ACTIVE VERSUS PASSIVE

DOES INCREASED INVESTMENTS IN PASSIVE FUNDS HAVE AN EFFECT ON ACTIVE FUND PERFORMANCE?

ELLIOT CORELL

JACOB HJELMSTEDT

Bachelor Thesis

Stockholm School of Economics

2022

Active versus Passive

Does increased investments in passive funds have an effect on active fund performance?

Abstract:

In recent decades, there has been a shift towards investing in passively managed funds. Grossman and Stiglitz (1980) describe in their theoretical paper how the share of uninformed investors is positively correlated to the utility of being informed. Our hypothesis, based on Grossman and Stiglitz (1980) is that the excess returns of actively managed funds are positively correlated to the market proportion of passively invested capital. Using a bivariate regression model, we perform an empirical analysis of whether the excess performance of actively managed U.S. funds and the market proportion of passively invested capital are correlated. Our analysis did not find any significant results of the correlation existing when analyzing the entire sample of funds. When analyzing the correlation using smaller subsamples of funds, we found some results indicating a negative correlation, contrary to our hypotheses.

Keywords:

Index fund, Mutual fund, alpha, passive investment, active investment

Authors:

Elliot Corell (24604)

Jacob Hjelmstedt (24685)

Tutors:

Peter Alexander Alekszíev, Visiting Teacher, Department of Finance

Examiner:

Adrien d’Avernas, Assistant Professor, Department of Finance

Bachelor Thesis

Bachelor Program in Business and Economics

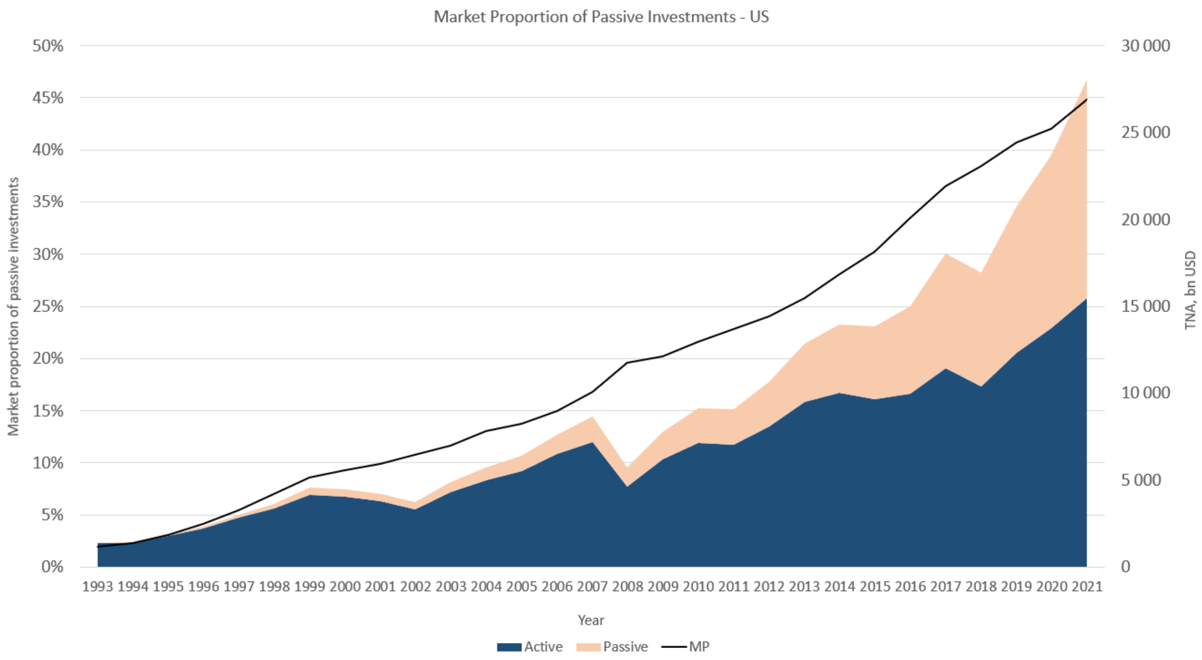
Stockholm School of Economics

1. Introduction

1.1. Background

Recent decades have seen a shift in the allocation of capital invested in actively managed funds compared to passively managed funds. In the U.S., passively managed funds went from making up only 2% of the market in 1993 to over 40% in 2021. The trend has been ongoing for a long time following influential research (Carhart, 1997) (Malkiel, 1995) (Jensen, 1968) being critical to the value added by active fund managers. After the 2008 financial crisis where actively managed funds faced especially large losses (Morningstar, 2019), the market proportion of passive investments increased more rapidly.

Figure 1:



Authors' visualization of morningstar direct data

Figure 1 illustrates the market proportion of passively invested capital in the U.S. market. On the left axis is the market proportion of passively invested capital expressed in percent. On the right axis is the aggregate TNA of all actively and passively managed funds respectively each year expressed in billions of dollars. The analysis in this paper will cover the period 2011-2021.

One of the most influential papers on active fund performance provides data suggesting that the managers' stockpicking talents have no explanatory value for the persistent success of funds but possibly for persistent underperformance of funds (Carhart, 1997). Other research shows that active fund managers have stockpicking skills that allows them to hold stocks that outperform the market, although the actively managed funds still underperform the market due to costs such as high transaction fees and other expenses (Wermers, 2000).

Under the assumption that the efficient market hypothesis holds, there should be no possible value to be added by active fund managers as the markets are informationally efficient, prices perfectly reflect all available information, and there is nothing to be gained by gathering information. It has however been reasoned that such markets are impossible due to the Grossman-Stiglitz paradox as there would be no reason to gather information, and prices could as such not perfectly reflect all available information (Grossman & Stiglitz, 1980). Instead, it is possible that while the financial markets are mostly informationally efficient, price irregularities can appear for shorter periods of time for investors gathering information to exploit (Malkiel, 2003) in accordance with the equilibrium for disequilibrium model (Grossman & Stiglitz, 1980).

The recent trend of increasingly passive investing could result in the market prices not fully reflecting all available information if it meant that the proportion of informed investors compared to uninformed investors decreased (Grossman & Stiglitz, 1980). There would then be a less efficient market where exploitable price irregularities may appear more often. Arbitrage opportunities emerging from passively managed funds has also been the topic of more recent research. Ben-David, Franzoni and Moussawi (2018) explore in a recent article how passively managed ETFs increase the volatility of the underlying assets and how active investors may exploit the arbitrage opportunities that emerge from this practice.

2. Literature & Theoretical background

2.1. Active Management Ability

There has been a lot of research done investigating whether or not managers of actively managed funds add any value. One of the most influential early papers on the subject is Jensen (1968), where, for the first time, fund managers are not only compared to other fund managers in order to find out if some managers perform better than others, but are instead compared to an absolute measure to find out if they, on aggregate, add any value at all. The results of this paper are very critical towards the value added by fund managers as it finds that, not only, do fund managers as an aggregate underperform the market, but there is in addition no individual manager who significantly outperforms what would be expected from a random chance (Jensen, 1968).

Following Jensen's research, other papers came to similar critical conclusions regarding the value added by fund managers. Malkiel (1995), after a wave of research positive to active fund management, researched the subject taking into account survivorship bias, which he deemed more important than others had estimated. The paper reaches the conclusion that active fund managers on aggregate underperform and that even though considerable performance persistence was observed in the 70s, there was no consistency in the 80s (Malkiel, 1995). In the cases where actively managed funds outperformed the market, it was found that other factors besides manager talents could explain the persistent overperformance and that only persistent underperformance remained an anomaly (Carhart, 1997).

Another strand of the research, on the other hand, provides evidence in support of active fund managers possessing some stock-picking talents. Grinblatt and Titman (1989, 1993), Grinblatt, Titman and Wermers (1995) and Daniel et al. (1997) all find that mutual funds tend to select stocks that outperform both the market and passive benchmarks of stocks with similar characteristics. What differs between this strand of research and research critical of active fund management is that those critical of active fund management analyze the entire fund portfolio return whilst those showing stock-picking skills of fund managers analyze only the individual equity holdings of funds.

Finally, Wermers (2000) attempts to bridge the gap between the two aforementioned strands of research analyzing both equity holdings of mutual funds, their gross returns, as well as their net returns to investors. By merging the CRSP database with the CDA Investment Technologies database, Wermers (2000) is able to decompose mutual fund returns into stock-picking talent, characteristic selectivity, timing of the stocks held, trading costs and expenses. The main finding of the paper is that the gross returns of equity in mutual funds do outperform the market, but the net returns to investors still underperform due to transaction costs, expenses and other holdings (Wermers, 2000).

2.2. Efficient Markets & Passive Investing

A common assumption in finance is that the efficient market hypothesis holds, that markets are informationally efficient and prices of securities perfectly reflect all available information. Such markets are proven to be an impossibility by the Grossman-Stiglitz paradox as there would be no profit to be gained by gathering information and as such no new information would be gathered (Grossman & Stiglitz, 1980). They instead propose a model where there is an equilibrium for disequilibrium so that prices only partially reflect the information of informed individuals so that those who spend resources to obtain information are compensated (Grossman & Stiglitz, 1980). This prediction is supported by Malkiel (2003) where he concludes that pricing irregularities may appear and even persist under short periods of time for arbitrageurs to take advantage of as markets cannot be perfectly informationally efficient (Malkiel, 2003). In their model, Grossman and Stiglitz (1980) predict that as the share of uninformed investors rises, the utility of gathering information to become an informed investor increases.

The upwards trend in passive investing’s effect on market efficiency has been investigated by Ben-David et al. (2018). In their paper they provide data suggesting that exchange traded funds (ETFs) increase the volatility of the underlying asset (Ben-David et al., 2018). In addition, they explain how this might be exploited by investors to earn arbitrage profits (Ben-David et al., 2018). The research by Ben-David et al. (2018) suggests a way in which an increase in passive investing may benefit active investors, as reasoned by Grossman and Stiglitz (1980). Other research made on the same topic has taken a theoretical approach and has arrived at an opposite conclusion (Bhattacharya & O’Hara, 2018) (Malamud, 2016). Bhattacharya and O’Hara (2018) find that while ETF trading may lead to pricing distortions in individual securities, it will, on aggregate, move prices closer to fundamentals. In a similar vein, Malamud (2016) reasons how ETFs may reduce volatility and comovement of some assets.

Another recent paper written on the subject of the impact of shifting towards passive investing is Anadu et al. (2020). Their paper describes the effect that the shift from active to passive investing seen in recent decades has on different index-inclusion effects and financial stability. Anadu et al. (2020) elaborate on the effects on volatility, described by Ben-David et al. (2018), Bhattacharya and O'Hara (2018), and Malamud (2016), that ETFs have and also discuss the effects that a shift to passive investing has on valuation, liquidity, comovement of returns and comovement of liquidity. Anadu et al. (2020) find mixed results on the impact that the shift towards passive investing has had on these different index inclusion effects. Some mechanisms associated with increased passive investing have increased the index inclusion effects or affected financial stability negatively, whilst others have done the opposite (Anadu et al., 2020).

3. Research Question

In this paper we examine historical data on the excess performance of actively managed U.S. mutual funds in relation to the performance of the two U.S. stock indices S&P 500 and NYSE/AMEX/NASDAQ/ARCA which, like in previous research, are used as an estimate for the performance of passive investing (Wermers, 2000). This data is examined in correlation to the proportions of U.S. capital that are invested actively and passively respectively during the same time period.

While the performance and value addition of active fund managers has been thoroughly researched (Jensen, 1968) (Carhart, 1997) (Wermers, 2000) as well as how increased passive investing may affect the underlying assets (Ben-David et al., 2018) (Anadu et al., 2020) (Bhattacharya & O'Hara, 2018), there has, to our knowledge, not been any research examining if actively managed funds perform better when the market proportion of passively invested capital increases. This leads us to in this paper attempt to answer the research question:

Is the performance of actively managed funds correlated to the market proportion of passively invested capital?

In investigating this question, we add to existing literature by testing whether the positive relationship between the share of uninformed, passive, investors and utility for becoming informed (Grossman & Stiglitz, 1980) exists in practice. We also add a nuance to the extensive previous research done on the subject of performance of actively managed funds by examining whether an outside factor has an effect on their performance. The results of our research could also shed light on the observed cyclical nature of the outperformance of active funds relative to passive funds (Hartford Funds, 2022).

Grossman and Stiglitz (1980) reason that as the share of uninformed investors in the market increases, the utility of spending resources to become informed increases. As passively managed funds do not gather any information about their investments, but instead only respond to market prices, they represent the uninformed investors from Grossman and Stiglitz (1980) in this paper. In our paper, we estimate the share of uninformed investors by the share of passively invested capital in the market. Additionally, Malkiel (2003) argues, contrary to the efficient market hypothesis, that exploitable price irregularities appearing for short periods of time are possible. That view is substantiated by the results from the empirical work of Ben-David et al. (2018). As such, we hypothesize that an increasingly large market proportion of passively invested capital should result in actively managed funds reaping higher rewards for staying informed.

4. Methodology

To examine if a relationship exists between the market share of passively invested capital and the performance of actively managed funds, our main analysis consists of two bivariate regression between the two factors. The first model regresses the excess return of actively managed funds relative to S&P 500 over the market proportion of passively invested capital and the second model regresses the excess return of actively managed funds relative to the NYSE/AMEX/NASDAQ/ARCA index from CRSP over the market proportion of passively invested capital. In the regressions, the standard errors for the excess returns of active funds are likely related to specific years and as such we cluster the standard errors yearly in accordance with Petersen (2009).

In addition to the regressions in our main analysis, we run additional regressions where our sample has been divided into different subsamples. In the first one of the additional analyses the sample is divided based on historical returns, and in the second one, the sample is divided by Morningstar category. In the third additional analyses, the sample is divided based on whether or not the fund is a fund of funds. Like our main analysis, additional analyses are done twice where the excess returns of active funds are first relative to the S&P 500 index and then the NYSE/AMEX/NASDAQ/ARCA index. The standard errors are clustered by year in the additional analyses as well.

4.1. Market Proportion of Passively Invested Capital

$$MP_T = \frac{IF_T}{(AF_T + IF_T)}$$

$$IF_T = Average\ aggregate\ TNA\ index\ funds\ for\ the\ year\ T$$

$$AF_T = Average\ aggregate\ TNA\ active\ funds\ for\ the\ year\ T$$

$$MP_T = Market\ proportion\ of\ TNA\ invested\ in\ index\ funds\ for\ the\ year\ T$$

The first variable needed to run our main regression is the average annual market proportion of passively invested capital in the U.S. (MP_T). Using the data retrieved from Morningstar Direct, we calculate this proportion by dividing the average annual aggregate total net assets (TNA) in passively managed U.S. funds (IF_T) by the average annual aggregate TNA of all funds. In this equation, the total TNA of all U.S. funds is made up of IF_T and AF_T, the average annual aggregate TNA of all actively managed U.S. funds.

4.2. Returns of Indices

$$R_{IF,T} = [(1 + r_{IF,t}) * (1 + r_{IF,t+1}) * (1 + r_{IF,t+2}) * \dots * (1 + r_{IF,t+11})] - 1$$

$$R_{IF,T} = Annual\ return, monthly\ compounded, for\ year\ T$$
$$r_{IF,t} = monthly\ return\ for\ month\ t$$

To obtain the annual returns of the index, the monthly returns have been compounded for each year. The annual returns from passive investing have, based on Wermers (2000), been estimated by using the annual returns ($R_{IF,T}$) from the S&P 500 index and the NYSE/AMEX/Nasdaq/ARCA index as proxies. Like Wermers (2000) we use simple returns in our analysis. The returns of the indexes, downloaded from CRSP, are expressed in monthly returns ($r_{IF,t}$) , and have been compounded monthly to obtain the annual returns ($R_{IF,T}$).

4.3. Performance of Actively Managed Funds

$$R_{Excess,T} = R_{AF,T} - R_{IF,T}$$

$$R_{AF,T} = Return\ of\ active\ fund\ for\ year\ T$$

$$R_{Excess,T} = Excess\ return\ of\ active\ fund\ in\ relation\ to\ Index$$

The performance of actively managed funds is estimated by the excess returns of actively managed funds ($R_{Excess,T}$). $R_{Excess,T}$ is the actively managed funds return ($R_{AF,T}$) above the returns of passive investing ($R_{IF,T}$). The excess return ($R_{Excess,T}$) is calculated for each active fund existing in year T. The returns from the actively managed funds are expressed annually when downloaded from Morningstar Direct.

4.4. Historical Returns of Actively Managed Funds

$$R_{AF,Historical} = [(1 + R_{AF,2011}) * (1 + R_{AF,2012}) * (1 + R_{AF,2013}) - 1]$$

$$R_{AF,Historical} = Compounded\ historical\ return\ for\ the\ years\ 2011 - 2013$$

The compounded historical return ($R_{AF,Historical}$) used in one of our additional analyses is calculated by compounding the annual return of an individual actively managed fund ($R_{AF,T}$)

over the years 2011, 2012 and 2013. The historical return is calculated for all actively managed funds in our dataset which have returns in all three of those years.

5. Data

5.1. Data Collection

We collected the historical data on U.S. mutual funds from Morningstar Direct. The dataset from Morningstar Direct covers annual data on fund returns as well as what percentage of their assets are allocated to U.S. equity, U.S. bonds, bonds and cash in 2021 for U.S. mutual funds between January 1, 2011 and December 31, 2021. In addition, the dataset shows if the fund is actively or passively managed, whether or not the fund is a fund of funds as well as what Morningstar category the fund has. The annual returns given by Morningstar Direct are expressed net of any management, administrative or 12b-1 fees, and other costs taken out of fund assets but they do not account for any sales charges. The Morningstar Direct dataset used for data on mutual funds does have a survivorship bias, meaning that it only includes funds still surviving at the end of 2021.

As our estimation of the performance of passive investing is based on Wermers (2000), the data collection for the returns of passive indices was obtained from the same source, CRSP. From CRSP we obtained value weighted monthly returns including dividends for both the S&P 500 Index and the NYSE/AMEX/NASDAQ/ARCA index between January 1, 2011 and December 31, 2021.

The data on the proportions of actively and passively invested capital in the market was gathered from Morningstar Direct. From Morningstar Direct, we collected data on the average annual aggregate TNA of all U.S. funds between 1993 and 2021, split into actively and passively managed funds.

5.2. Sample

To only include relevant funds, some of the mutual funds gathered from Morningstar Direct have been excluded from our analysis.

To only obtain funds that are actively managed, we have excluded all mutual funds that are not categorized as having an active management approach on Morningstar Direct. For the mutual funds to be comparable to the U.S. stock indices, and for it to be possible to evaluate the performance of the managers, we have excluded funds such as international funds and bond funds that typically hold or trade only small amounts of U.S. equity (Wermers, 2000).

Those funds were excluded by removing all funds with less than 50% of assets allocated to U.S. equity in 2021. After excluding these funds from our dataset, we arrive at a number of funds existing for every year during the period January 1, 2011 to December 31, 2021 displayed in Table 1.

Table 1:
Sample Size By Year

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Number of Funds	3759	3878	4010	4125	4235	4321	4403	4469	4524	4626	4660

Table 1: Presents the number of funds in the sample each year of the analysis after cleaning the data.

5.2.1. Sample Split 1: Historical performers

To extend our analysis we analyzed the historical performance of all funds existing 2011-2021. We calculate the compounded, historical return for the period 2011-2013 and conduct regressions with the four (4) new subsamples for the period 2014-2021. 1) The full 2014-2021 sample for reference, 2) The top 25 % historical performers of active funds existing 2011-2021. 3) The bottom 25 % historical performers of active funds existing 2011-2021, and 4) the middle 50 % performers of active funds existing 2011-2021. The full sample in this analysis differs from the sample in the main regressions in two ways. First of all, it only covers the period 2014-2021 instead of 2011-2021. Secondly, It only contains funds which have historical returns in the years 2011-2013. After excluding these funds from our dataset, we arrive at a number of funds existing for every year during the period January 1, 2014 to December 31, 2021 displayed in Table 2. As only funds with returns in 2011, 2012 and 2013 are included in this sample and since our dataset has a survivorship bias, the number of funds in each subsample does not change over the years.

Table 2:
Sample Divided By Historical Returns, 2014-2021

	Number of Funds
Full Sample	3759
Top 25%	939
Bottom 25%	939
Middle 50%	1881

Table 2: Presents the number of funds in each of the subsamples based on the historical performance of the fund. Bacuse only funds with returns in 2011-2013 are included and since the dataset has survivorship bias, the number of funds in each subsample is identical each year.

5.2.2. Sample Split 2: Morningstar categories

In the second additional analysis, our sample is divided into subsamples based on their Morningstar categories. After exclusions, our dataset contains actively managed funds with 43 different Morningstar categories. These categories were grouped into seven (7) different Morningstar category groups based on how they invest. The groups are 1) US Large-Cap Equity, 2) US Medium-Cap Equity, 3) US Small-Cap Equity, 4) US Unknown Equity, 5) World Equity, 6) Sector and 7) Target Date. The funds in the US Large-Cap Equity, US Medium-Cap Equity and US Small-Cap Equity groups invest in the equity of large, medium and small U.S. companies respectively. The funds in the US Unknown Equity group invest in the equity of U.S. companies without having a certain size criteria. The funds in the World Equity group invests in equity from companies from all over the world, although due to our exclusion criterias, they all have over 50% of funds allocated to U.S. equity. The funds in the Sector group invest in equity within a certain sector. The funds in the Target Date group all invest to achieve an optimal risk return blend with a certain future target date in mind. After excluding these funds from our dataset, we arrive at a number of funds existing for every year during the period January 1, 2011 to December 31, 2021 displayed in Table 3.

Table 3

Sample Divided by Morningstar Category

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
US Large-Cap Equity	1590	1635	1688	1707	1730	1759	1791	1808	1814	1832	1841
US Mid-Cap Equity	686	701	716	724	742	755	760	770	782	798	800
US Small-Cap Equity	803	820	836	861	893	916	928	942	955	969	969
US Unknown Equity	115	120	151	163	174	188	194	208	218	218	218
World Equity	269	280	288	306	323	328	355	362	364	366	381
Sector	296	322	331	364	373	375	375	379	391	395	395
Target Date	0	0	0	0	0	0	0	0	0	48	56

Table 3: Presents the number of funds each year in each of the subsamples based on Morningstar categories.

5.2.3. Sample Split 3: Fund of Funds

As a third additional analysis, we divide our original sample into two (2) different subsamples based on whether or not the fund is a fund of funds. Using the binary fund of funds status provided by Morningstar Direct, the full dataset is divided into 1) Funds of Funds and 2) All Funds Excluding Funds of Funds. After excluding these funds from our dataset, we arrive at a number of funds existing for every year during the period January 1, 2011 to December 31, 2021 displayed in Table 4.

Table 4											
Sample Divided By Fund of Funds Status											
	2011	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**
Funds of Funds	25	30	51	54	54	72	73	82	85	133	141
Excluding Funds of Funds	3734	3848	3959	4071	4181	4249	4330	4387	4439	4493	4519

Table 4: Presents the number of funds in the sample each year that are funds of funds and not funds of funds respectively.

5.3. Variables

We have performed our bivariate regressions using a formula defined as following:

$$R_{Excess,T} = \beta_0 + \beta_1 MP_T + \varepsilon_T$$

The dependent variable in our bivariate regressions, the excess return of active funds relative to the proxies for index funds (the first proxy is the S&P 500 index and the second proxy is the CRSP NYSE/AMEX/NASDAQ/ARCA index) is expressed as an annual percentage. The excess return in regression one is calculated as the annual return of each active fund subtracted with the annual return of the S&P 500 Index return for each period, T. The excess return in regression two is calculated as the annual return of each active fund subtracted with the annual return of the NYSE/AMEX/NASDAQ/ARCA index from CRSP each period, T. The independent variable is the market proportion of passive investments and is the same in both regressions. It is calculated as the amount invested in passively managed funds divided with the sum invested in actively managed funds and passively managed funds for each period. In both bivariate regressions, the dependent and explanatory variables are expressed as percentages.

5.4. Descriptive Statistics

Table 5

Summary Statistics Distributet by Year

	Number of Actively Managed Funds	Market Proportion	Average Excess Return Over S&P 500	Average Excess Return Over NYSE/AMEX/NASDAQ/ARCA
2011	3759	22,78%	-0,28%	2,71%
2012	3878	24,05%	-2,31%	-2,05%
2013	4010	25,83%	-5,38%	-3,46%
2014	4125	28,07%	10,35%	13,46%
2015	4235	30,25%	8,66%	11,73%
2016	4321	33,49%	2,34%	1,46%
2017	4403	36,54%	-16,63%	-15,30%
2018	4469	38,46%	2,16%	3,91%
2019	4524	40,68%	-0,84%	1,24%
2020	4626	42,05%	-8,79%	-10,61%
2021	4660	44,81%	2,49%	7,71%

Table 5: Presents the summary statistics for the variables used in the main regression for each respective year. The Market Propotion is the market proportion of passively invested capital in the market. The Average Excess Return Over S&P 500 and NYSE/AMEX/NASDAQ/ARCA respectively is the average excess return of individual actively managed funds over the passive index each year.

Table 5 displays the summary statistics for the variables used in our bivariate regression model, distributed over the time period 2011-2021. The market proportion of passively invested capital increases each year of our analysis with it holding a 22,78% share in 2011 and a 44,81% share in 2021. The number of actively managed funds in our sample has also increased each year of our analysis from 3759 in 2011 to 4660 in 2021. In our regressions, fund level excess returns are used. The averages of the fund level excess returns for each given year are presented in table 5. The average excess return of the actively managed funds in relation to the stock indices used for passive investments showed mixed results with some years showing the actively managed funds yielding higher returns and some years showing them yielding lower returns.

6. Empirical Analysis

6.1. Main Regression

Table 6

Main Regression Results

	Observations	Adjusted R2	MP Estimate	MP P-value	Intercept Estimate	Intercept P-value
S&P 500	47010	0,014	-0,198	0,419	0,059	0,481
NYSE/AMEX/NASDQ/ARCA	47010	0,012	-0,199	0,543	0,076	0,476

Table 6: Presents the results from the main regressions. S&P 500 and NYSE/AMEX/NASDAQ/ ARCA indicates that the results are for the regression where the excess returns are relative to that respective passive index. MP stands for the market proportion coefficient.

*Significant at 20% **Significant at 10% ***Significant at 5%

The results of the bivariate regression analysis for actively managed funds' excess return in relation to both the S&P 500 index and the NYSE/AMEX/NASDAQ/ARCA index against the market proportion of passively invested capital for the period 2011-2021 are shown in table 6. The regression generates similar results when the market proportion of passively invested capital is regressed with the excess return of active funds relative to the two respective indices. Regression 1 (S&P500), generates an estimate for the MP coefficient of -0,198, similar to the -0,199 estimate of regression 2 (NYSE/AMEX/NASDAQ/ARCA). The adjusted R-squared of regression 1 is 0,014 and the regression 2 adjusted R-squared is 0,012. Finally, both regressions have a slightly positive estimate for the intercept, 0,059 in regression 1 and 0,076 in regression 2. However, both regressions have high p-values for both the estimate of the MP coefficient as well as the estimate of the intercept. As such, the results from our two main regressions are not significant and there are no takeaways from the results other than that it is not significant.

6.2. Additional Regressions

6.2.1. Subsample: Historical Returns, 2014-2021

To extend our analysis, we divide our main sample into subsamples based on the funds historical performance i.e. the funds compounded return for the period 2011-2013. We create four subsamples, 1) all active mutual funds existing 2014-2021, 2) top 25 % historical performers, 3) Bottom 25 % historical performers, 4) Middle 50 % historical performers.

	Observations	Adjusted R2	MP Estimate	MP P-value	Intercept Estimate	Intercept P-value
Full Sample	30072	0,090	-0,723	0,096**	0,267	0,088**
Top 25%	7512	0,071	-0,639	0,027***	0,245	0,031***
Bottom 25%	7512	0,094	-0,829	0,0176*	0,292	0,170*
Middle 50%	15048	0,101	-0,711	0,105*	0,266	0,091**

Table 7

Regression Results, Historical Returns Sample Split, S&P 500

Table 7: Presents the results from the regression where the sample is split into the subsamples 1) Full Sample, 2) Top 25%, 3) Bottom 25% and 4) Middle 50% based on historical returns. The excess returns in this regression are relative to the S&P 500 index. MP stands for the market proportion coefficient.

*Significant at 20% **Significant at 10% ***Significant at 5%

In the S&P 500 regressions of the subsamples based on historical returns, presented in table 7, we find that the market proportion of passively invested capital is negatively correlated with the excess return of active funds for the period 2014-2021 in the cases where the results are significant. The MP coefficient is -0,723 for the full sample and it is significant at the 10 % level. The regression for the full sample had an adjusted R squared of 0,09. The results from subsamples 2 and 3 suggests that the top 25 % historically performing funds are negatively correlated to the market proportion of passive investments, as their coefficient is -0,639 significant at the 5 % level with an adjusted R squared of 0,071, and the bottom 25 % are also negatively correlated with a coefficient of -0,829 significant at the 20 % level with an adjusted R squared of 0,094. Finally the middle 50% of funds based on historical performance had an MP estimate of -0,711 significant at the 20% level and an adjusted R squared of 0,101. All four of the regressions generated significant positive estimates for the intercept of the model. The estimate for the intercept of the full sample is 0,267 and is significant at the 10% level. The estimate for the intercept of the top 25% is 0,245 and is significant at the 5% level. The estimate for the intercept of the bottom 25% is 0,292 and is significant at the 20% level. The estimate for the intercept of the middle 50% is 0,266 and is significant at the 10% level. These estimates for the intercept being positive suggest that on average, active funds outperform passive indices when MP is 0.

Table 8

Regression Results, Historical Returns Sample Split, NYSE/AMEX/NASDAQ/ARCA

	Observations	Adjusted R2	MP Estimate	MP P-value	Intercept Estimate	Intercept P-value
Full Sample	30072	0,082	-0,728	0,218	0,286	0,175*
Top 25%	7512	0,067	-0,645	0,127*	0,264	0,100**
Bottom 25%	7512	0,084	-0,835	0,281	0,311	0,247
Middle 50%	15048	0,090	-0,717	0,229	0,285	0,178*

Table 8: Presents the results from the regression where the sample is split into the subsamples 1) Full Sample, 2) Top 25%, 3) Bottom 25% and 4) Middle 50% based on historical returns. The excess returns in this regression are relative to the NYSE/AMEX/NASDAQ/ARCA index. MP stands for the market proportion coefficient.

*Significant at 20% **Significant at 10% ***Significant at 5%

In the NYSE/AMEX/NASDAQ/ARCA subsample regressions, presented in table 8, the results are not as significant as in the regressions of active funds excess returns relative to the S&P 500 index. The subsample for top 25 % performers is significant at the 20% level, with results similar to those of the S&P 500 regression, with a MP coefficient of -0,645 and an adjusted R squared of 0,067. All other estimates for the MP coefficient in this set of regressions are not significant. However, in these regressions we find three significant positive estimates for the intercept as well. The full sample has an estimate for the intercept of 0,286, significant at the 20% level. The top 25% sample has an estimate for the intercept of 0,264, significant at the 10% level. The middle 50% sample has an estimate for the intercept of 0,285, significant at the 20% level. Once again, these results of the intercepts suggest that on average, active funds outperform passive indices when MP is 0.

6.2.2. Subsample: Morningstar Categories, 2011-2021

For our second additional regression, our original sample has been divided into different subsamples based on their Morningstar category. Our original sample contained funds belonging to 43 different Morningstar categories which have been grouped based on their investment strategies to form 7 subsamples 1) US Large-Cap Equity, 2) US Medium-Cap Equity, 3) US Small-Cap Equity, 4) US Unknown Equity, 5) World Equity, 6) Sector, 7) Target Date.

Table 9						
Regression Results, Morningstar Category Sample Split, S&P 500						
	Observations	**Adjusted R2**	**MP Estimate**	**MP P-value**	**Intercept Estimate**	**Intercept P-value**
US Large-Cap Equity	19195	0,011	-0,155	0,562	0,056	0,545
US Medium-Cap Equity	8234	0,010	-0,155	0,399	0,049	0,456
US Small-Cap Equity	9892	0,025	-0,267	0,176*	0,078	0,209
US Unknown Equity	1967	0,030	-0,321	0,384	0,062	0,645
World Equity	3598	0,010	-0,137	0,469	0,024	0,751
Sector	3996	0,008	-0,240	0,712	0,052	0,809
Target Date	104	0,914	4,251	2,82E-12***	-1,921	2,74E-12***

Table 9: Presents the results from the regression where the sample is split into the subsamples 1) US Large-Cap Equity, 2) US Medium-Cap Equity, 3) US Small-Cap Equity, 4) US Unknown Equity, 5) World Equity, 6) Sector and 7) Target Date based on the funds Morningstar Category. The excess returns in this regression are relative to the S&P 500 index. MP stands for the market proportion coefficient.

*Significant at 20% **Significant at 10% ***Significant at 5%

Table 10						
Regression Results, Morningstar Category Sample Split, NYSE/AMEX/NASDAQ/ARCA						
	Observations	**Adjusted R2**	**MP Estimate**	**MP P-value**	**Intercept Estimate**	**Intercept P-value**
US Large-Cap Equity	19195	0,009	-0,157	0,654	0,074	0,526
US Medium-Cap Equity	8234	0,009	-0,157	0,516	0,066	0,421
US Small-Cap Equity	9892	0,021	-0,248	0,357	0,086	0,294
US Unknown Equity	1967	0,026	-0,318	0,472	0,078	0,616
World Equity	3598	0,009	-0,134	0,537	0,040	0,630
Sector	3996	0,007	-0,241	0,744	0,069	0,771
Target Date	104	0,965	6,802	2,98E-11***	-3,012	2,74E-12***

Table 10: Presents the results from the regression where the sample is split into the subsamples 1) US Large-Cap Equity, 2) US Medium-Cap Equity, 3) US Small-Cap Equity, 4) US Unknown Equity, 5) World Equity, 6) Sector and 7) Target Date based on the funds Morningstar Category. The excess returns in this regression are relative to the NYSE/AMEX/NASDAQ/ARCA index. MP stands for the market proportion coefficient.

*Significant at 20% **Significant at 10% ***Significant at 5%

When divided into these subsamples, most of the generated results are not significant and there is as such nothing to interpret based on those results. However, table 9 shows that when the excess return is relative to the S&P 500 index, the estimate for the MP coefficient in the US Small-Cap Equity subsample is significant at the 20% level. That MP estimate has a value of -0,248 and the regression has an adjusted R-squared of 0,021. This indicates that the market proportion of passive investments and the excess returns of funds investing in U.S. Small-Cap Equity relative to the S&P 500 index have a negative relationship. These results should however be interpreted carefully as a significance on the 20% level is relatively weak.

In addition, both table 9 and table 10 show a strong significance for the estimates of the MP coefficient and the intercept in the Target Date subsample. In both regressions, this subsample has a positive MP coefficient and a negative intercept, both of which have p-values showing that the estimates are significant at the 5% level. In addition, the adjusted r-squared of the first regression is 0,914 and that of the second regression is 0,965, indicating a strong explanatory value of the market proportion on the excess returns. However, while both regressions have 104 observations, the funds present in this subsample only have returns documented in 2020 and 2021. This is an issue as we have clustered the standard errors by year, and even if clustered on the correct dimension, too few clusters will create biased standard errors and a too small confidence interval (Petersen, 2009). As there are only two clusters in the target date category, the bias will be very large (Petersen, 2009).

6.2.3. Subsample: Fund of Funds, 2011-2021

For our third additional regression analysis, we have divided our original sample based on whether or not the fund is a fund of funds. Based on the fund of funds status in the Morningstar Direct database, our sample has been divided into 2 subsamples 1) Funds of Funds and 2) All Funds Excluding Funds of Funds.

Table 11

Regression Results, Fund of Funds Status Sample Split, S&P 500

	Observations	Adjusted R2	MP Estimate	MP P-value	Intercept Estimate	Intercept P-value
FoF	800	0,020	-0,212	0,581	0,017	0,909
Excl. FoF	46210	0,013	-0,194	0,428	0,058	0,484

Table 11: Presents the results from the regression where the sample is split into the subsamples 1) Funds of Funds (FoF) and 2) All funds excluding funds of funds (Excl. FoF) based on the funds FoF status. The excess returns in this regression are relative to the S&P 500 index. MP stands for the market proportion coefficient.

*Significant at 20% **Significant at 10% ***Significant at 5%

Table 12

Regression Results, Fund of Funds Status Sample Split, NYSE/AMEX/NASDAQ/ARCA

	Observations	Adjusted R2	MP Estimate	MP P-value	Intercept Estimate	Intercept P-value
FoF	800	0,020	-0,212	0,581	0,017	0,909
Excl. FoF	46210	0,013	-0,194	0,428	0,058	0,484

Table 12: Presents the results from the regression where the sample is split into the subsamples 1) Funds of Funds (FoF) and 2) All funds excluding funds of funds (Excl. FoF) based on the funds FoF status. The excess returns in this regression are relative to the NYSE/AMEX/NASDAQ/ARCA index. MP stands for the market proportion coefficient.

*Significant at 20% **Significant at 10% ***Significant at 5%

When the original sample is divided based on whether or not the fund is a fund of funds, the regressions do not generate any significant results in either of the subsamples, no matter which of the two passive indexes the returns of the active funds are relative to. As such there are no conclusions to be drawn from these regressions apart from that the results are not significant.

7. Discussion & Implications

Due to the low significance in both of the main regression analyses, we cannot reject that the share of passively invested capital in the market is negatively or not correlated to excess returns of actively managed funds. In our additional regression analyses, we do however find some significant results. In our subsamples based on historical returns, we find significant results indicating a negative relationship between the market proportion of passive investments and the excess return of actively managed funds relative to passive indices. In the regression analysis where the sample has been divided by morningstar category, however, we find that that same relationship is positive for actively managed funds investing to achieve a perfect blend of risk and return for a target date in the future. As previously discussed, due to the target date funds only having returns for 2020 and 2021, those results are biased and have a too small confidence interval (Petersen, 2009).

The simplest answer as to why our main regression is not significant is that the market proportion has no effect on the excess returns of actively managed funds. There are, however, other possible reasons as to why the results were not significant which leave room for the market proportion of passively managed investments to have an impact on the excess returns of actively managed funds. One aspect of our research that could cause our main regression to not be significant is the broad scope of the dependent variable, the excess return of actively managed funds. There are many factors, such as momentum (Jegadeesh & Titman, 1993), turnover ratio (Wermers, 2000), size and book-to-market equity (Carhart, 1997), affecting the excess returns. It is as such possible that the market proportion of passively invested capital does in fact affect the excess returns but that the magnitude of the effect is much smaller relative to the other factors, resulting in its effect not being distinguishable as the other variables are not included. If that were the case, having dependent factors directly representing the stock picking talents of the active fund managers, such as the characteristic selectivity measure and characteristic timing measure from Wermers (2000) might more accurately capture the effect that the market proportion has on the performance of the fund. Alternatively, the other factors could be included into the regression to distinguish what each factor's effect on the excess return is.

The significant estimates for the MP coefficient from our first subsample tests are negative, contrary to our hypotheses as well as what would be expected based on Grossman and

Stiglitz (1980). The results do however raise the question of causality in the relationship between the market proportion and excess returns. As we, with our data, were unable to research causality in our investigated relationship, we were limited to examining the correlation. With this in mind, it is possible that, rather than the increasing market proportion causing the excess returns to decrease, there is a reverse causality where the decreasing excess returns of active funds are what drives the increasing market proportion of actively invested capital. If actively managed funds underperform as described by, amongst others, Carhart (1997), then more investors would opt to instead invest their capital through passive vessels for investments.

Grossman and Stiglitz (1980) do however provide some possible instances that could result in a scenario such as the one observed in our first subsample tests. First of all, if the cost of information increased, the equilibrium share of informed investors would decrease and more investors would then choose to be uninformed. If the cost of information increased at one, or multiple, points of our time period, you would observe a transitional period where the share of uninformed investors increases and the excess returns of informed investors decreases. In addition, if the quality of the information gathered by informed individuals increases, the whole price system would become more informative, increasing the utility of both informed and uninformed investors, which has an ambiguous effect on the equilibrium share of informed investors where it could both increase and decrease. Finally, the degree of noise interfering with the information conveyed by the price system has an inverse relationship with how informative the price system is, meaning that if the degree of noise decreases, the price system will become more informative and the utility of being uninformed increases. If any of these three scenarios did occur, it could have affected the equilibrium share of informed investors during the analyzed time period and our observations could be explained within the equilibrium for disequilibrium model (Grossman & Stiglitz, 1980).

Finally, our dataset having a survivorship bias may have had major effects on our results. Malkiel (1995) establishes that survivorship bias in the data used for analysis of mutual fund performance can have a significant impact on the results. The survivorship bias will create a positive bias on the historical returns. With our analyzed time period stretching from 2011 to current day (2021), the bias will be the greatest in the early years and decrease as it approaches the present day. This may explain two of the odd findings in our analysis. First of all we find that the intercept for our regressions is positive when it is significant. This implies that, contrary to what is commonly accepted and found in previous research (Jensen, 1968) (Carhart, 1997) (Wermers, 2000), on average, actively managed funds net of management and administration fees outperform passive indices when there is no effect from the market proportion. With the survivorship bias in mind, this result is however not as odd as it creates a bias in favor of actively managed funds. Secondly, it is also possible that the survivorship bias may have affected the MP coefficient being negative in our significant results as the magnitude of the survivorship bias would be expected to be greater in the earlier years of the analysis. With the market proportion of passively managed capital increasing every year of our analysis compared to the previous year, the magnitude of the survivorship bias would be expected to have a negative relationship with the market proportion.

8. Limitations

As established in the discussion, there are multiple factors affecting the excess return apart from the market proportion. With these factors not being included in our regression, there is a strong possibility that our analysis has an omitted variable bias where the effect of these omitted variables are attributed to the market proportion. Including factors such as the momentum (Jegadeesh & Titman, 1993) or turnover ratio (Wermers, 2000) into our model could have decreased this bias. Although, due to lack of data we are unable to include them in this paper. However, it is not certain that the inclusion of these factors would result in a lower omitted variable bias, as it is also a possibility that the omitted variable bias would increase as we do not know the complete set of variables affecting the excess return (Clarke, 2005).

The results in our paper should be interpreted very carefully as the survivorship bias in our data could have a significant impact on the results (Malkiel, 1995). The Morningstar Direct database from which the data on the mutual funds is collected, is not survivorship bias free for the time period that we analyze which likely has an impact on our results.

The time periods we could analyze in our research were limited by what data the databases we had access to had. We used Morningstar Direct to gather the data needed to calculate the market proportions of passively invested capital each year. That dataset, however, only extended back to 1993 for its first year of data and then continued on up until 2021. This meant that we were unable to analyze the same time period as Wermers (2000). The timeframe of Morningstar Direct datasets limited us further as their data on the returns of U.S. mutual funds only stretched back to 2011, meaning that in the end our analysis covered the time period 2011-2021.

We were further limited as we did not have access to the CDA database used by Wermers (2000) to create the combined database of mutual funds. Without the data from this database, we could not recreate the characteristic selection and characteristic timing measures (Wermers, 2000) which more accurately represents the stock picking talents. As we could not analyze the same time period as Wermers (2000), the lack of access to the CDA database meant that we were limited to using excess returns of actively managed funds relative to the returns of passive investing as an estimate of the performance of active fund managers.

The annual returns given by Morningstar Direct are expressed net of any management, administrative or 12b-1 fees, and other costs taken out of fund assets but they do not account for any sales charges. Fees such as these are important to take into account when evaluating active fund performance as active funds outperform passive investments before they are applied but not after in Wermers (2000) paper. While it does have an impact on our data that these fees are included, as they undermine the performance of active fund managers, we assume that these fees remain constant over our 11 year period. As we do not evaluate whether or not active fund managers outperform the market, but rather whether or not their performance varies depending on the market proportion of passive investments, the fees should not affect our analysis as long as they remain constant.

9. Further Analysis

To extend upon our studied variable, it could be interesting to calculate the measures used in Wermers (2000), i.e. CS, CT and AS and determine whether any of those measures specifically can be explained by the market proportion of passive investments. Furthermore, an interesting extension of the analysis could be to test if the Carhart Measure is correlated to the market proportion of passively invested capital. The Carhart alpha is not as broad of a measure as the excess return of active funds which could increase the significance in an analysis of the market proportion of passive investments effect.

Our research only calculating the correlation between the two factors leaves some questions unanswered. One of the questions raised from our research is whether the market proportion of passive investments in fact affect the excess return negatively or if the inverse is true that the decreasing excess returns are what causes the market proportion to rise. New research using more extensive data over a longer time period aimed to answer this question would be a very interesting addition to our research.

According to Grossman and Stiglitz (1980), an increased proportion of uninformed investors would only affect the performance of informed investors indirectly through the decreased informational efficiency of the market. It would therefore be an interesting extension to instead of examining excess returns, a symptom of informational efficiency, examine a factor measuring informational efficiency directly in relation to the market proportion of passive investments.

To further investigate whether or not the equilibrium for disequilibrium model (Grossman & Stiglitz, 1980) holds despite our observations, examining factors such as cost of information, noise in the price system and quality of information would be interesting. All of these three factors could move the equilibrium share of informed investors (Grossman & Stiglitz, 1980) and could explain why some of our observations are not in accordance with the equilibrium for disequilibrium model.

References

Anadu, K., Kruttli, M., McCabe, P & Osambela, E., The Shift From Active to Passive Investing: Potential Risks to Financial Stability? *Financial Analysts Journal*, May. Vol 76(4).

Bhattacharya, A. & O'Hara, M., 2020. ETFs and Systemic Risks. *CFA Institute Research Foundation Brief*, January, 2020.

Ben-David, I., Franzoni, F. & Moussawi, R., 2018. Do ETFs Increase Volatility? *The Journal of Finance*, December, Vol. 73(6).

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

Clarke, K., 2005. The Phantom Menace: Omitted Variable Bias in Econometric Research. *Conflict Management and Peace Science*, Vol. 22(4).

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

Grinblatt, M. & Titman, S., 1989. An Analysis of Quarterly Portfolio Holdings. *The Journal of Business*, July, Vol. 62(3).

Grinblatt, M. & Titman, S., 1993. Performance Measurements Without Benchmarks: An Examination of Mutual Fund Returns. *The Journal of Business*, January, Vol. 66(1).

Grinblatt, M., Titman, S. & Wermers, R., 1995. Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior. *The American Economic Review*, December, Vol. 85(5).

Grossman, S. J. & Stiglitz, J. E., 1980. On the Impossibility of Informationally Efficient Markets. *The American Economic Review*, June, Vol. 70(3).

Hartford Funds (2022, February). *The Cyclical Nature of Active & Passive Investing*: https://www.hartfordfunds.com/dam/en/docs/pub/whitepapers/WP287.pdf

Jegadeesh, N & Titman, S., 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, March, Vol 48(1).

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

Malamud, S., 2016. A Dynamic Equilibrium Model of ETFs. CEPR Discussion Paper No. DP11469.

Malkiel, B. G., 1995. Returns from Investing in Equity Mutual Funds. *The Journal of Finance*. June, Vol. 50(2).

Malkiel, B. G., 2003. The Efficient Market Hypothesis and Its Critics. *Journal of Economic Perspectives*, Winter, Vol. 17(1).

Morningstar (2019, June 12). *A Look at the Road to Asset Parity Between Passive and Active U.S. Funds*: https://www.morningstar.com/insights/2019/06/12/asset-parity

Petersen, M., 2009. Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, *The Review of Financial Studies*, January, Vol. 22(1).

Wermers, R., 2000. Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses. *The Journal of Finance*, August, Vol. 55(4).