INDEX FUND FLOWS AND FUND DISTRIBUTION CHANNELS

2025

BANCO DE ESPAÑA
Eurosistema

Documentos de Trabajo
N.º 2518

Ricardo Barahona

INDEX FUND FLOWS AND FUND DISTRIBUTION CHANNELS (*)

Ricardo Barahona (**)

BANCO DE ESPAÑA

(*) This paper was previously circulated under a different title: "Why can't investors buy the right index fund?". I thank Rik Frehen and Joost Driessen for providing me with great feedback, guidance and endless support. I am also grateful to Dallin Alldridge, Ao Wang and Cristina Leal for their helpful discussions and to Clemens Sialm, Stefano Cassella, Kristy Jansen and Mancy Luo for their helpful comments. I would also like to express my gratitude to the participants at the 2023 Portuguese Finance Network, 2020 FMA, 2020 Western Economics Association and World Finance Conference meetings as well as the faculties of Tilburg University and Erasmus University for their valuable and constructive feedback. Finally, I thank Alberto Orts who greatly contributed to the 401(k) plan hand collection effort as a research assistant and Iman Dolatabadi for kindly sharing data on 401(k) plans from 2013.
(**) e-mail: ricardo.barahona@bde.es

Documentos de Trabajo. N.º 2518

April 2025

https://doi.org/10.53479/39443

The Working Paper Series seeks to disseminate original research in economics and finance. All papers have been anonymously refereed. By publishing these papers, the Banco de España aims to contribute to economic analysis and, in particular, to knowledge of the Spanish economy and its international environment.

The opinions and analyses in the Working Paper Series are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

The Banco de España disseminates its main reports and most of its publications via the Internet at the following website: http://www.bde.es.

ISSN: 1579-8666 (on line)

Abstract

In the United States, investors leave large amounts of money on the table when investing in index 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, defined contribution.

JEL classification: G23.

Resumen

En Estados Unidos, los inversores minoristas frecuentemente gastan mucho dinero en comisiones cuando invierten en fondos índice, incluso cuando existen alternativas mucho más baratas. En este estudio se demuestra que, aunque unas comisiones elevadas predicen de manera consistente un bajo rendimiento, los inversores muestran una escasa sensibilidad a ellas. Este fenómeno puede explicarse por la intermediación de fondos en el sector minorista y por la regulación a la que están sujetos los intermediarios financieros que venden fondos. Los flujos netos hacia fondos con comisiones elevadas son mayores cuando los asesores financieros cobran retribuciones por ventas pagadas por la gestora de dichos fondos. Sin embargo, esta relación desaparece cuando los fondos se comercializan en canales de distribución donde los minoristas están más protegidos, como aquellos brindados en planes de pensiones de aportación definida ofrecidos por empleadores. En conjunto, la evidencia sugiere que la imposición de deberes fiduciarios a los intermediarios de fondos contribuye a mejorar el bienestar de los inversores.

Palabras clave: fondos índice, fondos de inversión, aportación definida.

Código JEL: G23.

1 Introduction

Index funds have become a popular investment product for investors to participate in the stock market over the past two decades, amounting to 57% of the US equity fund market[1]. Since index funds with the same benchmark have nearly identical holdings, it is no surprise that fees for funds tracking the same index predict a large part 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, many investors choose funds that charge fees an order of magnitude higher than their direct competitors[2].

In this paper, I use index funds to study the role of different intermediation channels in the US mutual fund industry and how these contribute or alleviate the misallocation described above. Specifically, I study the impact on investor flows' sensitivity to fees in the two largest intermediation channels for mutual funds in the US. These are funds sold via brokers and to 401(k) defined contribution retirement plans[3]. Since skill in index mutual funds is quite limited, the main difference across funds that affects performance are the fund's total expense ratios[4].

My main finding is that while investors are not very sensitive to index fund fees among broker sold funds, the reverse is true for funds sold to employer sponsored defined contribution plans. For the broker channel these findings are consistent with brokers recommending high fee funds to clients in exchange for higher sales commissions. This adds to evidence on the limited advice provided by brokers in Bergstresser et al. (2008) and extends findings of Boldin and Cici (2010) for the broader index fund industry.

However, flows to funds sold to defined contribution retirement plans are much more sensitive to fees. To demonstrate this, I investigate how the composition of 401(k) plans of 859 listed firms across 5 years can result in higher fee sensitivity. Employees in firms that offer these plans are often presented with a limited menu of funds to choose from. (Pool et al., 2016) find that there are significant conflicts of interest present due to asset management firms involved in these plans

[1]Index funds represent 57% of total net assets of the CRSP domestic equity fund universe by the end of 2021.

[2]Elton et al. (2004), Hortaçsu and Syverson (2004) and Boldin and Cici (2010) document this for S&P 500 index funds.

[3]Almost 50% of total net assets in my domestic merged CRSP-Morningstar sample belong to funds that can be sold via brokers and in their 2021 annual fact book, the Investment Company Institute (2021) estimates that 83% of U.S. households that own mutual funds, have bought at least some of them through an employer sponsored retirement plans, of which the majority of assets lie in 401(k) plans.

[4]Although Crane and Crotty (2018) evidence of skill in index funds, I find that dispersion in skill is not as large in index funds and that expense ratios are still the main indicator investors should be concerned about when buying an index fund.

favoring their own (often under-performing) funds. However, I find that low cost index funds are more likely to be present in 401(k) plans, even in the presence of conflicts of interest.

These results highlight how intermediation and financial advice in mutual fund markets can affect investor choice. In particular, this paper investigates two large distribution channels for mutual funds where intermediaries face conflicts of interest that could guide investors to high fee funds, although I only find evidence for this in the broker channel. One key difference between the two channels is the standard of care that intermediaries have towards investors. Plan administrators have a fiduciary duty towards investors, which includes selecting a menu of funds with appropriate fees[5], whereas brokers are only subject to a suitability standard of care[6]. In the defined contribution channel, I find evidence consistent with the higher standard of care making investors more sensitive to index fund fees. First, I find that low cost index funds are more likely to be added to 401(k) plans and high cost funds more likely to be deleted. Second, when an asset management firm is involved a 401(k) plan, I do not find evidence of the firms favoring their own high fee index funds.

By studying two large distribution channels for mutual funds with different levels of investor protection, this paper contributes to our understanding of how constraints and conflicts of interest of intermediaries influence allocation to mutual funds. Additionally, by focusing solely on index funds, I can study how intermediaries extract value in a mutual fund market where managerial skill in stock picking and market timing is of little relevance. Furthermore, this paper also extends the index rationality puzzle, where there is over-investment in high cost index funds as described in Elton et al. (2004) and Boldin and Cici (2010), to the broader index fund market beyond S&P 500 index funds. In fact, as we can see in Figure 1, out of the domestic index funds from my CRSP-Morningstar sample, by the end of 2021, S&P 500 index funds only represent around 35% of the domestic equity index fund market whereas the other covered indices in this paper represent 50%.

This paper is related to a literature on how mutual fund investors can make sub-optimal choices. Sirri and Tufano (1998), Hortaçsu and Syverson (2004) and Roussanov et al. (2021) show that high fees charged by mutual funds can be sustained by investors' search costs. Other papers find evidence that asset management firms are able of extracting fees from investor mistakes, either through efforts

[5]This standard has been enforced in this setting as Mellman and Sanzenbacher (2018) document a large number of lawsuits related to 401(k) fiduciaries due to excessive fees and inappropriate investment choices.

[6]As I explain later, brokers must recommend investments of appropriate risk but don't need to consider fees in their recommendations.

of misguiding investors (Elton et al., 2004; Cooper et al., 2005; Barber et al., 2005; Cronqvist, 2006) or due to investors lack of understanding that index funds can be seen as commodities Choi et al. (2009). Alternatively, 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), although 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) and Boldin and Cici (2010) 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 40 basis points (b.p.) compared to the 47 b.p. of actively managed funds[7]. Throughout my sample ranging from January 2000 to June 2023, investors could have saved an average of $300 million every year in fees by following a simple rule of thumb of investing in an index fund with a fee in the lowest 30th percentile[8]. These are all lower bound estimates as larger benefits could be obtained by simply investing in the cheapest fund tracking a given index.

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 Gârleanu and Pedersen (2018). Index funds can be thought of as actively managed funds but with a very narrow investment objective. Indeed, 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. However, fees are a large component of investor returns and are highly persistent (Cooper et al., 2021). From this perspective, investors should react strongly to persistent signals of a fund's future performance net of fees, which is in 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) and Roussanov et al. (2021) 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

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

[8]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. See Section 4.1 of this paper for more details.

suggests the problem is becoming more severe. At the same time there is an argument to be made that search costs have decreased significantly over the past decades[9].

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 and institutional I find that most of the lowest sensitivity to index fund fees 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. 75% of U.S. households that own mutual funds outside their employer sponsored plans rely on such professionals[10]. 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[11]. Until recently, fund brokers were only held to the suitability standard of care. This standard 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. Consistently, 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, I also find evidence that transparency in broker compensation can improve investors' choices into low cost funds. I show 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 theoretical 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 of the intermediary's incentives, consistent with front-end loads being less effective than 12b-1 fees in steering investors to expensive funds.

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

[10]Estimates from the 2021 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.

[11]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.

These findings contribute to the discussion of 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. Additionally, Guercio and Reuter (2014), find that sensitivity of active mutual fund flows to performance is lower for broker sold funds. In this paper, I expand on these findings by providing evidence consistent with brokers recommending expensive index funds that pay them commissions, even when lower-cost and better-performing alternatives are available 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 2021, 83% 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 flows to funds offered to these plans are much more sensitive to fees. Using a hand collected dataset of 401(k) menus, I show that this is because low-fee index funds are more likely to be selected and kept in these retirement plans.

Despite the conflicts of interest also present in this channel, contrary to Pool et al. (2016), I find little evidence that this conflict of interest is exploited. 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 to evaluate them against comparable options, unlike the case with actively managed funds. For example, it might be easier for plan participants to successfully present legal challenges 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%. My findings also contribute to prior literature on 401(k) menu composition (Elton et al., 2006, 2007), where I document that index funds included in 401(k) on average charge fees that are 4 times lower than the average index fund.

Finally, this paper also adds to the discussion on what limits should be placed on fund intermediaries. Roussanov et al. (2021) show that eliminating conflicts of interest in brokers by shutting down fund 12b-1 marketing fees can generate investor welfare, however in their model the size of the mutual fund market is exogenously given. While financial advice given by brokers can be compromised by their conflicts of interest, the absence of this advice can also have the unintended consequence of investors taking less risk and staying out of the equity market (Gennaioli et al., 2015). Alternatively, by combining the results from the two intermediation channels studied in this paper, I offer some initial evidence on the role of standards of care in limiting the extent to which intermediaries exploit the full extent of their conflicts of interest.

2 Data and variable construction

2.1 Index fund selection

Throughout this paper I rely on the CRSP mutual fund and the Morningstar Direct databases to gather data on individual index funds and 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. Since I use daily return data to estimate some variables, the sample starts in January 2000 and ends in June 2023.

I focus on U.S. domestic equity funds and to identify which funds are index funds and which benchmark they track, I combine the CRSP mutual fund and Morningstar Direct following a procedure similar to Berk and van Binsbergen (2015) and Pástor et al. (2015)[12]. I use funds flagged in CRSP as *pure index funds* to avoid any sort of leveraged or smart beta products and from Morningstar Direct, I'm able to obtain the benchmark index for each fund. Finally, because we also need daily return data on the benchmark indices that each fund is targeting we focus on index popular providers such as S&P Global (S&P, Dow and Wilshire brands) and Russell which are also covered in Cremers and Petajisto (2009) as well as Nasdaq, MSCI and CRSP.

Figure 1 provides a snapshot of the total net assets of our covered sample, totaling $6.4 trillion of assets under management by in December 2021 which is roughly 42% of the assets of all domestic equity funds listed in CRSP and 85% of all index funds. In Table IA1 of the Internet Appendix, I also list how many funds in the sample are tracking what indices. In total, the main sample consists of 308 index funds, tracking 51 indices.

[12] I provide a summary of the procedure in Internet Appendix IA1.

2.2 Index fund performance and flows

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, $R_{f,b,t}$, is given as net of the expense ratio, in the most simple decomposition, the return of fund f tracking index b can be decomposed into:

$$R_{f,b,t} = IndexRet_{b,t} - ExpRatio_{i,b,t} + \alpha_{f,b,t}. \quad (1)$$

In most analyses of fund flows I use this simple decomposition where I back out $\alpha_{f,b,t}$ from the other 3 variables which are observable. $\alpha_{f,b,t}$ can be understood as the index fund's out-performance relative to its benchmark gross of fees. This means track $\alpha_{f,b,t}$ 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 $\alpha_{f,b,t}$ is -5 basis points per year with a standard deviation of 1.1%. Unless otherwise stated, I use yearly return measures updated every month for the these variables.

To obtain a fund's exposure to its benchmark, I also run the following OLS regression for every fund using a rolling window of 1 year of daily data to estimate β_I for every fund in a given month[13]:

$$R_{f,b,t} - r^F_t = a_{i,b} + \beta_I(IndexRet_{b,t} - r^F_t) + e_{f,b,t}, \quad (2)$$

where r^F_t is the daily risk free rate and $IndexRet_{b,t}$ is the daily return of the fund's benchmark.

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:

$$AbsBeta_{i,b,t} = |\beta_{If,b,t} - 1|. \quad (3)$$

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 $\alpha_{f,b,t}$ net

[13] While in principle we could also use β_I to obtain estimates of $\alpha_{f,b,t}$, I find that similarly to Elton et al. (2004), this does not impact the results of this paper.

of fees. Once again, I calculate this every month for each fund using a rolling window of daily data. Finally I also obtain two other risk metrics at the fund level, (i) β_{MKT}, measuring the fund's exposure to the CRSP value weighted market portfolio, which is estimated in a similar way as β_I, but instead we replace the index return for returns on the market portfolio and (ii) the 1 year rolling total volatility of the fund returns using daily data which I then annualize.

As is common in the mutual fund literature[14] 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 monthly frequency as this is the frequency at which we have data on fund TNA.

2.3 Aggregating share classes

Fund returns and other fund descriptive data from CRSP and Morningstar 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 Morningstar Fund ID variable for this. 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.

In Table 1 we can see the main descriptive statistics at the fund level. While most funds seem to have gross alpha (gross α_f) between -10 and 10 basis points per year, the dispersion in fees is twice that amount sitting between 10 and 50 basis points when looking at the difference between the 25th and 75th percentiles[15].

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 and institutional fund share

[14] 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.

[15] In contrast, for actively managed domestic equity funds, the year CAPM (Capital Asset Pricing Model) alpha ranges from -4.5% to 4.3% and the total expense ratio from 90 to 135 basis points for the 25th and 75th percentiles respectively.

classes and the Morningstar share type variable to identify retirement funds[16]. 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.

2.4 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 understand how index funds get selected as menu options in these retirement plans.

I 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 from 2013 to 2017, providing me with 401(k) plan menu's of 859 different employers and a total of 1033 plans[17].

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 service providers that deliver record keeping, consulting educational and other services to the 401(k) plan participants. These service providers 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 service provider is an asset management firm and which company is providing this service. In total, I find that in my sample 67% of plans are affiliated

[16] Morningstar reports that these are R, K or J type share classes which are fund shares typically available for defined contributions retirement plan participants to buy.

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

with an asset management firm and 16% of index funds included in a 401(k) plan belong to a fund family affiliated with that plan.

In Table 2 we can see the descriptive statistics of some key variables on the data collected for 401(k) plans. In Panel A, we can see the unconditional probability of an index fund being selected into a plan is quite low at 0.2% while roughly 2% of funds eligible to be added to a plan belong to an asset management that provides services to the plan. In Panel B we can also see some of the main characteristics of index funds already on 401(k) menus. The unconditional probability of a fund being removed from a plan on a given year is 13.5% while almost 16% of index funds are affiliated to a plan service provider. It's also interesting to note that one average, index funds in these retirement plans are cheaper than those in the full sample with an average expense ratio of 10 basis points compared to 40. Finally, Panel C reports some plan level statistics where we can see that 67% of plans in our sample have an asset management firm providing services such as record keeping or acting as plan trustees and that the average plan has 2 index fund options on its menu.

3 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 an index fund investor to analyze when choosing a fund tracking a given index is the fund's expense ratio. While this is shown to be true for S&P 500 funds in Elton et al. (2004), in this section I generalize this finding for the index funds in this sample, tracking a variety of different indices.

To understand what best predicts fund performance I run panel regressions at the monthly frequency of fund alphas net of fees on several variables to predict future fund benchmark adjusted returns net of fees in the cross section of funds. The main predictor variables analyzed are past fund alphas gross of fees, the most recent expense ratio, lagged *AbsBeta* and tracking error volatility. Given that this is a cross sectional analysis, I include Index Style-Time or Benchmark-Time fixed effects[18]. The first compares the cross section of funds tracking an index of a similar style, for example funds tracking the S&P 500 and the Russell 1000, while the latter compares funds that

[18]See Table IA1 for a description of which funds are grouped into which index style.

track the same index. To account auto-correlation in fund performance, standard errors are clustered at the fund level.

Table 3 shows that index fund fees are highly predictive of fund alphas net of fees in the following period. In columns 1 and 2, the dependent variable is monthly net alphas (annualized) in the following month, and in 3 and 4 yearly net alphas in the following year. Across all specifications, we can see that there is almost a reverse one to one relationship between fees and following period alphas, meaning a one percentage point increase in the expense ratio results in the same decrease in net alpha. This should not be surprising and in fact is a similar finding to what Elton et al. (2004) find for S&P 500 funds. 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.

However, I do not find evidence that skill is persistent in index funds as gross alphas do not positively predict future net alpha. In Table 3, all coefficients on gross alpha are insignificant. There is however limited evidence that funds that deviate from their exposure to the index may earn some additional net alpha. This effect is small however, from columns 1 and 2 of Table 3, a one standard deviation increase in $AbsBeta$ results in 1 basis point increase in alpha over the following month, which is economically small. Furthermore, at the yearly horizon this effect disappears. Taken together, these results suggest that two funds tracking the same index, or even indices of a similar style, are very close substitutes differing only in their price.

4 Index fund flows and distribution channels

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 documented by Elton et al. (2004) and Boldin and Cici (2010) is present for funds marketed to different types of investor. I then investigate how broker incentives may influence investor sensitivity to fees.

4.1 Baseline results

We start with a set of baseline results to then compare against the different distribution channels. In the previous section we established that expense ratios 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.

As an initial exercise, I start by calculating the amount of dollars saved in index funds' total expense ratio if they move from the most expensive funds to some of the lower cost funds available. Given the results from Section 3, a simple rule of thumb for an investor would be to simply select a low cost index fund for the index of choice. Specifically, in Panel A of Figure 2, I estimate how much dollars (annualized) would be saved by moving from the most two thirds of expensive S&P 500 funds to the fund sitting at the 33rd percentile. Here we can see that by following this simple strategy could save $100 million per year, roughly 10% of revenues of the S&P 500 market obtained only through fund expense ratios.

However, for non-S&P 500 funds from my sample, the misallocation can be much larger. In Panel B of Figure 2, I perform a similar exercise for the remaining index funds from my sample. To ensure that we're comparing fees of similar funds, I estimate the savings from moving from the two thirds of most expensive funds in terms of the expense ratio of a given *index style*[19], and move them to the fund sitting at the 33rd percentile of that style. There the total fees saved amount to almost $1.5 billion annually towards the end of the sample, or 30% of all fees. This of course is an conservative estimate as in both analyses, we could instead move investors to the cheapest fund of a given style.[20]

To analyze relationship of investors' sensitivity to fees in further detail, I regress monthly fund flows on fund past performance measures to test whether investors are responsive to these measures, especially fees:

[19] See Table IA1 from the Appendix to see which indices I assign to each index style.

[20] Alternatively, we could also measure the dollar amount of index fund alpha generated from implementing such a rule and I find similar results. Cooper et al. (2021) study this question in further detail and reach similar conclusions.

$$flow_{f,b,t} = \delta_1 ExpRatio_{f,b,t-1} + \delta_2 Gross\alpha_{f,b,t-1} + \delta_3 IndexRet_{b,t-1} + Controls + \eta_{b,t} + \epsilon_{i,b,t}, \quad (5)$$

where the main parameter of interest is δ_1, the coefficient on the fund expense ratio.

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 benchmark-time fixed effects which control for any unobserved demand for funds tracking any given index at any month. Under this specification, δ_1 can be interpreted as how investors respond to expense ratios of funds that track the same benchmark index. As I show earlier, funds tracking the same index are very close substitutes so we should expect a negative δ_1.

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 we can see that investors do respond to expense ratios to some extent. In column 4 which includes benchmark-time fixed effects, a one standard deviation decrease in a fund's total expense ratio is associated to an increase in monthly flows by 0.4 percentage points (4.9 p.p. in annual terms). These results are in line with past estimates (Elton et al., 2004; Boldin and Cici, 2010), however results from Figure 2 tell us that many investors are invested into to high fee index funds. Nevertheless, these results give us baseline estimates for fee sensitivity to compare against funds sold through particular distribution channels.

Despite not being very predictive of future fund performance, tracking error volatility is significant and with the expected sign despite its lower influence on future return as shown in the previous section. We can see this as the $TrackErrorVol$ coefficient is negative and statistically significant. This relationship suggests that investors value index funds that provide 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, in column 4 of Table 4, a 1 standard deviation decrease in tracking error volatility translates to monthly net flows growing 0.3 percentage points higher than the average fund (3.4 p.p. in annual terms). When looking within funds tracking the same index, there's also some evidence of investors valuing management skill, Gross α_f, however this effect is economically smaller than what I document for expense ratios and tracking error volatility. A one standard deviation increase results in a growth in net flows of 0.1 percentage points per year.

Also surprising is that for specifications without benchmark-time fixed effects, 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 0.04 and 0.11 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.

Additionally, I find that there is some evidence suggesting that funds providing additional services beyond the investment itself is valued. Proxies such as the number of share classes offered by the fund or fund family size have statistically and economically significant coefficients.

4.2 Retail and institutional investors

The source of the index fund puzzle becomes clearer when I disaggregate index funds into two different types of share classes. In this section, I perform a similar analysis as in Section 4.1, 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 than 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 two 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.

From Table 5, it becomes apparent that investors in retail share classes are primarily responsible for the lower of responsiveness to fees, as institutional investors are twice as responsive to index fund fees. 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, investors in institutional share classes are also sensitive to tracking error volatility and there's also some weak evidence for a preference for funds with higher exposure to market risk.

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 sensitive to fees than institutional investors as the former group is commonly 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. In fact, Hortaçsu and Syverson (2004) attribute a large portion of misallocation in S&P 500 index funds to search costs.

4.3 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 to 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 Act 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. Notably, under this new regulation, brokers are still allowed to market high cost funds even if identical lower cost alternatives exist.

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. In the first two columns we consider any broker sold funds, i.e. funds that have at least on of the following: 12b-1 marketing fees, front-end loads, level-loads or back-end loads. Here we find that the funds sold through brokers are less sensitive to the total expense ratio, suggesting that sales commissions dampen investor sensitivity to fees. Comparing broker and non-broker sold funds, an standard deviation decrease in the expense ratio results in a 0.4 p.p. increase in monthly flows relative to a 1.3 p.p. increase for non-broker sold funds.

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. Furthermore, on average, broker sold fund flows per month are 1 p.p. lower than those not sold through brokers. Nevertheless, 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 columns 3 to 6 of Table 6 we look at the individual effects of the two main broker compensation schemes, 12b-1 fees and front end loads. Here we see that most of the broker effect is driven by the 12b-1 marketing fees. One potential explanation is that investors are more aware that funds with front end loads are directly compensating brokers for their sales efforts since these fees are a one-off cost paid upfront making them more salient to the 12b-1 marketing fees which are part of the total expense ratio and are subtracted from the funds assets on an ongoing basis. In fact, Barber et al. (2005) find evidence of investors reacting to the salience of these fees for the universe of US actively managed equity funds, however in our setting for index funds, what we find is not a negative coefficient on the dummy variable indicating funds with front end loads, but instead a lack of a dampening effect for sensitivity to the total expense ratio.

An alternative explanation for the ineffectiveness of front end loads in distorting index fund allocation is that front end loads are high one-off fees that can be much larger than 12b-1 fees for short holding periods. This could affect investors with shorter horizons, as they may be responding to this additional fee. Unfortunately, we are not able to accurately observe investor horizons and details on front end load fees as these can vary for the same fund across different brokers.

5 Defined contribution plans

A large part of the assets under management of U.S. mutual funds come from defined contribution plans. In 2020, around 80% of US households with investments in mutual funds owned mutual funds through employer sponsored defined contribution pension plans and these plans manage $9.6 trillion dollars in assets[21]. 401(k) plans represent the majority of U.S. defined contribution assets.

401(k) plans are employer sponsored defined contribution plans that give employees of firms tax advantaged way of buying mutual funds. These plans are typically managed employers who often outsource part of this management to outside service providers. It is common for investment management companies to provide such services. These services typically come in the form of administrative, record-keeping or educational services to plan participants. While service providers receive direct compensation for these activities, the price at which they offer these is commonly jointly determined with the menu of funds that is offered to plan participants. When investment management companies act as service providers, there is an incentive for these firms to favor their

[21] See Investment Company Institute (2021).

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. In Table 2 I document that in my sample, that the average plan has on average 6 domestic equity funds and 2 equity index funds[22].

These plans are interesting to study in this context as not only does it represent one of the largest distribution channels for mutual funds in the US, but it also has two interesting characteristics relative to the broker channel. First, there are significant conflicts of interest between mutual fund investors and the distributors in these plans[23]. Second however is that unlike in the broker channel, fund intermediaries in 401(k) plans are subject to a much stricter standard of care with respect to mutual fund investors. As a result, from this we gain a better understanding on how such investment protections might affect investors. To study this I will use both fund level data as well as a hand collected sample of 401(k) plan menus described in Section 2.

While the presence of the conflict of interest may result in many investors being stuck to a restrictive 401(k) menu with high fee funds due to investment management firms favoring their own high cost 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 firms' power on setting menus. Plan sponsors (i.e. employers) and service providers, 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[24]. If this threat of lawsuits that plan fiduciaries face is large enough, we may see that low-fee index funds are more likely added

[22]These lower numbers reflect that in my sample, I only match 401(k) menu composition to U.S. domestic equity mutual funds.

[23]Pool et al. (2016) provide a detailed explanation and related empirical evidence on the conflicts of interest present in 401(k) plans.

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

to these menus, making 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 service providers and earn additional sources of revenue from higher margin services such as including higher cost actively managed mutual funds in these menus.

5.1 Fund level analysis

I start by looking mutual fund data both for fund shareclasses that are exclusively sold to employer sponsored pension plans as well as aggregate data at the fund level. For the fund shareclass analysis, I rely on a retirement fund indicator obtained from Morningstar that tells us whether a given shareclass is exclusively sold to defined contribution plans. I run a similar analysis as the one in Table 5. Here, I aggregate data at the fund level using only retirement shareclasses and estimate the regression from Equation 5.

In Table 7, we can see that funds exclusively offered to defined contribution plans are much more sensitive to fees when comparing to the baseline estimates from Table 4. For these funds, the coefficient on the expense ratio is 5 times larger that that of the baseline estimates[25].

However, funds do not need a special shareclass to be offered in 401(k) plans. To expand the sample, I consider that any fund family with this type of shareclass for a given fund, may also offer other funds without these specific shareclasses to these plans. In Table 8, I estimate a similar regression as 5 where I add an dummy variable that is equal to one for fund families that have at least one fund with such a shareclass, and its interaction with the fund's expense ratio. Through the negative coefficient on the interaction term in Table 8, we can see that funds of fund families more involved with defined contribution plans are more sensitive to fees. This evidence is suggestive that menu composition of 401(k) plans is sensitive to index fund fees, however the next section analyzes this in more detail.

5.2 Index fund menu options in 401(k) plans

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 deleted or added to a 401(k) plan.

[25]This difference is statistically significant at the 99% level.

For fund deletions, the dependent variable is a dummy that is equal to one if a fund that was observed in plan p in year $t-1$, is no longer present in that plan in year t. For new index funds added to a plan, the potential index fund additions on a given year a 401(k) plan are all the index funds from my CRSP-Morningstar merged sample, minus the funds that were already present in the previous year. Conversely, for these sets of funds I create a dummy variable that is equal to one if the fund is added to a menu on a given year and zero otherwise. From Table 2 Panel A we can see that the unconditional probability of an index fund being newly added to a 401(k) plan is 0.2% and in Panel B we can see that the unconditional probability of an index fund being deleted from a 401(k) plan is 13.5%.

In addition to the expense ratio, we are also interested in whether a fund is affiliated with the with the plan's largest service provider, when that provider is an investment management firm, as well as how that interacts with the expense ratio. Under the view that investment management firms 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, meaning that high fee affiliated funds would be less likely to be deleted from 401(k) menus and more likely to be added.

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 style being selected into a 401(k) plan while controlling for unobserved plan level variables.

In Table 9, I find that high fee index funds are more likely to be deleted from menus. In column 4, the specification with both index style-year and plan-year fixed effects, I find that a one standard deviation increase in the expense ratio results in a 26 p.p. higher probability of deletion, which is double the unconditional likelihood of 13.5%. I find no evidence of service providers favoring their own index fund offerings as the interaction term of affiliated funds and expense ratio is statistically insignificant. While this may be due to a limited sample size, the combined effect for affiliated funds would still result in a positive relationship between the expense ratio and deletion likelihood for affiliated funds.

Moving to plan additions, from Table 10 I find that funds with lower expense ratios are more likely to be added to 401(k) plans. Once again, when looking at the final column where we study the probability of a fund tracking the same index style being added to a plan, I find that a one standard deviation increase in the expense ratio reduces the likelihood of plan addition by 0.1 p.p. which despite seeming small, it's 5 times the unconditional probability of an index fund being added to a plan. Unlike in the case for fund deletions, here I do find some evidence of favoritism as affiliated funds are 1 to 1.7 percentage points more likely to be added to a 401(k) plans. Nevertheless, looking at the interaction terms between fund affiliation and expense ratio in columns 3 and 4, I don't find evidence that fund affiliation does not alter the sensitivity of addition rates to expense ratios.

While the fiduciary standard may drive the results in this section, an alternative explanation for the fund addition portion of this section is that investment management companies that act as service providers to 401(k) plans may use their cheap index funds as a way to become plan service providers to potentially offer the plan sponsor and plan participants higher margin services. However, I find unreported regressions, I find similar result for Table 10 when restricting the sample to plans where there's no investment management company acting as a service provider.

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. Nevertheless, given that index fund choice within a given menu is very restricted as the median plan only has 2 index funds, we can attribute most of the fund level evidence to the menu composition rather than investor flows within plans[26].

6 Conclusion

Even though index fund performance is remarkably predictable, I find that investors are no very responsive to fees, which strongly predict future performance. I extend evidence on the index fund

[26]I confirm this in unreported regressions, where I find that fund flows within 401(k) plans are not sensitive to fees when looking exclusively at index funds. This can be driven by the fact that index funds within a given plan may already be the cheapest option.

puzzle to funds tracking multiple indices and by showing it still persists today. 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.

When attempting to understand retail mutual fund investors' low 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 brokers and investment management firms. In the 401(k) market however, 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

Barber, Brad M., Xing Huang and Terrance Odean. (2016). "Which factors matter to investors? Evidence from mutual fund flows". *The Review of Financial Studies*, 29(10), pp. 2600-2642. https://doi.org/10.1093/rfs/hhw054

Barber, Brad M., Terrance Odean and Lu Zheng. (2005). "Out of sight, out of mind: The effects of expenses on mutual fund flows". *The Journal of Business*, 78(6), pp. 2095-2120. https://doi.org/10.1086/497042

Bergstresser, Daniel, John M. R. Chalmers and Peter Tufano. (2008). "Assessing the costs and benefits of brokers in the mutual fund industry". *The Review of Financial Studies*, 22(10), pp. 4129-4156. https://doi.org/10.1093/rfs/hhp022

Berk, Jonathan B., and Richard C. Green. (2004). "Mutual fund flows and performance in rational markets". *Journal of Political Economy*, 112(6), pp. 1269-1295. https://doi.org/10.1086/424739

Berk, Jonathan B., and Jules H. van Binsbergen. (2015). "Measuring skill in the mutual fund industry". *Journal of Financial Economics*, 118(1), pp. 1-20. https://doi.org/10.1016/j.jfineco.2015.05.002

Berk, Jonathan B., and Jules H. van Binsbergen. (2016). "Assessing asset pricing models using revealed preference". *Journal of Financial Economics*, 119(1), pp. 1-23. https://doi.org/10.1016/j.jfineco.2015.08.010

Boldin, Michael, and Gjergji Cici. (2010). "The index fund rationality paradox". *Journal of Banking & Finance*, 34(1), pp. 33-43. https://doi.org/10.1016/j.jbankfin.2009.06.018

Choi, James J., David Laibson and Brigitte C. Madrian. (2009). "Why does the law of one price fail? An experiment on index mutual funds". *The Review of Financial Studies*, 23(4), pp. 1405-1432. https://doi.org/10.1093/rfs/hhp097

Collins, Sean. (2005). "Are S&P 500 index mutual funds commodities?". *Perspective, Investment Company Institute*, 11(3), p. N3. https://www.idc.org/doc-server/pdf%3Aper11-03.pdf

Cooper, Michael J., Huseyin Gulen and P. Raghavendra Rau. (2005). "Changing names with style: Mutual fund name changes and their effects on fund flows". *The Journal of Finance*, 60(6), pp. 2825-2858. https://doi.org/10.1111/j.1540-6261.2005.00818.x

Cooper, Michael J., Michael Halling and Wenhao Yang. (2021). "The persistence of fee dispersion among mutual funds". *Review of Finance*, 25(2), pp. 365-402. https://doi.org/10.1093/rof/rfaa023

Crane, Alan D., and Kevin Crotty. (2018). "Passive versus active fund performance: do index funds have skill?". *Journal of Financial and Quantitative Analysis*, 53(1), pp. 33-64. https://doi.org/10.1017/S0022109017000904

Cremers, K. J. Martijn, and Antti Petajisto. (2009). "How active is your fund manager? A new measure that predicts performance". *The Review of Financial Studies*, 22(9), pp. 3329-3365. https://doi.org/10.1093/rfs/hhp057

Cronqvist, Henrik. (2006). "Advertising and portfolio choice". Working Paper. https://doi.org/10.2139/ssrn.920693

Egan, Mark. (2019). "Brokers versus retail investors: Conflicting interests and dominated products". *The Journal of Finance*, 74(3), pp. 1217-1260. https://doi.org/10.1111/jofi.12763

Ellison, Glenn, and Sara Fisher Ellison. (2009). "Search, obfuscation, and price elasticities on the internet". *Econometrica*, 77(2), pp. 427-452. https://doi.org/10.3982/ECTA5708

Elton, Edwin J., Martin J. Gruber and Christopher R. Blake. (2006). "The adequacy of investment choices offered by 401(k) plans". *Journal of Public Economics*, 90(6-7), pp. 1299-1314. https://doi.org/10.1016/j.jpubeco.2005.08.001

Elton, Edwin J., Martin J. Gruber and Christopher R. Blake. (2007). "Participant reaction and the performance of funds offered by 401(k) plans". *Journal of Financial Intermediation*, 16(2), pp. 249-271. https://doi.org/10.1016/j.jfi.2006.12.002

Elton, Edwin J., Martin J. Gruber and Jeffrey A. Busse. (2004). "Are investors rational? Choices among index funds". *The Journal of Finance*, 59(1), pp. 261-288. https://doi.org/10.1111/j.1540-6261.2004.00633.x

Gârleanu, Nicolae, and Lasse Heje Pedersen. (2018). "Efficiently inefficient markets for assets and asset management". *The Journal of Finance*, 73(4), pp. 1663-1712. https://doi.org/10.1111/jofi.12696

Gennaioli, Nicola, Andrei Shleifer and Robert Vishny. (2015). "Money doctors". *The Journal of Finance*, 70(1), pp. 91-114. https://doi.org/10.1111/jofi.12188

Guercio, Diane Del, and Jonathan Reuter. (2014). "Mutual fund performance and the incentive to generate alpha". *The Journal of Finance*, 69(4), pp. 1673-1704. https://doi.org/10.1111/jofi.12048

Hortaçsu, Ali, and Chad Syverson. (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), pp. 403-456. https://doi.org/10.1162/0033553041382184

Inderst, Roman, and Marco Ottaviani. (2012a). "Financial advice". *Journal of Economic Literature*, 50(2), pp. 494-512. https://doi.org/10.1257/jel.50.2.494

Inderst, Roman, and Marco Ottaviani. (2012b). "How (not) to pay for advice: A framework for consumer financial protection". *Journal of Financial Economics*, 105(2), pp. 393-411. https://doi.org/10.1016/j.jfineco.2012.01.006

Investment Company Institute. (2021). *2021 Investment Company Institute Factbook*. https://www.ici.org/system/files/2021-05/2021_factbook.pdf

Kronlund, Mathias, Veronika Krepely Pool, Clemens Sialm and Irina Stefanescu. (2020). "Out of sight no more? The effect of fee disclosures on 401(k) investment allocations". *Journal of Financial Economics*, 141(2), pp. 644-688. https://doi.org/10.1016/j.jfineco.2021.04.008

Mellman, George S., and Geoffrey T. Sanzenbacher. (2018). "401(k) lawsuits: What are the causes and consequences?". Issue in Brief, 18-8, Center for Retirement Research at Boston College. http://hdl.handle.net/2345/bc-ir:108010

Pastor, Lubos, Robert F. Stambaugh and Lucian A. Taylor. (2015). "Scale and skill in active management". *Journal of Financial Economics*, 116(1), pp. 23-45. https://doi.org/10.1016/j.jfineco.2014.11.008

Pool, Veronika K., Clemens Sialm and Irina Stefanescu. (2016). "It pays to set the menu: Mutual fund investment options in 401(k) plans". *The Journal of Finance*, 71(4), pp. 1779-1812. https://doi.org/10.1111/jofi.12411

Roussanov, Nikolai, Hongxun Ruan and Yanhao Wei. (2021). "Marketing mutual funds". *The Review of Financial Studies*, 34(6), pp. 3045-3094. https://doi.org/10.1093/rfs/hhaa095

Sirri, Erik R., and Peter Tufano. (1998). "Costly search and mutual fund flows". *The Journal of Finance*, 53(5), pp. 1589-1622. https://doi.org/10.1111/0022-1082.00066

Figure 1: **Index fund assets by index tracked.** This figure plots the total net assets managed by funds in my final sample that (1) track the S&P 500 index, (2) track other popular indices covered in the paper listed in Table IA1 of the Internet Appendix and (3) cover remaining indices.

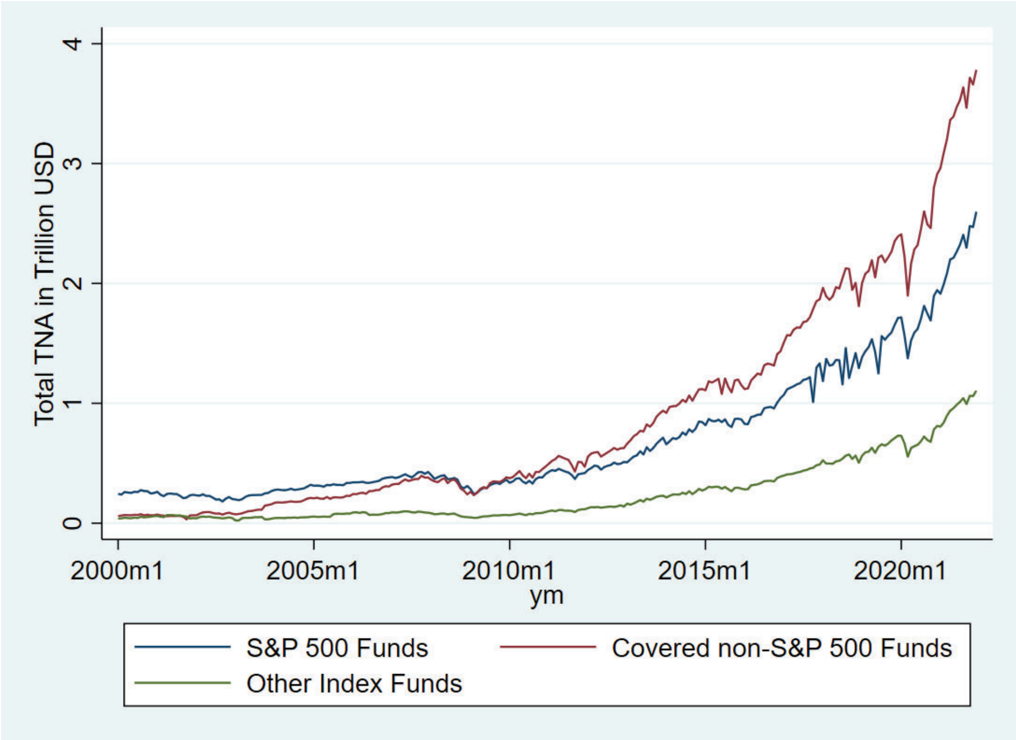
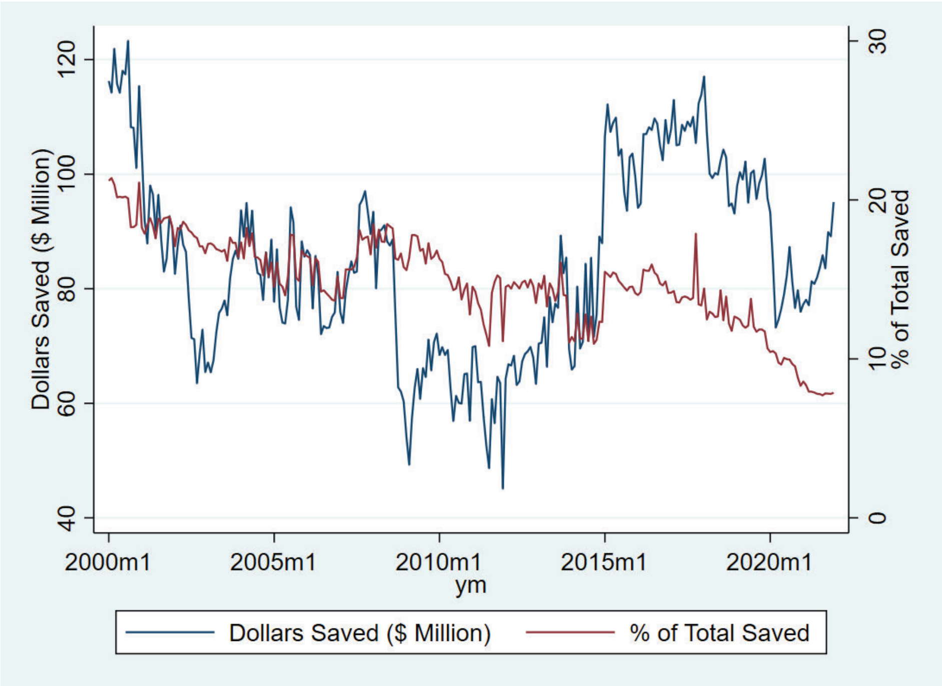


Figure 2: **Fees saved from moving to cheaper funds.** This figure plots how much investors could save in fees by switching to lower cost funds. In both panels, I calculate how much of the total expense ratio in annualized dollars could be saved every month by moving from a fund in the top 2 terciles in fees to the fund that sits at the 33rd percentile in fees. Panel (a) does this for S&P 500 funds and Panel (b) for non-S&P 500 funds covered in our sample. In Panel (b) distributions are calculated for each index type in a given quarter, for example, we calculate the fees saved from moving from the top 2 thirds of most expensive large growth index funds to the fund on the 33rd percentile for this style.

(a) S&P 500 fees saved



(b) Non-S&P 500 fees saved

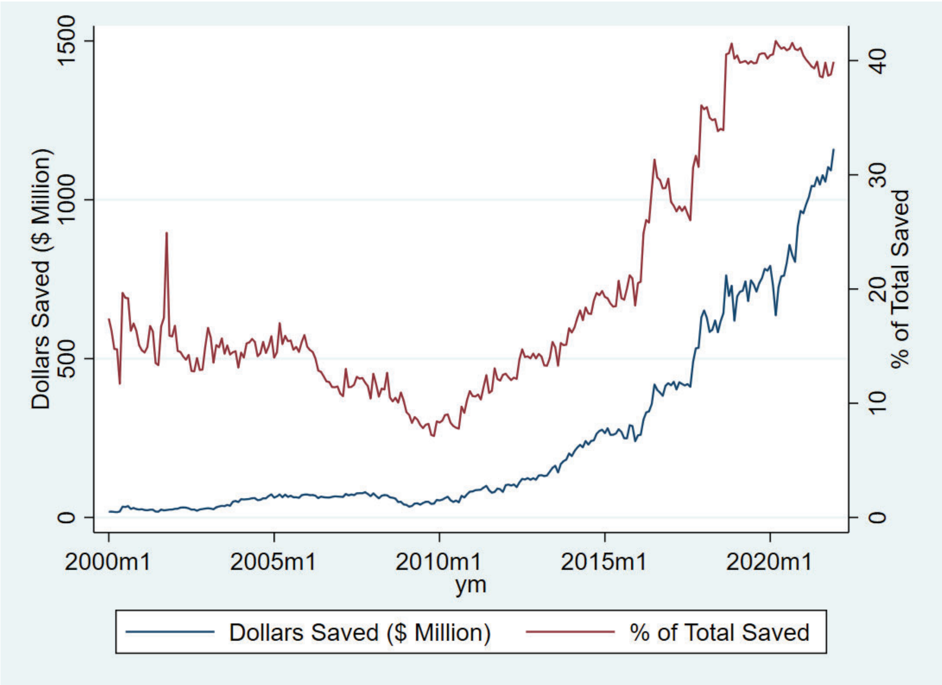


Table 1: **Summary statistics.** This table shows summary statistics of the variables for index funds used in this paper. For fund flows, fund returns and its components: fund gross returns are the fund yearly returns gross of the expense ratio, gross α_f is the fund alpha (Equation 1) gross of the expense ratio, index return is the annual return of the index that a fund targets, expense ratio is the fund's total expense ratio and fund flow is the fund's monthly flows defined in Equation 4. For fund risk variables: tracking error vol is the volatililty or standard deviation the fund's α_f net of fees, Abs Beta is the absolute of a fund's loading on it's target index and one and β_{MKT} is the fund's loading on the CRSP value weighted index of the US stock market. For other continuous fund level variables: Fund TNA is the fund's total net assets in millions of dollars, Fund age is the age of the fund in months, Turnover is the fund turnover ratio defined in CRSP, Num share classes is the total number of share classes of a given fund. Continuous variables for a fund's family: 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, fam num funds is the total number of funds offered by the fund family and family flows is the net monthly flows to a fund family. For binary variables we have: broker fund which is equal to one if the fund has 12b-1 marketing fees, front, level or rear loads; fund 12b1 which is equal to one if the fund has 12b-1 marketing fees; front load which is equal to one if the fund has front load fees; institutional, retail and DC fund are equal to one if the fund has at least one share class destined for institutional, retail or defined contribution plan investors respectively; DC family is on if the fund is part of a family that has at least one fund offered to defined contribution plans; and family w/ 5 Star Fund is one if the fund is part of a fund family that has at least one domestic equity with a 5 star Morningstar rating.

Variable	**Obs**	**Mean**	**Std**	**Min**	**25th Pct**	**Median**	**75th Pct**	**Max**
Fund Gross Returns	42705	0.089	0.186	-0.388	-0.017	0.106	0.193	0.621
Gross α_f	42705	0.000	0.011	-0.064	-0.001	0.000	0.001	0.059
Index Return	42705	0.095	0.193	-0.625	-0.012	0.111	0.197	1.805
Expense Ratio	42705	0.004	0.004	0.000	0.001	0.003	0.005	0.017
Fund flow	42705	0.009	0.085	-0.301	-0.009	0.001	0.015	0.617
Tracking Error Vol	42705	0.008	0.014	0.000	0.001	0.004	0.007	0.083
Abs Beta	42705	0.008	0.018	0.000	0.001	0.002	0.004	0.114
β_{MKT}	42705	1.040	0.127	0.785	0.961	0.999	1.099	1.460
Fund TNA	42705	8908	27690	5	180	919	3788	197315
Fund Age	42705	146.812	89.963	4	72	135	206	404
Turnover	42705	0.399	1.049	0.010	0.060	0.150	0.300	7.440
Num Share Classes	42705	2.248	1.625	1	1	2	3	9
Family TNA	42705	858956	1583941	218	27930	151545	767927	7648384
Family Asset Classes	42705	8.919	2.203	2	8	10	11	12
Family Num Funds	42705	66.808	60.459	1	22	51	103	292
Family Flows	42705	0.004	0.020	-0.055	-0.004	0.003	0.010	0.092
Broker Fund	42705	0.535	0.499					
Fund 12b1	42705	0.450	0.498					
Front Load	42705	0.191	0.393					
Institutional	42705	0.776	0.417					
Retail	42705	0.525	0.499					
DC Fund	42705	0.172	0.378					
DC Family	42705	0.528	0.499					
Family w/ 5 Star Fund	42705	0.704	0.456					

Table 2: **401(k) Summary Statistics.** This table reports relevant summary statistics for the collected sample for 401(k) plans. Panel A reports summary statistics of index funds eligible be added to a plan on a given year: new fund is an binary variable equal to one for funds that get added to a plan on a given year and affiliated is equal to one if the eligible fund and the 401(k) plan's main service provider are the same fund family or asset management firm. Panel B reports statistics of index funds already in 401(k) plans: deleted fund is a binary variable equal to one for funds deleted from a plan in a given year, affiliated is equal to one if the fund and the 401(k) plan's main service provider are the same fund family, option size is the total fund assets held by the 401(k) plan participants on a given year in millions of dollars, the expense ratio is the total expense ratio of the fund, and fund TNA is the total net assets of the fund in millions of dollars. In Panel C we report plan level descriptive statistics: Inv. Mgr. affiliated is a binary variable equal to one when the plan lists an asset management firm as a service provider, No. of affiliated is the number of funds in the plan's menu that are affiliated to the the plans largest asset management firm service provider, No. of options is the total amount of funds offered in the plan's menu, No. of Idx Options in the amount of domestic equity index funds offered in the plan's menu and plan size is the total assets of the plan in millions of dollars.

Panel A: Potential Index Fund Additions

	Obs	**Mean**	**Std**
New Fund	484364	0.002	0.042
Affiliated	484364	0.020	0.141

Panel B: Index Funds in 401(k) Plans

	Obs	**Mean**	**Std**	**Min**	**25th Pct**	**Median**	**75th Pct**	**Max**
Deleted Fund	4602	0.135	0.342					
Affiliated	4602	0.158	0.365					
Option Size	4602	48.6	201.4	0.0	1.1	5.8	27.1	5134.0
Expense Ratio	4602	0.001	0.001	0.0002	0.0007	0.0009	0.0011	0.0163
Fund TNA	4602	68765	70353	13	11164	46724	117341	211496

Panel C: 401(k) Plan Descriptive Statistics

	Obs	**Mean**	**Std**	**Min**	**25th Pct**	**Median**	**75th Pct**	**Max**
Inv. Mgr. Affiliated	3290	0.670	0.470					
No. of Affiliated	3290	0.647	1.574	0	0	0	0	12
No. of Options	3290	6.348	5.850	1	4	6	8	239
No. of Idx Options	3290	2.001	1.67	0	1	2	3	31
Plan Size	3290	1067	3599	1	61	219	653	56361

Table 3: **Index fund performance.** This table reports panel regressions of index funds' monthly and yearly α_f net of fees on components of the funds' past return, expenses, risk and individual characteristics. In columns 1 and 2 the dependent variable is the monthly α_f, and in columns 3 and 4 yearly α_f. Independent variables are described in Table 1. Standard errors are in brackets and are clustered at the fund level. Significance: ***99%, **95%, *90%.

Dep Var:	Net α_f (t)		Net α_f $(t \rightarrow t+12)$	
	(1)	(2)	(3)	(4)
Expense Ratio $(t-1)$	-1.064***	-1.063***	-1.092***	-1.097***
	[0.073]	[0.084]	[0.079]	[0.094]
Gross α_f $(t-1)$	0.056	0.034	-0.013	0.000
	[0.036]	[0.041]	[0.033]	[0.035]
Tracking Error Vol $(t-1)$	-0.024	-0.016	-0.001	0.016
	[0.033]	[0.040]	[0.046]	[0.060]
Abs Beta $(t-1)$	0.049*	0.064*	0.038	0.025
	[0.026]	[0.034]	[0.031]	[0.047]
β_{MKT} $(t-1)$	0.006	0.024	0.004	-0.005
	[0.004]	[0.023]	[0.004]	[0.020]
Fund Flows $(t-1)$	-0.006***	-0.003	-0.002**	-0.001
	[0.002]	[0.002]	[0.001]	[0.001]
Log fund TNA $(t-1)$	0.000	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]
Log fund Age $(t-1)$	0.001	0.001	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]
Turnover Ratio $(t-1)$	-0.001***	-0.001***	-0.001***	-0.001***
	[0.000]	[0.000]	[0.000]	[0.000]
Fixed effects:				
Index Style × Time	Yes	No	Yes	No
Benchmark × Time	No	Yes	No	Yes
Observations	42403	39300	39326	36285
R^2	0.137	0.2785	0.281	0.3759

Table 4: **Index fund flows.** This table reports panel regressions of index fund's net fund flows on components of the funds past return, expenses, risk, individual characteristics and characteristics of its mutual fund family. Independent variables are described in Table 1. Fam Num Funds has been rescaled such that it measures the total number of funds in a fund family in hundreds. Standard errors are in brackets and are clustered at the fund level. Significance: ***99%, **95%, *90%.

	(1)	(2)	(3)	(4)
Expense Ratio $(t-1)$	-1.494***	-1.636***	-1.402***	-1.080***
	[0.391]	[0.430]	[0.389]	[0.374]
Gross α_f $(t-1)$	0.056	0.058	0.076	0.102*
	[0.052]	[0.052]	[0.050]	[0.053]
Index Return $(t-1)$	0.041**	0.040**	0.108***	
	[0.018]	[0.018]	[0.039]	
Tracking Error Vol $(t-1)$	-0.042	-0.045	-0.127*	-0.197***
	[0.070]	[0.073]	[0.067]	[0.069]
Abs Beta $(t-1)$	-0.083	-0.08	-0.038	0.096
	[0.053]	[0.056]	[0.058]	[0.066]
β_{MKT} $(t-1)$	0.001	0.001	0.002	0.087**
	[0.006]	[0.007]	[0.015]	[0.042]
Front Load $(t-1)$	0.000	0.000	0.000	-0.001
	[0.002]	[0.002]	[0.002]	[0.002]
12b1 Fund	-0.007***	-0.006***	-0.006***	-0.005***
	[0.002]	[0.002]	[0.002]	[0.002]
Log Fund TNA $(t-1)$	-0.003***	-0.004***	-0.003***	-0.002***
	[0.001]	[0.001]	[0.001]	[0.001]
Log Fund age $(t-1)$	-0.011***	-0.011***	-0.011***	-0.013***
	[0.002]	[0.002]	[0.002]	[0.002]
Turnover Ratio $(t-1)$	0.009***	0.009***	0.009***	0.008***
	[0.001]	[0.002]	[0.001]	[0.002]
Num Share Classes $(t-1)$	0.003***	0.003***	0.003***	0.002***
	[0.001]	[0.001]	[0.001]	[0.001]
Log Fam TNA $(t-1)$	0.004***	0.004***	0.004***	0.003***
	[0.001]	[0.001]	[0.001]	[0.001]
Fam Asset Classes $(t-1)$	-0.002***	-0.002***	-0.002***	-0.001**
	[0.001]	[0.001]	[0.001]	[0.001]
Fam Num Funds $(t-1)$	-0.001	0.000	-0.001	-0.001
	[0.001]	[0.002]	[0.001]	[0.002]
Fam Flows $(t-1)$	0.124***	0.123***	0.109***	0.102***
	[0.041]	[0.041]	[0.040]	[0.038]
5 Star Fund Fam $(t-1)$	0.001	0.001	0.001	0.000
	[0.001]	[0.001]	[0.001]	[0.001]
Institutional $(t-1)$	-0.003	-0.003*	-0.003	-0.004**
	[0.002]	[0.002]	[0.002]	[0.002]
Retail $(t-1)$	0.001	0.000	0.000	-0.001
	[0.002]	[0.002]	[0.002]	[0.002]
DC Fund $(t-1)$	-0.005**	-0.004*	-0.005**	-0.003
	[0.002]	[0.002]	[0.002]	[0.002]
Fixed effects:				
Time	Yes	Yes	No	No
Benchmark	No	Yes	No	No
Index Style × Time	No	No	Yes	No
Benchmark × Time	No	No	No	Yes
Observations	42195	42195	42136	39072
R^2	0.052	0.053	0.157	0.377

Table 5: **Index fund flows by share class type.** This table reports panel regressions of index funds' net fund flows on components of the funds past return, expenses, and risk. In columns 1 and 2 I use fund level variables by aggregating only share classes destined for retail investors and in columns 3 and 4 we do the same but for share classes destined for institutional investors. Independent variables are described in Table 1, and the same control variables from Table 4 are included except for the binary variables for share class type. Standard errors are in brackets and are clustered at the fund level. Significance: ***99%, **95%, *90%.

Dep Var:	Retail		Institutional	
	(1)	(2)	(3)	(4)
Expense Ratio $(t-1)$	-1.639***	-1.400***	-2.949***	-2.617***
	[0.476]	[0.433]	[0.423]	[0.470]
Gross α_f $(t-1)$	0.023	0.243*	0.104**	0.076
	[0.137]	[0.140]	[0.050]	[0.055]
Index Return $(t-1)$	0.006		0.164***	
	[0.034]		[0.020]	
Tracking Error Vol $(t-1)$	-0.093	-0.077	-0.146**	-0.233***
	[0.121]	[0.118]	[0.060]	[0.064]
Abs Beta $(t-1)$	-0.087	0.147	0.085*	0.194***
	[0.063]	[0.092]	[0.050]	[0.057]
β_{MKT} $(t-1)$	-0.035*	0.068	0.024*	0.073*
	[0.019]	[0.078]	[0.013]	[0.038]
Fixed effects and controls:				
Controls	Yes	Yes	Yes	Yes
Index Style × Time	Yes	No	Yes	No
Benchmark × Time	No	Yes	No	Yes
Observations	21655	16817	32389	27538
R^2	0.319	0.180	0.173	0.279

Table 6: **Index fund flows: Broker sold funds.** This table reports panel regressions of index funds' net fund flows on components of the expenses, whether funds are sold through brokers, have 12b-1 marketing fees or front end loads and their interaction with the fund's expense ratio. The same control variables from Table 4 are included. Standard errors are in brackets and are clustered at the fund level. Significance: ***99%, **95%, *90%.

	(1)	(2)	(3)	(4)	(5)	(6)
Expense Ratio $(t-1)$	-3.536***	-3.188***	-3.438***	-3.108***	-1.842***	-1.393***
	[0.690]	[0.710]	[0.574]	[0.595]	[0.547]	[0.518]
Broker sold	-0.012***	-0.010***				
	[0.002]	[0.002]				
12b1 Fund			-0.014***	-0.013***	-0.007***	-0.006***
			[0.002]	[0.003]	[0.002]	[0.002]
Front Load			-0.003	-0.003	-0.005	-0.004
			[0.002]	[0.002]	[0.004]	[0.004]
Expense Ratio $(t-1)$	2.146***	2.097***				
$\times$ Broker sold	[0.634]	[0.653]				
Expense Ratio $(t-1)$			2.506***	2.468***		
$\times$ 12b1 Fund			[0.546]	[0.570]		
Expense Ratio $(t-1)$					0.845	0.596
$\times$ Front Load					[0.677]	[0.638]
Fixed effects and controls:						
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Index Style $\times$ Time	Yes	No	Yes	No	Yes	No
Benchmark $\times$ Time	No	Yes	No	Yes	No	Yes
Observations	42354	39278	42354	39278	42354	39278
R^2	0.158	0.378	0.158	0.379	0.158	0.378

Table 7: **Index fund flows: DC plan funds.** This table reports panel regressions of index funds' net fund flows on components of the funds past return, expenses, and risk. Fund level observations are aggregated using only share classes exclusively destined for defined contribution plan investors. Independent variables are described in Table 1, and the same control variables from Table 4 are included except for the binary variables for share class type. Standard errors are in brackets and are clustered at the fund level. Significance: ***99%, **95%, *90%.

	(1)	(2)	(3)	(4)
Expense Ratio $(t-1)$	-7.108***	-6.561***	-5.546***	-5.118***
	[0.948]	[0.928]	[1.048]	[1.105]
Gross α_f $(t-1)$	0.073	0.065	0.154	0.143
	[0.120]	[0.126]	[0.142]	[0.154]
Index Return $(t-1)$	0.012		0.277**	
	[0.020]		[0.114]	
Tracking Error Vol $(t-1)$	-0.468***	-0.359**	-0.387	-0.293
	[0.163]	[0.143]	[0.231]	[0.234]
Abs Beta $(t-1)$	-0.098	-0.071	-0.016	-0.246
	[0.138]	[0.162]	[0.267]	[0.320]
β_{MKT} $(t-1)$	-0.011	-0.004	0.178**	0.003
	[0.016]	[0.017]	[0.073]	[0.154]
Fixed effects:				
Controls	Yes	Yes	Yes	Yes
Time	Yes	Yes	No	No
Benchmark	No	Yes	No	No
Index Style × Time	No	No	Yes	No
Benchmark × Time	No	No	No	Yes
Observations	7060	7060	6261	5490
R^2	0.184	0.197	0.292	0.339

Table 8: **Index fund flows: DC family funds.** This table reports panel regressions of index funds' net fund flows on components of the expenses, whether funds are part of a fund family that has at least one fund with a share class destined exclusively to defined contribution plan investors and the respective interaction with the fund's expense ratio. The same control variables from Table 4 are included. Standard errors are in brackets and are clustered at the fund level. Significance: ***99%, **95%, *90%.

	(1)	(2)	(3)	(4)
Expense Ratio $(t-1)$	-0.977**	-1.148**	-0.942**	-0.661*
	[0.395]	[0.447]	[0.392]	[0.386]
DC Fam	0.011***	0.012***	0.011***	0.010***
	[0.002]	[0.002]	[0.002]	[0.002]
Expense Ratio $(t-1)$	-2.064***	-2.004***	-1.905***	-1.638***
× DC Fam	[0.545]	[0.593]	[0.534]	[0.549]
Fixed effects and controls:				
Controls	Yes	Yes	Yes	Yes
Index Style × Time	Yes	No	Yes	No
Benchmark × Time	No	Yes	No	Yes
Observations	42407	42407	42354	39278
R^2	0.053	0.054	0.159	0.379

Table 9: **401(k) Plan menu deletions** This table reports regressions of a binary variable on whether a index fund is deleted from a plan menu in a given year on components of the funds' expenses, affiliation to plan service providers, past return, risk and other plan level characteristics. Independent variables not described in Table 1 are *affiliated* which takes the value of 1 when the fund is affiliated to plan's largest asset management service provider, fund vol which is the fund's total annualized volatility, log option size which is log of the total assets of the fund held in a given plan-year and other index is a binary variable equal to one if the plan has another index fund of the same style. Standard errors are in brackets and are clustered at the fund level. Significance: ***99%, **95%, *90%.

	(1)	(2)	(3)	(4)
Expense Ratio $(t-1)$	32.884	60.835***	36.156	66.520***
	[28.465]	[19.025]	[31.593]	[20.714]
Afiiliated	-0.086	-0.110*	-0.013	-0.022
	[0.076]	[0.059]	[0.122]	[0.082]
Expense Ratio $(t-1)$			-33.202	-39.471
× Afiiliated			[43.252]	[29.009]
Gross α_f $(t-1)$	-1.834	-7.502	-1.705	-6.980
	[8.292]	[5.796]	[8.184]	[5.878]
Index Return $(t-1)$	-0.751	0.034	-0.745	0.065
	[0.951]	[1.179]	[0.945]	[1.180]
Fund vol $(t-1)$	-3.339**	0.043	-3.286**	-0.225
	[1.367]	[3.712]	[1.326]	[3.741]
Log Fund TNA $(t-1)$	-0.023	0.006	-0.023	0.008
	[0.033]	[0.016]	[0.033]	[0.016]
Log Fund age $(t-1)$	-0.009	-0.044	-0.009	-0.047
	[0.047]	[0.046]	[0.047]	[0.046]
Turnover Ratio $(t-1)$	-0.085	-0.057	-0.089	-0.065
	[0.096]	[0.064]	[0.097]	[0.066]
Log Option Size $(t-1)$	0.014	-0.011	0.014	-0.011
	[0.014]	[0.009]	[0.014]	[0.009]
Other Index $(t-1)$	0.023	-0.008	0.024	-0.008
	[0.080]	[0.059]	[0.080]	[0.059]
Fixed effects:				
Plan × Year	Yes	Yes	Yes	Yes
Index Style × Year	No	Yes	No	Yes
Observations	3414	3403	3414	3403
R^2	0.603	0.669	0.603	0.670

Table 10: **401(k) Plan menu additions** This table reports regressions of a binary variable on whether a index fund is added to a plan menu in a given year on components of the funds expenses, affiliation to plan service providers, past return, risk and other plan level characteristics. Independent variables not described in Table 1 are *affiliated* which takes the value of 1 when the fund is affiliated to plan's largest asset management service provider, fund vol which is the fund's total annualized volatility, and other index is a binary variable equal to one if the plan has another index fund of the same style. Standard errors are in brackets and are clustered at the fund level. Significance: ***99%, **95%, *90%.

	(1)	(2)	(3)	(4)
Expense Ratio $(t-1)$	-0.197**	-0.287**	-0.184*	-0.271**
	[0.096]	[0.131]	[0.094]	[0.128]
Afiiliated	0.010**	0.010***	0.017**	0.017**
	[0.004]	[0.004]	[0.008]	[0.008]
Expense Ratio $(t-1)$			-5.534	-5.625
$\times$ Afiiliated			[3.606]	[3.479]
Gross α_f $(t-1)$	-0.027	-0.022	-0.025	-0.019
	[0.028]	[0.023]	[0.027]	[0.022]
Index Return $(t-1)$	-0.008	-0.014*	-0.008	-0.013
	[0.006]	[0.008]	[0.006]	[0.008]
Fund vol $(t-1)$	0.018	0.009	0.018	0.006
	[0.019]	[0.036]	[0.019]	[0.036]
Log Fund TNA $(t-1)$	0.001***	0.002***	0.001***	0.001***
	[0.000]	[0.001]	[0.000]	[0.001]
Log Fund age $(t-1)$	0.002*	0.001*	0.002**	0.001*
	[0.001]	[0.001]	[0.001]	[0.001]
Turnover Ratio $(t-1)$	0.001**	0.002***	0.001**	0.002***
	[0.000]	[0.001]	[0.000]	[0.001]
Other Index $(t-1)$	-0.002**	-0.004***	-0.002**	-0.004***
	[0.001]	[0.002]	[0.001]	[0.002]
Fixed effects:				
Plan $\times$ Year	Yes	Yes	Yes	Yes
Index Style $\times$ Year	No	Yes	No	Yes
Observations	331617	316839	331617	316839
R^2	0.019	0.021	0.020	0.022

Internet Appendix

IA1 Additional dataset details

IA1.1 Merging CRSP and Morningstar

Here I provide a short guide on merging the CRSP and Morningstar databases, but more extensive descriptions can be found in Berk and van Binsbergen (2015) and Pástor et al. (2015), which I follow. I start by retrieving all US Domestic Equity funds from each database. From both databases I download monthly returns, total net assets, tickers and cusip (security identifier) numbers from January 2000 until June 2023. These are the main matching variables and from both datasets, I exclude funds with missing monthly total net assets. Note that at this stage, all observations are at the month-fund share class level, which will be aggregated to the fund level after the merge is complete. Ultimately, the goal of this merge is to match the *crsp_fundo* and *crsp_portno* unique share class and fund identifiers to the Morningstar unique share class and fund identifiers, *Sec ID* and *Fund ID* respectively.

In CRSP I have to do a few changes to some of these matching variables to be able to match observations to Morningstar. Morningstar only has the latest ticker and cusip number for each share class while in CRSP these may be time-varying. Additionally, in CRSP these security identifiers are missing early in the sample or in a fund's history but are populated later on. To deal with these two issues, I use the the most recent ticker listed in CRSP for each *crsp_fundo* and backfill it to the beginning of its history. There are cases where the same ticker is asigned to more than one fund in a given month, in those cases I follow Pástor et al. (2015) and set those values to missing. This way, each ticker-month combination uniquely identifies a share class-month observation. I then apply these same steps for fund cusips, another security identifier that is common in both databases, resulting in the cusip-month combination also uniquely identifying a share class-month observation.

For Morningstar, prior to merging, only a few steps are necessary. Unlike CRSP, Morningstar only has static values of cusip numbers and tickers for each share class, and since these can get reused, occasionally there are tickers or cusip numbers that are assigned to more than one *Sec ID* each month. In those cases I set those values to missing.

Before merging there is one final step done on both CRSP and Morningstar datasets: I assign fake cusip numbers and tickers by creating a string of that shareclass original ticker or cusip with the *crsp_fundo* or Morningstar *Sec ID*. This way we won't match firms with missing a cusip or ticker when matching on these variables but may be able to match using monthly TNA values or returns.

The final setp is to then merge the two datasets to map *crsp_fundo-month* to *Sec ID-month*. I use the four variables previously listed and do a sequential merge on these variables:

1. Match both prepared CRSP and Morningstar databases on exact ticker-month combinations.
2. Then match the remaining unmatched samples on exact cusip-month combinations.
3. Then match the remaining unmatched samples on an exact match based on month, total net assets and a difference in month returns that is at most two basis points.
4. Then match the remaining unmatched samples on an exact match based on month, monthly returns and a difference in total net assets that is at most 20,000 USD.

Once these steps are concluded, I perform a few additional checks to ensure good match quality. Following Berk and van Binsbergen (2015), I only keep funds where the *crsp_fundo* matches the Morningstar *Sec ID* for at least 60% of their history. In the cases where this is not 100%, I only keep the observations where the match is correct. Finally, I only keep observations where I'm able to match every shareclass of the fund.

Once this mapping is completed, this allows me to match fund information from CRSP which is the basis for most of this paper with some crucial variables from Morningstar: Benchmark Index, index fund indicator and indicators for share classes destined to defined contribution retirement plans.

IA1.2 Index and index fund selection

Here I describe the index and index fund selection for the US domestic fund universe. The starting point is to identify which funds from the merged CRSP-Morningstar database are index funds. To identify index funds I use the index fund flag in CRSP and define index funds under the *Pure index fund* category. However, the CRSP indicator occasionally has missing values early in the sample even though later on that fund is coded as an index fund. In those cases I backfill the CRSP index fund flag. Furthermore, some shareclasses have missing values for the index fund flag even though other share classes of the same fund have one. In such cases, I attribute the index fund flag of the largest share class of the fund. Finally in cases different classes from the same fund have flags that are inconsistent (for example if one share class is an active fund and the other an index fund), I classify these as undetermined.

From the funds identified as index funds, I then retrieve a list of all the primary benchmark indices listed in Morninsgstar. While there are hundreds of different indices I focus on the subset of indices that are: 1) popular and 2) can be fully understood as passive investments. As a result I select indices from providers that have at least 2 funds tracking one of their indices, there are: Dow Jones S&P (and Wilshire), Russel, Nasdaq, CRSP and MSCI. There are some indices with a few funds tracking them that are excluded as they either have ESG objectives or have some kind of underlying active strategy (commonly called smart beta products).

Finally I make two important corrections. First I estimate Equation 2 every month, using a rolling yearly window of daily fund returns. From there I retrieve the β_I and estimate $AbsBeta$ from Equation 3. This tells me the exposure of the index fund to the index. I exclude funds that have a value greater than 0.2 as these either likely incorrectly assigned to the index listed in Morninstar or are funds that provide negative or leveraged exposure to the target index which is not in the scope of this paper.

The other correction is that Morningstar only lists the most recent primary benchmark index. I assume the fund has always tracked that index, but some funds are older than the indices they track. This is the case for most funds tracking the CRSP indices. For those I assume that prior to the launch of the index, the fund tracks the index with the highest correlation from my chosen list. For example, funds tracking the CRSP US Mega Cap index, prior to its launch in 2012, they're assumed to track the S&P 500.

The resulting index selection and the number of funds from the sample that track those indices is described in Table IA1. One small observation that suprised me when assembling this dataset is the low number of funds tracking MSCI indices. It turns out that MSCI indices are popular for index funds tracking global equity or equity from non-US regions which is not within the scope of this paper.

Table IA1: **Benchmark Indices:** This table lists the indices tracked by the index funds analyzed in this paper. The second column describes the type of index which is tells us what type of stocks composed the index. The final column lists the total number of funds tracking each index in the sample.

Index Name	Index Style	Number of Funds
Russell 3000	All Stocks	10
Wilshire 5000		6
CRSP US Total Market		2
S&P 1500		2
S&P US TMI		2
S&P 500	Large Cap	95
Russell 1000		7
DJ Industrial Average		3
S&P 100		3
CRSP US Large Cap		1
CRSP US Mega Cap		1
Russell 1000 Growth	Large Growth	10
NASDAQ 100		7
S&P 500 Growth		4
S&P 500 Pure Growth		2
NASDAQ Composite		2
NASDAQ-100 Equal Weighted		2
CRSP US Large Cap Growth		1
CRSP US Mega Cap Growth		1
Russell 1000 Value	Large Value	12
S&P 500 Value		4
S&P 500 Pure Value		2
CRSP US Large Cap Value		1
CRSP US Mega Cap Value		1
S&P MidCap 400	Mid Cap	22
Russell Mid Cap		7
CRSP US Mid Cap		1
Russell Mid Cap Growth	Mid Growth	5
S&P MidCap 400 Growth		4
S&P MidCap 400 Pure Growth		2
CRSP US Mid Cap Growth		1
Russell Mid Cap Value	Mid Value	4
S&P MidCap 400 Value		3
S&P MidCap 400 Pure Value		2
CRSP US Mid Cap Value		1
Russell 2000	Small Cap	22
S&P SmallCap 600		15
Russell Small Cap Complete		5
Russell 2500		3
Wilshire 4500 Completion		3
S&P Completion		2
CRSP US Small Cap		1
Russell 2000 Growth	Small Growth	5
S&P SmallCap 600 Growth		3
S&P SmallCap 600 Pure Growth		2
CRSP US Small Cap Growth		1
Russell 2000 Value	Small Value	5
S&P SmallCap 600 Value		3
MSCI US Small Cap Value		2
S&P SmallCap 600 Pure Value		2
CRSP US Small Cap Value		1

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS

2401 LAURA HOSPIDO, NAGORE IRIBERRI and MARGARITA MACHELETT: Gender gaps in financial literacy: a multi-arm RCT to break the response bias in surveys.

2402 RUBÉN DOMÍNGUEZ-DÍAZ, SAMUEL HURTADO and CAROLINA MENÉNDEZ: The medium-term effects of investment stimulus.

2403 CLODOMIRO FERREIRA, JOSÉ MIGUEL LEIVA, GALO NUÑO, ÁLVARO ORTIZ, TOMASA RODRIGO and SIRENIA VAZQUEZ: The heterogeneous impact of inflation on households' balance sheets.

2404 JORGE ABAD, GALO NUÑO and CARLOS THOMAS: CBDC and the operational framework of monetary policy.

2405 STÉPHANE BONHOMME and ANGELA DENIS: Estimating individual responses when tomorrow matters.

2406 LAURA ÁLVAREZ-ROMÁN, SERGIO MAYORDOMO, CARLES VERGARA-ALERT and XAVIER VIVES: Climate risk, soft information and credit supply.

2407 JESÚS FERNÁNDEZ-VILLAVERDE, JOËL MARBET, GALO NUÑO and OMAR RACHEDI: Inequality and the zero lower bound.

2408 PABLO BURRIEL, MAR DELGADO-TÉLLEZ, CAMILA FIGUEROA, IVÁN KATARYNIUK and JAVIER J. PÉREZ: Estimating the contribution of macroeconomic factors to sovereign bond spreads in the euro area.

2409 LUIS E. ROJAS and DOMINIK THALER: The bright side of the doom loop: banks' sovereign exposure and default incentives.

2410 SALOMÓN GARCÍA-VILLEGAS and ENRIC MARTORELL: Climate transition risk and the role of bank capital requirements.

2411 MIKEL BEDAYO and JORGE E. GALÁN: The impact of the Countercyclical Capital Buffer on credit: Evidence from its accumulation and release before and during COVID-19.

2412 EFFROSYNI ADAMOPOULOU, LUIS DÍEZ-CATALÁN and ERNESTO VILLANUEVA: Staggered contracts and unemployment during recessions.

2413 LUIS FÉRNANDEZ LAFUERZA and JORGE E. GALÁN: Should macroprudential policy target corporate lending? Evidence from credit standards and defaults.

2414 STÉPHANE BONHOMME and ANGELA DENIS: Estimating heterogeneous effects: applications to labor economics.

2415 LUIS GUIROLA, LAURA HOSPIDO and ANDREA WEBER: Family and career: An analysis across Europe and North America.

2416 GERALD P. DWYER, BILJANA GILEVSKA, MARÍA J. NIETO and MARGARITA SAMARTÍN: The effects of the ECB's unconventional monetary policies from 2011 to 2018 on banking assets.

2417 NICOLÁS FORTEZA, ELVIRA PRADES and MARC ROCA: Analysing the VAT cut pass-through in Spain using web-scraped supermarket data and machine learning.

2418 JOSÉ-ELÍAS GALLEGOS: HANK beyond FIRE: Amplification, forward guidance, and belief shocks.

2419 DANIEL ALONSO: Stabilisation properties of a SURE-like European unemployment insurance.

2420 FRANCISCO GONZÁLEZ, JOSÉ E. GUTIÉRREZ and JOSÉ MARÍA SERENA: Shadow seniority? Lending relationships and borrowers' selective default.

2421 ROBERTO BLANCO, MIGUEL GARCÍA-POSADA, SERGIO MAYORDOMO and MARÍA RODRÍGUEZ-MORENO: Access to credit and firm survival during a crisis: the case of zero-bank-debt firms.

2422 FERNANDO CEREZO, PABLO GIRÓN, MARÍA T. GONZÁLEZ-PÉREZ and ROBERTO PASCUAL: The impact of sovereign debt purchase programmes. A case study: the Spanish-to-Portuguese bond yield spread.

2423 EDGAR SILGADO-GÓMEZ: Sovereign uncertainty.

2424 CLODOMIRO FERREIRA, JULIO GÁLVEZ and MYROSLAV PIDKUYKO: Housing tenure, consumption and household debt: life-cycle dynamics during a housing bust in Spain.

2425 RUBÉN DOMÍNGUEZ-DÍAZ and SAMUEL HURTADO: Green energy transition and vulnerability to external shocks.

2426 JOSEP GISBERT and JOSÉ E. GUTIÉRREZ: Bridging the gap? Fintech and financial inclusion.

2427 RODOLFO G. CAMPOS, MARIO LARCH, JACOPO TIMINI, ELENA VIDAL and YOTO V. YOTOV: Does the WTO Promote Trade? A Meta-analysis.

2428 SONER BASKAYA, JOSÉ E. GUTIÉRREZ, JOSÉ MARÍA SERENA and SERAFEIM TSOUKAS: Bank supervision and non-performing loan cleansing.

2429 TODD E. CLARK, GERGELY GANICS, and ELMAR MERTENS: Constructing fan charts from the ragged edge of SPF forecasts.

2430 MIGUEL GARCÍA-POSADA and PETER PAZ: The transmission of monetary policy to credit supply in the euro area.

2431 KLODIANA ISTREFI, FLORENS ODENDAHL and GIULIA SESTIERI: ECB communication and its impact on financial markets.

2432 FRUCTUOSO BORRALLO, LUCÍA CUADRO-SÁEZ, CORINNA GHIRELLI and JAVIER J. PÉREZ: "El Niño" and "La Niña": Revisiting the impact on food commodity prices and euro area consumer prices.

2433 VÍCTOR CABALLERO, CORINNA GHIRELLI, ÁNGEL LUIS GÓMEZ and JAVIER J. PÉREZ: The public-private wage GAP in the euro area a decade after the sovereign debt crisis.

2434 LIDIA CRUCES, ISABEL MICÓ-MILLÁN and SUSANA PÁRRAGA: Female financial portfolio choices and marital property regimes.

2435 RODOLFO G. CAMPOS, ANA-SIMONA MANU, LUIS MOLINA and MARTA SUÁREZ-VARELA: China's financial spillovers to emerging markets.

2436 LUDOVIC PANON, LAURA LEBASTARD, MICHELE MANCINI, ALESSANDRO BORIN, PEONARE CAKA, GIANMARCO CARIOLA, DENNIS ESSERS, ELENA GENTILI, ANDREA LINARELLO, TULLIA PADELLINI, FRANCISCO REQUENA and JACOPO TIMINI: Inputs in Distress: Geoeconomic Fragmentation and Firms' Sourcing.

2437 DANIEL DEJUAN-BITRIA, WAYNE R. LANDSMAN, SERGIO MAYORDOMO and IRENE ROIBÁS: How do changes in financial reporting standards affect relationship lending?

2438 ALICIA AGUILAR and RICARDO GIMENO: Discrete Probability Forecasts: What to expect when you are expecting a monetary policy decision.

2439 RODOLFO G. CAMPOS, JESÚS FERNÁNDEZ-VILLAVERDE, GALO NUÑO and PETER PAZ: Navigating by Falling Stars: Monetary Policy with Fiscally Driven Natural Rates.

2440 ALEJANDRO CASADO and DAVID MARTÍNEZ-MIERA: Local lending specialization and monetary policy.

2441 JORGE ABAD, DAVID MARTÍNEZ-MIERA and JAVIER SUÁREZ: A macroeconomic model of banks' systemic risk taking.

2442 JOSEP PIJOAN-MAS and PAU ROLDAN-BLANCO: Dual labor markets and the equilibrium distribution of firms.

2443 OLYMPIA BOVER, LAURA HOSPIDO and ANA LAMO: Gender and Career Progression: Evidence from the Banco de España.

2444 JESÚS FERNÁNDEZ-VILLAVERDE, GALO NUÑO and JESSE PERLA: Taming the curse of dimensionality: quantitative economics with deep learning.

2445 CLODOMIRO FERREIRA and STEFANO PICA: Households' subjective expectations: disagreement, common drivers and reaction to monetary policy.

2446 ISABEL MICÓ-MILLÁN: Inheritance Tax Avoidance Through the Family Firm.

2447 MIKEL BEDAYO, EVA VALDEOLIVAS and CARLOS PÉREZ: The stabilizing role of local claims in local currency on the variation of foreign claims.

2501 HENRIQUE S. BASSO, MYROSLAV PIDKUYKO and OMAR RACHEDi: Opening the black box: aggregate implications of public investment heterogeneity.

2502 MARCO BARDOSCIA, ADRIAN CARRO, MARC HINTERSCHWEIGER, MAURO NAPOLETANO, LILIT POPOYAN, ANDREA ROVENTINI and ARZU ULUC: The impact of prudential regulations on the UK housing market and economy: insights from an agent-based model.

2503 IRINA BALTEANU, KATJA SCHMIDT and FRANCESCA VIANI: Sourcing all the eggs from one basket: trade dependencies and import prices.

2504 RUBÉN VEIGA DUARTE, SAMUEL HURTADO, PABLO A. AGUILAR GARCÍA, JAVIER QUINTANA GONZÁLEZ and CAROLINA MENÉNDEZ ÁLVAREZ: CATALIST: A new, bigger, better model for evaluating climate change transition risks at Banco de España.

2505 PILAR GARCÍA and DIEGO TORRES: Perceiving central bank communications through press coverage.

2506 MAR DELGADO-TÉLLEZ, JAVIER QUINTANA and DANIEL SANTABÁRBARA: Carbon pricing, border adjustment and renewable energy investment: a network approach.

2507 MARTA GARCÍA RODRÍGUEZ: The role of wage expectations in the labor market.

2508 REBECA ANGUREN, GABRIEL JIMÉNEZ and JOSÉ-LUIS PEYDRÓ: Bank capital requirements and risk-taking: evidence from Basel III.

2509 JORGE E. GALÁN: Macroprudential policy and the tail risk of credit growth.

2510 PETER KARADI, ANTON NAKOV, GALO NUÑO, ERNESTO PASTÉN and DOMINIK THALER: Strike while the Iron is Hot: Optimal Monetary Policy with a Nonlinear Phillips Curve.

2511 MATTEO MOGLIANI and FLORENS ODENDAHL: Density forecast transformations.

2512 LUCÍA LÓPEZ, FLORENS ODENDAHL, SUSANA PÁRRAGA and EDGAR SILGADO-GÓMEZ: The pass-through to inflation of gas price shocks.

2513 CARMEN BROTO and OLIVIER HUBERT: Desertification in Spain: Is there any impact on credit to firms?

2514 ANDRÉS ALONSO-ROBISCO, JOSÉ MANUEL CARBÓ, PEDRO JESÚS CUADROS-SOLAS and JARA QUINTANERO: The effects of open banking on fintech providers: evidence using microdata from Spain.

2515 RODOLFO G. CAMPOS and JACOPO TIMINI: Trade bloc enlargement when many countries join at once.

2516 CORINNA GHIRELLI, JAVIER J. PÉREZ and DANIEL SANTABÁRBARA: Inflation and growth forecast errors and the sacrifice ratio of monetary policy in the euro area.

2517 KOSUKE AOKI, ENRIC MARTORELL and KALIN NIKOLOV: Monetary policy, bank leverage and systemic risk-taking.

2518 RICARDO BARAHONA: Index fund flows and fund distribution channels.