Why can't investors pick the right index fund?

Barahona, Ricardo*

November, 2020

Job Market Paper

Abstract

Investors leave large amounts of money on the table when investing in index funds, a popular investment product that accounts for 40% of equity funds. I show that even though high fees strongly predict poor performance, investors have little sensitivity to fees. This can be explained by fund intermediation in the retail sector and the legal standard of care that intermediaries have towards their clients. Net inflows to high-fee funds are higher when brokers and financial advisors receive sales commissions from the investment management company. When funds are sold through intermediaries held to higher standard of care, such as those sold to employer sponsored defined contribution pension plans, this is no longer the case. Together, this evidence suggests imposing fiduciary duties on fund intermediaries improves investor welfare.

Keywords: Index Funds, Mutual Funds
JEL classification: G23

*Department of Finance, Erasmus School of Economics and Tilburg University. I particularly thank Rik Frehen and Joost Driessen for providing me with great feedback, guidance and endless support. I also thank Dallin Alldridge, Ao Wang and Cristina Leal for their helpful discussions; Clemens Sialm, Stefano Cassella, Kristy Jansen and Mancy Luo for their helpful comments. I also thank the participants at the 2020 FMA, 2020 Western Economics Association and World Finance Conference meetings as well as the faculties of Tilburg University and Erasmus University for valuable and constructive feedback. Finally, I thank Iman Dolatabadi for kindly sharing data and for long conversations on this topic.

Email address: barahona@ese.eur.nl; Erasmus University Rotterdam - Burgemeester Oudlaan 50, 3062 PA Rotterdam

1 Introduction

Index funds have become a popular investment product for investors to participate in the stock market over the past two decades, growing from 10% to 40% of the US equity fund market[1]. Since index funds with the same benchmark have extremely similar holdings, it is no surprise that fees predict a large part of of the fund's future returns to investors. Yet, I find that even though funds with the same benchmark indices are close to perfect substitutes, there are funds charging fees an order of magnitude higher than their direct competitors that are able to attract significant amounts of capital.

An index fund seeks to replicate the performance of a basket of stocks. In this paper I show that funds that track the same index are close to perfect substitutes and that for a given benchmark index, simply choosing the index funds with the lowest fees is a good rule of thumb that results in a set of funds that outperform the competition. This is made clear in Figure 1 where I rank funds into expense ratio deciles and plot their risk adjusted returns net of fees over the following year. With such a strong predictor of future performance, it is puzzling why competition has not driven high fee funds out of business. Instead, we observe multiple funds tracking that the same index, charging fees that range from 0.05% to 1.5% of the invested amount per year.

This paper studies how the different intermediation channels of the index fund market explain how investors become either more or less responsive to index fund fees. I find that the weakest response to fees is concentrated in index mutual funds sold to retail investors. This is most evident in funds that provide sales commissions to investment professionals such as brokers and financial advisors. This is an important distribution channel, the majority of U.S. households rely on investment professionals when investing in mutual funds[2].

On the other hand, I find evidence that employer sponsored retirement plans are guiding investors towards lower cost funds. In this channel, retirement plan sponsors (employers) and trustees choose a limited set of funds in which employees can invest in with generous tax breaks. When setting

[1]Index funds represent 40% of total net assets of the CRSP domestic equity fund universe at the end of my sample in 2017.

[2]In their 2017 annual fact book, the Investment Company Institute estimates that 80% of U.S. households that own mutual funds rely on investment professionals when buying mutual fund shares outside of their employer sponsored retirement plan.

menus, sponsors and trustees are held to the fiduciary standard, a legal standard of care that requires these parties to put the interests of their clients above their own. Funds offered to these retirement plans are much more responsive to fees. I find evidence that this is explained by low-fee index funds having a higher likelihood of being selected into these retirement plans.

Studying the index fund space presents a unique opportunity to shut down managerial skill in delegated asset management and understand from what sorts of activities investment managers are able to extract value other than stock picking and market timing. Sirri and Tufano (1998) and Hortaçsu and Syverson (2004) argue that the high fees charged by mutual funds are sustained by investor's search costs. Other researchers suggest that mutual funds are able to extract fees from investor mistakes, either through efforts of misguiding investors (Elton et al. (2004), Cooper et al. (2005), Barber et al. (2005) and Cronqvist (2006)) or due to investors lack of understanding that index funds can be seen as commodities Choi et al. (2009). Investment management industry proponents defend that mutual fund companies extract such large fees from S&P 500 funds because mutual fund families provide ancillary services such as financial advice (Collins (2005)), however Elton et al. (2004) find little evidence of these services impacting fund choices.

Despite the potential explanations listed above, it is surprising that the level of fee dispersion in index funds is similar to that of actively managed funds. I extend the findings of Hortaçsu and Syverson (2004) that this is also the case for passively managed funds with benchmarks other than the S&P 500. Despite having lower fees on average, index fund fees have a standard deviation of 37 basis points (b.p.) compared to the 47 b.p. of actively managed funds[3]. Throughout my sample ranging from January 1999 to December 2017, investors could have saved an average of $270 million every year in fees by following a simple rule of thumb of investing in an index fund with a fee in the lowest 20th percentile. As is apparent in Figure 2, this value has only increased over time as the industry has grown, and towards the end of the sample this represents almost $1.2 billion annually or 50% of all fees charged by index funds. Alternatively, if we instead measure the benefits of this strategy in terms of benchmark adjusted returns, investors would still earn an average benefit of $150 million per quarter. These are all lower bound estimates as much larger benefits could be obtained by simply investing in the cheapest fund.

[3]Index and active funds considered here are from the CRSP domestic equity fund universe

The contribution of this article is twofold. I start by documenting that the S&P 500 fund puzzle discussed in Elton et al. (2004) extends to funds tracking 21 other popular American stock indices. I then show new empirical evidence on how fund intermediation and the standard of care to which intermediaries are held can explain a large piece of this puzzle.

The index fund puzzle is one where we find a high dispersion in fees among investment products that are close to perfect substitutes, in a market with low barriers to entry. I find that among funds tracking the same index, fees are the best and most important predictor for the future performance of index funds, across funds tracking 22 indices that represent 60% of all passively managed equity funds. Furthermore, managerial skill and tracking error volatility, a popular measure of index fund risk, vary little across funds that track the same index. Together, this suggests that index funds with the same benchmark index are close to perfect substitutes and that rational investors should simply pick low cost funds. I find that this is not the case for a large segment of the index fund market, in spite of the high dispersion in fees among index funds, investors are not responsive to these fees.

These findings stand in contrast to the theoretical models of the investment management industry with rational investors such as those of Berk and Green (2004) and Garleanu and Pedersen (2015). Index funds can be thought of as actively managed funds but with a very narrow investment objective. Since indices are not traded financial instruments, index funds do have some discretion when it comes to replicating indices such as optimizing trading costs, whether to fully replicate the index or selectively sample a portion of the index constituents among many others. It is therefore not surprising that there is some dispersion when it comes to fund returns gross of fees. If we insert our narrow investment objective actively managed funds into Berk and Green (2004), investors should react strongly to persistent signals of a fund's future performance net of fees, which is in stark contrast to what I find empirically.

One explanation that is consistent with the previous observations is that investment management companies have market power. Hortaçsu and Syverson (2004) argue that even small search frictions can result in a level of market power that is consistent with what we find in the data. Nevertheless, the rising market share of fees collected by high cost index funds suggests the problem is only becoming more severe while at the same time there is an argument to be made that search costs have decreased significantly over the past decades[4].

I find empirical evidence for an alternative explanation, where the incentives of fund intermediaries guide investors to high cost funds. After separating index fund class shares into retail, institutional and ETF shares, I find that most of the puzzle is found in the retail segment. In this segment, investment professionals such as brokers and financial advisors, play a large role in helping investors choose funds, 80% of U.S. households that own mutual funds outside their employer sponsored plans rely on such professionals[5]. As it is common for funds to pay sales commissions to these professionals, this results in a conflict of interest, and one that has been under higher scrutiny since the Dodd-Frank Act passed[6]. Currently fund brokers are only held to the suitability standard of care, one that requires brokers to recommend investments of appropriate risk to their clients, but allows them to recommend more expensive versions of the same product, even if a lower cost perfect substitute is available. As a result, I find that much of investor's lack of sensitivity to fees is concentrated in funds that provide brokers with monetary sales incentives.

While the fiduciary standard might help align broker incentives to those of their clients by imposing a legal duty on a broker to act solely in their client's best interest, I also find evidence that transparency in broker compensation can improve investors' choices into low cost funds. I argue that 12b-1 fees, which are commonly paid out as broker compensation and hidden in the fund expense ratio are much less salient than front-end load fees, one time fees at the time of purchase that the investor pays directly to the broker. As a result, 12-b1 fees are more effective in reducing investors' responsiveness to fees. This is consistent with the theory models for financial advice of Inderst and Ottaviani (2012a) and Inderst and Ottaviani (2012b). When fund intermediaries are compensated via sales commissions, it generates a conflict of interest that results in brokers and advisors guiding investors to high cost funds. This effect is weakened when fund investors are aware

[4]For example, Ellison and Ellison (2009) show how price search engines have dramatically increased demand sensitivity to prices across several product categories

[5]Estimates from the 2017 Investment Management Company fact book. Investment professionals include registered investment advisers, full-service brokers, independent financial planners, bank and savings institution representatives, insurance agents, and accountants.

[6]Section 913 of the Dodd-Frank Act gave the SEC the authority to establish whether or not brokers should have fiduciary responsibility when advising their clients. Since then, the SEC has passed the Regulation Best Interest on the June 5, 2019, strengthening the duty of brokers to act in their client's best interest but stopping short of giving them formal fiduciary duties.

of the intermediary's incentives, consistent with front-end loads being less effective than 12b-1 fees in steering investors to expensive funds.

These findings add to the discussion on what is the value added from receiving investment advice from a broker. Bergstresser et al. (2008) study this question for actively managed funds and find that brokers do little to help investors pick funds that outperform the market nor do they help investors time the market. In this paper, I find evidence that is consistent with brokers suggesting expensive index funds that pay them commissions, despite the availability of better and cheaper options for their clients. This is also consistent with the evidence found in Egan (2019), where broker compensation is crucial to understand why investors allocate so much money to strictly dominated reverse convertible bonds, a popular retail fixed-income product.

To study the importance of intermediary incentives, I also look at employer sponsored retirement plans, commonly known as 401(k) plans. By the end of 2016, 81% of U.S. households owning mutual fund shares, owned funds through this channel. These plans typically give employees a restricted menu of funds to choose from, where they can invest their savings and benefit from several tax breaks. In contrast to broker sold funds, the parties setting these menus are held to a fiduciary standard of care, meaning they are legally required to put their client's interests ahead of their own. As a result, I find that funds offered to these plans are much more sensitive to fees. Using a hand collected dataset of 401(k) menus, I show that a likely explanation is that low-fee index funds are more likely to be selected into these retirement plans. Index funds affiliated with the plan trustee are much more likely to be included in a 401(k) plan when they charge low fees. This is in contrast to the evidence that Pool et al. (2016) find, where mutual fund companies that act as pension plan trustees, set menus in detriment of pension plan participants.

One potential reason for my different findings is that I focus solely on index funds. Since I show index funds that track the same index are close to perfect substitutes, it is easier evaluate them against comparable options, unlike the case with actively managed funds. As a result, it might be easier for plan participants to successfully present a legal challenge to plan fiduciaries when complaining about an S&P 500 index fund that charges yearly fees of 1% when near perfect substitutes exist that charge less than 0.1%.

As a robustness test, I rule out that this result may be driven by investor preferences for additional services that could be bundled with high fee funds. Using several proxies for these additional services, I don't find evidence that this effect is driven by investor preference for fund families that provide such services. These findings hold when looking solely at mutual funds marketed to retail investors, where these services might matter the most.

2 Data and Sample Selection

2.1 Index fund selection and matching

Throughout this paper I rely on the CRSP mutual fund database to gather data on individual index funds as well as on their respective fund families. I also use Thomson Reuters Datastream to obtain the relevant index return data as well as Kenneth French's website for data on risk factors. The sample starts at the earliest time that daily data is available on fund returns which is January 1999 and ends in December 2017.

To study index funds I use the CRSP mutual fund database to analyze the funds that track some of the most popular U.S. indices. These are the major indices of the 3 main index families (S&P, Russel and Wilshire) also covered in Cremers and Petajisto (2009) plus the Nasdaq 100, Nasdaq composite and the Dow 30 as there are a few passive funds tracking these indices. In Table 1 I provide a list of all the indices which are tracked by my sample of passive funds with a short description of each index.

Identifying which index funds track which index requires a few steps as this information is not available on CRSP. I first select only the funds that invest exclusively in U.S. domestic equities as these indices only include U.S. stocks. I then rely on both the CRSP index fund flags and the text filters used in Appel et al. (2016) to identify which domestic equity funds are passively managed. The final step is to identify which index each fund is tracking.

Given the large number of passively managed funds in the sample, I rely on an algorithm to determine this. One straightforward method would be to use the active share measure from Cremers and Petajisto (2009), however this requires having data on historical index composition which is hard to obtain. I develop an alternative method that yields reasonable results. Taking advantage of the availability of daily return data for each fund every quarter, I calculate the tracking error volatility of each fund relative to all the selected indices. Each fund is then assigned to the index to which it has the lowest tracking error volatility.

Given the high correlation between some of these indices as well as with others not included here, it might be the case that a fund is wrongly attributed to one of the 22 indices analyzed here. To mitigate these concerns I drop funds that have an average tracking error volatility greater the 2%. I decide on this cutoff point based on results from Cremers and Petajisto (2009) as they find in their sample that there are very few funds with 0 to 10% portfolio deviation from their benchmark that have a tracking error above 2%. Furthermore, I also only allow funds to track a single index over their lifetime. Since I calculate tracking errors every quarter, I also drop funds that are not consistently assigned to the same index, dropping funds that are not assigned to the same index in less than 80% of the quarters. This alogirthm allows me to assign 294 index funds to one of the 22 indices from Table 1.

As a final step I check whether all index funds are correctly matched to their true benchmark index by comparing against the stated benchmark on the Morningstar Direct database[7]. Since I'm only able to match 194 funds to Morningstar, I manually check the remaining matches, correcting funds that are matched to the wrong index and dropping funds that do not track any of the 22 selected indices. The final sample is then composed of 264 index funds that represent roughly 60% of all total net assets of CSRP domestic equity index funds.

2.2 Index fund performance decomposition

To determine what drives index fund performance, it is useful to first decompose the return of the fund into multiple components. Given that CRSP fund return data is given as net of the expense ratio, in the most simple decomposition, the return of fund i tracking index b can be written as:

$$R_{i,b,t} = IndexRet_{b,t} - ExpRatio_{i,b,t} + TrackDif_{i,b,t}. \tag{1}$$

[7] I use the fund's CUSIP number which is both available on CRSP and Morningstar.

In most analyses of fund flows I use this simple decomposition where I back out $TrackDif$ from the other 3 variables which are observable. $TrackDif$ can be understood as the fund's return gross of fees and the target index. This means track $TrackDif$ will capture a certain element of managerial skill. While manager skill might seem like a strange concept for an index fund, these funds must make multiple discretionary decisions such as how to manage trading costs, cash management policies or even small but intentional deviations from the index. Nevertheless, the skill portion is only a small part of the fund return, the average fund $TrackDif$ is 2 basis points per quarter with a standard deviation of 0.25%.

To get a cleaner measure of skill however, it is possible to further decompose the fund return to calculate its gross alpha. By running an OLS regression of the fund returns on its benchmark index we can estimate a fund's net alpha and beta with respect to its index. As CRSP fund returns are net of fees, if we want to obtain gross alphas we simply need to add fees from the previous period.

$$R_{i,b,t}^{Net} = \alpha_{i,b} + \beta_{i,b,t} IndexRet_{b,t} + e_{i,b,t}. \tag{2}$$

From these regressions it's possible to extract other useful information regarding how well the fund tracks the index. Given that aim to replicate an index, their beta with respect to that index should be close to 1. In that sense I follow Elton et al. (2004) and calculate $AbsBeta$ as:

$$AbsBeta_{i,b,t} = |\beta_{i,b,t} - 1|. \tag{3}$$

Another common measure that funds usually provide clients is the fund's tracking error, or tracking error volatility which can be calculated by computing the standard deviation of $e_{i,b,t}$. This is the definition of tracking error volatility I also use in the index fund matching algorithm described earlier.

2.3 Index fund flows

As is common in the mutual fund literature[8] I calculate flows as the amount of new money going into funds as a percentage of total net assets (TNA) of the previous period, controlling for the fund's return over the period in question

$$flow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}(1 + R_{i,t})}{TNA_{i,t-1}}. \tag{4}$$

Throughout my empirical analysis I focus on flows at the quarterly frequency as this is the frequency at which funds report many of its characteristics while at the same time giving me a large enough time frame to accurately estimate fund betas and alphas using daily data.

While this definition is common throughout the mutual fund literature, it is important to note that when studying mutual fund flows to S&P 500 funds, Elton et al. (2004) use an alternative measure of flows[9]. This comes out of the concern that S&P 500 funds in their sample from 1997 to 2002 differ dramatically in size, where a very large fund will have a much harder job in achieving the same growth rates of smaller funds. While this is less of a concern in my sample, as the market share of the top 10% largest index funds declines from 80% to 50%, in unreported regressions I find that my results hold when using this alternative measure.

2.4 Aggregating share classes

Fund returns and fund descriptive data from CRSP is provided at the share class level, not at the fund level. To perform my analysis I follow most of the mutual fund literature in aggregating share class data at the fund level using the CRSP share class group identifier. For funds with multiple share classes, I add the total net assets of all share classes to obtain the fund's total net assets. Fund expense ratios and returns are computed as the TNA weighted average across share classes.

[8] Berk and Van Binsbergen (2016) and Barber et al. (2016) are recent examples that do this although many more can be provided from the extensive mutual fund literature.

[9] Their measure is similar to the one I describe in Eq.(4) except that in the numerator, they subtract the fitted values of a regression of fund dollar flows on mutual fund size. This measure can be interpreted as the surprise dollar flow given a fund's size, as a percentage of it's past net assets.

Further on in the analysis I also look at specific share classes that are only marketed to certain investors. I use the CRSP flag for type of share class to identify retail, institutional and exchange traded funds (ETF) share classes and aggregate fund data. I then aggregate share class data at the fund level like before, i.e. if a fund has two retail share classes the expense ratio of this fund's retail shares would be the weighted average of these two.

Another concern is that all ETF's are marked as both ETF and institutional funds. This means that to retrieve data pertaining only to institutional funds, I first drop all funds marked as ETF's to identify mutual funds exclusively sold to institutional investors. All of these steps allow me to cleanly identify fund flows and other fund level data from each of these 3 segments independently from one another.

2.5 Defined contribution plans

For many index funds, a large source of clients comes from U.S. defined contribution plans, especially 401(k) individual retirement accounts. In short, 401(k) retirement accounts are a popular investment vehicle that employers can provide to their employees, giving them an opportunity to invest in financial assets at lower tax rates. Employees are generally restricted to a menu of mutual funds and other financial assets that is set by their employer and other third parties. To study how the market for U.S. defined contribution plans interacts with the index fund market, I gather data from multiple sources described below.

To retrieve aggregate fund and fund family data I rely on the Pensions & Investments Defined Contribution Survey from 2012 until 2016. This is a yearly survey that lists the total assets under management each U.S. fund family manages in defined contribution plans. In addition to that, it also provides a list of the top 10 funds of each of these families in terms of defined contribution assets. I then match this data to the CRSP mutual fund database to obtain measures of a fund's defined contribution exposure as well as that of the fund family, defined as the ratio of DC assets to total assets.

I also hand collect a data set of multiple 401(k) menus from SEC 11-K forms. All publicly listed firms in the U.S. that offer their own stock as an investment option must release this form every year. From here I collect the all mutual fund investment options and the amount invested in each of these options. Funds collected this way are then matched to the CRSP mutual database by name using a fuzzy text matching algorithm and matches are then manually verified. In total, data is collected for the forms filed in 2013 and 2014, providing me with 401(k) plan menu's of 731 different employers and a total of 1082 plans[10].

To obtain additional plan information, I match these plans to the Department of Labor 5500 forms. From Schedule H I obtain information on the total size of plan. From Schedule C I obtain information on plan trustees, third party service providers that provide record keeping, consulting educational and other services to the 401(k) plan participants. Plan trustees are frequently asset management firms that may also influence the menu options as described in Pool et al. (2016). I use a fuzzy match algorithm to match all service providers by name against all the mutual fund family names on CRSP to determine whether a plan trustee is an investment management company and which company is providing this service. In total, I find that in my sample 50% of plans are affiliated with and asset management firm and 13.5% of funds included in a 401(k) plan belong to a fund family affiliated with that plan.

3 Predicting index fund performance

One standout feature of index funds is that their performance relative to their competitors is very predictable. Since we're comparing funds holding extremely similar asset portfolios, the key metric for a index fund investor to analyze when choosing a fund tracking a given index is the fund's expense ratio. This is clearly shown in Figure 1b, where picking the fund with the lowest fee delivers the best net of benchmark adjusted return, a yearly difference of almost 1.5% between top and bottom deciles.

To understand what best predicts fund performance I follow Elton et al. (2004) and run Fama and MacBeth regressions to understand which variables predict future fund benchmark adjusted returns net of fees in the cross section of funds. The main predictor variable analyzed are benchmark adjusted returns gross of fees, the most recent expense ratio and lagged *AbsBeta* and tracking error

[10]Plans for 2013 were kindly shared by Iman Dolatabadi.

volatility.

I find results similar to those of Elton et al. (2004), for a longer period sample and extending it to index funds that track other indices besides the S&P 500. Table 3 shows that index fund fees are highly predictive of fund alphas net of fees in the following period. Here it is clear that there is almost a reverse one to one relationship between fees and following period alphas. Note that expense ratios are expressed as a yearly fee, so from column 5 we can see that at the yearly frequency, a one percentage point increase in the expense ratio results in a 1 percentage point decrease in yearly alpha. This should not be surprising, if two funds are holding the same assets, the fund with the lowest expense ratio will naturally outperform the other in terms of net of fee returns. Furthermore, each basis point increase in fees should represent a one basis point decrease in net of fee alpha per year. This is clearly shown in Table 3 at multiple horizons.

There is also some evidence that skill is somewhat persistent in index funds as gross alphas also positively predict future net alpha. Despite this, the economic magnitudes are not very large, in the first column of Table 3 for example, a one standard deviation increase in gross alpha (25 b.p.) results in an increase of 3 b.p. in the net performance over the next quarter. There is however limited evidence that tracking error and AbsBeta are predictive of future performance. One potential reason is that there is little dispersion in these variables as funds seem to do a reasonably good job at tracking their target benchmark as is evident in Table 2. Note that by including index dummies I am also controlling for unobservable differences of funds tracking different indices. These results suggest that two funds tracking the same index are near perfect substitutes differing only in their price.

4 Do investors care about fees?

In the previous section we established that both expense ratios and index fund skill have significant predictive power over future fund performance. Given the strength of these results, it would be expected that rational investors would react to this and as a result, funds with lower expense ratios would demonstrate higher growth rates. This much would be expected in a competitive market where rational investors strongly react to persistent signals of fund performance.

To analyze this relationship, I start by regressing fund flows on fund past performance measures to test whether investors are responsive to these measures, especially fees:

$$flow_{i,b,t} = \delta_1 IndexRet_{b,t-1} + \delta_2 TrackDif_{i,b,t-1} + \delta_3 ExpRatio_{i,b,t-1} + \delta_4 FLoad_{i,b,t-1} + \eta_{b,t} + \epsilon_{i,b,t}, \quad (5)$$

where the main parameter of interest is δ_3, the coefficient on the fund expense ratio. In addition to the remaining return components from Eq. 2, it is also important to include front load fees, $FLoad$, which are one-off fees that retail investors pay to brokers. This is expressed as a dummy variable as we do not observe the actual fees charged, only whether or not a fund charges this sort of fee[11].

One concern in this analysis is that different indices may be more costly to replicate or that investors might have a preference for a given stock index that they may want to invest in. To deal with this concern, I include index-quarter fixed effects which control for any unobserved demand for funds tracking any given index at any quarter. Under this specification, δ_3 can be interpreted as how investors respond to prices of funds that track the same index. As I show earlier, funds tracking the same index are near perfect substitutes so we should expect a negative δ_3.

To further control for potential endogneity issues, I also include several fund family level variables that could proxy for unobserved services that funds may offer such as the size of the fund family, the number of funds a certain fund family offers and the number of share classes a given fund offers. I also include the net inflows to fund family, excluding the fund in question, to proxy for other unobserved effects such as family level marketing efforts.

In Table 4 I show what I call the index fund puzzle, where I find that expense ratios have little predictive power for index fund growth. This is quite surprising given how predictive fees are of a fund's net of fee performance as well as how persistent these are. Given how predictive fees are for future index fund performance, we would expect a negative and significant coefficient on expense ratios, as we would expect low fee funds to display higher growth rates that their high fee counterparts. This is not the case, coefficients on expense ratios are statistically insignificant in all

[11]In the CRSP database, we observe the maximum fee that funds allow brokers to charge but not the actual fee. Maximum fees can be as high as 5% of total investment and FINRA rules cap these fees at 8.5%.

specifications.

Despite not being very predictive of future fund performance, tracking error volatility is significant and with the expected sign despite it's lower influence on future return as shown in the previous section. We can see this as the $TrackError$ coefficient is negative and statistically significant. This relationship suggests that investors put a high value on an index fund provides a consistent exposure to the benchmark index, as a high tracking error volatility means the fund has a higher risk of deviating from its objective. The effect is economically meaningful, a 1 standard deviation decrease in tracking error volatility translates to quarterly inflows growing 0.9 percentage points higher than the average fund (4.2% quarterly growth rate). I also find no relationship between manager skill, $TrackDif$, which is not surprising given how economically small skill is in index funds.[12].

Also surprising is that past index returns seem to predict mutual fund growth as funds tracking indices exhibiting high recent past returns grow at a faster pace than their peers. The coefficients ranging between 0.30 and 0.39 are economically large given the volatility of the underlying indices that the funds in my sample track. One potential interpretation for these results is that investors may be extrapolating past index returns and chase after funds tracking these indices even though index returns are volatile and have little predictive power of future fund returns in the cross section of funds. Alternatively, some indices may have exposure to different risk factors that exhibit different average returns. For example, Russel 2000 funds are more exposed to small firms, earning higher returns in the long term and attracting more capital. Nevertheless, this pales in comparison to the opportunity of earning higher returns by simply investing in low fee funds, which is both more certain and does not carry any additional exposure to risk.

There is some evidence that investors are responsive to fees as funds with front end load fees display lower growth rates. As these fees can vary for the same fund across different brokers, I use a dummy variable for whether a fund has these fees or not. Despite the increased marketing effort, it does seem that ultimately investors are more sensitive to the price they have to pay for these funds. This is consistent with Barber et al. (2005) who argue that investors are more sensitive to *in your face fees* as opposed the ongoing and less transparent expense ratios. Later analysis however

[12] This result is robust to other measures of fund still such as gross CAPM alphas, Carhart 4 factor alphas and gross alphas with respect to the fund's benchmark.

suggests that this result is not so robust, at least in the context of index funds.

Finally, there is some evidence suggesting that funds belonging to fund families capable of providing clients with additional services to grow faster. Proxies such as fund family size, the number of funds in a fund family, and whether a fund family has a star fund show up as positively related to fund growth and with statistically and economically significant coefficients.

The results I find in Table 4 are significantly different than those found in Elton et al. (2004). I rule out that this is due to my choice of mutual fund flow measure, which is different than theirs, as in unreported regressions, I find that the results presented in Table 4 hold when using their measure of flows. Instead, I point to the fact that there is very little overlap between our two samples, their sample ranges from 1997 to 2002 where mine is between 1999 and 2017 across multiple indices. When restricting to the overlapping years, I do indeed find a negative but statistically insignificant coefficient on expense ratios.

5 Index fund flows and intermediation

In this section I analyze the different channels through which index funds are marketed to investors. While most of the mutual fund literature focuses on the question of how mutual funds generate or extract value through their investment activities, index funds provide a great setting to explore how mutual funds extract value from investors for reasons that have little to do with investment skill. As I show earlier in the paper, there is nothing special about the index funds in this sample, for a given benchmark index, most funds are following very similar investment strategies.

To study what explains investors' low responsiveness to fees, I study the different distribution channels for index funds, with a focus on intermediary incentives. I first investigate whether the puzzle is present for funds marketed to different types of investor. I then investigate how broker incentives may influence investor sensitivity to fees. Finally, I also investigate how the structure of U.S. defined contribution plans may influence investor responsiveness to fees.

5.1 Retail, Institutional and ETF investors

The source of the index fund puzzle becomes much clearer when I disaggregate funds into 3 different types of share classes. In these next regressions I perform a similar analysis as in Table 4, however I look at funds by isolating share classes that are marketed to different investors. This allows us to study mutual fund flows by looking at different segments of investors in isolation. This is better that simply including dummy variables indicating whether funds are marketed to retail or institutional clients. Many of the larger index funds sell to both types clients simultaneously, so there is a lot of detail lost when aggregating flows across all share classes.

Analysis at this level can be valuable as these three markets have very different characteristics. When comparing institutional to retail investors, we expect the former to have better financial literacy while at the same time enjoying large economies of scale when searching for the best funds. ETFs on the other hand are also marketed to retail clients, but have a different structure than traditional mutual funds. ETFs are more easily tradable as they are traded just like stocks listed on an exchange. This means that they are accessible to all investors with a brokerage account on any broker. Mutual funds on the other hand are either offered directly through the mutual fund providers or through brokers, however mutual fund menus will vary by broker, unlike ETFs. This means that switching mutual funds requires a larger effort and is more costly than for ETFs.

From Table 5, it becomes apparent that investors in retail share classes are primarily responsible for the lack of responsiveness to fees. More than that, higher fees are predictive of higher growth, however this effect is not statistically significant in all specifications. Furthermore, only retail share classes are subject to front end loads and here it is not as clear that these types of fees do affect investors, as the coefficient on front end loads is statistically insignificant, albeit with the expected negative sign. There is also no evidence that retail investors do care about the fund's correlation with its objective as both tracking error and *AbsBeta* are statistically insignificant and come with much smaller coefficients.

In contrast, both investors in institutional share classes and ETFs are extremely price sensitive. While ETF flows have non-significant coefficients on expense ratios under index-time fixed effects, these are negative and large so no statistical significance can be driven by the smaller sample size.

More surprising however is that ETF investors are chasing past index returns which I find little evidence of predicting future returns. While part of this can be driven by the fact that ETF's are easier and cheaper to trade at high frequencies, it is not clear as to why a rational investor would behave this way.

Breaking down the sample into separate types of share classes gives us a better understanding of investor responses to fees, however at this point it is difficult to generalize. It might come as no surprise to some that retail investors are less price sensitive than institutional investors as the former group can commonly be thought of as unsophisticated investors. Nevertheless, it may also be that retail investors have larger constraints than those faced by institutional investors such as having little time and resources to find the best funds available. On the under hand, institutional investors have economies of scale in search costs, making it less costly for an institutional investor to search for the cheapest funds.

On the other hand, lack of sophistication may also be less plausible as retail investors can also invest in ETFs. ETFs also have certain advantages as trading them is much simpler. Unlike mutual funds which are either only sold directly or might only be available in certain brokers, ETFs are traded on stock exchanges so in theory they should be accessible from any brokerage account. These lower frictions may make investors much more sensitive to prices. Nevertheless, there may be a selection effect and I can not rule out that more sophisticated investors self select into buying index ETFs while less sophisticated investors invest in index mutual funds.

5.2 Broker sold funds and incentives

One important feature of the mutual fund industry is how mutual funds get distributed and marketed to clients. Funds can either be sold directly from the investment management companies to their clients or alternatively they can rely on brokers to sell their funds. In the case investment management companies opt for the latter, they can choose to provide incentives to brokers by paying them sales commissions. Funds can do this through two main mechanisms, either by charging their investors 12b-1 marketing fees or by allowing brokers to charge front-end loads. 12b-1 marketing fees are a part of a fund's expense ratio and get deducted from the fund's assets, making these fees quite hidden from investors. On the other hand, front-end loads are one-off fees that brokers can charge to mutual fund investors and are paid up-front, making these fees very salient. It is important to note that these two broker compensation schemes are not mutually exclusive, some funds make use of both.

Funds that compensate brokers for their sales efforts, can lead brokers aggressively sell these same funds to investors with the intent of earning higher commissions. Furthermore, these incentives may be amplified due to evidence that many investors receive financial advice from their brokers. This concern has resulted in increased pressure to endow brokers with fiduciary duties towards their clients, making them legally obliged to put their clients' interest before their own when giving financial advice. The 2010 Dodd-Frank gave the SEC the authority to make this change and while there has been pressure on the SEC to make this change, it has instead introduced *Regulation Best Interest* in 2019. This new regulation gives brokers more responsibilities regarding their clients interests, but it stops short of giving them fiduciary responsibility.

To study how broker incentives may affect index fund flows, I regress flows on expense ratios and an interaction term of the expense ratio and whether a fund compensates brokers through 12b-1 fees or through front-end loads.

$$flow_{i,b,t} = \delta_1 ExpRatio_{b,t-1} + \delta_2 Broker_{i,b,t-1} + \delta_3 ExpRatio \times Broker_{i,b,t-1} + \eta_{b,t} + \epsilon_{i,b,t}. \quad (6)$$

In Table 6, I test what effect broker sold funds impact a fund's sensitivity to fees. We can see from Panel A, that 12-b1 marketing fees are very effective at reducing investor's sensitivity to index fund fees. While non-12-b1 index fund growth is inversely related to fees, the contrary is true for funds that engage in this practice. This effect is large and economically significant. When focusing on retail funds, columns 4 and 5 show that while a 1 percentage point increase in non 12b-1 funds results in quarterly outflows of -5.5%, that same increase results in a an additional 3.8 percentage points of inflows for 12b-1 funds. By including fund fixed effects, we can also estimate the impact of a fund introducing 12b-1 fees, where introducing broker compensation through 12b-1 fees results in 4.8 additional percentage points of inflows when increasing fees.

However, it isn't clear how much of this benefits the mutual fund company, as it is likely that a large portion of the expense ratio is used to compensate brokers in these cases. This evidence is in line with Egan (2019), who also finds that broker incentives also distort the market and lead brokers to recommend fixed income products that are strictly inferior to other available options to their clients.

In Table 6 Panel B, I also find a significant reduction in fee responsiveness when focusing on the retail share class. While on average, funds with front-end loads attract less flows, increasing the expense ratio by 1 percentage point actually increases flows by around 6 percentage points. This goes against the hypothesis of Barber et al. (2005), that investors are extremely sensitive to salient fees such as front-end loads. However, when studying the specification with fund fixed effects in column 6, there is no evidence that investors become less responsive to fees when a fund starts using front-end loads.

When compared to the more opaque 12b-1 fees where an investor might not be as aware of broker compensation, this evidence suggests that front end loads are less effective at guiding investors towards more expensive funds. An alternative explanation is that if front end loads are much larger than 12b-1 fees, then it might mean that investors are responding to this additional fee. Unfortunately, we are not able to accurately observe front end load fees as these can vary for the same fund across different brokers.

5.3 Defined contribution plans

An increasing part of the assets under management of U.S. mutual funds come from 401(k) plans. In my Pensions & Investments dataset, this amounts to 12% to 15% of the sample[13]. This market segment is interesting to study in this context due to its increasing size and the fact that its design has the potential to partially explain the index fund puzzle.

401(k) plans are employer sponsored defined contribution plans, that give employees of firms that offer such a plan a tax advantaged way of buying mutual funds. These plans are typically managed by two main entities, a plan sponsor which is the employer and a trustee. It is common

[13]This is likely underestimated as the dataset only covers the top 10 funds in assets of each fund family.

for investment management companies to act as trustees, providing administrative and educational services to plan participants as well as setting fund menus together with the plan sponsor. When investment management companies act as trustees, there is an incentive for these firms to favor their own funds when setting menus. These menus are quite restrictive, Pool et al. (2016) find that the average plan has 20 funds on offer and find evidence that investment management companies do indeed favor their own funds.

Despite the potential of many investors being stuck to a restrictive menu with high fee funds, it may still be rational for investors to opt for high fee funds due to the large tax benefits involved. For index funds, it can be the case that investment management companies may take advantage of this market to steer clients into high fee index funds. If this were the case, then we should expect index fund flows of funds with higher dependency of 401(k) assets to be less sensitive to fund fees.

There is however a strong counteracting force on investment management companies power on setting menus. Plan sponsors (i.e. employers) and trustees, those responsible for setting plan menus, also have a fiduciary duty to plan participants. This means that employers have a legal obligation to act in plan participants' best interest. One important implication is that this puts pressure on employers to make sure fund menus offered in the retirement plans are fair. This legal obligation has already resulted in several successful lawsuits against employers, with high fee fund menus being one of the primary reasons why these lawsuits are filed[14]. If this threat of lawsuits that plan fiduciaries face is large enough, we may see that low-fee index funds are more likely added to these menus. As a result, funds offered to 401(k) plans may be more responsive to fees. In addition to this, fund families may also be willing to use index funds as loss leaders as a way to become a plan trustee and earn additional sources of revenue from higher margin services.

5.3.1 Fund level analysis

To study this question, I look at both aggregate data at the fund level as well 401(k) plan level data. For fund level data, I analyze whether funds offered to defined contribution plans are more or less sensitive to fees. I study this both at the extensive margin, i.e. whether a fund is offered at all to defined contribution (DC) plans, as well as the intensive margin. For the extensive margin

[14]See Mellman and Sanzenbacher (2018) for a summary of 401(k) lawsuits over the past two decades.

I construct a binary variable at the fund level as well as at the family level, of whether a fund is offered to a DC plan or whether a fund family offers at least one fund to a DC plan, respectively. For DC intensity, I follow Kronlund et al. (2020) and define it as the fund or fund family's total DC assets as a proportion of their total net assets at a given quarter.

In Table 7 Panel A, we see that funds offered to DC plans or funds that belong to a family that manages DC assets are much more sensitive to fees. To the extent that DC plans can represent a significant portion of flows, this suggests that funds with low fees are more likely to be selected to become a part of 401(k) plans. While an index fund that is offered to a 401(k) attracts an additional 3 percentage points of quarterly flows, I also find that for funds offered to a 401(k) plan, a one standard deviation decrease in a fund's expense ratio results in an additional 4 percentage points of quarterly flows.

In Panel B we see this result is also borne out when looking at DC intensity, however only at the fund level. As the DC intensity variables are standardized, we can interpret the interaction term as funds with a DC intensity one standard deviation above the average benefit from an fund flow increase of roughly 8 percentage points when decreasing their expense ratio by 1 percentage point.

5.3.2 Plan level analysis

To understand whether the previous result is driven by the fact that funds with low expense ratios are more likely to be selected into 401(k) plans, I use plan level data to estimate a linear probability model analyzing what drives the probability of an index fund being added to a 401(k) plan. From the 1082 fund menus I collected, I allow each of these menus to choose any of the index funds from my CRSP sample. From Table 8 Panel A, we can see that the unconditional probability of an index fund being included in a 401(k) plan is 0.5%.

In addition to the expense ratio, we are also interested in whether a fund is affiliated with the plan trustee as well as how that interacts with the expense ratio. Under the view that investment management companies have a conflict of interest and want to place their funds in a given menu, we would expect that these funds are less sensitive to expense ratios.

In addition to these main variables, I also include a set of fund and plan level controls. More importantly, I include Plan-Year and Index-Year fixed effects, which allows me to interpret results as the probability of a fund tracking a specific benchmark being selected into a 401(k) plan while controlling for unobserved plan level variables.

From Table 8 Panel B, I find that affiliated index funds are on average 12 percentage points more likely of being selected into a 401(k). Compared to the unconditional probability of addition of 0.5%, this means these funds are 24 times more likely to be selected. However, from the results I don't find evidence that this comes as a result of a conflict of interest from a plan trustee, as the interaction term between affiliated funds and expense ratio is negative and significant: from column 4, a one standard deviation decrease in the expense ratio of an affiliated fund results in a 15 percentage point increase in the probability of being added to the plan. As plan trustees are also fiduciaries, this suggests that the legal threat is strong enough such that when fund families are trustees of a given plan, they only add their own index funds to the menu when these have low fees.

An alternative explanation for this is that investment management companies that act as trustees may use their cheap index funds to become trustees, potentially charging the plan sponsor and plan participants higher margin services. While I do not have data on additional revenues that investment management companies may make through their trustee activities, if this effect is stronger in large 401(k) plans, that would be consistent with this story. In fact this is what I find in Table 9 where I split the sample into the top and bottom 20% plans by total assets under management. For large plans, a one standard deviation decrease in expense ratio for an affiliated fund results in a 25 percentage point increase in the probability of addition, 66% higher than in Table 8. In Table 9, I also look at whether plans that are not affiliated with an investment management company are sensitive to expense ratios when selecting funds into their plan. Surprisingly, expense ratios are positively related to selection meaning that expensive index funds are more likely to be selected into these plans. Nevertheless, compared to earlier results, these are economically smaller, a one standard deviation increase in fees only increases likelihood of the fund being added to the plan by 0.46 percentage points.

These results suggest that when retail customers invest their money through 401(k) plans, they are more likely to invest in low cost rather than high cost index funds. However, this comes mostly as a result not of their own choice but through the employers and trustees providing them with a menu that's more likely two include one or two cheap index funds. Of course for this to happen, there needs to be a likely probability that plan participants sue their employers if they only provide them with expensive funds. To the extent that these complaints exist and several 401(k) plans include cheap options, retail investors cannot be assumed to be that naive. In fact, Kronlund et al. (2020) document that 401(k) investors become very sensitive to fund fees within a plan menu when the Department of Labor introduces regulations on more transparent fee and performance disclosure.

6 Robustness checks

6.1 Index choice

There are two points related to my sample selection that I wish to discuss in this section. The first is regarding the disproportionately high amount of S&P 500 funds compared to other index funds in my sample. While this may put in question whether we are able to generalize the results of this paper to other non S&P 500 trackers, I address this issue by re-doing the analysis for only non S&P 500 funds. While this would be too extensive to include in this paper, in unreported regressions I find that the main results of this paper hold, even when excluding S&P 500 funds.

Although with different magnitudes, I find the same qualitative results for non S&P 500 index funds for both the broker channel and 401(k) menu additions. Regarding the DC fund level analysis, I find statistically insignificant results but always with the same coefficient signs, potentially due to the fact that the sample is too small to achieve accurate coefficient estimates.

6.2 Non-portfolio services

In this section, I discuss to what extent non-portfolio services play a role in investors' insensitivity to index fund fees. One popular hypothesis among industry proponents (Collins, 2005) is that investors may choose an index fund due to other services that an investment management company might bundle when offering funds to its clients. While we can not observe these services directly, I find little evidence that this is happening by using proxies for these potential services.

To test whether non-portfolio services may explain why investors are so unresponsive to index fund fees, I estimate a model where I interact several proxies for services that mutual fund families may offer in addition to the index fund in question. I use four proxies for these: fund family size, the number of asset classes offered by the fund family, the number of funds offered by the fund family and finally, whether or not a fund family has a star fund[15]. With the exception of the star fund variable which is a binary variable, the other 3 are standardized to mean zero and unit standard deviation for ease of interpretation.

In Table 10, I explore whether larger fund families reduce sensitivity to fees. I find that the interaction coefficients of interest are in general negative, meaning funds from larger fund families that offer with more asset classes are more sensitive to price. This suggests that funds part of a richer service portfolio are typically more price sensitive. I do not however argue that this is a causal effect as variables such as fund size correlate with many other unobserval factors. In fact, fund family size has been also been a common proxy for search costs in the mutual fund literature, with larger fund families facing lower search costs, which could explain why index funds in these families are more price sensitive.

An alternative explanation is that large fund families in this space follow a high volume strategy by offering cheap index funds. There is also some anecdotal evidence that large fund families are using index funds as loss leaders to then attract investors to other offerings. Examples of this can be found in news coverage of Fidelity Investment's offering of a zero fee index fund[16].

I also find funds that are a part of fund families that offer many funds or that offer star funds obtain higher than average quarterly growth rates, I do not find that this alters price sensitivity. Like Elton et al. (2004), but I find little evidence that additional services offered by mutual fund families affect investors price sensitivity to these.

Because retail clients are more likely to benefit from bundled services, I also carry out this analysis for mutual funds sold exclusively to retail clients. Nevertheless, I find similar results in

[15] A fund family has a star fund in a given quarter if it ranks in the top 10% of funds of its CRSP category in a given quarter.

[16] See Financial Times article: https://www.ft.com/content/d8569037-98fa-35bd-b3e5-861e8168161d

Table 11, further suggesting that bundling is not such a very important feature in the index fund market.

7 Conclusion

Even though index fund performance is remarkably predictable, it is striking to find that investors fail to use variables that strongly predict future performance to their advantage, especially fees. I extend evidence on the index fund puzzle to funds tracking multiple indices and by showing it still persists today, even as index funds have become a very popular investment product that now accounts for 40% of the domestic equity mutual fund assets.

By looking at different markets where index funds are sold, I find this effect is mostly driven by investors in retail share classes of mutual funds, as investors in institutional share classes seem to understand index funds well enough to make use of these predictors. The fact that ETF shares are also somewhat price sensitive and also available to retail clients suggests that the way ETFs are offered either significantly reduce frictions or alternatively ETF investors form a different clientele of more sophisticated investors.

When attempting to understand retail mutual fund investors' slow response to fees, I find evidence that some forms of broker compensation steer investors to high fee funds. On the other hand, the rising popularity of 401(k) plans may be increasing investors sensitivity to index fund prices, as employers have clear incentives to provide their employees with menus comprised of funds with reasonable fees. The lack of fiduciary duties in the former, and their presence in the latter suggests that this legal standard provides strong incentives for advisors, brokers and employers to provide investors with good financial advice.

Finally, this paper also shows evidence that retail investors' investments are largely shaped by external influence. When investing on their own, many investors are exposed to the influence of brokers that guide them to high fee funds that benefit the brokers. In the 401(k) market, investors are more likely to invest in low fee funds, but only because they have already been chosen for them in a 401(k) menu with few options.

References

Appel, I. R., Gormley, T. A., and Keim, D. B. (2016). Passive investors, not passive owners. Journal of Financial Economics, 121(1):111–141.

Barber, B. M., Huang, X., and Odean, T. (2016). Which factors matter to investors? evidence from mutual fund flows. The Review of Financial Studies, 29(10):2600–2642.

Barber, B. M., Odean, T., and Zheng, L. (2005). Out of sight, out of mind: The effects of expenses on mutual fund flows. The Journal of Business, 78(6):2095–2120.

Bergstresser, D., Chalmers, J. M., and Tufano, P. (2008). Assessing the costs and benefits of brokers in the mutual fund industry. The Review of Financial Studies, 22(10):4129–4156.

Berk, J. B. and Green, R. C. (2004). Mutual fund flows and performance in rational markets. Journal of political economy, 112(6):1269–1295.

Berk, J. B. and Van Binsbergen, J. H. (2016). Assessing asset pricing models using revealed preference. Journal of Financial Economics, 119(1):1–23.

Choi, J. J., Laibson, D., and Madrian, B. C. (2009). Why does the law of one price fail? an experiment on index mutual funds. The Review of Financial Studies, 23(4):1405–1432.

Collins, S. (2005). Are s&p 500 index mutual funds commodities? Perspective, Investment Company Institute, 11(3):N3.

Cooper, M. J., Gulen, H., and Rau, P. R. (2005). Changing names with style: Mutual fund name changes and their effects on fund flows. The Journal of Finance, 60(6):2825–2858.

Cremers, K. M. and Petajisto, A. (2009). How active is your fund manager? a new measure that predicts performance. The Review of Financial Studies, 22(9):3329–3365.

Cronqvist, H. (2006). Advertising and portfolio choice. Available at SSRN 920693.

Egan, M. (2019). Brokers versus retail investors: Conflicting interests and dominated products. The Journal of Finance, 74(3):1217–1260.

Ellison, G. and Ellison, S. F. (2009). Search, obfuscation, and price elasticities on the internet. Econometrica, 77(2):427–452.

Elton, E. J., Gruber, M. J., and Busse, J. A. (2004). Are investors rational? choices among index funds. the Journal of Finance, 59(1):261–288.

Garleanu, N. B. and Pedersen, L. H. (2015). Efficiently inefficient markets for assets and asset management. Technical report, National Bureau of Economic Research.

Hortaçsu, A. and Syverson, C. (2004). Product differentiation, search costs, and competition in the mutual fund industry: A case study of s&p 500 index funds. The Quarterly Journal of Economics, 119(2):403–456.

Inderst, R. and Ottaviani, M. (2012a). Financial advice. Journal of Economic Literature, 50(2):494–512.

Inderst, R. and Ottaviani, M. (2012b). How (not) to pay for advice: A framework for consumer financial protection. Journal of Financial Economics, 105(2):393–411.

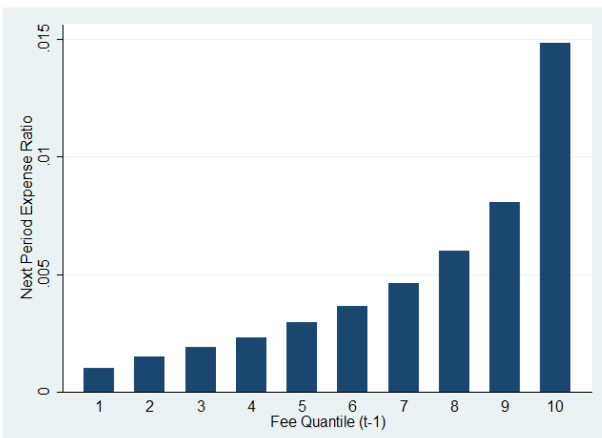
Kronlund, M., Pool, V. K., Sialm, C., and Stefanescu, I. (2020). Out of sight no more? the effect of fee disclosures on 401 (k) investment allocations. Journal of Financial Economics, 401:19–6.

Mellman, G. S. and Sanzenbacher, G. T. (2018). 401 (k) lawsuits: What are the causes and consequences? Issue in Brief, pages 18–8.

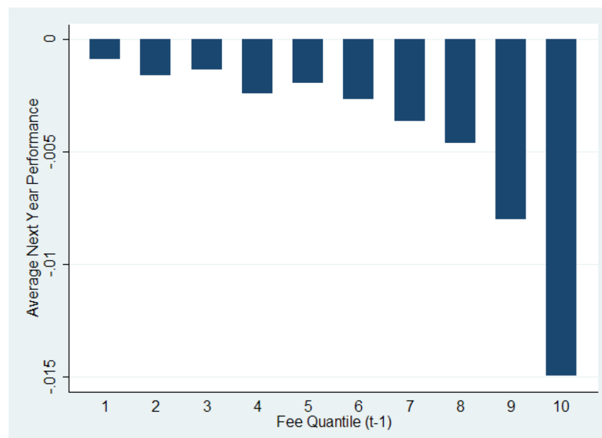
Pool, V. K., Sialm, C., and Stefanescu, I. (2016). It pays to set the menu: Mutual fund investment options in 401 (k) plans. The Journal of Finance, 71(4):1779–1812.

Sirri, E. R. and Tufano, P. (1998). Costly search and mutual fund flows. The journal of finance, 53(5):1589–1622.

Figure 1: Index fund deciles sorted by fees in the last quarter



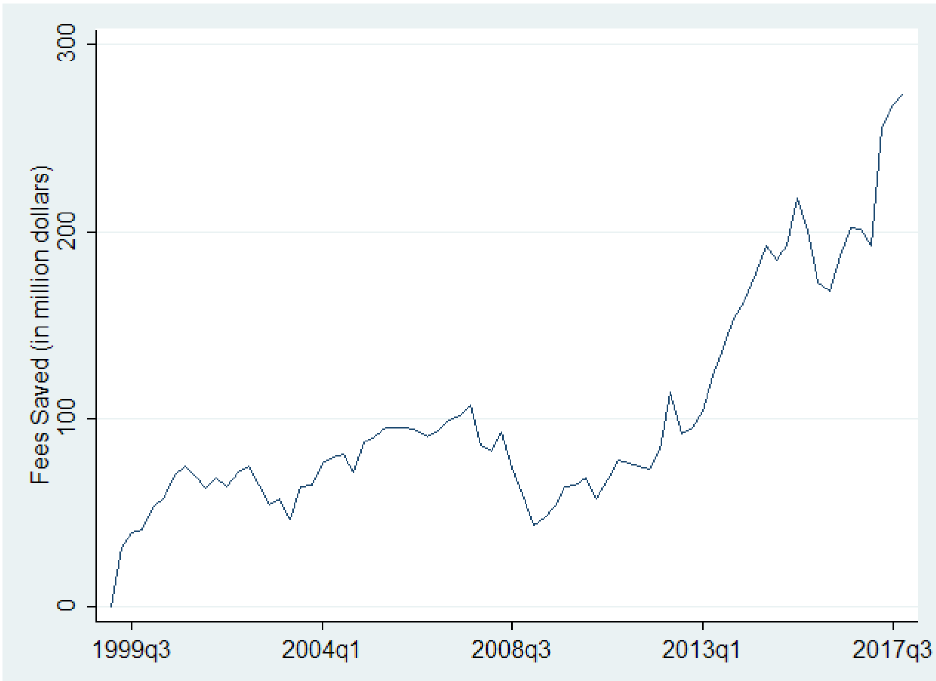
(a) Average next quarter fees



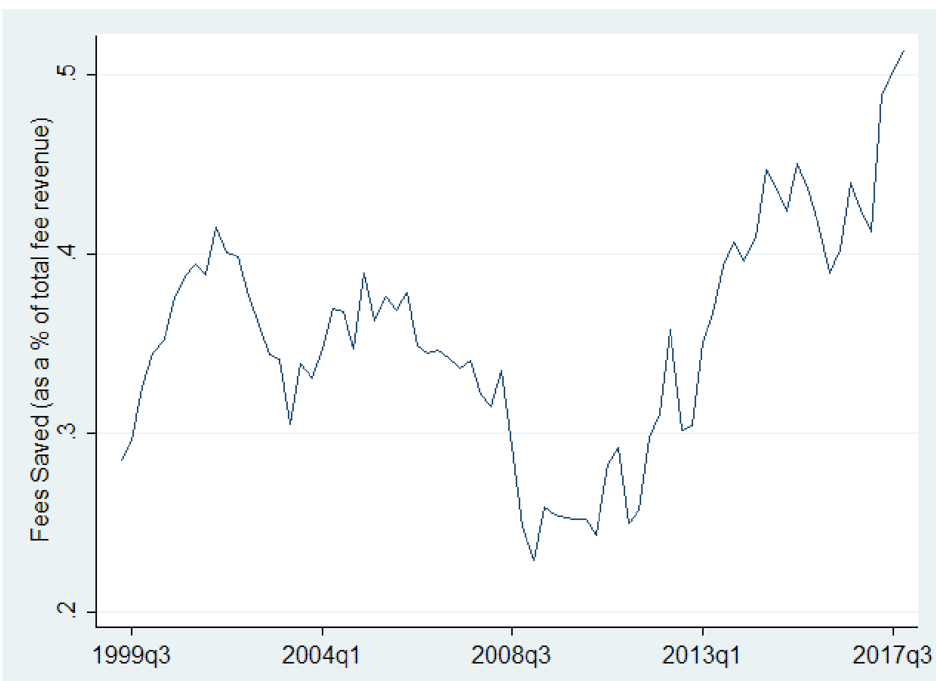
(b) Average next year net alpha

In this figure, I use the total expense ratio to rank funds into decile bins and plot the average expense ratio of each decile in the following year in Panel A. In Panel B I plot the average risk adjusted return (with respect to the fund's benchmark index) of each decile, net of fees, over the following year.

Figure 2: Fees saved by investing the 20% cheapest funds



(a) Fees saved per quarter in $Million



(b) Fees saved per quarter as a percentage of total fees

This figure shows how much investors would save in fund expenses if those that hold funds in the 80% most expensive funds would instead invest their money in the fund that sits at the 20th percentile ranked on fees. Panel A shows this amount in dollar terms and Panel B shows this amount as the percentage of all fees charged in terms of fund expense ratios.

Table 1: Index Descriptions

Index Name	Short Description	Funds
S&P 500	This index includes the largest and most liquid 500 companies, with approximately 80% coverage of available market capitalization. Style: Large Cap.	114
S&P Growth	Growth segment of the stocks in the S&P 500. Style: Large Growth.	3
S&P Value	Value segment of the stocks in the S&P 500. Style: Large Value.	5
S&P MidCap 400	This index measures the performance of 400 mid-sized companies, smaller than the ones in the S&P 500. Style: Mid Cap.	26
S&P SmallCap 600	This index measures the performance of 600 small cap companies, smaller than the ones in the S&P 400. Style: Small Cap.	17
Russell 1000	This index includes the largest and most liquid 1000 companies. Style: Large Cap.	7
Russell 2000	This index includes the next 2000 largest companies in the US, with no overlap with the Russel 1000. Style: Small Cap.	26
Russell 3000	This index includes the the companies present in the Russell 1000 and Russell 2000 indices. Style: Broad market index.	5
Russell 1000 Growth	Growth segment of the stocks in the Russell 1000. Style: Large Growth.	5
Russell 1000 Value	Value segment of the stocks in the Russell 1000. Style: Large Value.	6
Russell 2000 Growth	Growth segment of the stocks in the Russell 2000. Style: Small Growth.	4
Russell 2000 Value	Value segment of the stocks in the Russell 2000. Style: Small Value.	3
Russell 3000 Growth	Growth segment of the stocks in the Russell 3000. Style: Growth.	2
Russell 3000 Value	Value segment of the stocks in the Russell 3000. Style: Value.	1
Russel Mid-Cap	Performance of the 800 smallest companies in the Russell 1000. Style: Mid Cap.	8
Russel Mid-Cap Growth	Growth segment of the stocks in the Russell MidCap index. Style: Mid Growth.	4
Russel Mid-Cap Value	Value segment of the stocks in the Russell MidCap index. Style: Mid Value.	4
Wilshire 5000	A total market index including all actively traded US stocks. Style: Broad market index.	3
Wilshire 4500	Index composed by the Wilshire 5000 companies excluding the S&P 500 constituents. Style: Small - Mid Cap Index.	5
Dow Jones Industrial Average	Price weighted index of 30 US blue-chip companies. Style: Large Cap.	5
Nasdaq 100	Index tracking the largest companies traded on the Nasdaq stock exchange. Style: Heavily tilted towards large technology stocks.	9
Nasdaq Composite	Index tracking all companies traded on the Nasdaq stock exchange. Style: Heavily tilted towards technology stocks.	2

Table 2: Summary Statistics

This table shows summary statistics of the variables used in this analysis. Fund flows are the mutual fund flows in millions of dollars while percentage fund flow is a fund's dollar flow divided by the total net assets of the previous period. Tracking error is the fund's tracking error volatility, FLoad Dummy is a binary variable that is equal to 1 if the fund has front end load (sales charge), Marketing Dummy is a binary variable that is equal to 1 if the fund has 12b-1 marketing fees, Fund TNA is the total the fund's total net assets, Share Classes is the total number of share classes of a given fund, Fund age is the age of the fund in years, Family TNA is a fund family's total assets in millions of dollars, Family Asset Classes is the total number of asset classes offered at the fund family level as measured by the first two letters of CRSP style code, Family funds is the total number of funds offered by the fund family (in hundreds) and family flows is the net quarterly dollar flows to a fund family. DC Fund and family intensity is the percentage of total net assets of a fund or fund family that is owned by defined contribution (DC) plans respectively. DC Fund and DC Fund Family is a dummy variable that is equal to one is the fund is offered to DC plans or if the family offers a fund to a DC plan respectively.

Variable	**Observations**	**Mean**	**St. Dev.**	**1st Pct**	**25th Percentile**	**Median**	**75th Percentile**	**99th Pct**
Fund flow	10057	35.4	214.1	-628.1	-12.9	1.2	31.3	858.7
Percentage fund flow	10057	0.0419	0.2122	-0.4381	-0.0241	0.0054	0.0510	1.6654
TrackDif	9782	0.0002	0.0025	-0.0195	-0.0002	0.0001	0.0005	0.0193
Index return	10284	0.0223	0.0846	-0.3616	-0.0186	0.0311	0.0672	0.3498
Expense Ratio	10040	0.0042	0.0037	0.0006	0.0018	0.0030	0.0054	0.0253
F Load Dummy	10284	0.2104	0.4076	0	0	0	0	1
Tracking Error	10284	0.0052	0.0045	0.0003	0.0018	0.0038	0.0070	0.0200
Abs ($\beta - 1$)	10284	0.0064	0.0110	0.0000	0.0012	0.0027	0.0064	0.1688
Marketing Dummy	10284	0.3410	0.4741	0	0	0	1	1
Fund TNA	10284	2735.5	5279.2	1.7	132.6	590.2	2287.0	21706.1
Share Classes	10284	2.16	1.53	1	1	2	3	9
Fund age	10284	7.35	5.42	0	3	6	11	27
Family TNA	10284	309241.7	532407.0	12.5	15034.7	71204.0	273052.7	1969620.0
Family asset classes	10284	7.99	2.24	1	7	8	10	12
Family Funds	10284	1.03	0.98	0.01	0.34	0.68	1.43	4.8
Total family flows	10057	4.5	11.5	-19.3	-0.2	0.3	4.9	40.3
Fund Return	10284	0.0214	0.0840	-0.2492	-0.0190	0.0306	0.0669	0.2408
Fund alpha wrt Index	9782	0.0002	0.0024	-0.0227	-0.0002	0.0001	0.0005	0.0207
DC Intensity Fund	609	0.4330	0.2546	0.0036	0.2442	0.4713	0.5678	1.0000
DC Intensity Family	1505	0.1526	0.1472	0.0000	0.0514	0.1093	0.1876	0.7832
DC Fund	2270	0.2683	0.4432	0	0	0	1	1
DC Fund Family	2270	0.6630	0.4728	0	0	1	1	1

Table 3: Return predictability of index funds

These are Fama-MacBeth regressions where the dependent variables are index fund alphas with respect to the index they track over multiple horizons. The independent variable are alphas (gross of fees), the last observed expense ratio, the fund's tracking error volatility and absolute difference of the fund's beta w.r.t. to its index and one. These regressions are at the quarterly frequency, each t represents a quarter. t-statistics are displayed in square brackets and are calculated using heteroscedasticity and autocorrelation consistent Newey-West standard errors (with 4 lags in column 5 and 12 lags in column 6).

	(1)	(2)	(3)	(4)	(5)	(6)
Dependend Vars	Next Q	Next Q	Next Q	Next Q	Next Year	Next 3 Years
Gross Alpha (t-1)	0.141**			0.125**	0.212	0.42
	[2.24]			[2.40]	[1.52]	[1.34]
Gross Alpha (t-2)	0.023			-0.011	0.109	0.271
	[0.49]			[-0.29]	[0.61]	[1.31]
Gross Alpha (t-3)	0.058			0.079**	0.004	0.205**
	[1.62]			[2.38]	[0.04]	[2.21]
Gross Alpha (t-4)	0.027			0.022	0.055	0.142
	[0.72]			[0.69]	[0.63]	[1.33]
Expense Ratio (t-1)		-0.240***		-0.238***	-0.918***	-2.679***
		[21.30]		[-40.29]	[-24.72]	[-16.75]
Tracking Error (t-1)			-0.013	0	-0.026	-0.118
			[-0.82]	[0.02]	[-0.41]	[-1.40]
Abs Beta (t-1)			0	0.01	0.057*	0.065
			[0.02]	[1.43]	[1.89]	[1.27]
Index Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8741	9782	10006	8735	7973	6278
Avg. R-squared	0.5319	0.3762	0.3301	0.6772	0.7428	0.7729

Significance: ***99%, **95%, *90%

Table 4: Mutual fund flows: Aggregate funds

This table shows panel regression coefficient estimates of percentage fund flows on the indicated regressor variables. Track error is the fund's tracking error volatility, FLoad Dummy is a binary variable that is equal to 1 if the fund has front end loads, Marketing Dummy is a binary variable that is equal to 1 if the fund has 12b-1 marketing fees, Log TNA is the log of the fund's total net assets, Share Classes is the total number of share classes of a given fund, Fund age is the age of the fund in years, Log(Family TNA) is the log of the fund family's total assets, Family Asset Classes is the total number of asset classes offered at the fund family level as measured by the first two letters of CRSP style codes, family funds is the total number of funds offered by the fund family and family flows is the dollar flows to the fund family. Standard errors are clustered at the fund level.

	(1)	(2)	(3)	(4)
Index Ret (t-1)	0.328***	0.286**	0.263**	
	[2.76]	[2.52]	[2.35]	
TrackDif (t-1)	0.119	-0.403	0.238	-0.665
	[0.11]	[-0.37]	[0.23]	[-0.58]
Expense Ratio (t-1)	2.311	0.772	0.489	0.38
	[1.24]	[0.49]	[0.31]	[0.24]
FLoad Dummy (t-1)	-0.037***	-0.024**	-0.023**	-0.022**
	[-3.92]	[-2.15]	[-2.18]	[-1.99]
Track Error (t-1)	-0.591	-2.344***	-2.344***	-1.26
	[-0.68]	[-2.70]	[-2.63]	[-1.24]
Abs $(\beta - 1)$ (t-1)	0.236	-0.052	-0.051	0.053
	[0.68]	[-0.17]	[-0.17]	[0.15]
Marketing Dummy (t-1)		0	0.008	0.01
		[0.03]	[1.04]	[1.28]
Log(TNA) (t-1)		-0.017***	-0.017***	-0.015***
		[-5.57]	[-5.44]	[-4.92]
Share Classes (t-1)		-0.002	-0.001	-0.002
		[-0.62]	[-0.38]	[-0.50]
Fund age (t-1)		-0.003***	-0.003***	-0.003***
		[-5.72]	[-4.44]	[-4.29]
Log(Family TNA) (t-1)		0.011***	0.010***	0.009**
		[3.17]	[3.12]	[2.52]
Family Asset Classes (t-1)		-0.008***	-0.009***	-0.007***
		[-3.72]	[-3.91]	[-2.91]
Family Funds (t-1)		0.012**	0.014***	0.015***
		[2.59]	[2.89]	[3.12]
Star Fund (t-1)		0.029***	0.025***	0.021***
		[3.88]	[3.46]	[2.63]
Family Flows		0.001***	0.001***	0.001***
		[3.83]	[2.62]	[2.65]
Time FE	Yes	Yes	Yes	No
Index FE	No	No	Yes	No
Index × Time FE	No	No	No	Yes
Observations	8427	8427	8427	7979
Adj. R-squared	0.0517	0.0924	0.0957	0.127

Significance: ***99%, **95%, *90%

Table 5: Mutual fund flows by share class type

This table shows panel regression coefficient estimates of percentage fund flows on the indicated regressor variables as well as a set of control variables: log of fund size, number of fund share classes, fund age, log of family size, number of asset classes offered by the fund family, number of funds offered by the fund family and the total amount of dollar flows to the mutual fund family. In each of the Retail, Institutional and ETF specifications, fund level variables are calculated only taking into account the data on the respective share classes. Standard errors are clustered at the fund level.

	Retail			Institutional			ETF		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Index Ret (t-1)	0.238	0.263		0.288	0.248		0.466***	0.441***	
	[1.02]	[1.10]		[1.31]	[1.21]		[3.16]	[3.27]	
TrackDif (t-1)	-1.978	-0.381	-2.349	4.863	4.885	4.253	1.687	1.886	2.067
	[-1.17]	[-0.22]	[-1.35]	[1.50]	[1.47]	[1.00]	[0.87]	[0.92]	[1.05]
Expense Ratio (t-1)	3.206**	2.947	2.674	-13.418***	-14.327***	-12.894***	-17.965**	-25.702*	-12.763
	[2.00]	[1.33]	[1.25]	[-3.16]	[-2.92]	[-2.89]	[-2.04]	[-1.86]	[-1.06]
FLoad Dummy (t-1)	-0.023	-0.023	-0.017						
	[-1.52]	[-1.49]	[-1.12]						
Track Error (t-1)	-0.607	0.2	0.483	-1.857	-2.325	-0.739	-0.609	-1.458	-1.895
	[-0.54]	[0.17]	[0.43]	[-0.79]	[-0.85]	[-0.24]	[-0.42]	[-0.76]	[-1.06]
Abs ($\beta - 1$) (t-1)	-0.293	-0.498	0.061	-0.596	-0.324	0.325	-0.627	-0.7	-1.118**
	[-0.76]	[-1.17]	[0.12]	[-0.55]	[-0.30]	[0.21]	[-1.02]	[-0.97]	[-2.31]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Index FE	No	Yes	No	No	Yes	No	No	Yes	No
Index × Time FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	3681	3680	3410	2803	2803	2412	1753	1753	942
Adj. R-squared	0.1063	0.1104	0.1221	0.0564	0.0558	0.0642	0.2055	0.2262	0.1264

Significance: ***99%, **95%, *90%

Table 6: Mutual fund flows: Broker Compensation

These tables show panel regression coefficient estimates of percentage fund flows on the indicated regressor variables as well as a set of control variables: log of fund size, number of fund share classes, fund age, number of asset classes offered by the fund family, number of funds offered by the fund family, whether a fund family offers a star fund and the total amount of dollar flows to the mutual fund family. Marketing dummy is a binary variable that is equal to 1 if one of the share classes of the fund has 12b-1 marketing fees. FLoad dummy is a binary variable that is equal to 1 if one of the share classes of the charges front end loads. In the Retail specifications, fund level variables are calculated only taking into account the data on mutual fund retail share classes. Standard errors are clustered at the fund level.

Panel A: Marketing fees

	All			Retail		
	(1)	(2)	(3)	(4)	(5)	(6)
Expense Ratio (t-1)	-9.227***	-8.602***	-18.016***	-5.258*	-5.466*	-18.675*
	[-4.24]	[-3.86]	[-2.88]	[-1.66]	[-1.73]	[-1.77]
Marketing Dummy	-0.047***	-0.042***	-0.085***	-0.037*	-0.036**	-0.170*
	[-4.17]	[-3.76]	[-2.75]	[-1.92]	[-2.27]	[-1.98]
Expense Ratio	12.610***	11.752***	21.011***	9.386***	9.331***	23.555**
× Marketing Dummy	[5.73]	[5.39]	[3.09]	[2.75]	[2.99]	[2.04]
Index Ret (t-1)	0.277**			0.268		
	[2.51]			[1.13]		
TrackDif (t-1)	0.571	-0.28	-0.759	-0.283	-2.171	-2.429
	[0.55]	[-0.24]	[-0.68]	[-0.16]	[-1.23]	[-1.37]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	No	Yes	No	No
Index FE	Yes	No	No	Yes	No	No
Index × Time FE	No	Yes	Yes	No	Yes	Yes
Fund FE	No	No	Yes	No	No	Yes
Observations	8427	7979	7960	3410	3681	3402
Adj. R-squared	0.1037	0.1346	0.2107	0.1263	0.0708	0.1895

Significance: ***99%, **95%, *90%

Panel B: Front-end loads

	All			Retail		
	(1)	(2)	(3)	(4)	(5)	(6)
Expense Ratio (t-1)	0.618	0.255	-8.346	-0.731	-1.800	-7.427
	[0.34]	[0.15]	[-1.27]	[-0.31]	[-0.85]	[-0.72]
FLoad Dummy (t-1)	-0.021	-0.024	-0.046	-0.072**	-0.074**	-0.12
	[-0.99]	[-1.09]	[-1.00]	[-2.23]	[-2.55]	[-1.18]
Expense Ratio (t-1)	-0.299	0.278	8.486	6.598*	7.873**	12.966
× FLoad Dummy (t-1)	[-0.10]	[0.09]	[1.19]	[1.77]	[2.29]	[1.18]
Index Ret (t-1)	0.263**			0.266		
	[2.35]			[1.14]		
TrackDif (t-1)	0.242	-0.667	-0.904	-0.624	-2.599	-2.494
	[0.23]	[-0.58]	[-0.85]	[-0.35]	[-1.45]	[-1.39]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	No	Yes	No	No
Index FE	Yes	No	No	Yes	No	No
Index × Time FE	No	Yes	Yes	No	Yes	Yes
Fund FE	No	No	Yes	No	No	Yes
Observations	8427	7979	7960	3410	3681	3402
Adj. R-squared	0.0956	0.1269	0.2067	0.1271	0.0639	0.1869

Significance: ***99%, **95%, *90%

Table 7: Mutual fund flows: Funds listed in defined contribution plans

These tables show panel regression coefficient estimates of percentage fund flows on the indicated regressor variables as well as a set of control variables: log of fund size, number of fund share classes, fund age, log of family size, number of asset classes offered by the fund family, number of funds offered by the fund family and the total amount of dollar flows to the mutual fund family. In Panel A columns 1 to 3, DC Fund is a dummy variable equal to 1 if a fund is offered to DC plan, in columns 4 to 6 DC Fam is equal to 1 if a given fund family offers at least one fund to a DC plan. In Panel B columns 1 to 3, DC intensity is measured as the percentage of a mutual fund's TNA owned by defined contribution plans. In columns 4 to 6, DC intensity is measured as the percentage of a fund family's TNA owned by defined contribution plans. Both DC intensity measures are standardized to simplify the interpretation of the interaction terms. Data used for these tables ranges from June 2012 until June 2017. Standard errors are clustered at the fund level.

Panel A: Defined contribution fund inclusion

	DC Fund			DC Fund Family		
	(1)	(2)	(3)	(4)	(5)	(6)
Expense Ratio (t-1)	0.961	0.523	0.271	3.458*	2.988	2.867
	[0.54]	[0.29]	[0.14]	[1.90]	[1.55]	[1.46]
DC Fund/Fam (t-1)	0.023	0.029**	0.032**	0.081***	0.075***	0.080***
	[1.58]	[2.05]	[2.18]	[3.83]	[3.41]	[3.50]
Expense Ratio (t-1)	-10.178***	-10.498***	-11.041***	-15.061***	-14.092***	-15.299***
× DC Inclusion (t-1)	[-2.99]	[-3.32]	[-3.49]	[-3.37]	[-3.16]	[-3.35]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	No	Yes	Yes	No
Index FE	No	Yes	No	No	Yes	No
Index × Time FE	No	No	Yes	No	No	Yes
Observations	2143	2143	2055	2143	2143	2055
Adj. R-squared	0.0758	0.0729	0.1087	0.0833	0.0784	0.1157

Significance: ***99%, **95%, *90%

Panel B: Defined contribution fund intensity

	DC Intensity Fund			DC Intensity Fund Family		
	(1)	(2)	(3)	(4)	(5)	(6)
Expense Ratio (t-1)	-8.362***	-9.734***	-9.305***	-14.001***	-19.980***	-21.700***
	[-3.68]	[-5.20]	[-4.51]	[-2.84]	[-3.68]	[-3.74]
DC Intensity (t-1)	0.032***	0.031***	0.031***	0.027**	0.025**	0.023*
	[3.47]	[3.38]	[3.19]	[2.08]	[2.09]	[1.90]
Expense Ratio (t-1)	-9.085***	-8.250***	-7.984***	-4.525	-1.786	-0.775
× DC Intensity (t-1)	[-3.58]	[-3.18]	[-3.03]	[-1.41]	[-0.53]	[-0.22]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	No	Yes	Yes	No
Index FE	No	Yes	No	No	Yes	No
Index × Time FE	No	No	Yes	No	No	Yes
Observations	227	227	215	624	624	595
Adj. R-squared	0.212	0.2118	0.1873	0.1937	0.2102	0.2783

Significance: ***99%, **95%, *90%

Table 8: 401(k) Plan menu additions

Panel A shows summary statistics relevant for the linear probability model in Panel B, where plan size is reported in millions of dollars. Panel B shows the coefficient estimates of a linear probability model where the dependent variable is equal to 1 whenever an index fund is offered in a given 401(k) plan menu. Expense ratio is the expense ratio of a fund and Affiliated is a dummy variable equal to one when the fund belongs to the same family as a plan trustee. I control for fund return over the previous year, Log(Family TNA) is the log of the fund family's total assets, Log(Fund TNA) is the log of the fund's total assets, Fund age in years, Other index is a dummy equal to 1 if the plan includes another index fund, No. of options is the number of investment options offered in a given plan and Log(Plan Size) is the log of the total plan value. Standard errors are clustered at the fund level.

Panel A: 401(k) Plan Summary Statistics

Variable	Observations	Mean	St. Dev.	Min	Median	Max
Fund Addition	395543	0.0049	0.0697	0	0	1
No. of Options	395543	23.9	8.8	1	24	138
Plan Size	395543	860.2	2770.0	0.9	199.0	48151.0
No. of Index Fund Options	395543	1.8	1.4	0	2	17
Other Index	395543	0.86	0.35	0	1	1
Affiliated Fund	395543	0.02	0.12	0	0	1
Inv. Manager Affiliated	395543	0.50	0.50	0	0	1

Panel B: 401(k) Plan Additions

	(1)	(2)	(3)	(4)
Expense Ratio (t-1)	-11.245*	-12.551**	-32.216	-39.079**
× Affiliated	[-1.91]	[-2.16]	[-1.58]	[-1.99]
Expense Ratio (t-1)	0.389**	0.367**	1.554**	1.300**
	[2.22]	[2.33]	[2.45]	[2.44]
Affiliated	0.067***	0.067***	0.111**	0.119**
	[2.82]	[2.87]	[2.17]	[2.38]
Fund Return (t-1)	0	-0.002	0.127	0.106
	[0.00]	[-0.19]	[1.57]	[1.53]
Log(Family TNA) (t-1)	0.001**	0.001**	0.003**	0.002**
	[2.58]	[2.52]	[2.26]	[2.09]
Log(TNA) (t-1)	0.004***	0.004***	0.007***	0.006***
	[3.49]	[3.45]	[2.93]	[2.98]
Fund age (t-1)	-0.001**	-0.001***	-0.001**	-0.001**
	[-2.33]	[-2.65]	[-2.20]	[-2.51]
Other Index	-0.003*	-0.986***	-0.007**	-0.978***
	[-1.68]	[-168.31]	[-2.33]	[-79.43]
No. of Options	0.000***		0.000***	
	[4.52]		[3.28]	
Log (Plan Size)	0.000		0.000	
	[-0.59]		[-0.64]	
Year	Yes	No	No	No
Plan × Year FE	No	Yes	No	Yes
Index × Year FE	No	No	Yes	Yes
Observations	346949	346949	154257	154257
Adj. R-squared	0.0259	0.1931	0.0448	0.2401

Significance: ***99%, **95%, *90%

Table 9: 401(k) Plan menu additions

This table shows the coefficient estimates of a linear probability model where the dependent variable is equal to 1 whenever an index fund is offered in a given 401(k) menu. Expense ratio is the expense ratio of a fund and Affiliated is a dummy variable equal to one when the fund belongs to the same family as a plan trustee. I also include the following controls: fund return over the previous year, Log(Family TNA) is the log of the fund family's total assets, Log(Fund TNA) is the log of the fund's total assets, Fund age in years and Other index. Columns 1 and 2 estimate the model for the 20% largest plans, columns 3 and 4 estimate the model for the 20% smallest funds and columns 5 and 6 estimates the model for 401(k) plans not affiliated with an investment management company. Standard errors are clustered at the fund level.

	Large Plans		Small Plans		No Affiliation	
	(1)	(2)	(3)	(4)	(5)	(6)
Expense Ratio (t-1)	-47.081***	-64.101***	-6.717**	-13.417		
× Affiliated	[-3.11]	[-2.84]	[-2.11]	[-0.86]		
Expense Ratio (t-1)	0.401**	1.479**	0.412**	1.113**	0.318**	1.150**
	[2.44]	[2.45]	[2.44]	[2.06]	[2.16]	[2.38]
Affiliated	0.115***	0.148***	0.073***	0.092*		
	[3.34]	[2.98]	[2.78]	[1.67]		
Fund Return (t-1)	0.003	0.13	0.005	0.077	-0.003	0.113
	[0.32]	[1.43]	[0.86]	[1.16]	[-0.30]	[1.57]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Plan × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Index × Year FE	No	Yes	No	Yes	Yes	Yes
Observations	69580	30936	68939	30651	174140	77424
Adj. R-squared	0.1826	0.2361	0.1768	0.2128	0.1926	0.2317

Significance: ***99%, **95%, *90%

Table 10: Mutual fund flows: Non-portfolio Services

This table shows panel regression coefficient estimates of percentage fund flows on the indicated regressor variables: Fam TNA is log of fund family size, log of fund size, number of fund share classes, fund age, number of asset classes offered by the fund family, number of funds offered by the fund family, a dummy equal to 1 if a fund family has a star funds and the total amount of dollar flows to the mutual fund family. Fam TNA, Fam asset classes and Fam funds have been standardized to simplify the interpretation of the interaction terms. Standard errors are clustered at the fund level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expense Ratio (t-1)	-1.792	-1.583	-0.53	-0.504	0.219	0.217	4.827*	1.917
	[-1.34]	[-1.09]	[-0.37]	[-0.34]	[0.15]	[0.14]	[1.66]	[0.72]
Fam TNA (t-1)	0.038***	0.033***						
	[5.09]	[4.53]						
Expense Ratio (t-1)	-3.863***	-3.173***						
× Fam TNA	[-3.81]	[-3.37]						
Fam Asset Classes (t-1)			-0.009	-0.008				
			[-1.60]	[-1.34]				
Expense Ratio (t-1)			-2.051***	-1.793**				
× Fam Asset Classes (t-1)			[-2.91]	[-2.57]				
Fam Funds (t-1)					0.015***	0.012**		
					[3.15]	[2.51]		
Expense Ratio (t-1)					-0.533	0.475		
× Fam Funds (t-1)					[-0.43]	[0.44]		
Star Fund							0.052***	0.032*
							[2.87]	[1.95]
Expense Ratio (t-1)							-4.755	-1.729
× Star Fund (t-1)							[-1.50]	[-0.59]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	Yes	No	Yes	No	Yes	No
Index FE	Yes	No	Yes	No	Yes	No	Yes	No
Index × Time FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	9244	8833	9244	8833	9244	8833	9244	8833
Adj. R-squared	0.0957	0.1196	0.0936	0.1184	0.0926	0.1184	0.0926	0.1172

Significance: ***99%, **95%, *90%

Table 11: Mutual fund flows: Non-portfolio Services — Only retail mutual funds

This table shows panel regression coefficient estimates of percentage fund flows of retail mutual funds on the indicated regressor variables: Fam TNA is log of fund family size, log of fund size, number of fund share classes, fund age, number of asset classes offered by the fund family, number of funds offered by the fund family, a dummy equal to 1 if a fund family has a star funds and the total amount of dollar flows to the mutual fund family. Fam TNA, Fam asset classes and Fam funds have been standardized to simplify the interpretation of the interaction terms. Standard errors are clustered at the fund level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expense Ratio (t-1)	1.776	2.173	2.161	2.485	3.004	2.599	5.106*	3.092
	[0.84]	[1.06]	[1.01]	[1.19]	[1.41]	[1.24]	[1.82]	[1.08]
Fam TNA (t-1)	0.047***	0.031***						
	[3.39]	[2.90]						
Expense Ratio (t-1)	-3.800**	-1.554						
× Fam TNA	[-2.19]	[-1.18]						
Fam Asset Classes (t-1)			-0.005	-0.014				
			[-0.53]	[-1.55]				
Expense Ratio (t-1)			-1.896*	-0.638				
× Fam Asset Classes (t-1)			[-1.88]	[-0.65]				
Fam Funds (t-1)					0.016**	0.008		
					[2.32]	[1.37]		
Expense Ratio (t-1)					-1.703	0.944		
× Fam Funds (t-1)					[-0.94]	[0.68]		
Star Fund							0.032	0.022
							[1.63]	[1.09]
Expense Ratio (t-1)							-2.491	-0.482
× Star Fund (t-1)							[-0.97]	[-0.18]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	Yes	No	Yes	No	Yes	No
Index FE	Yes	No	Yes	No	Yes	No	Yes	No
Index × Time FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3680	3410	3680	3410	3680	3410	3680	3410
Adj. R-squared	0.1137	0.1227	0.1116	0.1222	0.1105	0.1222	0.1104	0.1218

Significance: ***99%, **95%, *90%