition he goes as far as to claim that any empirical analysis that comes to a different conclusion is guilty of improper measurement. (Sharpe 1991.)

Pedersen (2018) points out that within this aforementioned Sharpe's framework, the market portfolio is taken as given and is a fixed set of securities, eventhough in real life this set of assets is constantly changing based on new companies being created and old ones disbanding. Also in reality we need a representation for the abstract market portfolio, for which we can use an index such as the S&P 500. In these cases, the constituents of the index do change over time, when companies pass and fall under the index inclusion line and get added or deleted from the index. This forces the passively managed fund to rebalance their market-weighted portfolios, in which cases they're often forced to trade under more unfavourable prices than active managers would. As active managers have more flexibility around these index changes, for example not being tied to following the effective dates, they can take advantage of the price pressure and other inefficiencies and provide increased returns compared to passive investors.

Another aspect in how the actively managed funds gain an advantage over the passive ones, is when capital markets are raising capital through methods such as IPO's (initial public offering), SEO's (seasoned equity offering) and share repurchases. In these instances active investors can perform due diligence and participate in the opportunities that they find undervalued. (Pedersen 2018). According to Ljungqvist (2007) in 1990's United States the underpricing of IPOs averaged over 20%, and according to more recent studies, such as Boulton et al. (2020) with IPO data spanning from 1998 to 2018, this profit opportunity and relation still stands.

The way active investors as informed actors can choose which IPOs to participate in, they can benefit from the underpricing of certain offerings and gain advantage over the passive investors who cannot distinguish the good opportunities from the bad ones, due to the lack of security analysis. Active investors who take advantage of the underpriced IPO can sell the shares to the passive investors in the secondary markets, who need to buy them into their market portfolio. Even if the passive investor were to participate in every IPO available, they would be suspect to adverse selection. In cases where the active investors have done their research and have evaluated the stock to be a good buy, it would become oversubscribed and the passive investors would still need to supplement their market portfolios from the secondary market. On the contrary, when the active investors deem the IPO to be overpriced, they don't participate in it, allowing the passive investors get all the shares they've requested. With the low demand on these shares, they will eventually end up in the secondary market at a lower price, from where the active investors can buy them at a discount. In both scenarios, passive investors end up being at a disadvantage. (Pedersen 2018.)

The previous also holds for other instances on capital markets, such as SEOs and share repurchases. Passive investors not performing their own research on these issues causes the same adverse selection than in the case of IPOs, from which the active investors can benefit of. The fact that there are actively managed funds doing the security analysis on these assets is crucial so that the markets can function at fair values. Without them, any offering regardless of their fundamental value would be bought out by the passive investors enticing opportunistic behaviour for example to bring companies public at any price. Active investors thus create a positive impact on the economy and allow companies to raise capital. (Pedersen 2018.)

The fact that Sharpe's framework holds for any time period has come under scrutiny, as it assumes that the passive investor has obtained the market portfolio at the start of any time period for the midquote price. This assumption does not take into account that the passive investor has had to trade in order to obtain this portfolio in the first place. As previously argued, if the passive investor in average pays a premium for their securities, and sells them at discount compared to active investors, then it is possible that the active investor can gain higher returns compared to passive investors before fees. (Pedersen 2018.) With these arguments in mind, the assumption of passive investor being able to obtain their portfolio for favourable prices can seem misplaced and weaken Sharpe's point.

As actively managed funds pursue creating value by performing security analysis and within their perception picking undervalued stocks and benefitting of overvalued ones, it is true that some funds do manage to outperform the competition. But this still turns out as a zero-sum game when we observe the active investors as a complete segment, as the active managers who seem "informed" about the market gain profits off the "non-informational" investors who withstand the losses. For example within the same active investor segment there exists investors who pick glamour stocks versus those who choose value stocks. The gains and losses within the same active investor group are averaged and net to zero. This outcome holds because how Sharpe defines "active" being anyone who is not "passive", grouping together informed and non-informed investors. (Pedersen 2018.)

Even if the assumption is held that before costs, the average active fund performs as well as the average passive fund, there exists active funds that do outperform the passive indices alongside with the other active funds. Pástor et al. (2015) find significant evidence of fund factors that play a role in the active mutual funds ability to outperform the passive competition. They find that the increase in size of the assets under the management of the fund decreases their ability of outperforming the passive indices. They also find that the managers in these active funds have become more skilled over time, but this has also coincided with the industry growth. Despite of this, their results suggest that the longer the fund has existed, the less likely it is to outperform the passive indices.

In their paper, Sheng et al. (2023) provide a theory which they call valuation cost hypothesis as one of the explanations on why funds that have to spend a large amount of effort and resources on research focused on hard to value stocks, can exhibit larger fees. As actively managed funds have freedom to create opportunities in the market and are not tied to certain benchmark indices, they can create value in ways that passively managed funds can not. Attempting to perform valuation and investing into smaller, less liquid, highly specialized or otherwise hard to value firms requires increased efforts. But this in contrast can provide vastly higher returns in accordance with the higher risk compared to the larger, safer and more established companies. These harder to value firms are often high-growth companies which currently have very low profitability, and their their strategy is focused primarily on growth before establishing profitability. But as these companies currently are not very profitable, their valuation is mostly based on their future potential, and this is very hard to evaluate and pinpoint precisely. One issue with attempting to evaluate such effect is that it is not entirely clear how one would quantify how opaque or hard a fund is to value correctly. Kumar (2009) discusses value this uncertainty and attempts to determine which factors play a role in the effect. He finds that idiosyncratic volatility, as in the volatility in the stock's price which is not dependent on the overall market volatility, is one aspect that creates value uncertainty. One explanation for this is disposition effect in which people hold higher reference values for their stocks than they otherwise would. These people are less likely to sell when they are down to realize their losses due to their distorted price anchoring. This hesitance in turn introduces friction in the market price of the stock and makes valuation more difficult. Additionally, asset intangibility has been shown to affect the valuation sentiment as found by Baker and Wurgler (2006), as with the analyst forecast dispersion in IBES database which aggregates analyst estimates on the stock prices. Sheng et al. (2023) used these three predictors and their results support the hypothesis that the expense ratio relates positively to these three proxy variables that indicate the difficulty of valuation.

2.3 Index inclusion and exclusion effect

When using an index as a representation of the market portfolio that the passive investors hold, becomes the impact of the changes in this index significant to the returns they are able to obtain. As previously mentioned, the passive investor has limited options to time their transactions around the changes in the index, as the passively managed funds attempt to replicate the performance of the index as closely as possible. A measure called "tracking error" is very important indicator for these passive funds, as it indicates how little they have deviated from the underlying index's performance. As the funds tracking the same index are rebalancing their portfolios simultaneously around the effective date, it is shown in research such as Petajisto (2011) to cause price pressure upwards on index additions, and downwards on index deletions. The effective date is a pre-announced date when the changes, additions and deletions in the index come into effect.

According to Petajisto's (2011) research on the index premium on S&P 500 and Russell 2000 indices for the years 1990 to 2005, he found that market-adjusted price impact for these two indices on average were +8.8% and +4.7% for additions, and −15.1% and −4.6% for deletions respectively. These changes are calculated from the date the changes to the index are announced to the public to the date when they become effective. The price effect partially reverses over time, but not completely. It is observed that over half of the price effect is reversed after the following two months, but long-term estimates on whether it reverses completely are hard to do accurately. Taking into account how often these passive funds following the S&P 500 and Russell 2000 have to rebalance their portfolios, he was able to calculate how much this cost, index premium, actually affects a passive investor. For the S&P 500 he observed an annual cost of 21–28 basis points (bp) and 38–77 bp for Russell 2000 which the passive investor loses having to trade with unfavourable prices. The cost estimation was done based on comparing the index fund's performance to an index-neutral strategy, in which a portfolio with a similar profile and characteristics is built but which is not tied to an index. The range between the lower and upper bound of the cost estimation comes from the uncertainty of how much of the index premium is reversed over time, as this is subject to discussion. (Petajisto 2011.) The issue of index premium is very important from the passive investor's perspective as despite of being "passive" and avoiding trading, they still have to participate in it if they are using an index as a representation of the market portfolio. This strategy allows the passive funds to obtain a very low expense ratio, for example 0.03% for Vanguard S&P 500 ETF and 0.10% for Vanguard Russell 2000 ETF following their respective indices (Vanguard). Contrasting these low 3–10 bp expense ratios with the proposed index turnover cost of 21–28 bp for S&P 500 and 38–77 bp for Russell 2000 annually seems very significant consideration for the passive investor.

Many different theories for the observed price effect on index inclusions and exclusions have been proposed in the literature, which mostly differ in whether the reversal of this premium is complete or only partial. This distinction is what potentially the difference consists of between the prices a passive investor and an active investor would pay. The smaller the reversal and the more permanent the observed positive abnormal returns following an index change are, the better it turns out to be for the passive investor. This is due to the passive investor having to trade with unfavourable prices around the index change, and if the effect is just temporary and it reverses, the premium paid for the stock vanishes. Next I will present the generally applied theories explaining the observed effect and how they view its permanence as it has great significance in comparing performance of passive and active fund management.

2.3.1 Price pressure and imperfect substitutes hypotheses

Price pressure hypothesis (PPH) suggests that the price effect observed with big sales or purchases is caused solely by the inelasticity of the short-term demand curve and is absent of actual new information and does not signify the future prospects of the firm. In these cases, the market participants who are willing to accommodate the market's liquidity needs demand compensation in a form of a premium. PPH states that the demand shock is only temporary and in the long-term the stock's value will return to their fair full informational value. This is in contrast to imperfect substitutes hypothesis (ISH) in which it is assumed that the securities do not represent perfect substitutes to each other and that the long-term demand curve shifts permanently and creates a new equilibrium price level. (Harris & Gurel 1986.)

Evidence in support of the price pressure phenomenom hypothesis includes Elliott and Warr's (2003) study from 1989 to 2000 on NYSE and Nasdaq in which they observed positive cumulative market-adjusted returns of 0.73% for additions and 3.41% for Nasdaq. This impact was measured from the announcement date to the effective date. Both of these were statistically significant and different from zero. One trading day later after the effective date, they observed a −0.61% abnormal return on NYSE and −1.15% for Nasdaq firms. This shows that for NYSE firms most of the price effect had reversed within one trading day, but for Nasdaq only a partial reversal is observed for the first trading day post-effective date.

2.3.2 Long-term downward-sloping demand hypothesis

A lot of finance theories rely on the assumption of investors being able to sell and buy an asset without causing a significant impact on the price level. Firms with different characteristics such as dividend policy or how they are financed should not affect its value. This is stated in the Modigliani-Miller theorem, however it assumes a world without taxes which changes the balance when applied in the real world where interest is tax-deductible and in which they also cause other market frictions. Also theories based on the efficient market hypothesis such as CAPM or APT assume that the stocks are near perfect substitutes to each other, thus facing horizontal or nearly horizontal demand curve. This means that regardless of the selling or buying pressure on the stock, the price should not move up or down significantly and stays near its fundamental value. Financial economists have sought ways to measure the slope of the demand curve and have done so by inspecting large block trades measuring whether these large supply or demand shocks cause a move in price. This method of estimation though is in a way flawed, as the large block trades could also be consistent with the information hypothesis, in which these trades act as a signal to the market that the buyers or sellers have private information and that the trade is a good deal. (Shleifer 1986.)

Shleifer (1986) is one of the first to research this phenomenom and proposes that examining the stock inclusions and exclusions from an index as a better way to measure the slope of the long-term demand curve. Index changes are based on publicly available information such as market capitalization and are usually predictable. From this we could assume that index changes do not signal a real change in the value of a stock, and this way the observed changes would provide a good research point in measuring the effect of how much buy or sell pressure affects the stock price. In his study of 246 firms from 1966 to 1983 he found no evidence of positive cumulative abnormal returns (CAR) approaching the announcement date, which would indicate that the markets were not pricing in the inclusion before the announcement in anticipation of the index addition. In the sample from September 1976 onwards he finds that the shareholders see around 3% capital gain compared to the announcement date's valuation. This premium seems to persist for at least 10 to 20 trading days. This effect could be attributed to the information hypothesis, which Shleifer tests by comparing the abnormal announcement date returns between the groups with different bond ratings. The firms without previous bond ratings did not gain higher return than other firms included into the index, which he claims casts doubt into the applicance of information hypothesis in this case.

Taking a look at other indices apart from S&P 500 is important to expand our viewpoint, as they might behave differently and have different characteristics. The case of Nikkei 500 provides a great opportunity of studying the effect of downward-sloping demand curve (DSDC) hypothesis, as the Nikkei 500 index is restructured and balanced annually simultaneously, instead of over the year which is the case for S&P 500. In a study with data from 1991 to 1999 they found an increase in trading volume for index additions, increasing their demand. This demand is bound to drive the price upwards in support of the DSDC hypothesis. Similarly, the opposite reaction is true for the index deletions. The permanent event-day price effects combined with the observed short and long-term volume impacts lend credence to the hypothesis that firm stocks are not perfect substitutes, and that the long-term demand curve for stocks might actually be downward sloping. (Liu 2000.) However reaching statistically significant conclusions on the slopeness of long-term demand curves is difficult as the estimation window required for this is so long.

Another issue with measuring the slope of the long-term demand curve is that most research done previously on this topic relies on events such as S&P 500 index changes, which themselves are not exogenous, and can appear informational. A study released in 2019 attempts to rectify this issue by examining an event in Indian securities market in 2010, in which a regulatory change forced all listed companies, except government-owned ones to have at least 25% public shareholding. Public shareholding here being classified as holding of equity shares by parties which are not promoters or subsidiaries of said company. This provides an event which by its nature cannot hold any new information, but yet causes large blockholders to sell large portion of their companies' shares. They rule out the possibility of these events being informational by examining whether the price reaction is observed after the sale announcement or on the day of the actual sale. The price reaction to the former would imply that the markets are reacting to the information of the sale, and the latter implies that the effect is solely due to the sale volume. They find a significant decline in the share prices on the day of the sale, but with a quick reversal. Within 16 days post-sale, the price level was already at pre-event levels ruling out the existence of a downward-sloping demand curve in an exogenous supply shock. (Jain et al. 2019.) With the evidence from exogenous shocks available, the explanation of index change price effects having less than full reversal being due to downward-sloping demand curve seems unlikely assuming the index change being a non-informational event. Competing theories such as information, investor awareness or liquidity hypothesis provide more credible and fitting evidence to explain the price effect observed on index changes.

2.3.3 Information or certification hypothesis

The notion that the observed price effect around index changes is due to the positive or negative informational signal given by the action of inclusion or exclusion, is very often offered as an explanation. Chen H. et al. (2004) propose that despite of the index change decisions being made based on public information, they have the incentive to minimize how often this needs to be done. For this, they attempt to select companies into the index which have longevity in their business, great future prospects and exceptional leadership. However this observed effect should be symmetrical for both upside and downside based on the its nature, as the choice of which firm to exclude from the index would be subject to similar inspections on its future prospects. From their samples from before and after 1976, when the procedure of announcing index changes was made publicly accessible, they found discrepancies between the two samples. There was no evidence of positive abnormal returns before the 1976 change, which puts into question the informational value of the index addition. Also prior to the change, investors could call S&P 500 to inquire about the upcoming changes in the index, but according to Harris and Gurel (1986) this opportunity was not popular, and only about five to ten investors contacted them about this per year. These were mainly index fund managers. They suggest that if the index change provided a certification signal, it would have been expected that this opportunity was more taken advantage of.

A study by Hrazdil (2010) from 1987 to 2004 focused on the factors by which the stocks are chosen into the S&P 500 index and whether these include more than the eight criteria stated by Standard & Poor’s. Typically there are multiple eligible candidates when changes are proposed into the S&P 500. In the model they incorporate the firms which were eligible but not added to find how they differed from the ones which managed to get in. They also investigate whether the criteria found provides information that is not already in the price, and would explain some of the positive abnormal returns that are often observed on index inclusions. Through their model they find that excess to the eight given criteria, it seems that S&P committee also considers the risk variables and historial performance through publicly available data. Evidence further indicates that along with the historial measures, S&P may incorporate some private value-relevant information about the company’s future performance in the decision but that its role is very small. The highest explanatory power appears to be on the current performance variables. The evidence for information hypothesis they find is limited, as the future operating performance seems independent of the decision whether the firm was included in the S&P 500 and to be dependent mostly on the firm’s historical publicly available measures and information.

When discussing the information hypothesis, it is important to distinguish the assumption that important information leads to an inclusion, versus the reverse where the index inclusion leads to the improved performance of the company. To find the difference between these two, it is possible to compare the company’s expected future earnings before and following the index inclusion. Analyzing the company performances between 1987 and 1999 for firms that were added into the S&P 500 index, it is observed that along with increased investor expectations for these firms, these firms also see an improvement in their actual earnings compared to the benchmark companies. They conclude that index inclusion does indeed lead to better corporate performance and object to the position often held in previous studies of index changes being information-free events. (Denis et al. 2003.)

Not all indices behave similarly to S&P 500 and this allows us to see whether we can find evidence for this hypothesis from other sources. Russell 3000 tracks the 3000 largest public equity companies in the US, weighted by capitalization. As a subset of Russell 3000, the Russell 2000 is a small-cap stock index which tracks 2000 of its smallest firms. Contrary to S&P 500 whose changes are irregular, Russell 2000 balances on an annual basis and it sees about 500 companies enter and leave the index each year. It also provides a better way to inspect the firms leaving an index, as usually index exclusions from S&P 500 are caused by bankruptcies, mergers or reorganizations. This differs from Russell 2000 which firms are often removed from but still continue in business. Firms are added and removed from the index based on market capitalization, and they can either move in or out of Russell 1000 or from Russell 2000 altogether. It is observed that that due to these changes being clearly predictable and announced way before the reconstitution date, the price effect found to be a temporary event, and that it has no permanent effect on the valuation of the firm. (Biktimirov et al. 2004.) The fact that the decisions made on Russell 2000 are based solely on objective measure such as market capitalization, and not by a subjective committee like in S&P 500 could partly explain why it has no permanent effect, and why for S&P 500 it could be seen as a positive signal.

2.3.4 Investor awareness and information cost hypotheses

Investor awareness hypothesis was originally brought up by Merton (1987) in his market segmentation model, when he started looking into the markets reaching an equilibrium pricing considering that all participants have incomplete information. In their model all investors are only aware of a subset of all available securities, and transact only within that group. Their models main assumption is that the investor builds their optimal portfolio consisting of only securities they are aware of.

In order for an investor to become aware of a stock, two things are required. First, someone needs bear the cost and effort to do the research, become aware of the stock and then transmit that information onwards for example as an analysis. Second, the receiver of the information needs to withstand the cost or effort to gather and process it. Merton compares these information costs to Arbel-Carvell-Strebel theory of “generic” or “neglected” stocks. In this theory, stocks that are smaller or less known have smaller amount of analysts looking into them and the quality of the information available is lower compared to companies that are well known. As a consequence of all of these factors, it is logical for the investors to require higher expected returns for these securities that have higher information gathering costs and lower quality information. (Merton 1987.) To continue from Merton's work, Chen H. et al. (2004) study the significance of shadow cost. Shadow cost is a premium that the investors require for the non-systematic risk they bear by holding insufficiently diversified portfolio. In the event of S&P 500 index inclusions, the newly found exposure alerts the investors about the company, bringing it into their awareness. The larger pool of securities the investor is aware of, the better diversified they are, and the lower their required return is.

By this market segmentation model, the investor awareness effect would be asymmetrical for index inclusions and exclusions. Unlike how the investors become aware of a stock, it is not logical for investors to become "unaware" of it, despite of it being removed from the index. From this follows that the reduction in shadow cost for inclusions must be larger, than the shadow cost for deletions. This asymmetrical effect is what mostly seperates this theory from other theories related to price effects such as downward-sloping demand curve hypothesis or price pressure hypothesis where the effect is similar to both directions. By analyzing samples ranging periods from 1962 to 1999, the evidence concurs with the theory that the effect is asymmetrical. With the advantage of having samples from multiple time periods with different market conditions, the study was able to find supportive evidence for investor awareness theory. It was found that prior to 1976 when it became a standard for index changes to be announced publically, there was no significant price effect for index changes. Without the announcements for S&P 500 index changes pre-1976, there would have been no new investors to become aware of the stock. (Chen et al. 2004.)

The issue with studying theories on investor awareness is the lack of accurate ways to measure it directly. There are some ways to measure it indirectly, namely via amount of registered shareholders, amount of institutions, and by institutional ownership. In the same study from 1962 to 1999 by Chen H. et al. (2004) it appears that a significant increase of shareholders is observed on index additions, but the decrease observed for deletions is small to none. The result on deletions is however prefaced with caution of a small sample size. It is pointed out that with index funds being major participant in the transactions, they would be buying large blocks on additions, and selling large blocks on deletions. As the counterparty to these trades would most likely be non-institutional investors, this would actually decrease the amount of shareholders in index additions and increase on deletions. The amount of institutional investors is also expected to increase with the index addition. Eventhough institutions are more sophisticated and are most likely already "aware" of the companies before the index announcements, they however are not necessarily that interested in the companies until they are in the index. They find that institutional ownership does in fact increase on index additions and vice versa. Biktimirov and Xu (2019) in their study on Nasdaq 100 find that even after controlling for other factors such as liquidity, the proxies for investor awareness are still highly related to the cumulative abnormal returns observed around index changes. Their recent findings give additional support to the investor awareness hypothesis.

Assessing the investor recognition of companies, it is logical that the effect is more pronounced when it comes to companies and markets in which the awareness is lower in the first place, such as emerging markets and their companies. So studying these alternative markets could give a good data point in how it affects markets other than S&P 500 and alike. A study on emerging markets between 1996 and 2008 was done on MSCI EM index, and analyst recommendations data was used as a proxy for investor recognition in this case due to limited data availability in these markets. They find an asymmetric price effect, which is contrary to what a pure demand or information-driven theory alone would expect. This is consistent with the investor awareness hypothesis, but does not alone confirm the theory, so another proxy for investor recognition, the analyst recommendations is also used. A causality can be observed, that the stocks given higher amount of coverage witness larger price increases, which is even more amplified if the stock previously had no access to foreign markets. The authors argue that the permanent price increase observed on index inclusions can be attributed to the increase in investor awareness on emerging markets. (Hacıbedel 2014.)

The results in the same study by Hacıbedel also divert from the previous research regarding the price effect around index inclusions on S&P 500. Unlike the previous research where there is evidence of price reversal in studies on S&P 500 index, in the case of emerging markets there was no sight of this phenomenom. There are two explanations, one is that the index announcement contains new information, but this would be expected to apply for both inclusions and exclusions, thus rendered incorrect. The second one is that the investor awareness is increased by the publicity that comes with the inclusion, and as this is asymmetric effect it is consistent with the observed results. The author emphasizes that these results are specific to emerging markets, and cannot necessarily be generalized to more advanced markets. This is due to the companies in emerging markets having very limited visibility and access to foreign markets being very significant for them. (Hacıbedel 2014.) The reversal or the lack of it is very important when discussing the costs the passive investor has to endure due to index turnover. As the index has to transact with the inflated prices around index changes, it is more beneficial to the passive investor if the effect was permanent and thus the value would not erode in the near future, or until the index has to sell the stock. Also the more permanent it is, the smaller the gains are the active funds can take advantage of in the form of significant price effects.

2.3.5 Liquidity hypothesis

The basis for liquidity hypothesis is that following index inclusions the increased liquidity in the asset reduces the trading costs for the investor, thus reducing their expected returns. It is generally observed following an inclusion in the index that there is significant increase in trading volume initially right after the inclusion, but also in long-term. Another way to measure the liquidity on the stock, that also would reduce the trading costs, is the bid-ask spread. This is the difference between the prices for which the participants are offering to buy, and those willing to sell. The stock also witnesses increase in open interest on futures, which further increases the stock's value. (Edminster et al. 1996.) Though liquidity is generally seen as a positive, too widely diffused stockholding by some accounts could cause problems as none of the parties have the incentive to perform internal monitoring of the firm, as they only gain very little benefit from it. Also the increase of liquidity is said to lower the "exit cost" of unhappy stockholders, which would further reduce the amount of internal monitoring. (Bhide 1993.)

Hedge and McDermott (2003) in a similar manner find that the index inclusions increase liquidity for added stocks and that the effect lasts long-term, past the three month window they observed. The liquidity appeared as increase in quoted dollar market depth, trading volume, and trade frequency. Their analysis shows that the decrease in trading costs comes mainly from reduction in effective spread and from the fall of variability in order flow. The behaviour common to buy-and-hold index trading is reduction in larger block orders in order flow. This in itself helps the market-makers to improve the management of their inventory, with the more predictable and frequent order flow reducing the overall costs. They also offer an alternative view, in which the asymmetric information costs in trading could also in theory increase. This can happen if the frequency of trades and variance of uninformed trades were to decrease. However, in the study they found evidence of liquidity decreasing for deletions in the following months, which is seen as increase in effective spread. Biktimirov and Xu (2019) study expanded on the previous research by differentiating between new and old additions into the index. They found consistent evidence of liquidity decreasing on new index deletions from Nasdaq 100 index.

2.4 Impact of rise in passive investing

The popularity of passive investing has increased greatly in the past decade, which has enabled the ease of access for anyone to get into investing, whether it be mutual funds or exchange-traded funds (ETFs). Sushko and Turner (2018) consider that the consequences of this shift in investing have not been sufficiently researched and may be the cause of some problems in the securities markets in long-term. The major issue stems from the fact that cash flows into passive mutual funds and passive ETFs remained persistent regardless of the market sentiment, as opposed to active mutual funds which faced significant fund outflows during market stress periods. In a way, these passive funds act as stabilizers in the market during times of stress preventing the natural price movement to their fair values.

Nature of passive investing is that the funds are allocated based on the index constituents weighted by the market capitalization, regardless of the fundamental value of the assets themselves. These mechanical investment rules allow distortions in the pricing of these underlying assets. As passive portfolio managers do not seek security-specific information when deciding their asset allocation, they're taking advantage of the work done by active investors who keep the prices fair. When the share of passive investors increase compared to the active ones, the information value the prices represent is bound to decrease. Another observed effect is the co-movement of assets, as passive funds trade with the full range of the whole portfolio. This causes the prices of these assets in the same basket to move in parallel. How this is counter-productive and actually hurts the passive investor is, that the co-movement of assets in this index reduces the diversification benefit of holding a portfolio with large amount of assets. In a study of 462 additions into S&P 500 index a significant increase in correlation between the stock and other index constituents was observed after the inclusion. When measuring the correlation of daily returns from 200 days before to inclusion and 200 days after, the correlation coefficient increased by 15.6% from 0.45 to 0.52. (Sushko & Turner 2018.)

Claessens and Yafeh (2013) also studied the comovement effect and gathered data from forty developed and emerging countries spanning over 10 years. Focusing on the major indices, they found that firms that get added into the index experience an increase in their beta, and this effect is even more pronounced if the initial beta was low to begin with. They also find that the less concentrated the index is into a selected few stocks, the less comovement was observed. This is said to imply that if an index has only few large constituents, it is possible to just buy these few stocks separately and not an index. Subsequently, the effect was more prominent in indices with a large number of smaller companies. In the cases where the effect was observed, it means that the after inclusion market returns become a higher predictor of the specific firms returns than prior to the index addition. The results were fairly consistent for both developed and emerging markets, but for some countries the effect was weaker, with zero to even negative comovement effect, these included countries such as Finland, India and Portugal. With the assumption of index changes not including any new real information, the effect has to be derived from something else. It has been suggested that it can be thanks to the increase of liquidity with increase in volume, tightening of bid-ask-spread, and also in the increase of informational oversight and analyst coverage into the firm.

Some previous crises have been blamed on the rise of passive investing, and the "constant allocation strategy regardless of fundamental value". The disastrous 2008 oil bubble has been attributed by some to investors that can be called "index investors" or "index speculators" which invested heavily in commodity ETFs. With these ETFs being heavily weighted with crude oil futures, it drove up the price further than warranted by the market. According to Michael Burry, the investor who predicted the 2008 housing crisis, the same passive investment strategies that played their part in the 2008 oil bubble were partly responsible for the equity price bubble in 2019 (Tokic 2020). The equity markets always keep hitting market highs, but it is unclear how much of it due to the increase in passive investing. Tokic (2020) further points to emergence of positive-feedback trading, pushing active funds taking positions in passive funds to take advantage of the major gains in major market indices, which itself is a sign of bubble-like behaviour. Positive-feedback trading in this case can be characterized as behaviour where the investor keeps buying during market advancements, and sells during market declines. He believes that during recessions this will lead to large fund flows from passive to active investments, causing a great selling pressure on ETFs, pushing the price down.

As the increase in popularity of passive investing is largely driven by retail investors, there are avenues in which this can be taken advantage of. According to Bradford de Long et al. (1990) famous speculative investor George Soros has used the behaviour of the public in his investment strategies. As he views that these opportunities, for example in 1960's conglomerate and 1970's Real Estate Investment Trust (REIT) booms provide an opportunity to seek to benefit off the uninformed investors on the market. The strategy he describes is that by purchasing these assets pre-emptively before the public, the following price increase and popularity will drive up the price even further, as in a positive-feedback loop. Finally, when the companies fail to perform to the high expectations and before the inevident drop in the asset price, the speculators have already sold their positions. These speculators could be categorized into two groups, the insiders and the outsiders, where the professional insiders are the ones causing market destabilization by exaggerating the upswings to lure in the outsiders who buy high during and sell low after the euphoria (Kindleberger & Aliber 2011). With this in consideration, it could be profitable even for the "smart money" to keep piling capital into the major passive indices regardless of the fundamental expectations as the euphoria of the bull market is bound to bring in the uninformed "outsiders" driving up the prices even further. When the markets eventually face a correction, it is more likely that the retail investors are the ones taking the larger hit, as the professional investors have pulled out.

Despite of their rise in popularity, passive funds held only a 15% share of the total securities in the US equity markets, and around 5% or less in other equity markets in 2018. But is expected to become of even greater significance in the future if the same trend continues. The measurement of how much assets are managed passively can be measured by reviewing the assets under management by index tracking funds, but this does not come without problems, as in practice classifying between active and passive investing is not always clear. There are a lot of active funds whose portfolios have high resemblance to the benchmark portfolios, and it is said that some them practice "closet indexing" in which they do use the benchmark index to assist them in portfolio creation. This means that some active funds, especially large ones like pension funds and insurance companies do incorporate aspects of passive investment strategies in their active portfolios. Also the increase of more complex ETFs such as "smart beta" strategies in which the ETF uses some other than traditional market value to weight the asset allocation have gained popularity. These can be factors such as value, volatility, or dividend yield. These increasingly complex strategies could be classified as more active than passive. With these factors in mind, the assets under passive management might have a bigger impact than the numbers might initially lead to believe. (Sushko & Turner 2018.)

Cremers et al. (2016) performed a study open-end equity mutual funds and exchange-traded funds in 32 countries from 2002 to 2010. They find that active funds have become more active and lowered their costs since the increased competitive pressure from low-cost passive funds. They also studied the phenomenom of closet indexing in active fund management and that how much of active management is actually active. They did this by calculating the active funds share of portfolio holdings that differ from the benchmark index holdings, giving an estimation of how much their portfolio strategy differs from the passive benchmark portfolio. They find that with the explicitly indexed funds providing a highly competitive and low cost option for gaining exposure to market beta, the active funds with higher costs are forced to differentiate themselves by becoming more active. This is supported by the evidence, as the countries with more explicitly indexed funds had actively managed funds with lower total expense ratios and with more differentiated products, as in having less overlap between their portfolio and the benchmark one. On the other hand, in countries with limited options for explicitly indexed funds, it was observed that actively managed funds practiced more closet indexing while still charging higher fees and underperforming the benchmark indices. Their findings suggest that the increase in explicit passive investing is good for the market efficiency, as the competition among the active funds provides better options for the investors. Nevertheless, they emphasize that continuous growth of index based passive investing warrants further research and can have large impact on the markets and asset prices.

Anadu et al. (2020) argue that with the increase of passive investing, very specialized strategies such as leveraged and inverse ETFs have also grown in popularity possibly having large negative effects as they can amplify the market swings up and down. This combines with the fact that the shift to passive investing has concentrated the funds between smaller amount of asset management firms, allowing them to bloat in size. They also emphasize that the large size of these index tracking funds have a great impact on the market phenomena such as index inclusion effect which has been extensively reviewed in this thesis. When it comes to specifically mutual funds, a large benefit the passive mutual funds have is the decreased risk of having to do liquidity transformation. This is a process in which illiquid assets are transformed from long-term securities to liquid assets such as cash, or cash-like assets like deposits. In general, actively managed mutual funds are more vulnerable to unexpected negative cashflows as they offer cash redemptions and they run the risk of investors pulling out their capital during market uncertainty. During high net outflows from the mutual fund, they might be forced to raise cash by selling their assets to pay back the investors. This can cause two problems, first, this would incur extra transaction fees and second, they might endanger the fund as whole if they have positions which cannot be liquidated quickly or at fair prices. In some way this can create bank run like scenarios in which the investors are incentivized to quickly to redeem their investments before the fund suffers great losses from fire selling their assets. Investors into active mutual funds are also more likely to be the type who chase profits by jumping from fund to fund based on their performance. This behaviour can also be detrimental to a fund's stability which has had a period of comparatively bad performance. This behaviour in itself would amplify the effect and worsen the losses even further. Though this effect is less pronounced among passive mutual funds and they appear to be less susceptible to redemptions risks during economic hardships.

Exchange-traded funds are largely passive investment vehicles, and they do their best to reduce liquidity transformation. This means that most of them set limits on their redemptions and these exact policies can be found from their prospectus. A common way to do this is to limit the cash redemptions to a certain amount, and the exceeding part will be paid in-kind. In-kind redemptions mean that instead of the investors being paid back in cash, they are paid in a way which doesn't compromise the liquidity of the fund. This can be for example by paying in pro rata shares or in other more illiquid securities. According to the statistics, as much as 92% of ETF assets as of March 2018 used exclusively in-kind redemptions. Those which offered a combination of cash redemptions and in-kind redemptions, also reserved the right to transition to in-kind redemptions if it was deemed necessary due to liquidity concerns. (Anadu et al 2020.) A study by Agarwal et al. (2023) found that open-end mutual funds can also greatly benefit from placing limitations on cash redemptions, and by offering redemptions in-kind, reduce fund runs and improve stability especially for illiquid funds. Though their findings suggest that implementing such policy does aid in the event of a large net outflows, it seems to reduce the inflows into the fund during good times compared to the funds which do not have such limitations on redemptions.

As mentioned, the rise in certain specialized passive investment vehicles such as leveraged or "geared" ETFs and inverse ETFs have amplified the market volatility by their mechanisms. This is because both of these options trade in the same direction as the market moves. For example when the asset the leveraged ETF is tracking increases in value, the LETF must in turn buy more or increase the exposure through another investment vehicle to the underlying asset in order to stay on the leverage target. In turn, the inverse ETF creates the counter effect by having a short position in the underlying asset, and when the tracked asset falls in value, the value of their short position increases. This forces them to reduce their short position to stay on the target. (Anandu et al 2020.) Unlike regular market movement which is usually unpredictable, the movement caused by these leveraged and inverse investment mechanisms can be anticipated. This creates an opportunity for front-runners which attempt to capitalize on the fact that these funds are forced to trade in order to balance their portfolios. Eventhough LETFs are not a huge percentage of the whole market, studies done for the Federal Reserve have noted that the effect is quite significant. The effect has been said to be visible during the 2008–2009 financial crisis and 2011 European sovereign debt crisis. One way the effect has been isolated, is by tracking the price movement after 3:00pm if the stock has already moved 1% previously that day. Prior phenomenom was not observed in data before 2007 and the introduction of LEFTs, but only after then. This is due to how the funds can wait until the last hour before the market closes to perform their portfolio rebalancing. Trading right before the market is closing has the risk of moving the price disproportionately, but on the other hand it has the weird benefit where a possible large fluctuation in the price were to trigger a circuit breaker and prevent trading on the asset for the rest of the day letting the investors digest the changes. Though it is feared that such large movements before the market close could erode the investor confidence. (Tuzun 2013.)

As more and more funds are concentrating in passively managed funds, it has created a trend of these funds growing very large in terms of assets under management. A common way to track the concentration is Herfindahl–Hirschman Index which has been averaging 2700 for passive and 460 for active management since 2004. A value of 2500 is considered to indicate that there is a high concentration in large funds. (Anandu et al 2020.) This is natural in the sense that passive management benefits more from the economies of scale than active management. Passive funds are able to easily scale up their size with the same strategy and are able to improve cost efficiency, while actively managed funds might not be able to apply their main strategy past a certain size. This could even lead to a situation where increase in assets might drag down the averages as the marginal profit opportunities are not as good. Passive funds also are more similar to each other, which makes people less incentivized to browse around instead of just selecting the most popular one with the lowest expense ratio.

The concentration of the asset management industry can also pose unforeseen risks in which even factors unrelated to the actual market conditions could cause large disruptions. These could be scandals within these large funds, or for example cybersecurity breaches which could cause large redemptions or kill the confidence in a certain fund manager. Though for example in the 2003 mutual fund trading scandal, the outflows witnessed from these problematic funds seemed to flow into alternate funds on the market, negating the market impact overall. (Anandu et al 2020.) Ben-David et al. (2018) find empirical evidence that stocks that have larger ownership by ETFs show higher volatility than comparable stocks with more diversified ownership. They suggest that the stock return volatility is most likely not attributed to the better price discovery due to the ETF, rather that the increased demand is transferred to the underlying stock price without any regard of fundamentals. The trading activity by arbitrageurs that take advantage of the pricing differences between the ETF and its asset basket, also drive the volatility on these stock prices. Their findings show around 50 bp greater alpha for portfolios with high ETF ownership, suggesting that the investors expect a risk premium, at least in the short term, due to the nondiversifiable risk that the ownership concentration creates through increased volatility.

Comovement is the effect where stocks move together in a correlated way. The fact that a selection of stocks are grouped together in a basket such as S&P 500, has been shown to cause the stocks to move together. With ETF data from 2006 to 2013, Da and Shive (2018) find evidence that the arbitrage between the ETFs and their asset baskets is the cause of the comovement between the index members' returns. One aspect that affects the strength of the effect appears to be the turnover ratio of the ETF, meaning how much it buys and sells compared to its size. Further, their results show that the ETF daily returns are negatively autocorrelated and the turnover being one explanatory factor. The negative autocorrelation could be a sign of some kind of price reversals that happens due to ETF arbitrage which follows the excessive comovement. However the results and analysis by Chen et al. (2018) on S&P 500 index additions and deletions data from 1976 upto 2012 contrasts this by attributing the comovement to fundamentals. From their view, there is no excess comovement and that the changes are due to the changing betas of these stocks, which could be classified as the "winners". The literature largely agrees on the existence of comovement, though causality and explanations remain mixed.

One of the large active management funds that started competing against the great value proposition that passive investing provides is Fidelity International. In 2017 they unveiled their new variable management fee which is partially linked to their performance. Their new structure lowered the base annual management fee by 10 bp but with a chance of increasing it by up to 20 bp in case their performance exceeds the benchmark by 2% or more. (Financial Times 2017.) As the passive investment funds share of the total market has increased, it is expected that more active fund managers follow the suit to move more towards performance fee models (Belton 2017).

2.5 Effects of fund characteristics on performance

The question on whether the funds performance can be predicted based on known features and characteristics about the fund has been a widely discussed topic for a long time. In addition to the fund characteristics, the manager's impact on the fund's performance has been largely researched. Some of these characteristics are easier to measure quantitatively but some, such as skill, is a difficult variable to proxy for. For example, when interested in how big of a role the fund manager's skill plays, or whether it matters at all, measuring it is very difficult when the timespans are very long and multiple variables wihin the fund change during that period. A study by Golec (1996) attempted to solve this issue by selecting a shorter period to look at, in an attempt to isolate its impact. He found that funds with managers who are under 46 years old, have over 7 year long tenure within that fund and have MBAs, perform better than those which don't. His findings suggest that the most skilled fund managers were able to keep higher administrative fees while still performing as well if not better than others. However, in this paper the only way to indirectly measure the skill is through age and its size, as is suggested often in literature, the size of the fund equates to how much resources it has access to and through that, better management. Fund age in turn can be an indication of learning-on-the-job effect. As there are only very few studies done on performance characteristics of ETFs specifically, literature for mutual funds, index funds and alike will be reviewed and applied.

2.5.1 Expense ratio

One of the most significant characteristics to look for when selecting an ETF is the expense ratio, and it is almost without fail the most used variable accounted for in these studies. It is often tied to other characteristics of the fund, and some of its effect can also come indirectly through other variables such as size, as they can be interlinked with some spillover effect. This inverse relationship of expense ratio and size to good performance was observed by Elton et al. (2012) regarding mutual funds. They showed that top half of the funds in the sample, sorted by size, have 15% lower total expenses compared to the bottom half. They also find that the well performing funds tend to decrease their expense ratios instead of increasing them. This is not consistent with the theory which proposes that the management tends to increase their fees to account for their better performance. Malkiel and Saha (2020) state that there is strong empirical evidence to back up expense ratio being tied to the fund returns, and to be a factor when forecasting its future returns. This evidence goes back to Sharpe (1966) when it was observed that management fees played a significant role in the mutual fund's returns. Malkiel and Saha (2020) continue with their findings that in addition to expense ratio, also fund turnover and Sharpe ratios seem to provide some significant information. Consistent with prior literature they find expense ratio to have a negative impact on the net returns. They notice that the expense ratio's elasticity differs for the international equity funds, showing an increase of 12 bp in the expense ratio which they believe to be caused by higher research costs on international assets. When they selected the funds with the combination of cheapest expense ratio, lowest turnover and largest Sharpe ratio, it produced significantly better returns compared to the average fund or any other combination of the characteristics. They find that using these selection criteria did not cause a large increase in the volatility of the returns and that a combination of the characteristics seemed to have a multiplier effect on the returns. Eventhough this selection criteria produces significantly better returns in the short-term, they highlight that as the overperformance is not consistent year-over-year, an active trading strategy would be needed to always stay invested in the correct funds. This active trading strategy with the increased costs could in themselves erode the extra returns that were gained through it. Their study finds it more efficient to use a passive strategy of selecting the securities based on certain criteria and then just holding them. Despite of not overperforming consistently, the selected funds will still stay among the better picks due to their better characteristics such as low expense ratio.

The effects of expense ratio are well researched when it comes to mutual funds, but the relationship specifically among ETFs is less studied. However, the negative relationship with fund performance appears to be as common among ETFs as it is with regular mutual funds. For example a study by Paudel and Naka (2023) found the negative relationship to exist within all size quantiles in their study of index-tracking passive ETFs on the US market from 2009 to 2018.

Expense ratio in itself doesn't tell the whole picture, as it does not include the general expenses of the funds which can vary greatly and erode the fund's performance more than anticipated. This was inspected in Bogle's (2014) commentary on Sharpe's (1966, 1991) articles. He notes that there are extra costs that incur for actively managed equity but also for index funds, which are not accounted for just by looking at the expense ratio. For example a fund can perform better than is predicted by its expense ratio if it does very little trading, thus minimizing its brokerage fees and transaction costs. As actively managed funds tend to trade more, these costs will impact them more. The amount of trading done has also generally increased greatly, when measured by portfolio turnover. It is stated that the turnover within actively managed funds has increased fivefold in 50 years starting from 1960s. Index funds also have the advantage of always being fully invested, whereas active funds generally have to hold a small allocation of cash balance. These all-in costs are very rarely if at all researched as measuring them accurately is very difficult. In his article Bogle (2014) gives very conservative estimations on how much transaction costs, cash drag and sales loads have an effect. With as favorable estimate as he could come up with, he estimated for his analysis that actively managed funds suffer a cost of around 50 bp annually for their trading. He estimated that the small portion of around 5% in cash the active funds hold to have a negative impact of around 15 bp on returns. This effect called cash drag is something that index funds don't suffer from, neither do they suffer from the large trading costs. The third point he highlights which is not included in the expense ratio, nor is often accounted for, is the now less common practice of front-loading. This is a now less common practice of having a up-front cost to buy into a fund. The large shift into commission based models has somewhat negated this effect, but for active funds many distribution drags still survive including broker and adviser fees. The overall advantage Bogle (2014) estimates to be over 2% in favor of index funds, with the expense ratio being only half of the equation and with the other less transparent costs accounting for the rest. This is where he criticizes Sharpe's methodology of only taking the expense ratio into account, stating that the differences between the actively managed funds and low-cost index funds are larger than originally suggested. These are even before the possible tax inefficiencies between these two instruments are considered, hurting active funds even further. These inefficiencies include the realization of capital gains when trading which in some cases can even be at a higher rate when classified as short-term gains. One benefit that favors the ETFs in comparison to mutual funds is that when an investor sells shares of the fund, the other party is usually not the fund itself, rather another investor. This means that the fund doesn't need to sell its holdings and be a counterparty in this transaction, leaving it unaffected of capital gains with the asset is still being considered liquid. Overall he concludes the difference between an active fund and an index fund to be around 1.8% with conservative estimates. The also highlights the importance of accounting for other costs that incur within the fund, and not only the expense ratio which is the cost the investor is directly facing. When comparing how these characteristics and their effects are comparable between mutual funds and ETFs, a study by Elton et al. (2019a) on data from 1994 to 2016 found that expense ratios seemed to have larger negative impact on ETFs performance compared to index funds in general. They find that an increase of 1% in the expense ratio reduces the gross performance by 4 bp among index funds but as high as 27 bp among ETFs.

One factor that directly affects the size of the expense ratio, but is less known about, are the fees that the funds have to pay to the index providers. Paper on index providers by An et al. (2023) reveals that the sector is highly concentrated measured by the Herfindahl-Hirschman index to the few major index providers. These large players have the name recognition and market power, controlling ca. 95% of the ETF market together, these players being S&P Dow Jones, CRSP, FTSE, Russell, MSCI and NASDAQ. For their services of managing the index, they demand a licensing fee based on the fund's AUM plus often an additional flat fee. For example an investor into SPDR S&P 500 ETF (SPY) pays an expense ratio of 9 bp, of which 3 bp goes straight to the index provider. They also estimate that only around 40% of the licensing fees are required for the marginal costs of maintaining the index, with the rest being a markup for the index provider. As these licensing fees are a component of the expense ratios, the index provider's fee and its impact is visible directly on this variable. Separating the index providers impact isn't trivial as the disclosure of the licensing fees is voluntary, and it is only a small portion of the funds that do so.

Morningstar (2019) compared different index providers and came to the conclusion that indexes are very undifferentiated when they capture a similar market, with the small differences coming mainly from weightings. They also find that these indices with similar market focus provide very similar returns, which suggests that the index providers with more name recognition and with higher licensing fees do not necessarily provide any higher returns. This could allow funds to possibly lower their expense ratios and thus increase the investors' returns by switching to an comparable but more affordable index provider. An example of this was the switch done by Vanguard in 2012, which in turn created huge savings and allowed them to decrease their expense ratios across the board.

A higher expense ratio you find among active funds does not necessarily guarantee the existence of superior research and effort from the part of the management. This is because plenty of active funds have changed their strategy to closely match the indices which their performance is often compared against. This “closet indexing” or “index hugging” is largely frowned upon, as these active funds continue to keep their fees high despite of not providing the services that these fees are expected to pay for. For example, the Financial Conduct Authority in the UK considered any active funds that have a tracking error less than 1.5 as funds that take very modest positions and mirror the market closely. They found there to be around 109 billion pounds in funds that they consider to fall in this category. The regulator considers the existence and popularity of these funds, charging high fees without adding any value, as a failure of the market competition. They suspect that there is a perceived “going rate” which is accepted among active funds and their investors, that is not related to their actual performance or effort to add value. This rate does not either seem to be dictated by the market competition as spikes can be observed in the amount of funds with an expense ratio of exactly 1% and 0.75%. (O’Dell 2016.)

One theory on the reason for the shift from high fee active funds to low fee passive investment vehicles is the possible change in the market environment which consequently has made active stock picking ineffective. De Franco (2021) fails to find any structucal changes in the US equity markets to explain how the the stock picking would have become less profitable within the last 10 years. He takes a look at how well the stock picking managers would need to perform, and finds that a very high success rate would be required from them to actually beat the benchmark and passive competition. They suggest this to not be due to any structural change in the landscape, rather because the passive competition has become so efficient, managing to lower their expense ratios so low. The author theorizes that this high success rate required to beat the passive alternatives has driven more active funds towards closet indexing, which ends up being a losing outcome from any investor’s perspective. The issue of active funds participating in closet indexing was researched when the phenomenom was still young, and even back then the evidence showed that this strategy is not beneficial to the investor and was classified as bad behaviour (Taylor 2004). Both issues, the investors being less sensitive to expense ratios among active funds and that the strategies among these funds might closely follow that of the passive funds, are both a strain on the expense ratio to performance relationship among active funds. These prop up the expense ratio that the investors end up paying but without any of benefits that normally come with the higher fees that are supposed to aid the fund in trying to bridge the performance gap caused by its own making.

2.5.2 Fund size and family

When attempting to explain the differences in performance by characteristics, fund size is among one of the most studied factors for its explanatory power. There are variety of different ways to measure the size factor including but not limited to assets under management, number of stocks the fund holds, net asset value and number of funds belonging to the same fund family.

Starting as early as with data from 1988 to 1990, Golec (1996) found that the fund's asset size was a significant factor in explaining the yield, showing that in general the larger funds outperformed the smaller ones. He believed that at least some of this could be explained by the assets coefficient capturing part of an indirect effect of the expense ratio, as larger funds are able to provide less fees, and this in turn would increase yields. For mutual funds in general there are contrasting results on whether there are constant economies or even diseconomies of scale. For example Pástor et al. (2015) find that there appears to be decreasing returns to scale for active mutual funds at a fund level, though their results were not statistically significant. The variable used here for the fund size was assets under management with a slight modification to account for what stocks the fund is limited to buying due to its size. At an industry level, the evidence for decreasing returns to scale was stronger, and it was more pronounced for funds with higher turnover, volatility and among small-cap funds. They suggest that even when the skills among active managers have improved over time, it hasn't shown in the fund performance over its lifetime. They suppose this is to be caused by the the industry level decreasing returns to scale. This would also explain the overperformance by the new more skilled entrants. Zhu (2018) continues based on this earlier study implementing improved methodology, and finds statistically significant decreasing returns to scale at the fund level that Pástor et al. (2015) were not able to. They also run simulations in an attempt to evaluate whether AUM or logarithmic scale of it produces better fit in their models, concluding that both of them produce very similar results. Zhu (2018) finds fund size to be a statistically significant factor with a negative impact on the performance among active funds, with the effect being slightly less prominent with the loglinear model for size over linear one. When these funds are sorted into decile groups by size, the strength of the effect seems to be about three-fold going from bottom decile to the top decile, while still remaining consistently at decreasing returns to scale.

Crany and Crotty (2018) argued that index tracking funds behave similarly by their characteristics to funds that are actively managed, as from their viewpoint, the management skill still plays a role in these index funds. This notion is the basis of the study by Paudel and Naka (2023), where they point that if the performance of the ETF depends on managerial skill and they behave similarly to active funds, then it follows that there should be also be a relationship between the fund's performance and its size. They focused specifically on index-tracking passive equity ETFs in the US markets from 2009 to 2018, with the model using various control variables including expense ratio, age and number of holdings. The findings suggest that as is the case with active mutual funds, there exists diseconomies of scale. So unlike the previous study, they were able to show evidence that the returns of the ETF do not stay constant scale to the size. Rather they find that when the asset base increases, the returns decrease as a consequence. They suggest that when a fund has positive alpha, it is an indication that the fund does not have enough assets under its management to be at the optimal fund size. Vice versa, a fund showing negative alpha would be an indication that the fund has become too large. Consistent with majority of the studies, they also conclude that expense ratio is a major negative factor on performance, which they measure in respect to size.

The ETF being part of a larger fund family was found to be a significant contributor to funds differential return in the study by Elton et al. (2019a). This was measured as a log of the number of funds in that family, to serve as an indicator of the fund's sophistication and skill. It was believed that a family that controls a larger number of funds would have the ability and resources to find ways to generate additional profits or cut costs, and this seemed to hold true. Other variables they used included fund size in comparison to the sector size and ratio of holdings which indicates how closely the fund tracks the index. However, neither of these were statistically significant. It is interesting that in this case the fund size didn't seem to add significant information, instead, the amount of funds in the family did even when they are both very similar characteristics and an indication of the management's resources.

If we assume that index funds do not participate in any kind of securities selection or unnecessary trading, their transaction costs should be measurable by comparing their performance shortfall from the benchmark they are tracking minus operating costs with profits from possible securities lending added. This is the basis of the method used by Adams et al. (2022) in their study to determine how the transaction costs are affected by fund and family sizes. As the transaction costs are not reported by the fund's themselves, studies often attempt to estimate these costs by computing estimates based on active fund trading, but these are not necessarily totally comparable. One reason is the inability for index funds to avoid high trading cost stocks, including index addition and deletion price effects. The estimation method used in this study avoids many generalizations and measures only the observable effect that ends up being a direct cost to the investor by lowering the returns. From 1993 to 2016, they find the fund's operating cost efficiency improving in relation with its scale. When the index funds are sorted by size quintiles, they observe a 59 bp difference in transaction costs between the top and bottom quintiles. One theory to explain these efficiency gains, that come with larger size, includes the increased bargaining power to lower commissions and the increase in resources to invest in costly systems that improve trade executions at better prices. These efficiencies extend to family sizes, as the negotations are often done family wide and the resources available to the family are usually available to all of its members.

Cici et al. (2018) brings up the access to a trading desk and its efficiency as a significant advantage the larger fund families have. The efficiency of the trading desk virtually determines the gap between the actualized return, and the return of the fund strategy on paper if it was able to trade continually at observed prices without any transaction costs. Their method measures the differences between the trading desks between different fund families during index composition changes by comparing the index fund's gross returns to the underlying index's return. After calculating ranking of the fund families based on their trading desk performance, they are able to use these rankings to determine whether the more efficient trading desks outperform the bottom ranked ones among the active trading funds. They find that the actively managed funds belonging to fund families with most efficient trading desks as per their ranking, significantly outperform the ones who are ranked as the least efficient. The difference is also economically significant being more than 120 bp annually. A positive relation between the trading desk efficiency and the frequency that they tend to trade is also found, showing an increase of 14 percentage points in portfolio turnover. This positive fund family effect is also consistent with studies on passive ETFs such as Elton et al. (2019a). In that study they used the number of index funds that belong in the same family as a proxy for family size, and found the impact to be statistically significant and positive by differential returns. If the apparent benefit of the fund belonging to a larger family is due to the better efficiency of the trading desk, you could assume that the family fund effect would be stronger among active funds which tend to do more trading and thus would have larger benefit from more efficient trade executions.

2.5.3 Fund age

The age of the fund is often used as one of the variables to explain the fund performance, generally measured in either months or years from the inception of the fund. The results are mixed, with a combination of the fund's performance worsening through the its lifetime, the age not mattering, and the age having a learning effect having a positive impact on performance. Webster (2002) examined this by limiting their study to funds that have been around already for a long time, minimum of 20 years and up to 31 years. This is quite selective sample and is quite unique compared to other studies due its long time period. From here the data was split to three year intervals with the goal of measuring whether there was difference between these timespans as the fund aged. These returns are then adjusted to the varying market conditions by deducting the S&P 500 or NYSE composite index returns from them. They find that for mutual funds, raw and objective adjusted returns (comparing against other funds from the same category) age is not a significant factor, but the market adjusted returns against an index benchmark show a negative relation with the age.

Pástor et al. (2015) propose that one of the reasons for the negative correlation between returns and age at the fund level is that over time the industry grows and through that, more skilled competition enters the market hurting the existing participants. They also note that this would not be explained by incubation bias as they also test it by excluding the first three years and with their sample only including funds with the minimum size of $15 million, with no significant change in results. As they find that the fund's skill level stays constant over the fund's lifetime, the observed increase in management skill would not be due to existing funds becoming more skilled, rather that the new entrants are more skilled than the incumbents. They also find younger funds overperforming measured in benchmark adjusted returns, when comparing three to six years old funds to ones over ten years old, with the results being statistically significant. It is also noted that this relation doesn't only hold across funds, but also within the same fund where the performance deteoriates over the fund's lifetime. Though they suggest that this fund level effect would turn to near zero or even slightly positive if you were to account for the industry growth but being only marginally statistically significant. This result contrasts with a lot of prior literature, but that it could be explained by considering there to be skill improvement within the fund over time. Study by Zhu (2018) also finds that despite of the younger funds having the tendency to perform better, their results show that the younger funds often do not receive the required fund flows from investors to perform at their optimal asset allocation. Though this misallocation is rectified as they age, to the point at which they grow past that optimal size and start becoming more inefficient. He supposes that such misallocations in the market should sort themselves out and eventually disappear. Malkiel and Saha (2020) also examine fund age as an explanatory characteristic for actively managed funds' performance and observe slightly negative to no impact on returns over the fund's lifetime, though their results were not statistically significant. Fund age was also included as one of the characteristics to explain performance in a study by Filip (2018) on Polish equity markets. They found that within the domestic market, the age had a negative impact on risk-adjusted returns with statistical significance. Ferreira et al. (2013) also finds evidence of the fund age having a negative impact on its performance among non-US funds, but among US funds, they did not find any statistically significant impact. Based on this, younger funds outside of the USA appear to be more capable at finding profitable investment opportunities compared to US funds.

The evidence is mixed on the relationship between fund age and future fund flows among actively managed funds. A decrease in sensitivity to past performance in relation with fund age is found by Brown and Wu (2016) indicating that investors might be more eager to put their money into funds that are older, with name recognition even if they do not perform the best. If the fund age has a positive effect on the fund flows, this would directly increase the fund's assets under management, suggesting that these two variables might have some correlation. When an investor is choosing an active fund to invest in, they have to rely more on their previous track record compared to one investing in passive funds which are very undifferentiated.

The literature and research regarding performance characteristics of ETFs specifically in comparison to studies done on mutual funds is more sparse, but a study by Paudel and Naka (2023) used the fund age as a factor in their model on US market data from 2009 to 2018. The paper was mainly focused on the fund size, but they also found the coefficient for fund age to be statistically significant and positive with a differential impact for the size-sorted quartiles.

2.5.4 Other common explanatory characteristics

One factor that has been found to have an impact on the performance is the liquidity of an ETF. This has been shown for example by Paudel and Naka (2023) in which they demonstrate how illiquid ETFs perform worse. Here it was measured as the normalized bid-ask spread and in general ETFs are viewed as quite liquid as they encounter plenty of intraday trading and contain very little information asymmetry compared to traditional open or closed-end funds. One way that low liquidity can propagate to the ETF is from the underlying securities, or through the creation and redemption mechanisms.

An interesting way for funds to increase their returns against their peers is the act of securities lending. This is a practice where the fund temporarily lends out its securities to other parties for a fee, generating additional income but not without some risks. These include counterparty risk if the given collateral for the borrowed securities doesn’t cover them, and possible losses on cash collaterals which were reinvested. Though this practice can be seen as controversial as the borrower usually is a short-seller and lending to them can cause downwards price pressure on their own holdings. Another concern is that during the loan, the asset’s ownership is transferred to the borrower, giving them voting rights and ability to have influence on the companies’ annual meetings. ETFs generally are conservative with how much they lend out their securities, and iShares reported their ratio to be 14% for their Europe domiciled ETFs for the year ending in 2023 (Blackrock 2024). In a study by Elton et al. (2019a) the securities lending ratio showed a statistically significant positive impact on performance.

One problem with researching the prevalence or impact of this practice, and why there are less studies about this, is that funds generally are not very transparent about their securities lending programs. This leads to investors often being inadvertently exposed to extra risks which they might not have been aware of. In the worst case the fund could lose the shares lent out to a hedge fund if they fall into bankruptcy, as was the case with Lehman Brothers in the 2008 financial crisis. In such cases if the collaterals are not enough to cover the losses, the lender may take a huge hit as happened in the 2008 crisis. Following these events the prevalency of securities lending fell by about 40%. Regarding ethical considerations of securities lending, most funds do not return these extra returns to the investors, which might contribute partly to the opaqueness of the information. (Dunham et al. 2016.) This means that the more securities lending the ETF does, the more it is able to gain additional profits. Whether it would show up as excess returns would depend on whether the they actually share these profits directly with the investors by lowering the expense ratio or not.

Whereas index funds do not generally lose opportunities due to their securities lending, this might not be the case for active funds as they lose their opportunity to actually sell their shares if necessary. This is in addition to also being the counter-party to the short-sellers who are selling their holding short and betting against them. Evans et al. (2017) find that actively managed mutual funds which participate in securities lending underperform compared to the ones which do not. The difference between the two is significant and between 0.5% and 0.7% annually in risk-adjusted net return. They theorize that the active funds tend to hold onto the their positions even when they have strong short-sell demand. Despite of them making some profits from lending them out, these gains are not enough to cover the losses caused by the holdings' downwards trajectory. It is interesting how these active funds do not register or process the signal of high short-sell demand as an indication to sell their positions. Rather they end up lending these assets for a small fee while possibly taking the hit with the decreasing asset prices. The authors suggest that the short-sellers in these situations are more skilled and understand the stock's market prospects better than the fund managers. If you were to generalize their results, it does not seem beneficial for active funds to lend out their securities as it restricts their options on when to sell, whereas these restrictions are not something that hurt passive funds who will hold their assets regardless. Among index funds, they did not find any negative effect caused by securities lending.

The life of a fund might come to an end by liquidation but this does not have to be due to its poor performance. Sherrill and Stark (2018) study what are the common characteristics indicating fund's likely failure. Unlike actively managed funds which have a strong dependence on producing excess returns compared to the market, passive ETFs do not rely on their performance. They find that the very early days of the fund are very critical on whether it'll survive, and that new active funds which focus on a hot topic are more likely to liquidate compared to ETFs with a similar objective. The size of the fund seems to be primary driver on whether a fund liquidates or not, though being part of a larger fund family seems to improve its chances of survival. According to their model, funds having large inflows and larger expense ratios were way less likely to close, by 61.7% and 67.4% respectively. Higher expense ratio creates more revenue to the fund family and appears to make the funds more stable. Fund liquidations are problematic from the perspective of the investors as generally the timing of the capital gains can be controlled, but in the event of a liquidation the tax consequences might be unexpected.

Trading costs are large cost that active funds face, which largely does not affect passive funds apart from portfolio rebalancing. The issue with trading costs in general is that they are very difficult to measure from the outside, and funds do not necessarily volunteer all of the information. One way analysts have managed to incorporate the transaction costs into their models is to proxy it with fund turnover ratio. This measures essentially how many times the fund rotates its portfolio within a given period, and it is calculated by dividing the smaller of total purchases or sales within typically a year, with the average holdings of the fund over the past 12 months. The higher this turnover value is, the more the fund has done trading. This makes it a good proxy of trading activity, and can be used in models as a possible differentiator between the performance of active and passive funds. Despite of being popular measure of transaction costs, it has been criticized of measuring the frequency of trading but not the total trading costs which include variety of brokerage fees and average spreads (Adams et al. 2022).

Generally excess trading and high fund turnover is associated with the erosion of fund performance due to trading costs. On the other hand it can also be viewed from the perspective where a fund has a very successful trading strategy. In these cases, the higher turnover could mean that the fund had abundance of opportunities to take advantage of within their strategy and performed well. Optimally the fund tries to reach a point where trading is only done when the marginal benefit of it exceeds the marginal costs from conducting it. The size of the tax burden is largely tied on what types of assets the fund is trading, with the burden being larger for funds focusing on small-cap stocks and value portfolios. For example low-turnover funds that largely hold large capitalization stocks have to liquidate their holdings very rarely, and when they do, these are the not very well performing ones. These types of funds would be quite tax efficient, making the portfolio turnover ratio a good proxy for tax consequences also. In a study on mutual funds between 1990 and 2016, the funds that marketed themselves as tax-efficient didn't perform any worse before-taxes compared to the competition. This means that the constraints and limitations put on by tax efficient strategy do not seem to have a negative impact on performance. The study finds that funds that create larger tax burdens on their investors, are not able to offset these costs through superior pre-tax performance. This shows that tax efficient funds and strategies outperform the tax inefficient ones. (Sialm & Zhang 2020.)

As mentioned, the higher fund turnover can indicate larger tax consequences. This is due to how when actively trading, the fund makes short-term capital gains, and may have to pay capital gains tax on these profits. This is in contrast to a fund which for example has very low turnover but makes its returns from the appreciation of the held assets, which is often the case for passively managed funds. A well managed index fund only has to add or remove assets based on changes in the underlying index, which is not often in the case of S&P 500 in which it is done quarterly.

Pástor et al. (2017) found that the funds actually performed better the more they traded when studying active US equity mutual funds over 30 years from 1979 to 2011. They suggest that funds are able to determine when it is beneficial to trade, and find that an increase of one standard deviation in portfolio turnover ratio was associated with an annual increase of 0.66% in performance. Their analysis also points out that this effect seems to be stronger for less liquid stocks, indicating the presence of more profit opportunities. This includes funds holding large amounts of small-cap stocks, which show less competition and higher bid-ask spreads making them more profitable. They also find that active funds trade more when the investor sentiment or volatility is high, and when liquidity is low. All of these are something that can produce mispricings. Lower liquidity also creates an upwards pressure on transaction costs due to increased spreads, which could discourage trading. This means that the fund has to take the increased transaction costs into account when deciding whether the trading is beneficial. Their research and evidence suggests that the funds still tend to trade more during illiquidity despite of the downsides that come with it. Despite of active funds historically not being able to outperform their passive competition, the authors highlight active funds' societal benefit of improving market efficiency by correcting market mismatches.

Similar positive relation between performance and portfolio turnover ratio was also observed by Vidal et al. (2015) when studying the effect of expense ratio on performance among mutual funds from a CRSP survivor-bias-free sample while using turnover ratio as one of the control variables. Their sample being limited to actively managed equity funds and being corrected for survivor-bias suggests that their findings is not due to the better performing funds being over represented or bad performing funds being under represented in the sample, preventing bias in the results of funds which do more transactions. Wermers (2000) notices that these funds with higher portfolio turnover exhibit substantially higher transaction costs and charge higher expenses. They find that the funds with high portfolio turnover tend to hold stocks with higher average returns compared to the low-turnover funds, suggesting that these funds' managers have superior stock-picking skills. The differences between the highest and lowest deciles within the sample are huge, from highest ones sporting an average total net assets-turnover of 155% when the lowest decile has an average of 14% per year. This shows that even among active management, there are varying strategies and that the difference between the ones trading the most and least can be ten-fold. Even if some investors might consider a high turnover ratio as an indicator of skill, a fund with higher turnover does not necessarily indicate that as this frequent trading can also be done by any manager, even when it is not profitable to do so due to the marginal costs. They find that despite of the characteristic-adjusted net returns showing negative and insignificant results, an average high turnover fund from the sample still significantly beats their comparison benchmark index ignoring any tax consequences.

It is likely that there are differences on how portfolio turnover affects the two fund management styles differently, as generally passive funds do not do trading in order to make profits, rather to comply with their indexing strategies. In these cases, most trading done ends up having an negative impact on the performance due to fees. For example Elton et al. (2019a) in their study including both passive mutual funds and ETFs found that portfolio turnover has a negative impact on differential returns with the result being highly significant. In another study by Elton et al. (2019b) they took a look at whether selecting active funds based on certain criteria can predict whether it will outperform passive ETFs. One of these criteria was low past portfolio turnover, and they didn't find the active funds based on this criteria to produce any better returns than their passive ETFs.

3 DATA AND RESEARCH METHODS

3.1 Data

For the initial sample, a selection of 100 most popular exhange-traded funds is selected for both passive and active management style categories based on their assets under management on the US markets. The restriction to US markets simplifies the model as we do not need to account for different market environments and comparison benchmarks. This also resolves the issue of duplicates i.e. funds which are essentially the same but in different currencies, eliminating them from the sample. Discerning factors between the funds such as assets under management, management style, lifetime of the fund and expense ratio will be collected. This information will be collected from publicly available sources, preferably from the original source being the fund themselves. If not available, other source such as Morningstar, Yahoo Finance or similar service will be used.

The required data for these funds' returns will be acquired from their total return indices, which are retrieved from LSEG Datastream. The total return index allows us to track all cash distributions such as dividends, interest and splits alongside the regular capital gains. This gives more accurate representation of the investor's actual returns compared to just measuring price returns. For the model a benchmark to compare the returns against is required, for which we will be using S&P 500. It is seen as the most prestigious and most tracked passive investment asset by market capitalization, and is also widely used as the benchmark of the US markets.

The service from which the list of active and passive funds is retrieved from is ETF Database, which is a website founded in 2009. Following their growth over the years they have become the largest independent ETF database with an aim to provide a truthful, unbiased and authorative source of information for various constituents such as financial advisors and individual investors (ETFdb 2021). Only the list of the ETFs were retrieved from ETF Database based on their market capitalization, and further data is collected from other sources. The list of the ETFs included in the sample are presented in Appendix 1 and Appendix 2 alongside with their expense ratios, fund inception dates and assets under management. "Ark ETF Trust - ARK Space Exploration & Innovation ETF" (ARKX) was available on the ETF Database list, but accurate data could not be found from Yahoo Finance or other sources, causing it to be excluded from the sample. "RiverFront Dynamic US Flex-Cap ETF" (RFFC) was added from the tail end of the list to replace it.

The additional information for the stocks such as expense ratio, inception date are retrieved from Yahoo Finance and then this data is processed into usable form. The list of active and passive ETFs and their information were retrieved for the date of 17th of April 2021.

The time series data for assets under management have been obtained from a subscription based service called YCharts. The service is founded in 2009 and claims to provide data of over 45,000 mutual funds, ETFs and CEFs (YCharts 2021). They provide monthly data for all of our ETFs for the past 5 years including assets under management. One limitation that has not been taken into account is the possible changes in the expense ratio within our time window. These changes are rare and there does not exist many services that provide data on these changes over time. However, these changes are unlikely to have a large impact on our results as these are quite rare and the possible changes will not be large. That means that the expense ratio used in this study for each of the ETFs is the one in place as of September 2021.

3.2 Methodology

The goal of this quantitative study is to research whether there exists characteristics for actively and passively managed exchange-traded funds which drive their performance, and whether there is a performance gap between active and passively managed ETFs. This topic has way more research among mutual funds, with the ETFs being a newer phenomenom and warranting further studies. The models and variables used for this study are in accordance to previous literature and what have been throughly considered. The main model of this study is the regression analysis which is ran on active, passive and also on combined data in an attempt to isolate whether there is significant factor that could attribute a difference between the two through a dummy variable. The variables used for the regression analysis are familiar from previous studies, being assets under management as a proxy for fund size, the age of the fund since its inception, and its expense ratio. Different options exist when choosing proxies for these different characteristics such as size, and it is worth considering which one to use, such as AUM versus total net assets. As previously mentioned, a dummy variable is selected such as that it captures the effect of the active management style. The dependent variable chosen is the abnormal returns against the benchmark index of S&P 500, as the funds in question all operate in the US markets. In some cases abnormal returns can also be referred to as market-adjusted returns. The goal is to compare the ETFs against the most followed index, and see whether any of these characteristics can produce returns that beat or lose to the index consistently. If this was the case, these characteristics would be something an investor should be on the look out for when making informed investment decisions.

4 EMPIRICAL RESULTS

The dependent and independent variables used in the main regression models are shown in Table 1. The choice of the independent variables and characteristics which are selected to explain the fund's abnormal returns is done by researching previous literature. Fund size and age are widely used characteristics in previous studies. These are easily accessible to any investor researching the funds in which to invest in, and as these have shown to be an indication of the fund's future performance. The usage of fund size can be proxied by variables such as assets under management, and is used for example in a study by Sheng et al. (2023) on how fund characteristics can be an predictor of the alpha it generates. In their study a natural logarithm of the fund size is used as an independent variable in their regression model instead of the absolute value. The reason for the choice of using the natural logarithm is not mentioned, but generally taking the logarithm of an independent variable can reduce the variance of the residuals thus improving the homoscedasticity of the data. When using linear regression models such as ordinary least squares (OLS), the heteroscedasticity in the data can cause inaccuracies in the coefficients standard errors. Khorana et al. (2009) explain their decision of using the logarithm of the fund size in their model as they expect the marginal effect of the size to diminish as the fund grows. Another reason to use logarithmic transformations of variables is the inherent property in natural logarithms which allows rough estimation of the

Table 1: Regression model variable definitions

Definition and explanation of the dependent and independent variables used in the regression models. Where applicable, the origin and instances of use in previous literature of a specific variable is mentioned.

Variable	Abbreviation	Explanation	Previous literature
Abnormal returns	AR	Returns that exceed the market, usually an index such as S&P 500 used for it, scaled by 2 orders of magnitude	Petajisto 2011
Assets under management	AUM	Measure of fund's scale, natural logarithm of the fund's total assets under management in millions	Sheng et al. 2023; Khorana et al. 2009; Zhu 2018; Pástor et al. 2015
Expense ratio	FEE	Expense ratio of the fund scaled by 2 orders of magnitude	Sheng et al. 2023; Elton et al. (2019a); Pástor et al. 2015; Paudel & Naka 2023
Fund age	AGE	The natural logarithm of the fund's age in days from inception	Sheng et al. 2023; Pástor et al. 2015; Paudel & Naka 2023
Bond (dummy)	BOND	Dummy variable of 1 or 0, depending if the fund is focused in bonds or not	
Active (dummy)	ACTIVE	Dummy variable of 1 or 0, depending if the fund is actively managed or not	

percentual change in the dependent based on the independent variables coefficient. This is due to the fact that for small numbers, the root of natural logarithm *e* to the power of a number α is roughly equivalent of 1+α. (Gelman & Hill 2007, 60.) The variables and their usage is further explained in Section 4.3.

4.1 Descriptive statistics

The initial samples from which the fund selection began consisted from a total of 200 funds of which 100 were passively managed and the other 100 were actively managed funds. These funds were chosen based on their market capitalization to capture the most popular options an investor were to choose from. From this, the selection had to be reduced to the funds which had available data for the required time period of minimum of 3 years. This left us with 39 active funds and 100 passive funds, of which 82 are equity and the rest bond funds. The descriptive statistics for the filtered samples of active and passive equity funds are listed in Table 2. The total assets under management for all of the active fund candidates was 66.8 billion dollars in AUM with the average

Table 2. Sample statistics

The initial sample size for both the active and passive categories is 100 which is then filtered based on criteria ending up with 39 active ETFs and 82 passive ETFs. The list of funds are retrieved from ETF Database, assets under management, expense ratio and fund inception dates are retrieved from Yahoo Finance and the timeseries data is collected from LSEG Datastream.

Summary statistics

	Active						
Variable	Mean	SD	Q1	Median	Q3	Min	Max
Excess return (%)	-0.238	0.0520	-2.425	-0.343	1.508	-42.25	52.78
Expense ratio (%)	0.645	0.00337	0.520	0.650	0.790	0.130	2.010
Size (AUM in $B)	0.539	2.102	0.0739	0.123	0.208	0.0032	27.89
Age of fund (days)	1508.36	934.24	857.5	1285	1976.5	43	5291
No. of obs.	1404						
No. of funds	39						

	Passive						
Variable	Mean	SD	Q1	Median	Q3	Min	Max
Excess return (%)	-0.359	0.0346	-1.922	-0.362	1.152	-21.77	33.87
Expense ratio (%)	0.158	0.00137	0.060	0.120	0.190	0.030	0.680
Size (AUM in $B)	33.514	45.756	12.075	18.585	35.315	0.0991	388.85
Age of fund (days)	5366.63	1997.00	3911	5555.5	6954	93	10436
No. of obs.	2952						
No. of funds	82						

fund being around 668 million dollars and median being around 156 million dollars. Within our selected sample, the average was around 539 million and median around 123 million dollars. The active ETF market is significantly smaller than the equivalent passive one by looking at the top 100 candidates from both groups. The passive candidates totaled almost up to 4.6 trillion dollars in AUM, with the average of 45 billion and median of 27 billion dollars. Within our passive equity sample, the average was around 33.5 billion with a median of 18.6 billion dollars. Even this small comparison gives us a good indication that the ETF market is dominated by passive investment vehicles with the difference between the two groups of 100 most popular options being almost 70-fold. Part of this can be attributed to the increased popularity of passive investing, but partly it could also be that active management generally attempts to capitalize on market inefficiencies, and the size of a fund makes these opportunities more difficult to take advantage of. A large active fund might for example have problems capitalizing on a mispricing in an asset that has low liquidity or low market capitalization, as investing in it might cause large pressure in its demand and price.

One of the most important factors in our model is the assets under management, which varies over time within the sample unlike some of the factors such as the fee ratio.

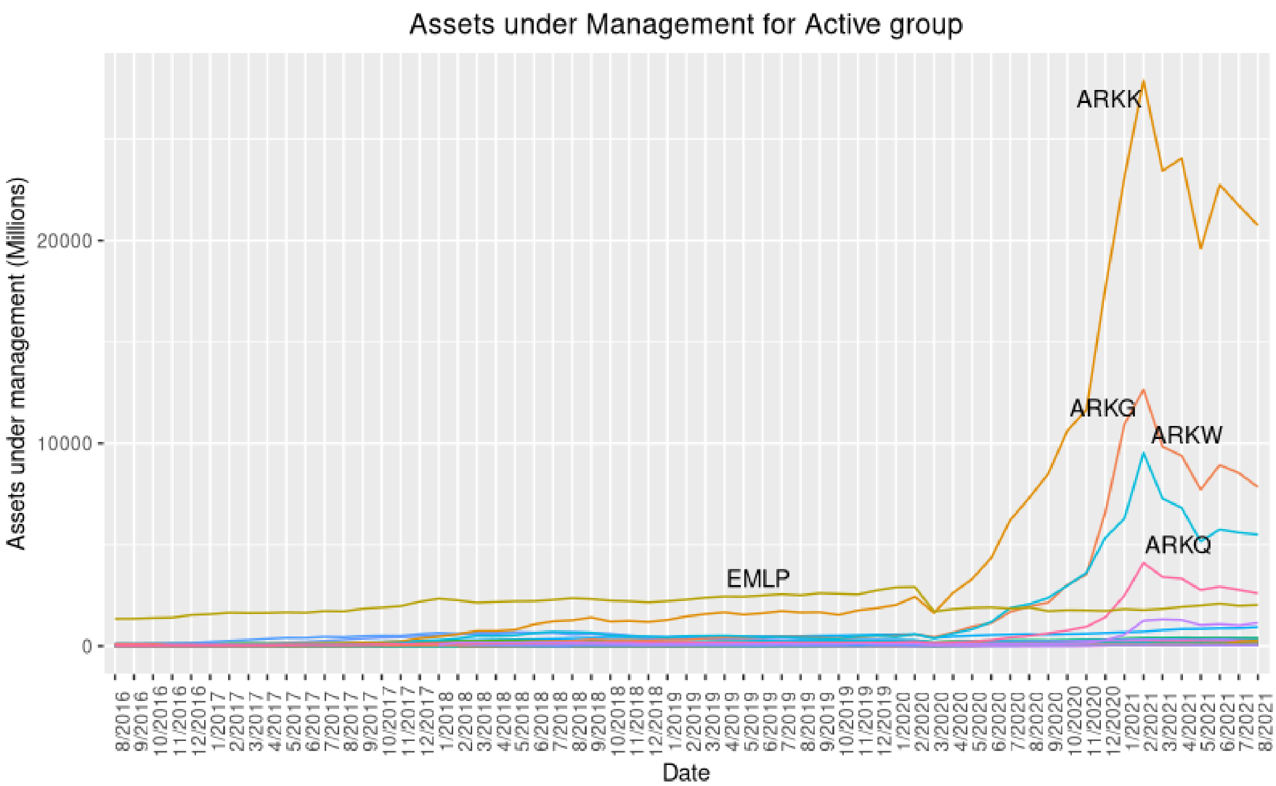

Figure 1: Assets under management for Active ETFs.

Assets under management of the active group plotted on a line graph from 8/2016 to 8/2021.

Out of the 39 active ETFs in our sample, four are owned by the same fund family and gained large inflows in early 2020. These four ETFs in our active group are part of the asset management firm

ARK Invest which was founded in 2014 by Catherine Wood. It markets itself as differing from the other available ETFs based on focusing on disruptive innovation and new technologies. (ARK Invest About 2021). ARK Invest has become very popular among investors and has risen to the ranks of the largest active ETFs and in February 2021 it peaked with over 50 billion USD in assets under management. Figure 1 shows that the ARK ETFs are way larger than any of the rivaling alternatives among active ETFs. Up until the end of the sample ARKK stayed as the largest of the group with around 20.75 billion USD in assets under management. Its focus point is on companies that use innovation for growth and some of its large increase in value can be attributed to Tesla's bullrun. Tesla was still the largest single holding of ARKK with a weight of 10.45% as of September 16th 2021. (ARK Invest Holdings 2021.) ARKG in return focuses on health care innovation and therapeutics, ARKW in artificial intelligence and blockchains, whereas ARKQ on robotics, automation and space exploration. The rest of the ETFs in our active group sample are smaller compared to the behemoths of ARK Invest, which could indicate that the active funds tend to stay smaller and more agile for their investing strategies.

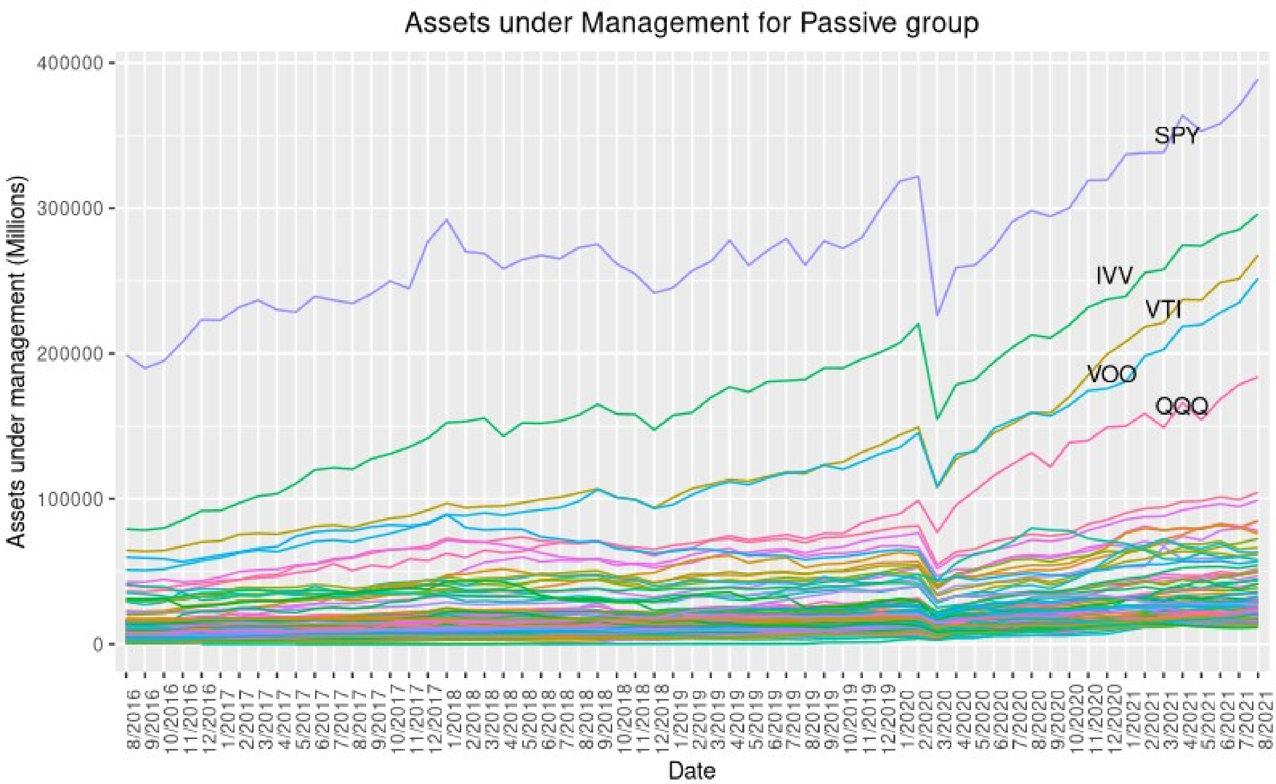

Figure 2: Assets under management for Passive ETFs

Assets under management for the passive group plotted on a line graph from the time range of 8/2016 to 8/2021.

Out of the 82 passive ETFs in the sample, the group has five clear ETFs that separate from the crowd by size, them being SPY, IVV, VTI, VOO and QQQ. SPY is the most popular passive ETF and along with VOO and IVV they aim to track the S&P 500 index with as little tracking error as possible. Following that, QQQ tracks the NASDAQ-100 index and VTI tracks the CRSP US Total Market Index. By visually observing the line graph, the market crash of February 2020 caused by the initial market reaction to corona virus is clearly visible as a significant drop in the assets under management across the board in the passive ETFs. By September 2021 the five largest passive ETFs in combination held almost 1.4 trillion USD in assets under management, of which the largest was SPY with 389 billion USD. Compared to the inflow and outflow behaviour of the active ETFs, the overall market sentiment seems to be more visible in the AUM graph for passive ETFs. This could indicate that it is less likely that investors move their capital out of the market or into other funds. This is something that was discussed as a possible pitfall in the chapter about the rise of passive investing, where the passive investor's behaviour creates a dampening effect. The fund flows into active ETFs seem more volatile and driven by trends or hype.

4.2 Sample normality and two-sample t-test

To gain a better understanding of our data, we can run statistical tests to determine from what kind of distribution our returns and abnormal returns originate from. As we are interested in the difference in performance between the active and passive ETFs, we can run some preliminary tests to see whether the differences are statistically significant as we will be using the management style as one of the dummy variables in the regression model later. A two-sample t-test, also known as independent or unpaired t-test, comes in two different forms with slight differences in the underlying assumptions of the samples and their characteristics. These two different variations are the standard Student's t-test and the Welch's t-test. Both of these tests have following null hypotheses:

$$H_0 : \mu_1 - \mu_2 = 0 \text{ (or } H_0 : \mu_1 = \mu_2\text{)}$$
$$H_a : \mu_1 - \mu_2 \neq 0 \text{ (or } H_a : \mu_1 \neq \mu_2\text{)}$$

In this case the null hypothesis H_0 says that the means μ_1 and μ_2 between the two groups are equal, which can also be expressed as their difference being zero. If we cannot accept the null hypothesis, meaning that the difference between the means is large enough to be statistically significant, we must reject the null hypothesis and revert to the alternative hypothesis of the means between the two groups being inequal.

There are prerequisites to these tests, such as that the groups must be independent of each other, the data points are randomly sampled from the population and for the standard Student's t-test it is assumed that the variances between the two groups are equal. However the strict requirement of variances being equal can be alleviated by choosing the Welch's t-test instead. However these two variations of the t-test produce very similar results as long as the group sizes and standard deviations don't have large differences. As the amount of observations differs quite a lot between the two groups as the sample contains way fewer active funds than passive funds, both of these tests will be used. The t-tests also assume that the data points are normally distributed.

For the following part we will use the sample spanning three years from 8/2018 to 8/2021. The selection of this timerange is to ensure that our selected ETFs have enough price data to provide enough datapoints. This does not become a problem in the case of the passive ETFs as most of them are older and more established having data dating back to 2000's. In the case of the active ETFs, a lot of them have been created in the mid 2010's and quite a bit even past that. This creates a problem that our sample of active ETFs becomes very small if we choose a longer time range. When increasing our time span, the size of the sample is reduced significantly. If we try restricting our sample to have data for one year, we end up having 96 active ETFs, 57 left at 2 years, 39 left at 3 years, 27 left at 4 years and 17 left at five years. With this in consideration, we need to make a decision regarding whether we value larger time span with fewer active ETFs versus larger amount of funds with a shorter time range of data and less datapoints.

To increase the accuracy of our tests and models, a dummy variable was added that specified whether the fund invests specifically in bonds. In our sample the only funds that invest only in bonds and are classified as such, are passively managed. None were found in the actively managed group. Generally bonds have significantly lower returns compared to stock and equity, mostly due to lower risk of losing on your investment. If we average the returns of the passively managed funds that invest in bonds we observe a return of around 0.29% monthly, and among the ones which invest in equities the corresponding number is 1.20% monthly. As apparent, the returns between these two classes of investments are very different, with the difference being 0.90% monthly and 11.77% annualized. Despite of this time period having unusually large annual returns, the discrepancy still stands. Having both of these asset classes in our sample mixed would skew the results greatly as these funds with bonds are only in passively managed funds category. If included, it would drag down the average returns of the passively managed funds compared to the active group. For this reason bond funds are excluded from this test. A total of 18 bond funds were removed in this exclusion from the data which is used for the tests and the combined models.

Thus the final sample for the t-test is gathered from the funds that fill the requirement of having price history for the last 3 years and that they are not investing in bonds. The total sample consists of 39 actively and 82 passively managed ETFs. As previously mentioned, the t-test requires the returns to be normally distributed to gain correct results. Even thought this normality is taken as given, there are ways we can evaluate the observed distribution. One of these ways is visual one where a histogram or a density graph is created and compared to what a normally distributed returns should look like. In Figure 3 these abnormal returns are plotted for both, the active and the passive groups. The density graph on the first row represents the distribution of the abnormal returns, where the frequencies are plotted on the Y-axis and the abnormal return percentage is plotted on the X-axis. The density graph is trimmed when datapoints are missing beyond a certain point. On the second row, in the quantile-quantile plot the abnormal returns are plotted on the Y-axis and the normal theoretical quantiles are plotted on the X-axis. Quantile-quantile is used to compare two probability distributions by plotting their quantiles against each other. The comparison is often done against a normal distribution which allows us to plot a line of $y=x$. When the plotted quantiles fall on this line, it indicates that the two distributions, in this case the abnormal returns and the normal distribution have similar distributions. A grey background is drawn around the plotted line which represents the confidence interval of 95%.

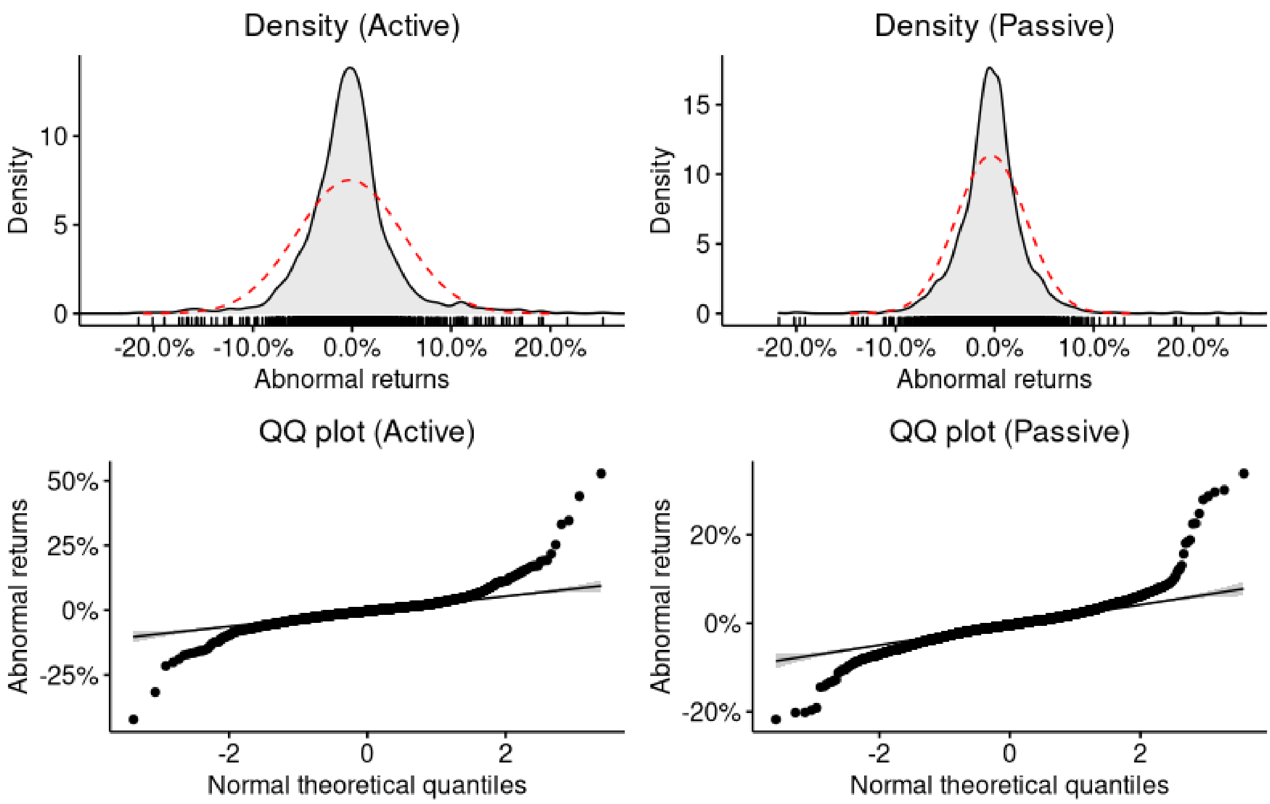

Figure 3. Density and Quantile-Quantile plots of monthly abnormal returns for both Active and Passive ETF groups from 8/2018–8/2021.

The density graph represents the frequency of monthly abnormal returns, with ticks on the X-axis representing a single observation. Normal distribution is overlayed on the density graph to provide a base to compare to. Quantile-quantile plots the quantiles of the monthly abnormal returns in the sample against the normal theoretical quantiles. The grey background on the plot marks the 95% significance level for normal distribution.

The abnormal returns for both active and passive are graphed into a density plot to view the spread of our data points. In this Figure 3 the points are represented as a continous black line with grey background. Density plot is a variation of a histogram and in it the continuous data points are smoothed via a method such as kernel smoothing. This creates a better representation of the distribution than a histogram as it gets rid of the noise and does not require us to choose an amount of bins as we would have to do for a histogram. The plot shows that there is a distinct shape in the distribution with higher density in the middle and the density decreasing the further away from the middle point you travel, approaching zero. A normal distribution is overlayed as a red dashed line on the graph which allows us to compare the observed distribution to something that would be expected if the abnormal returns were normally distributed as this is a prerequisite for most statistical tests. The center point of the distribution in both active and passive groups appears to be taller than the normal distribution. It could be said that the observed distribution is a leptokurtic distribution with a positive kurtosis. This means that the peak is higher and the tails are heavier than in a normal distribution. Even with just a visual observation, it seems clear that our sample's abnormal returns don't follow the normal distribution. However, it is also possible to come to this conclusion by using statistical methods to find whether we can say this at a certain confidence level.

The quantile-quantile graph plots the quantiles of the abnormal returns against the quantiles of a normal distribution. We can see that the data points around the center quantiles are fitting with the normal distribution within the 95% confidence level, but the further towards the tails the data points start deviating from the normal distribution significantly. This would indicate that by visual observation the tails would seem heavier than expected from something that is normally distributed.

It is generally assumed that returns are normally distributed allowing calculations and models to be used as this assumption is very critical for forecasting. There are several tests we can perform to assess whether our results could come from a sample that is normally distributed, one of which is a Jarque-Bera Test which is a goodness-of-fit test. This tells us whether the distribution's kurtosis and skeweness matches the one of a normal distribution. Another test that can be used to determine whether the sample comes from a normally distributed population is Shapiro-Wilk test. Both of these tests are ran for both of our samples and results shown in Table 3.

Table 3: Normality tests for active and passive groups' abnormal returns from 8/2018–8/2021.

The table includes both Jarque-Bera test and Shapiro-Wilk's normality test results which were ran on the abnormal returns in the sample for both active and passive groups.

Jarque-Bera Test	**Active**	**Passive**
X-squared	22289	29293
p-value	< 2.2e-16	< 2.2e-16
Shapiro-Wilk Normality Test		
W	0.82124	0.87485
p-value	< 2.2e-16	< 2.2e-16

Both Jarque-Bera Test and Shapiro-Wilk Normality Test have a null hypothesis that the picked sample originates from a normal distribution, and p-value represents risk that rejecting the null hypothesis would be incorrect. With a traditional confidence level of 95%, the p-value of under 0.05 would mean that we can reject the null hypothesis and assume that the sample is not from a normal distribution, and this is the case for both tests. This is the same conclusion that we could have come to when inspecting the distributions visually with the density and quantile-quantile plots. When using Shapiro-Wilk test and interpreting its results, it is important to account for their limitations such as its tendency to report significant deviations from normal distributions even from small differences when the sample size is very large, as it is in our case (Field 2009, 144). Using multiple different tests and plotting the observations visually in density and quantile-quantile plots can help us in combatting the weaknesseses that one test might have and aid in gaining a more complete picture.

Despite the usual prerequisite for running t-tests on data is the assumption of the difference between the two means being normally distributed, with a larger sample the impact of this onto our test results becomes smaller and smaller. With a large sample, the difference of means between these two distributions is said to approach normal distribution. The distinction of what is considered large enough is not clearly defined but the larger skeweness the distribution has, the larger sample should be used. (Bruce 2014, 159.) In our case through observation it is visible that the distribution is quite symmetrical and not very skewed which would improve the statistical explainability of the t-test with our data. Next we can perform two-sample t-test in order to determine whether the two samples, the returns of the active and the passive ETFs, could originate from the same population.

Table 4: Results of the two-sample t-test on active and passive ETF's abnormal returns.

The first sample consists of 39 active ETFs abnormal returns which give us a total of 1404 datapoints and the comparison sample consists of 82 passive ETFs abnormal returns with a total of 2952 datapoints. Both of these samples are from the same three year period of 8/2018 to 8/2021.

				95% Confidence Interval of the difference	
Two-Sample t-Test	**t**	**df**	**p-value**	**Lower**	**Upper**
Monthly data aggregated					
Equal variances assumed	0.2977	70	0.7668	-0.006873	0.009285
Equal variances not assumed	0.2977	66.38	0.7669	-0.006881	0.009230
Monthly data					
Equal variances assumed	0.9075	4354	0.3642	-0.001400	0.003811
Equal variances not assumed	0.7905	2014	0.4293	-0.001786	0.004198
Mann-Whitney U Test	**W**		**p-value**		
Monthly data aggregated	637		0.9059	-0.007871	0.006338
Monthly data	2070691		0.9668	-0.001933	0.001848

In Table 4 the results of the two-sample t-test are compiled. The two-sample t-test on daily data results to a p-value of 0.4293 being larger than the 0.05 which is the limit when it is safe to reject the null hypothesis that the samples from the two different populations have equal mean at a 95%

confidence level. This means that we cannot reject the null hypothesis and it cannot be said with a statistical significance that the mean abnormal returns between these two types of funds differ from each other in our sample data within this timespan.

The non-parametric version of the independent or non-matched two-sample t-test is called Mann-Whitney or Wilcoxon rank-sum test. This test does not have the assumptions that are normal for parametric tests such as the Student's t-test, that the data is from a normally distributed population or from one that resembles it approximately. (Corder & Foreman 1972, 1; 69–80.) When used on a sample with non-normally distributed data, the non-parametric tests often provide better statistical explainability compared to the parametric ones, though in the case that the sample size is larger the parametric test should suffice. The reason for larger samples being viable for parametric tests is that in probability theory, the central limit theorem states that means of samples through random sampling approach a normal distribution with the same mean, even if the original variables or population are not normally distributed. (Kwak & Kim 2017.) According to the results in Table 4, the Mann-Whitney U test produces a p-value of 0.9688 for monthly data, which would indicate that we can not reject the null hypothesis of the two samples having no significant difference in their distributions. The 95% confidence interval of difference can also be interpreted to come to the same conclusion. As the value of 0 is included in our difference interval, it implies that it is statistically significant that there is no difference between the abnormal returns of active and passive groups. Through statistical tests and visual inspection we have determined that abnormal returns do not follow a normal distribution. Similarly through the tests we have determined that there is no statistically significant difference between the two distributions, suggesting that within this sample we cannot say that the abnormal returns between active or passive ETFs differ from each other significantly.

4.3 Main model and results

The main model attempts to evaluate whether different factors, such as management style, assets under management in the fund, lifetime of the fund or the expense ratio have an effect on how the fund performs. On the left hand side of the model there is the abnormal returns on the asset *i* which is calculated from the asset's total return index. Total return index is a comprehensive measurement that takes into account all possible cash distributions, dividends, interest payments and splits for the components of the index and assumes that the gains from dividends will be reinvested back into it. This is a widely used way to track the actual returns from an asset in the financial literature, and not just the price changes as in price return index. Abnormal returns here is essentially a market-adjusted return, which is a measure that attempts to remove the effect of the overall market movement in either direction from the gains of the single asset. This is achieved by subtracting the market returns from the asset returns in order to determine whether the asset had any independent movement aside from the general market sentiment. In this study, the fund is expected to perform as well as the market, and any deviation from this will be measured as the abnormal returns. The market benchmark used here is the performance of the S&P 500, similarly measured from its total return index. Therefore, the abnormal returns are calculated as follows:

$$AR_{it} = R_{it} - E(R_{it}), where$$

AR_{it} is the abnormal returns on asset i at time t
R_{it} is the actual realized returns on asset i at time t
$E(R_{it})$ is the expected return on the asset i at time t

The choice of the market benchmark used here is due to its prevalence in the financial literature, though alternative benchmark indices which see some use exist such as Dow Jones and Russell 2000. If the measured abnormal return is positive, it indicates that the stock has performed better than the benchmark index meaning that the change in the asset's price is not solely attributed to the market moving as a whole. A measure often associated with abnormal returns is cumulative abnormal returns (CAR) in which abnormal returns are cumulatively added up over a period. For example if a stock were to provide average market returns, the abnormal returns added up over an extended period of time would equal to zero, as the positives and negatives cancel each other out. If cumulative abnormal returns were to be positive or negative over a reasonable period, it can be conclued that the asset would on average perform better or worse than the benchmark index depending on whether the CAR is positive or negative. CAR is most often used for measuring the total effect of a specific event during certain period. But in this study we are not necessarily interested in a single event and its impact. Rather we are interested in measuring the differences between the active and passive fund performance and how different fund characteristics affect it over a long period. The cumulative nature of CAR also assumes that the abnormal returns separately will be independent and void of autocorrelation, which we do not need to guarantee by using abnormal returns instead. Using abnormal returns themselves also simplifies the analysis of the results without the need to consider how the aggregate returns over longer time period are to be interpreted. It also lets us compare our results against other studies easier with similar measurements.

The main model is multivariate regression model created to capture how different independent factors affect the dependent variable of abnormal returns. The first variable for the model, AUM is assets under management within the fund and is here used as a proxy for the fund size which is one of the most commonly used explanatory factors in the literature. Instead of the nominal value, the natural logarithm of AUM will be used, as it has been used in the past and has been shown to provide as robust proxy for the size as the nominal value. This can be attributed to its logarithmic nature giving diminishing marginal effect for larger and larger funds. This approach also gives more weight to the differences in size for smaller funds with the differences between two large funds being less significant. To interpret this variable, if the coefficient of β_1 were to be negative, it would indicate that the fund size is inversely correlated with the abnormal returns. This would mean for example that smaller funds could be more agile and to be expected to perform better in general if all other factors were to remain same, which is often referred to as ceteris paribus as a concept.

The second variable and probably the one with the largest impact based on previous literature is the expense ratio of the fund as a percentage, denoted as *FEE*. This is a simple measure gives an indication of the ongoing fee the investor directly pays to have their funds managed by this fund. Though as explained previously in this thesis, this doesn't necessarily give a holistic view of how much the fund spends on fees overall. This part of the model will attempt to determine whether the fund's expense ratio is a burden to its performance, or whether it is warranted and an indication of superior skill or resources put into the research for example. The most recent literature on this is quite consistent in finding this effect being negative, so the expected slope coefficient of β_2 is negative. This would suggest that an increase in the expense ratio erodes the fund's performance.

The third factor is the lifetime of the fund *AGE*, which is calculated in days from the inception date of the fund. This variable attempts to explain whether the fund's age has an impact on its performance. As this aims to isolate the age effect in itself, we would find whether younger funds overperform due to their age, and not due to some other reasons such as them being smaller in size which is more common among younger funds. Instead of the nominal value, a natural logarithm of the fund age will also be used here, as it gives a higher emphasis on the difference between younger funds compared to older where it is less likely to make a huge difference.

The first dummy variable in the model, *ACTIVE* is a binary variable, given only the value of either 1 or 0. Here the value 1 denotes that the fund is actively managed, and value 0 denotes that it is a passive fund. This variable lets us assess whether factor of the fund management style has an impact on the fund's performance, namely the abnormal returns. This will be visible in the coefficient of β_4 in the model, and it is used only on the combined sample which includes both, the passive and the active funds. As the dummy variable gets the value of 1 for active funds, it means that any effect visible in the coefficient should be interpreted as the difference in the dependent variable between active and passive funds. So for example a positive coefficient for the dummy variable would indicate better performance of active funds compared to passive with all other factors remaining same. The second dummy variable *BOND* is similarly either 1 for bond funds, and 0 for equity funds. This dummy variable is used only for the passive sample in order to isolate the difference in performance of bond funds compared to the equity funds, which are expected behave and perform very differently. As the abnormal returns are adjusted based on the equity markets, the low returns of bond funds would sway the results of the model and pull down the average abnormal returns for the whole passive group. As the impact of other characteristics is also likely to be different for bond funds, and additional regression is ran for the passive group which excludes the bond funds completely.

4.3.1 Active sample over three timespans

The sample of actively managed funds is quite limited in quantity as they are way more likely to be younger which leads to a situation where they lack the required data going back multiple years. This means that out of the 100 largest active funds, only 39 of them have enough data for three years, 27 for four years and only 17 have enough data for a time span of five years. In the first multivariate regression model we have the abnormal returns as the dependent variable and as independent variables we have assets under management, the fund's expense ratio and its age. The regression can be formulated as:

$$\begin{gathered} E(AR)_{it} = c_{it} + \beta_1[\log_e(AUM_{it-1})] + \beta_2[FEE_i] + \beta_3[\log_e(AGE_{it-1})], \text{ where} \\ E(AR)_{it} \text{ is the abnormal return on asset } i \text{ at time } t, \\ \beta_{1,2,3} \text{ are the corresponding betas on each of these independent variables:} \\ AUM_{it-1} = \text{assets under management in the previous period} \\ FEE_i = \text{expense ratio} \\ AGE_{it-1} = \text{age of the fund in the previous period} \end{gathered} \tag{1}$$

This regression is ran on three different data samples and the results are reported in Table 5. None of our coefficient are statistically significant, which can be partly attributed to the smaller number of datapoints and funds overall in this group. Despite of this, some observations can be drawn from this such that the sign of the coefficient is consistent across all three timespans for assets under management and expense ratio. The coefficients for AUM in Table 5 are very small or near zero,

Table 5: Regression results of active ETF abnormal returns for three time spans

The samples consist of 39, 27 and 17 actively managed funds with 1365, 1269 and 1003 datapoints for years 3, 4 and 5 respectively. The samples consist of data starting from 8/2016 and all ending in 8/2021. Statistical significance represented as *p < .1; **p < .05; ***p < .01; Standard error in parenthesis.

	AR, 3 year	AR, 4 year	AR, 5 year
Constant c	-1.094 (1.52)	0.028 (1.603)	1.626 (2.006)
Assets under management $\log_e(AUM_{t-1})$	0.039 (0.122)	0.098 (0.115)	0.017 (0.123)
Expense ratio FEE	-0.386 (0.457)	-0.552 (0.441)	-0.393 (0.464)
Fund age $\log_e(AGE_{t-1})$	0.133 (0.236)	-0.056 (0.230)	-0.198 (0.282)
Observations	1365	1269	1003
Adj. R^2	-0.001	>-0.001	>-0.001

suggesting that there is very little to no relation between the independent variable to the abnormal returns. A positive coefficient would indicate that the actively managed funds which have more capital can take advantage of opportunities in the market, and would be able to create positive abnormal returns. Among mutual funds, one of the early studies was by Chen J. et al. (2004) who suggested and found evidence that the size of the fund would erode it's performance due to liquidity and organizational diseconomies of scale. Pástor et al. (2015) results among active mutual funds suggest decreasing economies of scale at the fund level, but their results did not meet the level of statistical significance. They attempt to improve their methodology in Pástor et al. (2020) by changing how they measure the fund size by incorporating the fund's activity and how they well they use their available resources. While incorporating the trade turnover and the trading activity, their findings are still consistent with the fund's size having negative returns to scale but now their results are statistically significant. The inclusion of turnover is likely a driving factor in the measurement, as it is shown in e.g. Malkiel and Saha (2020) and Da and Shive (2018) that fund's turnover ratio is negatively correlated with the returns. In our results, the impact of AUM on abnormal returns is little to none and statistically insignificant to suggest economies of scale either way. Zhu (2018) suggests that there is a point in the size of the funds in which they grow past their optimal asset allocation. As our sample contains largely smaller ETFs compared to the mutual fund markets, it is possible that these negative effects of scale prominent among mutual funds are not prominent enough among ETFs yet. The proposition of a fund's returns scaling with its size with no limitations is unfeasible, as in that world the most optimal fund would assume all of the wealth and become the whole market. Similarly, even constant returns to scale is unrealistic as it would imply that the fund's investment strategy is infinitely scalable. (Zhu 2018.) From this angle, it is reasonable to assume that possible the positive relationship of returns to scale is only temporary, and at certain point any further increase in size would no longer increase the returns.

The consistent negative coefficient for expense ratio would indicate that the higher the expense ratio is, the less likely it the actively managed fund is to produce abnormal returns ceteris paribus when benchmarked against the market. One interpretation for this is that the managers of these funds which demand higher fees do not actually deliver better results from the aspect of the investor. We do not measure the returns before the fees here, rather what the investor actually receives after the fees as this is the deciding factor when they're selecting a fund to invest in. Pástor et al. (2020) found strong empirical evidence among active mutual funds of negative correlation between fund size measured as AUM and expense ratio, supporting the general sentiment of larger funds being able to offer lower expense ratios. As both of these variables are accounted for in the model, the lower expense ratio benefit spillover from the increase in fund size should not be visible in the AUM effect as it is isolated in the expense ratio variable.

4.3.2 Passive sample with and without bonds over three timespans

The sample for the passively managed ETFs is considerably larger as these funds have generally been around longer and are more longer-lived than active ETFs which are quite a new phenomenom. Out of the 100 largest passively managed funds by market capitalization, in the first sample we have data for 82 of them for three years, 81 for four years and 79 for the whole five year period. In this multivariate regression we have the abnormal returns as the dependent variable and the same independent variables of assets under management, the fund's expense ratio and its age similarly to before. We split the sample in two due to how some of these passive funds are classified as bond instead of equity funds. These two fund types can not realistically be compared as equals, so in an attempt to isolate and measure their difference in abnormal return, we have one sample with only the equity funds, and one with both equity and bond funds. A new dummy variable *BOND* is added to the regression, having a value of either 1 or 0 indicating whether the fund is classified as bond fund, separating this effect and being measurable from the dummy variable's coefficient. The regression can be formulated as:

$$E(AR)_{it}=c_{it}+\beta_1\left[\log_e(AUM_{it-1})\right]+\beta_2\left[FEE_i\right]+\beta_3\left[\log_e(AGE_{it-1})\right]+\beta_4\left[BOND_i\right],\text{ where}$$
$$E(AR)_{it}\text{ is the abnormal return on asset } i \text{ at time } t,$$
$$\beta_{1,2,3,4}\text{ are the corresponding betas on each of these independent variables:}$$
$$AUM_{it-1}=\text{assets under management in the previous period} \tag{2}$$
$$FEE_i=\text{expense ratio}$$
$$AGE_{it-1}=\text{age of the fund in the previous period}$$
$$BOND_i=\text{dummy variable for bond funds}$$

The results of this regression are reported in Table 6. The dummy variable for the bond funds in our passive sample is significant at a 99% confidence level. The coefficients for this variable ranges from -0.801 to -0.949, with a decreasing standard error over the longer time spans. This result indicates that the bond funds witness 0.80 to 0.95 percentage point lower abnormal returns compared to their equity investing counterparts. When we consider how the abnormal or market-adjusted returns are calculated, the reason for this result becomes apparent. As the abnormal returns here are the returns that exceed the benchmark index, S&P 500 in our case, it is not unexpected that the bond funds perform worse compared to the market index which invests exclusively in equity companies with their higher risk and reward profiles. According to the historical data from Morningstar Direct, as of 2018, the average annual return for bond funds was 6.8% while being 10.4% for stocks over the past 30 years. If this annual difference was converted to monthly returns, the Bloomberg U.S. Aggregate Bond Index performed around 0.3 percentage points worse monthly than the S&P 500 Index. (Schwab 2019.) This means that within our sample, the passive ETFs which invest into bonds have performed worse during our sample time period than the 30 year average of the aggregate index. This can partly be attributed to the historically poor performance of

Table 6: Regression results of passive ETF abnormal returns for three time spans

The first samples consist of 82, 81 and 79 passively managed funds consisting of 2870, 3807 and 4661 datapoints for years 3, 4 and 5 respectively. The second set of samples also include bond funds on top of equity funds, and they include 100, 99 and 97 passively managed funds consisting of 3500, 4653 and 5723 datapoints for years 3, 4 and 5 respectively. The samples consist of data starting from 8/2016 and all ending in 8/2021. Statistical significance represented as *p < .1; **p < .05; ***p < .01; Standard error in parenthesis.

	AR, 3 year	**AR, 4 year**	**AR, 5 year**	**AR, 3 year with bonds**	**AR, 4 year with bonds**	**AR, 5 year with bonds**
Constant	-0.270	-1.804**	-1.324	0.429	-1.323	-0.892
c	(1.10)	(0.872)	(0.850)	(1.247)	(0.970)	(0.916)
Assets under management	-0.090	-0.124**	-0.087	-0.139	-0.139**	-0.095
$\log_e(AUM_{t-1})$	(0.083)	(0.063)	(0.057)	(0.093)	(0.070)	(0.061)
Expense ratio	-0.716	-1.434***	-1.345***	-0.640	-1.222***	-1.098***
FEE	(0.500)	(0.393)	(0.332)	(0.540)	(0.419)	(0.350)
Fund age	0.110	0.339***	0.242**	0.084	0.294**	0.195
$\log_e(AGE_{t-1})$	(0.142)	(0.126)	(0.116)	(0.160)	(0.137)	(0.124)
Bond (dummy)				-0.801***	-0.866***	-0.949***
BOND				(0.180)	(0.141)	(0.118)
Observations	2870	3807	4661	3500	4653	5723
Adj. R^2	>-0.001	0.004	0.003	0.005	0.009	0.012

the bond market during the sample period, especially during 2021 when it fell negative. Before this, the index had only dipped down to negative annual returns four times since its inception 46 years ago. (Morningstar 2023.) Another reason for the large, from 9.2% to 10.8% annualized difference in returns between the two asset types is this sample period in which equity returns were unusually high.

For both samples, with or without bonds included, the coefficient for assets under management is significant at α=0.05 but only in the four year time span, though the standard errors across all spans are quite small indicating that this coefficient is quite likely to be negative. As we are using a natural logarithm of the assets under management variable here, the interpretation of the coefficient is not as straightforward. Due to the logarithmic transformation the coefficient has to be interpreted as follows, a percentage increase in the independent variable causes an $\beta_1 \times \log(1.01) = \beta_1 \times 0.995\%$ or an increase of roughly one percent of the coefficient's nominal value in the dependent variable. So the coefficient which ranges from -0.09 to -0.12 for the equity ETFs means that a one percent increase in assets under management would lower the amount of abnormal returns the passive fund is able generate by 0.0009 to 0.0012 percentage points. This negative effect seems to be slightly larger on the sample which includes bonds, showing as high as 0.0014. If this were to be annualized, it would amount to around a 1.1 bp to 1.4 bp hit on abnormal returns for an equity passive fund if its AUM were to increase by 1%. As we compare the results from our regression on the passive sample against the active sample, the effect of assets under management on the abnormal returns seems to affect these two groups differently. Actively managed funds appear to gain from near zero to small benefit from the increase in available funds, which the active fund can use to capitalize on opportunities, while the passive funds usually have to allocate their funds based on the indices they track, appear to suffer from the diseconomies of scale. One explanation for the inverse relation, as reviewed before in this thesis and in previous literature, is the passively managed funds' need to track their benchmark indices. This has shown to cause an effect called index premium or benchmark effect. The root cause for this negative impact is the fund's need to balance their holdings according to the tracked index, while minimizing the tracking error. This rebalancing leads to the passive funds having to pay a larger price for the stocks entering the index, and selling them at a discount when leaving it, as other market entities can predict and take advantage of these events. The magnitude of this impact depends on which theory you subscribe to on how permanent the effect is and whether there is price reversal. The more permanent the effect is, the better it would be for the passive funds. In addition to Petajisto (2011), also Raddatz et al. (2017) found evidence of this effect. This could at least partially explain why larger funds tracking large indices appear to show worse performance. This result is consistent with the diseconomies of scale amongst passive ETFs which has been observed by many such as Paudel & Naka (2023). In their study for the period of 2009 to 2018 with total net assets as a proxy for size, their findings suggest that the beta coefficients are not constant return-to-scale as they observe a decrease in the coefficients when the model is ran on different size quartiles. They find that amongst the smaller funds, there is a positive relation with the size and risk-adjusted alpha, but that this relation deteoriates and eventually turns to negative when the fund size has grown large enough. This from their point of view enforces the theory where there exists an optimal fund size, and exceeding this point introduces diseconomies of scale.

As observed for the active sample, the passive ETFs similarly report negative coefficients for the expense ratio. These coefficients are significant at a 99% confidence level for both the four and five year time spans. The coefficients for the expense ratio in the sample with bonds excluded are -1.434 and -1.345 for four and five year periods respectively. This means that an increase of 1 percentage point in the expense ratio is expected to have a negative effect of -1.35% to -1.43% on the fund's abnormal returns. If we look at the coefficients of the sample with bonds included, the negative effect is slightly smaller with slightly larger standard errors. This could be due to how the bond funds' drag down the averages with their lower returns among funds with similar expense ratios. The observation that a 1 percentage point increase in expense ratio decreases the abnormal returns by more than 1 percentage point is an interesting observation. This would imply that an increase in the expense ratio would actually be a larger drawback than the expense ratio in itself would suggest. This is consistent with Elton et al. (2019a) who find similar relation with both passive mutual funds and passive ETFs. Their models show after expenses differential returns decreasing by 1.26% per year for ETFs and 1.04% for mutual funds on 1% increase in the yearly expense ratio. This relation in which the after-fees coefficient exceeds negative one has also been found among active funds in a study by Sheng et al. (2023). The returns being so negatively correlated with the expense ratio, even past the logical point of -1 could be due to variety of options. The relationship might not be completely linear skewing the result, there might be unaccounted inefficiencies among the funds with higher expense ratios causing them to perform worse, or for example investor behaviour in which they are very sensitive to small differences in expense ratio, causing the funds with higher fees to suffer from fund outflows or other issues.

The regression model shows a positive coefficients for the fund age variable across the passive ETFs in general. This indicates that the fund is able to generate positive abnormal returns based on the older it is. The coefficient for the fund age is significant at a 95% confidence level on the five year sample, and at a 99% confidence level on the four year sample with coefficients from 0.242 to 0.339 respectively. As the fund age variable is taken as a natural logarithm of the absolute value, the interpretation of the coefficient differs from the regular. According to the results, increasing the age of the fund by a one percent would be seen as a 0.0024 to 0.0034 percentage point increase in the abnormal returns monthly. Here it is very important to remember the distinction between a percent and a percentage point increase as they can affect the interpretation. According to this result, the more mature funds seem to do slightly better and are more likely to outperform the benchmark index than younger funds. This coefficient is only statistically significant in our regression on passive funds, though the coefficient still seems to also stay positive in our other regressions. Paudel & Naka (2023) similarly found positive relation with fund age and performance among passive ETFs, with statistically significant results at $\alpha=0.01$. With their methodology they were able isolate behaviour for different size quartiles, finding the effect to be greater among smaller fund sizes with the effect becoming less significant at the larger end, suggesting that the relation might not be just linear. As their study used the nominal value for fund age, there is a possibility that the logarithmic transformation that was done for our age variable rectifies this effect.

Sheng et al. (2023) also observe a positive coefficient for their proxy of the fund age among active funds, though this effect is very small and not statistically significant. Their study specifically sought to use survivor bias free database in an attempt to alleviate any biases which could affect their results. As our data does not take survivorship bias in account, it is possible that part of the observed effect can be biased by the survivors. This can skew the returns if older firms that performed well are more likely to be still around, while the worse performing ones have dissolved or merged into other funds. When attempting to account for multiple factors and fixed effects in their model, Pástor et al. (2015) find mixed results. At fund level, their results show that in active mutual fund management, the performance erodes within the fund's lifetime. They observed a 1.23 bp decrease in monthly gross benchmark-adjusted returns for each additional year the fund had existed. When adjusted for industry size and its growth, they found that the impact was insignificant and very little if any. They suggest this to be due to the newer entrants to the market being more skilled than the incumbents.

If there was a positive relation of performance with fund age, it would indicate some kind of learning and experience among the fund's management, or improved name recognition driving larger inflows and opportunities that comes with age. It is possible that some of the managerial skill effect will spill to this variable as it is not accounted for through any other proxy variable. Managerial skill is a difficult factor to measure and it is also widely debated topic on how much it plays a role in the fund's performance. One way to quantify the skill is the fund's ability to produce alpha in excess over the charged fees within active management, and this is what Berk and Green (2004) did. They believe that there are active funds which are able to beat other participants in the market by finding inefficiencies and profitable opportunities. Crane and Crotty (2018) expand the scope from active management to passive, specifically index funds. Their results imply that index funds can exhibit skill similarly to active funds, despite of index fund managers having less freedom to employ differential strategies. Some ways that passive and index fund managers can still stand out from others is the ability keep tracking error very low, minimizing expenses within the fund which might not show up in expense ratio, and optimizing rebalancing around index changes by minimizing costs caused by negative index price effects. This uncaptured skill is likely to show up as bias in one of the variables, most likely the fund age.

As is apparent from the comparison between the two passive samples of which one has only equity investing funds, and other has both equity and bond funds, there is clear difference between these two samples which affect the results of our regression. For this reason, in the following regressions we will be using only the subset of the passively managed funds that invest in equity. This should make the results against the actively managed counterparts more comparable as the active sample also only consists of equity funds.

4.3.3 Combined sample over three timespans

The regression for the combined sample which includes both the equity investing passive funds and actively managed funds differs from the previous regressions by its introduction of a new dummy variable in an attempt to capture the effect difference between the two fund management types. This *ACTIVE* variable is set to 1 if the fund is actively managed and 0 if not. As we are not including the bond focused passive funds in this sample, the *BOND* dummy variable is left out from this regression. The other independent variables are the same, being assets under management, the expense ratio and the fund's age. The regression can be formulated as:

$$E(AR)_{it} = c_{it} + \beta_1[\log_e(AUM_{it-1})] + \beta_2[FEE_i] + \beta_3[\log_e(AGE_{it-1})] + \beta_4[ACTIVE_i] \text{, where}$$
$$E(AR)_{it} \text{ is the abnormal return on asset } i \text{ at time } t,$$
$$\beta_{1,2,3,4} \text{ are the corresponding betas on each of these independent variables:}$$
$$AUM_{it-1} = \text{assets under management in the previous period} \quad (3)$$
$$FEE_i = \text{expense ratio}$$
$$AGE_{it-1} = \text{age of the fund in the previous period}$$
$$ACTIVE_i = \text{dummy variable for active funds}$$

The results of this regression are reported in Table 7. This regression attempts to isolate the difference between the two groups through the active dummy variable, with the assumption of them behaving similarly across the other variables. The results show similar trend for the expense ratio coefficient as the previous regressions, being negative and significant at a 99% confidence level for four and five year time spans this time. Based on the coefficients of -0.705 and -0.665, there appears to be around 0.67 to 0.71 percentage point decrease in abnormal returns when the independent variable of expense ratio increases by one percentage point. This negative relationship is consistent with the results from the previous regressions though the size of the effect seems to be somewhere between what was observed for the active and passive samples. As there is quite a difference between the coefficient between the active and passive samples, with the combined sample coefficient falling between the two, it can be assumed that active funds are less and passive funds are more sensitive to changes in expense ratio. As the impact of the expense ratio is not isolated with the *ACTIVE* dummy variable, it is expected that the combined sample exhibits the characteristics of both groups. The reason for the different sensitivity is most likely the investors'

Table 7: Regression results of Combined sample abnormal returns for three time spans

The 3 year sample consists of a combination of 39 active, 82 passive funds with 1365, 2870 datapoints respectively, the 4 year sample consists of a combination of 27 active, 81 passive funds with 1269, 3807 datapoints respectively and the 5 year sample consists of a combination of 17 active, 79 passive funds with 1003, 4661 datapoints respectively. The active funds are distinguished from the passive funds through a dummy variable in the regression model. The data for these samples is starting from 8/2016 for the longest range and all ending in 8/2021. Statistical significance represented as *p < .1; **p < .05; ***p < .01; Standard error in parenthesis.

	AR, 3 year	**AR, 4 year**	**AR, 5 year**
Constant	-0.997	-1.320	-0.449
c	(0.974)	(0.824)	(0.827)
Assets under management	-0.015	-0.0030	-0.017
$\log_e(AUM_{t-1})$	(0.066)	(0.053)	(0.050)
Expense ratio	-0.383	-0.705***	-0.665***
FEE	(0.295)	(0.257)	(0.232)
Fund age	0.103	0.126	0.044
$\log_e(AGE_{t-1})$	(0.124)	(0.108)	(0.109)
Active (dummy)	0.380	0.644**	0.661**
ACTIVE	(0.382)	(0.306)	(0.295)
Observations	4235	5076	5664
Adj. R^2	>-0.001	0.001	0.002

expectations of these two types, investors investing into passive funds are very concious of the expense ratio, whereas active fund investors are more willing to pay higher fees in return expecting to gain it back through superior performance in a fund with more active trading strategy. Sheng et al. (2023) find that using conventional CAPM, Fama-French three-factor and four-factor models gives the expected large and negative expense ratio coefficient after-fee, and slightly negative to zero before-fee coefficient among active funds. They suppose that active funds with drastically different expense ratios tend to invest in different types of stocks. By attempting to control for these differences using the Fama-French five-factor model, controlling for stocks with low operating profitability and high investment rates, the picture changes. With the FF5, the coefficient is large, positive and highly significant before-fees, and near zero after-fees. Their results under this model would suggest that before-fees the funds with larger expense ratios are actually able to produce alpha to cover for the increased costs and after-fees, the result is near zero. According to this result, the active funds would still be able to cover their increased costs after-fees but barely. These results would be consistent with the theory of Berk and Green (2004) who studied managerial skill and suggested that majority of the active fund managers are able to generate value excess their fees.

One factor that has been shown to increase the gap between active and passive funds that follow an index benchmark, are the effects caused by index inclusion and exclusion events. These events cause a negative effect, and the costs caused by it are usually referred to as index premium. Upwards price pressure can be observed following the index change announcement up until the effective date for index additions, and downwards price pressure for index exclusions. This effect in favor of active funds is something that would not be driven by any descriptive fund characteristic listed in this model or in general, as it caused by the inherent limitations and constraints that concern only passive funds. Any advantage in favor for active funds in trade timing flexibility or opportunities would show up as a positive effect on the dummy variable separating these two groups. Nie (2024) continues based on Petajisto (2011) previous work, and provides an updated view on the index premium from 2004 to 2012. He finds that this effect is more pronounced for index deletions and the year 2008 stands out with index deletions on average showing cumulative abnormal returns (CAR) as high as -31.79%. They find for S&P 500, on average the index additions show an average CAR of 3.98% and -9.90% for deletions between the years 2004 and 2012. As the index changes are usually initiated by something fundamental that warrants a removal of a member, it is logical for these to be a more significant negative signal to investors. In contrast, index additions are mostly done just replace the vacancies. These price effects are theoretically only temporary, and Nie (2024) shows that the price reversal among additions happens within 20 to 30 days, which is less than observed by Petajisto (2011) for 1990 to 2005. They find that the time between temporary misspricings' appearance and the time before they are fixed has shortened over the years, showing that markets have improved in efficiency. Patel and Welch (2017) use even longer time span from 1979 to 2015 and find the investor demand hypothesis to be most credible. They observed a short-term 3% to 4% increase in abnormal return for additions and -5% for discretionary or -1% for forced deletions, with the evidence suggesting a lessening trend over the years. Similarly another contemporary study by Bennett et al. (2020) finds a reduction in the effect when comparing the results between 1997 to 2007 and 2008 to 2017 separately, with the impact falling from ca. 4.8% to 0.7% over the years. These studies are consistent with the existence of the index price effect, and largely show that the strength of the effect has decreased over the years. The price reversal delay has shown to have reduced over the years suggesting improved market liquidity and efficiency.

The behaviour of active funds and arbitrageurs which try to benefit from the index changes in S&P 500 is the focus of the study by Liao et al. (2022). With the assumption that the stock betas will converge towards one over time due to the increased comovement after index addition, the beta arbitrageurs can create portfolios to take advantage of this effect. They highlight how active traders are able to buy for example these leveraged low beta stocks for cheap prior to index change implementation dates, while indexer is forced to follow. In the previous case, both of these parties are on the same side of the trade, but the active trader has the advantage. Index deletions on the other hand create excess supply, pushing down prices and making them attractive option for active traders going for leveraged long positions.

The coefficients for the fund age are positive for each year like previously in the passive sample, though the results are not significant at a high confidence level. The new active dummy variable is significant in the two longer samples at a 95% confidence level with coefficient values of 0.644 and 0.661. With the assumptions of the two groups behaving similarly across the other independent variables, the positive coefficient for *ACTIVE* dummy variable would suggest that within this model the active funds outperform the passive funds with the other characteristics kept the same. The difference in abnormal returns between these groups with all other variables staying equal would be a 0.64 to 0.67 percentage points monthly in favor of active funds. Eventhough this sounds large, this result can not be solely used to say that one should invest into an active fund as they perform better, as the other variables are rarely the same. Say the expense ratio is almost always higher for the active funds which in turn pushes the abnormal returns down. Also this regression does not take in account that the two groups most likely do not behave identically for all the other independent variables. This is looked further into later in alternative analysis section.

The next regression is ran on the same combined sample as the previous one, but two independent variables are added. The first of one of these is the previous periods abnormal returns, meaning it will show us if there is an effect where a fund’s previous overperformance compared to the index is predictive of future overperformance. The second variable is this previous abnormal returns multiplied by the *ACTIVE* dummy variable. This creates an interaction term which allows us to see if there is difference in the size of the effect between active and passive funds. The whole regression can be written out as:

$$\begin{gathered} E(AR)_{it} = c_{it} + \beta_1\left[\log_e\left(AUM_{it-1}\right)\right] + \beta_2\left[FEE\right] + \beta_3\left[\log_e\left(AGE_{it-1}\right)\right] + \beta_4\left[ACTIVE_i\right] \\ + \beta_5\left[AR_{it-1}\right] + \beta_6\left[ACTIVE \times AR_{it-1}\right], \text{ where} \\ E(AR)_{it} \text{ is the abnormal return on asset } i \text{ at time } t, \\ \beta_{1,2,3,4,5,6} \text{ are the corresponding betas on each of these independent variables:} \\ AUM_{it-1} = \text{ assets under management in the previous period} \\ FEE_i = \text{expense ratio} \\ AGE_{it-1} = \text{lifetime of the fund} \\ ACTIVE_i = \text{dummy variable for active funds} \end{gathered} \tag{4}$$

The results for this regression are reported in Table 8. The coefficient for expense ratio is significant at a confidence level of 99% for four and five year samples as before. The size of the negative effect is very similar to the previous regression, being -0.71% and -0.66% respectively for four and five year samples. The fund age coefficients are positive as is the trend, indicating a positive correlation between the abnormal returns and how long the fund has been around, though the amount of the effect isn't statistically significant. The effect of the dummy variable is very similar to the results of the previous regression being around 0.65 to 0.66. The first new independent variable of this regression is the abnormal returns with a lag of one period capturing the reference group of passive ETFs. It is significant only in the three year sample with a coefficient of -0.051 with a confidence level of 95%. A negative relationship between the previous periods overperformance to the next period's performance is definitely counterintuitive, but part of this can be explained with the next interaction term which captures some of this effect for the actively managed funds. Wang & Zheng (2024) find that there is little evidence of funds which trade on momentum gaining abnormal returns, rather that the success can be mostly attributed to the skill of timing the momentum trading correctly. Grinblatt et al. (2020) also find that despite of the funds following momentum strategies performing better than contrarian funds (funds that pick past losers), the difference disappears after all momentum related factors are controlled for. This would be in line with the "momentum" factor being very close to zero. Our coefficient being close to zero is also consistent the fact that passively managed funds have no option to do any active trading, so there shouldn't be much if any observed benefit from past performance. As the factor in this regression uses only the previous month's overperformance as an explanatory variable, it does not capture a longer trend in the fund's performance. As generally the momentum factor is crafted for a longer time span such as 3, 6 or 12 months, this result would not be able to accertain whether an effect exists for even longer time windows.

The next variable is an interaction term in which the previous period's abnormal returns are multiplied with the *ACTIVE* dummy variable separating the "momentum" effect between the active and passive funds. The coefficient is 0.118 for the three year sample at a confidence level of 99%, 0.062 and 0.059 at a confidence level of 95% for four and five years respectively. This statistically significant positive coefficient indicates that the actively managed funds are more likely to display the “momentum” compared to passively managed funds which are at the mercy of the overall markets. As the “momentum” factor is slightly negative in the reference group, it indicates that the passively managed funds tend to perform inversely to the previous periods performance. The interaction term of the “momentum” factor which is 0.059 to 0.118, would cancel the negative “momentum” effect observed in the passive group. This means that across all three time spans, the previous performance of the fund seems to be larger predictor of future performance for actively managed funds than passively managed funds.

It seems to be common for all of the ran regressions that the results for the three year sample are not highly significant even if the four and five year samples are at a very high confidence level.

Table 8: Regression results of Combined sample abnormal returns with “momentum” factor included for three time spans

The 3 year sample consists of a combination of 39 active, 82 passive funds with 1365, 2870 datapoints respectively, the 4 year sample consists of a combination of 27 active, 81 passive funds with 1269, 3807 datapoints respectively and the 5 year sample consists of a combination of 17 active, 79 passive funds with 1003, 4661 datapoints respectively. The active funds are distinguished from the passive funds through a dummy variable in the regression model. This regression also includes an independent variable for abnormal returns with lag and an interaction term for the active group. The data for these samples is starting from 8/2016 for the longest range and all ending in 8/2021. Statistical significance represented as *p < .1; **p < .05; ***p < .01; Standard error in parenthesis.

	AR, 3 year	**AR, 4 year**	**AR, 5 year**
Constant	-0.939	-1.352	-0.472
c	(0.974)	(0.824)	(0.827)
Assets under management	-0.024	-0.0060	-0.020
$\log_e(AUM_{t-1})$	(0.066)	(0.053)	(0.050)
Expense ratio	-0.370	-0.708***	-0.661***
FEE	(0.295)	(0.257)	(0.232)
Fund age	0.104	0.132	0.050
$\log_e(AGE_{t-1})$	(0.123)	(0.108)	(0.107)
Active (dummy)	0.362	0.657**	0.654**
$ACTIVE_i$	(0.382)	(0.306)	(0.295)
Abnormal return (lag 1)	-0.051**	-0.030	-0.015
AR_{t-1}	(0.022)	(0.019)	(0.017)
Active × Abnormal return (lag 1)	0.118***	0.062**	0.059**
$ACTIVE \times AR_{t-1}$	(0.031)	(0.028)	(0.027)
Observations	4235	5076	5664
Adj. R^2	0.003	0.002	0.002

For example, the expense ratio has consistently show storng explanatory value and has been statistically significant for the longer time spans, but never for the three year sample. The reason for this difference could possibly be attributed to the smaller sample size, or that this three year sample includes a section in the market movement where the fund characteristics such as expense ratio made less of an impact in a very turbulent market. This shorter time span of 2018 to 2021 is covered largely by the uncertainty in the markets caused by the pandemic from 2020 to 2021. This could explain why the expense ratio coefficient is less sensitive for the three year sample compared to the others in the tested regression models.

4.4 Alternative analysis and robustness

The results in Table 7 and Table 8 suggest that the sole difference in management of the fund being either actively or passively managed would be a statistically significant estimator and indicate a superior performance for actively managed funds ceteris paribus. This result rests on the model assuming that the other independent variables behave similarly across these two groups when we are estimating the impact of the *ACTIVE* dummy variable on abnormal returns. As has been highlighted from literature previously, there appears to be differences how these characteristics can have varying impact for active and passive funds. For example expense ratios are drastically different between the two groups, with median being 0.65% for active and 0.12% for passive funds within our sample. The active funds use the higher fees to pay for research analysts, managers, and market research while also having increased transaction costs due to active trading. The relatively small fees of passive funds are spent on licensing fees paid to index providers and other operational costs, but they save on the lack of research costs. As these two fee structures are very different, it is not unexpected that their relation to the produced abnormal returns are different.

In Table 9 an interaction term is added for the expense ratio variable in attempt determine whether the some of the difference can be attributed to the different behaviour within this variable between these two groups. Reason for this is that there appears to be very large deviation in the coefficients of expense ratio between the active and passive groups, and that based on prior literature the investors are more sensitive to changes in it among passive funds. By inspecting the adjusted R-squared between the two panels, the interaction terms seem to improve the model very slightly. Using adjusted R-squared here ensures sure that it is an actual improvement instead of just due to the increase in variables.

The addition of the interaction term for expense ratio removes the large and significant impact of the ACTIVE dummy variable, indicating that it is likely that the active management effect found on Table 7 and Table 8 is due to omitted-variable bias. This would suggest that the effect of the interaction terms on these characteristics between the two groups do explain some of the active management effect, and when these are included, the effect lessens and becomes insignificant. However, the effect still appears to be positive, though insignificant with large standard error. This suggests that the sole fact of a fund being either passive or active doesn’t have statistically significant predictive power. Rather it suggests that the characteristics on their own drive the performance, with expense ratio being the largest contributor. The interaction term is only slightly significant at α=0.1 level for the 5 year sample showing that active funds suffer 0.919 percentage points less in abnormal returns for each 1 percentage point increase in expense ratio compared to passive funds, though still staying negative. The adjusted R-squared increasing slightly suggests that this addition improves the model. Similarly the statistical significance of the *ACTIVE* dummy variable disappears when an interaction term is tested similarly for either assets under management or fund lifetime for robustness.

Table 9: Regression results of Combined sample abnormal returns with interaction term included for three time spans

The 3 year sample consists of a combination of 39 active, 82 passive funds with 1365, 2870 datapoints respectively, the 4 year sample consists of a combination of 27 active, 81 passive funds with 1269, 3807 datapoints respectively and the 5 year sample consists of a combination of 17 active, 79 passive funds with 1003, 4661 datapoints respectively. The active funds are distinguished from the passive funds through a dummy variable in the regression model. An interaction term is added for expense ratio separating the effect between the two groups. The data for these samples is starting from 8/2016 for the longest range and all ending in 8/2021. Statistical significance represented as *p < .1; **p < .05; ***p < .01; Standard error in parenthesis.

	AR, 3 year	**AR, 4 year**	**AR, 5 year**
Constant c	-0.933 (0.983)	-1.280 (0.824)	-0.498 (0.828)
Assets under management $\log_e(AUM_{t-1})$	-0.025 (0.069)	-0.025 (0.055)	-0.042 (0.052)
Expense ratio FEE	-0.636 (0.579)	-1.240*** (0.445)	-1.227*** (0.381)
FEE interaction term FEE×ACTIVE	0.347 (0.683)	0.825 (0.560)	0.919* (0.495)
Fund age $\log_e(AGE_{t-1})$	0.111 (0.125)	0.157 (0.110)	0.090 (0.110)
Active (dummy) ACTIVE	0.243 (0.468)	0.284 (0.392)	0.232 (0.374)
Observations	4235	5076	5664
Adj. R^2	>-0.001	0.001	0.002

Different functional forms of assets under management variable are explored to test for robustness. These include the nominal value as is, double logarithm of AUM and AUM squared. These variations are tested on the baseline regressions from Table 5 and Table 6. Regardless of the AUM scale chosen, for all periods and both management styles the expense ratio always remains highly statistically significant. The choice of scale makes a noticeable difference for the active group specifically, possibly explaining the insignificant results received in Table 5. The findings in Table 5 with the logarithmic AUM scale suggest that it has no statistically significant relation with the abnormal returns. Though if this scale is changed to nominal, the AUM coefficient is significant for 1 period, and if AUM squared scale is used, it is significant at α=0.01 for all three tested periods.

This also greatly increases the adjusted R-squared for the active model and is shown on Table 10. When this same scale is tested for the model with passive group data, the variable is insignificant for all periods. This stark difference between the two groups indicates to us that these two groups behave very differently with the change in how assets under management is measured. This suggests that AUM relationship does not fit linear or logarithmic scale the best, rather that it has some quadratic properties among the active ETFs. This relation means that the change in the effect increases or decreases more the larger it grows. As the coefficient is negative, it suggests that the performance suffers at an increasing rate as AUM increases. In that case, there exists a point at which it is no longer beneficial for the active fund to expand. This limit seems to come earlier for passive than active funds, after it has reached the benefits which are gained by its increase in size. This theory of an optimal fund size is widely connected to active mutual funds (Paudel & Naka 2023). The results are is similar whether the nominal value was used in addition with the squared value. As stated, among the passive group regardless of the AUM scale used, the expense ratio coefficient always stays statistically significant. The observation that for passive funds the log and loglog-scales show the best fit for assets under management enforces the finding that these two management styles have different relationships with AUM and performance. Further research is needed on which size scale should be used for each of the fund styles.

Table 10: Regression results of active ETF abnormal returns with modified assets under management scale

The samples consist of 39, 27 and 17 actively managed funds with 1365, 1269 and 1003 datapoints for years 3, 4 and 5 respectively. The samples consist of data starting from 8/2016 and all ending in 8/2021. This regression uses AUM squared scale instead of logarithmic. Statistical significance represented as *p < .1; **p < .05; ***p < .01; Standard error in parenthesis.

	AR, 3 year	**AR, 4 year**	**AR, 5 year**
Constant *c*	-1.565 (1.515)	-0.442 (1.601)	1.199 (1.98)
Assets under management $(AUM_{t-1})^2$	-1.06×10^{-8}*** (3.42×10^{-9})	-1.02×10^{-8}*** (3.16×10^{-9})	-1.07×10^{-8}*** (3.44×10^{-9})
Expense ratio FEE	-0.324 (0.437)	-0.416 (0.417)	-0.363 (0.439)
Fund age $\log_e(AGE_{t-1})$	0.227 (0.217)	0.070 (0.218)	-0.123 (0.263)
Observations	1365	1269	1003
Adj. R^2	0.006	0.007	0.008

To determine the robustness of the baseline models, correlation analysis is done for the used variables to determine whether there is significant multicollinearity between the variables and these are reported in Table 11. A strong correlation between the variables can hurt the accuracy of the model's coefficient estimates and how much each variable actually contributes on the dependent variable. Values of 1 or -1 indicate perfect relationship, and values above 0.7 for either direction would suggest high correlation. The data tested within their own groups suggest that there are no large correlations between the independent variables, with the highest being around 0.47 for the passive group between age and size. This suggest that there is some tendency for funds to grow over time.

Table 11: Correlation matrix

Table presents the correlation matrix for the variables in Table 5 and Table 6 with their respective group data from 2016-2021. Refer to Table 4 for variable explanations. First line shows the value of the active group, and second line of the passive group.

	$\log_e(AUM_{t-1})$	FEE	$\log_e(AGE_{t-1})$
$\log_e(AUM_{t-1})$	1		
	1		
FEE	0.2801	1	
	-0.1251	1	
$\log_e(AGE_{t-1})$	0.3358	-0.0374	1
	0.4689	0.1658	1

It is also worth to consider other limitations or biases that might affect the results of the thesis, which were slightly touched upon in Section 1.3. The funds were picked based on their size for both management styles out of the funds that were still in operation at that time. This means that the used data does not take into account survivorship bias by using datasets which correct for this issue. If we were to reason which way this would bias the results, it is likely that passively managed ETFs are more stable and do not go out of business due to bad performance. In contrast, actively managed funds are more likely to utilize more volatile strategies that can cause issues to the fund stability and force them out of business. This could bias the active fund sample to consist of funds that performed better while the bad performers have disbanded, driving up the performance of the average active ETF. The time range subject to the study also includes the partial effect of the pandemic which can limit the generalization of the results.

5 CONCLUSIONS

In this thesis, I document the role and impact the fund characteristics play in the performance of both active and passive exchange-traded funds. For passive ETFs the findings show a statistically significant positive relation for fund's age and a negative relation for expense ratio and assets under management against the fund's abnormal i.e. market-adjusted returns. The expense ratio with a large negative effect on performance is the most significant fund characteristic to look for, and this negative relation is consistent with nearly all of the previous literature e.g. Elton et al. (2019a) and Paudel and Naka (2023). Fund age is found to have a positive impact on the performance suggesting improvement through learning, and this effect has been found to be stronger for smaller funds by size (Paudel & Naka 2023). The passive ETFs also exhibit diseconomies of scale effect when measured by AUM, being statistically significant for one period. The active ETFs similarly show negative relation between expense ratio and abnormal returns with the effect being less than half of what was observed for passive, though the result remains statistically insignificant.

When the two fund management types are compared to each other and when they are assumed to react similarly in respect to the model variables, the results initially suggest that active ETFs were to perform better compared to passive ones, measured through a dummy variable being statistically significant at 95% confidence level. This result implies that actively managed ETFs are able to generate 0.64% to 0.66% higher monthly abnormal returns over the passive ones ceteris paribus. The combined sample and other variables with their relations mirror closely the results of the passive group from earlier, with the expense ratio again being statistically significant. Though when the assumption of the two groups behaving similarly in regards to the other variables in the model is relaxed and an interaction term is introduced, the dummy variable for active management loses its statistical significance. This finding would suggest that specifically among ETFs the management style in itself does not have significant predictive power, rather that it is driven by other characteristics of the fund which can be accounted for within the model. Even if statistically insignificant, the active dummy still appears to remain positive despite of a combination of interaction terms introduced. This suggests that there is could still be a positive effect which cannot be explained by the existing model variables. This thesis further provides insight into the relationship between exchange-traded fund characteristics, their performance, and the differences between active and passive management.

5.1 Future research

There are still aspects of the fund characteristics which are yet to be thoroughly researched or which would provide great additions to the analysis. Even though expense ratio covers most of the fund's expenses and is a statistically significant factor in predicting the fund's performance, its components such as index provider fees and their impact is less researched. These index licensing fees comprise roughly third of the overall expense ratio and have previously even caused shifts in large funds which found it beneficial to change their index providers, seeing large savings for their investors. It is suggested that competition among index providers could decrease the expense ratios by up to 30% and through that improve the performance of these index funds (An et al. 2023).

Some other aspects that warrant further research and which could be instrumental in explaining the gap between the two asset management styles and fund's performance are the impact of trading costs and securities lending. Something that is common between these two is the difficulty of measurement and obtaining the information from the funds. Often the trading activity costs are proxied by fund turnover, and they have shown promise in explaining some of the observed negative alpha (Malkiel & Saha 2020). When it comes to index funds specifically, they have very few ways to stand out against the competition. As these funds tend to stick to their assets for a long period of time, they can take advantage of this holding by participating in securities lending. In previous literature this has been shown to have a significant positive impact on the index fund's performance (Elton et al. 2019a). This would be a valuable addition as an explanatory factor, but obtaining this information is difficult as this data is not often reported anywhere by the funds. Also more studies are needed about the size of the index premium with recent data to determine how much it has changed over the years, whether it is permanent or sees reversal and whether it can still account for some of the active versus passive gap.

With the market around ETFs being larger than ever and growing, studies looking into indicators that predict performance are invaluable. And if these exist, the investors should know what the characteristics are to keep an eye out for when choosing the correct fund.

REFERENCES

Adams, J. – Hayunga, D. – Mansi, S. (2022) Index fund trading costs are inversely related to fund and family size. *Journal of banking & finance*, Vol. 140, Article 106257.

Agarwal, V. – Ren, H. – Shen, K. – Zhao, H. (2023) Redemption in Kind and Mutual Fund Liquidity Management, *The Review of Financial Studies*, Vol. 36, No. 6, p. 2274–2318.

An, Y. – Benetton, M. – Song, Y. (2023) Index providers: Whales behind the scenes of ETFs. *The Journal of Financial Economics*, Vol. 149, No. 3, p. 407–433.

Anadu, K. – Kruttli, M. – McCabe, P. – Osambela, E. (2020) The Shift from Active to Passive Investing: Risks to Financial Stability? *Federal Reserve Bank of Boston, Supervisory Research and Analysis Unit.*

ARK Invest Holdings, ARKK Holdings. <https://ark-funds.com/wp-content/fundsiteliterature/holdings/ARK_INNOVATION_ETF_ARKK_HOLDINGS.pdf>, retrieved 16.9.2021.

ARK Invest About, About ARK. <https://ark-funds.com/about-ark>, retrieved 16.9.2021.

Baker, M. – Wurgler, J. (2006) Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance*, Vol. 61, No. 4, p. 1553–2068.

Belton, P. (2017) Asset managers turn to pure performance fee model to fight off passive funds. *SNL Financial Services Daily; Charlotsville*.

Ben-David, I. – Franzoni, F. – Moussawi, R. (2018) Do ETFs Increase Volatility? *The Journal of finance (New York)*, Vol. 73, No. 6, p. 2471–2535.

Bennett, B. – Stulz, R. – Wang. Z. (2020) Does Joining the S&P 500 Index Hurt Firms? *NBER Working Paper Series*, 2020-07, Working Paper 27593.

Berk, J. – Green, R. (2004) Mutual Fund Flows and Performance in Rational Markets. *The Journal of political economy*, Vol. 112, No. 6, p. 1269–1295.

Bhide, A. (1993) The hidden costs of stock market liquidity. *Journal of Financial Economics*, Vol. 34, No. 1, p. 31–51.

Biktimirov, E.N. – Cowan, A.R. – Jordan, B.D. (2004) Do demand curves for small stocks slope down? *The Journal of Financial Research*, Vol. 27, No. 2, p. 161–178.

Biktimirov, E.N. – Xu, Y. (2019) Asymmetric stock price and investor awareness reactions to changes in the Nasdaq 100 index. *Journal of Asset Management*, Vol. 20, No. 2, p. 134–145.

Blackrock, iShares Activating the power of portfolios: Securities lending for iShares domiciled in Europe. <https://www.ishares.com/uk/professional/en/literature/brochure/ishares-securities-lending-en-emea-pc.pdf>, retrieved 24.9.2024.

Bogle, J. (2014) The Arithmetic of "All-In" Investment Expenses. *Financial analysts journal*, Vol. 70, No. 1, p. 13–21.

Boulton, T.J. – Smart, S.B. – Zutter, C.J. (2020) Worldwide short selling regulations and IPO underpricing. *Journal of Corporate Finance*, Vol. 62, 101596.

Bradford de Long, J. – Shleifer, A. – Summers, L.H. – Waldmann, R.J. (1990) Positive Feedback Investment Strategies and Destabilizing Rational Speculation. *The Journal of Finance*, Vol. 45, No. 2, p. 379–395.

Brown, D. – Wu, Y. (2016) Mutual Fund Flows and Cross-Fund Learning within Families. *The Journal of finance (New York)*, Vol. 71, No. 1, p. 383–424.

Bruce, P.C. (2014) Introductory statistics and analytics: a resample perspective. 1st ed. John Wiley & Sons, Inc, New Jersey.

Chen, H. – Noronha, G. – Singal, V. (2004) The Price Response to S&P 500 Index Additions and Deletions: Evidence of Asymmetry and a New Explanation. *The Journal of Finance*, Vol. 59, No. 4, p. 1901–1929.

Chen, H. – Singal, V. – Whitelaw, R. (2016) Comovement revisited. *Journal of financial economics*, Vol. 121, No. 3, p. 624–644.

Chen, J. – Hong, H. – Huang, M. – Kubik, J.D. (2004) Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization. *The American economic review*, Vol. 94, No. 5, p. 1276–1302.

Cici, G. – Dahm, L. – Laura, K. (2018) Trading efficiency of fund families: Impact of fund performance and investment behavior. *Journal of banking & finance*, Vol. 88, p. 1–14.

Claessens, S. – Yafeh, Y. (2013) Comovement of Newly Added Stocks with National Market Indices: Evidence from Around the World. *Review of Finance*, Vol. 17, No. 17, p. 203–227.

Corder, G.W. – Foreman, D.I. (1972) Nonparametric Statistics: A Step-by-Step Approach. 2nd ed. John Wiley & Sons, Inc, New Jersey.

Crane, A. – Crotty, K. (2018) Passive versus Active Fund Performance: Do Index Funds Have Skill? *Journal of financial and quantitative analysis*, Vol. 53, No. 1, p. 33–64.

Cremers, M. – Ferreira, M.A. – Matos, P. – Starks, L. (2016) Indexing and active fund management: International evidence. *Journal of Financial Economics*, Vol. 120, No. 3, p. 539–560.

Da, Z. – Shive, S. (2018) Exchange traded funds and asset return correlations. *European financial management : the journal of the European Financial Management Association*, Vol. 24, No. 1, p. 136–168.

De Franco, C. (2021) Stock picking in the US market and the effect of passive investments. *Journal of asset management*, Vol. 22, No. 1, p. 1–10.

Denis, D.K. – McConnell, J.J. – Ovtchinnikov, A.V. – Yu, Y. (2003) S&P 500 Index Additions and Earnings Expectations. *The Journal of Finance*, Vol. 58, No. 5, p. 1821–1840.

Dunham, L. – Jorgensen, R. – Washer, K. (2016) Securities Lending Activities in Mutual Funds and ETFs: Ethical Considerations. *Journal of business ethics*, Vol. 139, No. 1, p. 21–28.

Edminster, R.O. – Graham, A.S. – Pirie, W.L. (1996) Trading cost expectations: Evidence from S&P 500 index replacement stock announcements. *Journal of Economics and Finance*, Vol. 20, No. 2, p. 75–85.

Elliott, W. – Warr, R. (2003) Price Pressure on the NYSE and Nasdaq: Evidence from S&P 500 Index Changes. *Financial management*, Vol. 32, No. 3, p. 85–99.

Elton, E. – Gruber, M. – Blake, C. (2012) Does Mutual Fund Size Matter? The Relationship Between Size and Performance. *Review of asset pricing studies*, Vol. 2, No. 1, p. 31–55.

Elton, E. – Gruber, M. – de Souza, A. (2019a) Passive mutual funds and ETFs: Performance and comparison. *Journal of banking & finance*, Vol. 106, p. 265–275.

Elton, E. – Gruber, M. – de Souza, A. (2019b) Are Passive Funds Really Superior Investments? An Investor Perspective. *Financial analysts journal*, Vol. 75, No. 3, p. 7–19.

ETFdb, ETF Database - About Us. <https://etfdb.com/about/>, retrieved 14.4.2021.

Evans, R. – Ferreira, M. – Prado, M. (2017) Fund Performance and Equity Lending: Why Lend What You Can Sell? *Review of Finance*, Vol. 21, No. 3, p. 1093–1121.

Fama, E. (1965) The Behavior of Stock-Market Prices. *The Journal of Business*, Vol. 38, No. 1, p. 34–105.

Fama, E. (1970) Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, Vol. 25, No. 2, p. 383–417.

Ferreira, M. – Keswani, A. – Miguel, A. – Ramos, S. (2013) The Determinants of Mutual Fund Performance: A Cross-Country Study. *Review of Finance*, Vol. 17, No. 2, p. 483–525.

Field, A. (2009) Discovering Statistics Using SPSS. 3rd ed. SAGE Publications Ltd, London.

Filip, D. (2018) The impact of fund attributes on performance: Empirical evidence for Polish equity funds. *Zbornik radova Ekonomskog fakulteta u Rijeci*, Vol. 36, No. 2, p. 465–488.

Financial Times, Fidelity International outlines sliding management fee scale. 29/11/2017. <https://www.ft.com/content/7a7c37b9-7db7-3965-a432-3b21e450e5ff>, retrieved 19.9.2024.

Gelman, A. – Hill, J. (2007) Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York.

Golec, J. (1996) The Effects of Mutual Fund Managers' Characteristics on Their Portfolio Performance, Risk and Fees. *Financial services review*, Vol. 5, No. 2, p. 133–148.

Grinblatt, M. – Jostova, G. – Petrasek, L. – Philipov, A. (2020) Style and Skill: Hedge Funds, Mutual Funds, and Momentum. *Management science*, Vol. 66, No. 12, p. 5505–5531.

Hacıbedel, B. (2014) Does investor recognition matter for asset pricing? *Emerging Markets Review*, Vol. 21, p. 1–20.

Harris, L. – Gurel, E. (1986) Price and Volume Effects Associated with Changes in the S&P 500 List: New Evidence for the existence of Price Pressures. *The Journal of Finance*, Vol. 41, No. 4, p. 815–829.

Hrazdil, K. (2010) S&P 500 index inclusion announcements: does the S&P 500 committee tell us something new? *Managerial Finance*, Vol. 36, No. 5, p. 368–393.

Jain, A. – Tantri, P. – Thirumalai, R.S. (2019) Demand curves for stocks do not slope down: Evidence using an exogenous supply shock. *Journal of Banking and Finance*, Vol. 104, p. 19–30.

Khorana, A. – Servaes, H. – Tufano, P. (2009) Mutual Fund Fees around the World. *The Review of financial studies*, Vol. 22, No. 3, p. 1279–1310.

Kindleberger, C.P. – Aliber, R.Z. (2011) Manias, panics and crashes: a history of financial crises. 6th edition, Palgrave Macmillan, p. 46.

Kumar, A. (2009) Hard-to-Value Stocks, Behavioral Biases, and Informed Trading. *The Journal of Financial and Quantitative Analysis*, Vol. 44, No. 6, p. 1375–1401.

Kwak, S.G. – Kim, J.H. (2017) Central limit theorem: the cornerstone of modern statistics. *Korean Journal of Anesthesiology*, Vol. 70, No. 2, p. 144–156.

Liao, Y. – Coakley, J. – Kellard, N. (2022) Index tracking and beta arbitrage effects in comovement. *International review of financial analysis*, Vol. 83, Article 102330.

Liu, S. (2000) Changes in the Nikkei 500: New Evidence for Downward Sloping Demand Curves for Stocks. *International Review of Finance*, Vol. 1, No. 4, p. 245–267.

Ljungqvist, A. (2007) Handbook of Empirical Corporate Finance SET, Chapter 7 – IPO Underpricing. 1st edition, Elsevier 2007, p. 375–422.

Malkiel, B. – Saha, A. (2020) Mutual Fund Returns and Their Characteristics: A Simple Approach to Selecting Better Performing Actively-Managed Funds. *The Journal of investing*, Vol. 29, No. 3, p. 63–75.

Merton, R.C. (1987) A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, Vol. 42, No. 3, p. 483–510.

Morningstar (2019) Asset Management in an Era of Cost Pressure: An Index Response, Morningstar, November 2019.

Morningstar (2023) The Return of the Bond Market. <https://www.morningstar.com/bonds/return-bond-market>, retrieved 22.10.2024.

Morningstar (2024) Morningstar's US Active/Passive Barometer Mid-Year 2024. <https://www.morningstar.com/lp/active-passive-barometer>, retrieved 22.11.2024.

Nie, D. (2024) Index Premium Trends Resulting from Composition Changes to the S&P 500 and Its Implications for Market Efficiency. *Journal of accounting, business & management*, Vol. 31, No. 1, p. 1–10.

Nyakurukwa, K. – Seetharam, Y. (2023) Alternatives to the efficient market hypothesis: an overview. *Journal of Capital Markets Studies*, Vol. 7, No. 2, p. 111–124.

O'Dell, A. (2016) FCA sets its sights on index-hugging managers. *Global Investor*, 2016-11, ISSN 0951-3604.

Pástor L. – Stambaugh, R. – Taylor, L. (2015) Scale and skill in active management. *Journal of financial economics*, Vol. 116, No. 1, p. 23–45.

Pástor, L. – Stambaugh, R. – Taylor, L. (2017) Do Funds Make More When They Trade More? *The Journal of finance (New York)*, Vol. 72, No. 4, p. 1483–1528.

Pástor, L. – Stambaugh, R. – Taylor, L. (2020) Fund tradeoffs. *Journal of financial economics*, Vol. 138, No. 3, p. 614–634.

Patel, N. – Welch, I. (2017) Extended Stock Returns in Response to S&P 500 Index Changes. *Review of asset pricing studies*, Vol. 7, No. 2, p. 172–208.

Paudel, K. – Naka, A. (2023) Effects of size on the exchange-traded funds performance. *Journal of asset management*, Vol. 24, No. 6, p. 474–484.

Pedersen, L.H. (2018) Sharpening the arithmetic of active management. *Financial Analysts' Journal*, Vol. 74, No. 1, p. 21.

Petajisto, A. (2011) The index premium and its hidden cost for index funds. *Journal of Empirical Finance*, Vol. 18, No. 2, p. 271–288.

Raddatz, C. – Schmukler, S.L. – Williams, T. (2017) International asset allocations and capital flows: The benchmark effect. *Journal of international economics*, Vol. 108, p. 413–430.

Schwab, Why Diversification Matters, July 23, 2019. <https://www.schwab.com/learn/story/why-diversification-matters>, retrieved 5.5.2024.

Sharpe, W.F. (1966) Mutual Fund Performance. *The Journal of business,* Vol. 39, No. 1, p. 119–138.

Sharpe, W.F. (1991) The Arithmetic of Active Management. *Financial analysts journal*, Vol. 47, No. 1, p. 7–9.

Sheng, J. – Simutin, M. – Zhang, T. (2023) Cheaper Is Not Better: On the 'Superior' Performance of High-Fee Mutual Funds. *The Review of Asset Pricing Studies*, Vol. 13, No. 2, p. 375–404.

Sherrill, D. – Stark, J. (2018) ETF liquidation determinants. *Journal of empirical finance*, Vol. 48, p. 357–373.

Shleifer, A. – Summers, L. (1990) The Noise Trader Approach to Finance. *The Journal of Economic Perspectives*, Vol. 4, No. 2, p. 19–33.

Shleifer, A. (1986) Do Demand Curves for Stocks Slope Down? *The Journal of Finance*, Vol. 41, No. 3, p. 579–590.

Sialm, C. – Zhang, H. (2020) Tax-Efficient Asset Management: Evidence from Equity Mutual Funds. *The Journal of finance (New York)*, Vol. 75, No. 2, p. 735–777.

Sushko, V. – Turner, G. (2018) The implications of passive investing for securities markets. *BIS Quarterly Review*, March 2018.

Taylor, J. (2004) A note on closet-indexing. *Journal of economics and business*, Vol. 56, No. 6, p. 431–441.

Tokic, D. (2020) The passive investment bubble. *The Journal of Corporate Accounting & Finance*. Vol. 31, No. 1, p. 7–11.

Tuzun, T. (2013) Are Leveraged and Inverse ETFs the New Portfolio Insurers? *Finance and Economics Discussion Series (FEDS),* 2013.

Vanguard, Vanguard Russell 2000 ETF (VTWO). <https://investor.vanguard.com/etf/profile/VTWO>, retrieved 26.2.2021.

Vanguard, Vanguard S&P 500 ETF (VOO). <https://investor.vanguard.com/etf/profile/VOO>, retrieved 26.2.2021.

Vidal, M. – Vidal-Garciá, J. – Lean, H. – Uddin, G. (2015) The relation between fees and return predictability in the mutual fund industry. *Economic modelling*, Vol. 47, p. 260–270.

Wang, F. – Zheng, L. (2024) Do fund managers time momentum? Evidence from mutual fund and hedge fund returns. *European financial management : the journal of the European Financial Management Association,* Vol. 30, No. 1, p. 55–91.

Webster, D. (2002) Mutual Fund Performance and Fund Age. Available at SSRN: <https://ssrn.com/abstract=1764543>.

Wermers, R. (2000) Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transaction Costs and Expenses. *The Journal of finance (New York)*, Vol. 55, No. 4, p. 1655–1695.

YCharts, Ycharts applications. <https://get.ycharts.com/applications/>, retrieved 15.9.2021.

Zhu, M. (2018) Informative fund size, managerial skill, and investor rationality. *Journal of financial economics*, Vol. 130, No. 1, p. 114–134.

APPENDICES

Appendix 1. Sample of 39 largest actively managed ETFs sorted by assets under management as of August 2021.

Ticker	Name	AUM	Fee Ratio	Inception Date	Type
ARKK	ARK INNOVATION ETF	20.75B	0.75%	2014-10-30	Equity
ARKG	ARK GENOMIC REVOLUTION ETF	7.844B	0.75%	2014-10-30	Equity
ARKW	ARK NEXT GENERATION INTERNET ETF	5.498B	0.79%	2014-09-30	Equity
ARKQ	ARK AUTNS.TECH.& ROBOTICS ETF	2.603B	0.75%	2014-09-30	Equity
EMLP	FST.NAE.EN.INFR.ETF	2.028B	0.96%	2012-06-20	Equity
BLOK	AMPLIFY TRFM.DATA SHARING ETF	1.145B	0.71%	2018-01-16	Equity
SECT	MAIN SECTOR ROTATION ETF	935.65M	0.80%	2017-09-05	Equity
VFVA	VANGUARD US VALUE FACTOR ETF	398.33M	0.14%	2018-02-13	Equity
DUSA	DAVIS SELECT US EQUITY ETF	387.13M	0.62%	2017-01-11	Equity
DWLD	DAVIS SELECT WORLDWIDE ETF	372.45M	0.63%	2017-01-11	Equity
SYLD	CAMBRIA SHAREHOLDER YIELD ETF	275.55M	0.59%	2013-05-13	Equity
AMZA	INFRACAP MLP ETF	265.41M	2.01%	2014-10-01	Equity
PJUL	INNOVATOR S&P 500 POWER BUFFER ETF	261.40M	0.79%	2018-08-07	Equity
DINT	DAVIS SELECT INTERNATIONAL ETF	260.02M	0.65%	2018-03-01	Equity
CACG	CLEARBRIDGE ALL CAP GROWTH ESG ETF	223.23M	0.53%	2017-05-03	Equity
PHDG	INVESCO S&P 500 DOWNSIDE HEDGED ETF	218.45M	0.40%	2012-12-05	Equity
DFNL	DAVIS SELECT FINANCIAL ETF	215.65M	0.64%	2017-01-11	Equity
BATT	AMPLIFY LITHIUM & BATTERY TECHNOLOGY ETF	201.66M	0.59%	2018-06-04	Equity
CCOR	CORE ALTERNATIVE ETF	193.31M	1.09%	2017-05-23	Equity
TTAC	TRIMTABS US FREE CASH FLOW QUALITY ETF	192.43M	0.59%	2016-09-27	Equity
RFDI	FST.RIVERFRONT DYM. DEVD.INTL.ETF	167.67M	0.83%	2016-04-14	Equity
LRGE	CLEARBRIDGE LARGE CAP GROWTH ESG ETF	167.23M	0.60%	2017-05-22	Equity
AIEQ	AI POWERED EQUITY ETF	161.85M	0.80%	2017-10-17	Equity
VFMO	VANGUARD US MOMENTUM FACTOR ETF	151.35M	0.13%	2018-02-13	Equity
FLLV	FRANKLIN LIBERTY US LOW VOLATILITY ETF	136.14M	0.29%	2016-09-20	Equity
IETC	ISHARES EVOLVED US TECHNOLOGY ETF	130.65M	0.18%	2018-03-21	Equity
RFDA	RIFT.DYM.US DIV. ADVG. ETF	129.01M	0.52%	2016-06-06	Equity
OSCV	OPUS SMALL CAP VALUE ETF	125.27M	0.79%	2018-07-18	Equity
GVAL	CAMBRIA GLOBAL VALUE ETF	123.09M	0.65%	2014-03-11	Equity
HUSV	FIRST TRUST HORIZON MANAGED VOL DOMESTIC ETF	122.35M	0.70%	2016-08-24	Equity
VFQY	VANGUARD US QUALITY FACTOR ETF	110.86M	0.13%	2018-02-13	Equity
VFMF	VANGUARD US MULTIFACTOR FUND ETF	98.64M	0.19%	2018-02-13	Equity
RESP	WISDOMTREE US ESG FUND	96.62M	0.28%	2007-02-23	Equity

Ticker	Name	AUM	Fee Ratio	Inception Date	Type
DGRE	WISDOMTREE EMM.QLT.DIV. GROWTH FUND	89.91M	0.32%	2013-08-01	Equity
HDMV	FIRST TRUST HORIZON MANAGED VOL DEV INTL ETF	87.52M	0.80%	2016-08-24	Equity
AADR	ADVISORSHARES DORSEY WRIGHT ADR ETF	82.21M	1.10%	2010-07-20	Equity
EYLD	CAMBRIA EMERGING SHAREHOLDER YIELD ETF	73.67M	0.69%	2016-07-13	Equity
RFFC	RIVERFRONT DYNAMIC US FLEX-CAP ETF	63.36M	0.52%	2016-06-06	Equity
FTHI	FIRST TRUST BUYWRITE INCOME ETF	45.83M	0.85%	2014-01-06	Equity

Appendix 2. Sample of 100 largest passively managed ETFs sorted by assets under management as of August 2021.

Ticker	Name	AUM	Fee Ratio	Inception Date	Type
SPY	SPDR S&P 500 ETF TRUST	388.85B	0.09%	1993-01-22	Equity
IVV	ISHARES CORE S&P 500 ETF	296.00B	0.03%	2000-05-15	Equity
VTI	VANGD.TTL.STK.MIF. ETF	267.56B	0.03%	2001-05-24	Equity
VOO	VANGUARD 500 INDEX FUND ETF	251.62B	0.03%	2010-09-07	Equity
QQQ	INVESCO QQQ TRUST SERIES 1	183.82B	0.20%	1999-03-10	Equity
VEA	VANGUARD DEVELOPED MARKETS INDEX FUND ETF	104.42B	0.05%	2007-07-20	Equity
IEFA	ISHARES CORE MSCI EAFE ETF	99.11B	0.07%	2012-10-18	Equity
AGG	ISHARES CORE US AGGREGATE BOND ETF	88.98B	0.04%	2003-09-22	Bond
VTV	VANGUARD VALUE ETF	84.94B	0.04%	2004-01-26	Equity
VUG	VANGUARD GROWTH INDEX FUND ETF	84.70B	0.04%	2004-01-26	Equity
BND	VANGUARD TOTAL BOND MARKET INDEX FUND ETF	80.86B	0.03%	2007-04-03	Bond
VWO	VANGD.EMM.STK.IX. ETF	77.95B	0.10%	2005-03-04	Equity
IEMG	ISHARES CORE MSCI EMERGING MARKETS ETF	76.14B	0.11%	2012-10-18	Equity
IWF	ISHARES RUSSELL 1000 GROWTH ETF	72.55B	0.19%	2000-05-22	Equity
IJR	ISHARES CORE S&P SMALL- CAP ETF	66.37B	0.06%	2000-05-22	Equity
VIG	VANGD.DIV.APPREC. IX.ETF	63.64B	0.06%	2006-04-21	Equity
IWM	ISHARES RUSSELL 2000 ETF	62.81B	0.19%	2000-05-22	Equity
IJH	ISHARES CORE S&P MID-CAP ETF	61.76B	0.05%	2000-05-22	Equity
GLD	SPDR GOLD SHARES	58.15B	0.40%	2004-11-18	Equity
EFA	ISHARES MSCI EAFE	57.01B	0.32%	2001-08-14	Equity
IWD	ISHARES RUSSELL 1000 VALUE ETF	54.34B	0.19%	2000-05-22	Equity
VO	VANGUARD MID-CAP INDEX FUND ETF	51.88B	0.04%	2004-01-26	Equity
VGT	VANGD.ITECH.IX.ETF	50.21B	0.10%	2004-01-26	Equity
VXUS	VANGUARD TTL.INTL.STK. ETF	49.43B	0.08%	2011-01-26	Equity
VB	VANGUARD SMALL-CAP INDEX FUND ETF	46.73B	0.05%	2004-01-26	Equity
VCIT	VANGD.INTM.-TERM CBD.ETF	46.66B	0.05%	2009-11-19	Bond
BNDX	TOTAL INTERNATIONAL BOND ETF	45.12B	0.08%	2013-05-31	Bond
XLK	TECHNOLOGY SELECT SECTOR SPDR FUND	44.70B	0.12%	1998-12-16	Equity
VNQ	VANGUARD REAL ESTATE INDEX FUND ETF	44.02B	0.12%	2004-09-23	Equity
LQD	ISHARES IBOXX $ INV GRADE CORPORATE BOND ETF	42.94B	0.14%	2002-07-22	Bond
ITOT	ISHARES CORE S&P TTL.US STK.MKT.ETF	41.77B	0.03%	2004-01-20	Equity
XLF	FINANCIAL SELECT SECTOR SPDR FUND	41.51B	0.12%	1998-12-16	Equity
VCSH	VANGD.SHORT-TERM CBD.ETF	41.25B	0.05%	2009-11-19	Bond
BSV	VANGUARD SHORT-TERM BOND INDEX FUND ETF	41.00B	0.05%	2007-04-03	Bond
BIV	VANGD.INTERMEDIATE- TERM BD.IX.ETF	39.85B	0.05%	2007-04-03	Bond
VYM	VANGUARD HIGH DIVIDEND YIELD ETF	38.82B	0.06%	2006-11-10	Equity

Ticker	Name	AUM	Fee Ratio	Inception Date	Type
IVW	ISHARES S&P 500 GROWTH ETF	36.10B	0.18%	2000-05-22	Equity
VEU	VANGUARD FTSE ALL WORLD EX US ETF	34.85B	0.08%	2007-03-02	Equity
XLV	HEALTH CARE SELECT SECTOR SPDR FUND	33.07B	0.12%	1998-12-16	Equity
TIP	ISHARES TIPS BOND	32.33B	0.19%	2003-12-04	Bond
SCHX	SCHWAB US LARGE-CAP ETF	32.07B	0.03%	2009-11-03	Equity
IWB	ISHARES RUSSELL 1000 ETF	30.60B	0.15%	2000-05-15	Equity
DIA	SPDR DJ.INAG.ETF TST.	30.54B	0.16%	1998-01-13	Equity
IXUS	ISHARES CORE MSCI TOTAL INTL.STK.ETF	29.87B	0.09%	2012-10-18	Equity
IWR	ISHARES RUSSELL MID-CAP ETF	29.33B	0.19%	2001-07-17	Equity
EEM	ISHARES MSCI EMERGING MARKETS ETF	29.25B	0.68%	2003-04-07	Equity
RSP	INVESCO S&P 500 EQL WGHT ETF	28.95B	0.20%	2003-04-24	Equity
USMV	ISHARES MSCI USA MIN VOL FACTOR ETF	28.61B	0.15%	2011-10-18	Equity
IAU	ISHARES GOLD TRUST	28.40B	0.25%	2005-01-21	Equity
SCHF	SCHWAB INTERNATIONAL EQUITY ETF	27.36B	0.06%	2009-11-03	Equity
SCHD	SCHWAB US DIVIDEND EQUITY ETF	27.28B	0.06%	2011-10-20	Equity
VV	VANGUARD LGCP.ETF	25.94B	0.04%	2004-01-27	Equity
IGSB	ISHARES 1-5 YR.INV. GDE. CPRT.BD ETF	25.38B	0.06%	2007-01-05	Bond
MBB	ISHARES MBS	24.86B	0.06%	2007-03-13	Equity
QUAL	ISHARES MSCI USA QUALITY FACTOR ETF	24.76B	0.15%	2013-07-16	Equity
VBR	VANGUARD SMALL-CAP VALUE INDEX FUND ETF	24.32B	0.07%	2004-01-26	Equity
VT	VANGD.TTL.WLD.STK. IX. ETF	24.08B	0.08%	2008-06-24	Equity
MUB	ISHARES NATIONAL MUNI BOND ETF	23.37B	0.07%	2007-09-07	Bond
IVE	ISHARES S&P 500 VALUE ETF	22.73B	0.18%	2000-05-22	Equity
SCHB	SCHWAB US BROAD MARKET ETF	21.67B	0.03%	2009-11-03	Equity
VGK	VANGUARD EUROPEAN STOCK INDEX FUND ETF	21.30B	0.08%	2005-03-04	Equity
ESGU	ISHARES ESG AWARE MSCI USA ETF	21.14B	0.15%	2016-12-01	Equity
EMB	ISHARES JP MORGAN USD EMRG.MKT.BD.	20.82B	0.39%	2007-12-17	Bond
XLE	ENERGY SELECT SECTOR SPDR FUND	20.81B	0.12%	1998-12-16	Equity
MDY	SPDR S&P MIDCAP 400 ETF	20.45B	0.23%	1995-05-04	Equity
DGRO	ISHARES CORE DIVIDEND GROWTH ETF	20.38B	0.08%	2014-06-10	Equity
PFF	ISHARES PF.&.INSCS. ETF	19.93B	0.46%	2007-03-26	Equity
HYG	ISHARES IBOXX $ HIY.CBD. ETF	19.62B	0.49%	2007-04-04	Bond
SDY	SPDR S&P DIVIDEND ETF	19.61B	0.35%	2005-11-08	Equity
SHY	ISHARES 1-3 YR.TRSY.BOND	19.17B	0.15%	2002-07-22	Bond
SCHP	SCHWAB US.TIPS ETF	19.14B	0.05%	2010-08-05	Equity
XLY	CSM.DISCRTNY.SLT. SECT. SPDR FD.	18.97B	0.12%	1998-12-16	Equity
XLI	SECT.SPDR TST.SBI INTER INDS.	18.84B	0.12%	1998-12-16	Equity
DVY	ISHARES SELECT DIVIDEND	18.02B	0.39%	2003-11-03	Equity
JPST	JPMORGAN ULTRA- SHORT INCOME ETF	17.62B	0.18%	2017-05-17	Equity

Ticker	Name	AUM	Fee Ratio	Inception Date	Type
VXF	VANGUARD EXTENDED MARKET INDEX FUND ETF	17.23B	0.06%	2001-12-27	Equity
TLT	ISHARES 20+ YR.TRSY.BOND	17.14B	0.15%	2002-07-22	Bond
ACWI	ISHARES MSCI ACWI ETF	16.96B	0.32%	2008-03-26	Equity
VHT	VANGUARD HEALTH CARE INDEX FUND ETF	16.78B	0.10%	2004-01-26	Equity
SCHG	SCHWAB US LARGE-CAP GROWTH ETF	15.89B	0.04%	2009-12-11	Equity
VLUE	ISHARES MSCI USA VALUE FACTOR ETF	15.84B	0.15%	2013-04-16	Equity
VBK	VANGUARD SMALL-CAP GROWTH INDEX FUND ETF	15.83B	0.07%	2004-01-26	Equity
IWP	ISHARES RUSSELL MID-CAP GROWTH ETF	15.71B	0.24%	2001-07-17	Equity
SCHA	SCHWAB US SMALL-CAP ETF	15.63B	0.04%	2009-11-03	Equity
GOVT	ISHARES US TREASURY BOND ETF	15.44B	0.15%	2012-02-14	Bond
IWN	ISHARES RUSSELL 2000 VALUE ETF	15.02B	0.24%	2000-07-24	Equity
VMBS	VANGD.MGE.-BACKED SECS. ETF	14.90B	0.05%	2009-11-19	Equity
XLC	COMM.SVS.SLT.SECT. SPDR FD.	14.84B	0.12%	2018-06-18	Equity
EFV	ISHARES MSCI EAFE VALUE ETF	14.82B	0.39%	2005-08-01	Equity
VOE	VANGUARD MID-CAP VALUE INDEX FUND ETF	14.69B	0.07%	2006-08-17	Equity
SHV	ISHARES SHORT TRSY.BD.	14.63B	0.15%	2007-01-05	Bond
IEF	ISHARES 7-10 YR.TRSY.BD.	14.55B	0.15%	2002-07-22	Bond
MTUM	ISHARES MSCI USA MOMENTUM FACTOR ETF	14.38B	0.15%	2013-04-16	Equity
IWS	ISHARES RUSSELL MID-CAP VALUE ETF	14.09B	0.24%	2001-07-17	Equity
MINT	PIMCO ENH.SHT.MAT. ACV. EXCH TR	14.06B	0.36%	2009-11-16	Equity
GSLC	GLDS.ACTIVEBETA US LGCP. EQ.ETF	13.54B	0.09%	2015-09-17	Equity
SCZ	ISHARES MSCI EAFE SMCP.	13.38B	0.40%	2007-12-10	Equity
GDX	VANECK VECTORS GOLD MINERS ETF	13.13B	0.52%	2006-05-16	Equity
SLV	ISHARES SILVER TRUST	12.88B	0.50%	2006-04-21	Equity
EWJ	ISHARES MSCI JAPAN ETF	11.55B	0.51%	1996-03-12	Equity