Plan Design, Retirement Portfolios, and Investors' Welfare

Marco Loseto*

September 2023

Please click here for the latest draft

Abstract

Employer-sponsored retirement plans are a crucial component of the US savings system. However, despite their prevalence, plans differ widely in the quality and costs of the investment options available, with many plans including funds that are substantially more expensive than comparable alternatives available in the market. To uncover the factors contributing to the design of low-quality and high-cost plans, this paper develops an equilibrium framework of plan design and fee competition between fund providers. The model features a two-layer demand system where plan sponsors design their retirement plan menu and plan investors form their portfolios from the available options. The observed variation in plan inclusion probabilities and portfolio allocations identifies sponsors' and investors' preferences, respectively. Plan sponsors are less responsive to funds' fees than plan investors and are particularly so if compared to investors who actively form their retirement portfolio. Counterfactual exercises suggest that policies mandating the inclusion of low-cost default options and limiting the inclusion of expensive funds can generate significant welfare gains for plan investors.

KEYWORDS: 401(k), retirement plan design, portfolio choice, funds fees

JEL Classification: G11, G23, G28, G51, J32, L13, L51

*Kenneth C. Griffin Department of Economics, University of Chicago, mloseto@uchicago.edu. I am deeply grateful to Lars Hansen, Ali Hortaçsu, and Scott Nelson for their continuous guidance and support. I am also grateful to John Heaton, Adam Jørring, Ralph Koijen and Federico Mainardi for their detailed and thoughtful feedback. For their helpful comments I thank seminar participants at the Yiran Fan memorial conference, UChicago Economic Dynamics & Financial Markets working group and the UChicago IO lunch. Financial support from the Bradley fellowship, Stevanovich fellowship, Neubauer PhD fellowship and the Bank of Italy's Bonaldo Stringher scholarship is gratefully acknowledged. This research was also funded in part by the John and Serena Liew Fellowship Fund at the Fama-Miller Center for Research in Finance, University of Chicago Booth School of Business. All errors are my own.

1 Introduction

Employer-sponsored defined contribution (DC) retirement plans are a crucial component of the US savings system, holding nearly $11 trillion in assets as of 2021.[1] These plans allow employees to allocate a portion of their pre-tax income towards retirement savings through a range of investment options, typically mutual funds, to build long-term wealth. For many workers, the assets held in DC plans are among the most important components of their balance sheets and are a significant determinant of their future retirement security.[2]

However, despite their importance, many plans do not provide their investors with high-quality or cost-efficient investment options. For instance, in 2019, nearly 16% of sponsors (i.e., employers) did not provide a Target-Date-Fund (TDF), which automatically adjusts asset allocations over time and offers a set-it-and-forget-it solution for long-term retirement savings. Over half of the plans also failed to offer low-cost S&P 500 Index Funds or ETFs such as the Vanguard 500 Index Fund or the Vanguard S&P 500 ETF (Figure 1). Even more strikingly, one out of every five sponsors did not offer an equity fund with an expense ratio below 10 basis points.[3]

A closer look at plan expenses reveals substantial dispersion across sponsors, with the difference in the average expense between plans at the 75th and 25th percentiles of about 40 basis points (Figure 2). To put this in perspective: assuming an annual return of 6%, if an employee with an annual income of $70,000 contributes 10% to their 401(k) and shifts from a plan at the 75th percentile to one at the 25th percentile, they could save approximately $95,000 in investment fees.[4]

Beyond dispersion, plan expenses are also surprisingly high, with the asset-weighted average expense ratio for the median plan in the 2019 cross-section close to 40 basis points.[5] For context, in that same year, had a retail investor constructed a portfolio of Vanguard index funds to obtain exposure to all asset classes available in a typical retirement plan, the expense ratio would have been more than four times lower.[6]

[1]Among those 401(k) are the largest totalling $7.7 trillions. As a share of the US retirement market assets, employer-sponsored DC plans account for 30%. If one includes individual retirement accounts (IRA), DC plans account for 63% of the US retirement assets. https://www.icifactbook.org/.

[2]According to the 2019 Survey of Consumer Finances (SCF), for a working age household, the average account balance in a DC plan (including IRAs) was nearly $270,000. More generally, retirement accounts are the second-most commonly held type of financial asset after transaction accounts (https://www.federalreserve.gov/publications/files/scf20.pdf).

[3]These figures are even worse for years before 2019 and for retirement plans of smaller size.

[4]The calculation assumes a working period of 40 years.

[5]These patterns are not limited to the 2019 cross-section. Plan expenses were even higher before 2019 (Figure 6). At the same time, the dispersion in plan expenses has been roughly stable over time (Figures 3, 4) and remains even when comparing plans of similar size (Figure 5).

[6]In 2019, the expense ratio for a Vanguard equally-weighted portfolio of retail index funds, including its International equity index funds (VEMAX, VEUSX, VPADX), US Equity funds (VGSLX, VFIAX, VIMAX, VSMAX) and Bond Fund (VBTLX) is below 10 basis points. By retail, I mean that the minimum investment required is none or limited. Figure 6 compares the expense for the median plan against this portfolio of Vanguard index funds over time.

Additionally, more expensive plans do not seem to produce better investment performance for their investors (Table 1).[7] All things considered, it is unsurprising that employees have increasingly sought to hold plan sponsors accountable for violating their fiduciary duties, with high investment costs emerging as the common theme in many recent lawsuits.[8]

Why would plan sponsors include investment options that are not as cost-efficient as comparable alternatives in the marketplace? On the one hand, sponsors may not be sensitive to funds' fees and may value attributes other than fees when designing their plan menu. For example, agency frictions may push sponsors to favour the inclusion of funds affiliated with their plan recordkeeper (Pool, Sialm and Stefanescu (2016)) or to incorporate costlier options to reduce direct fees paid to the recordkeeper (Bhattacharya and Illannes (2022)).[9] On the other hand, considering the large number of investment options available, it is likely that for many sponsors, it is not feasible to consider all existing options and find the cheapest ones. In practice, sponsors hire plan providers (a.k.a recordkeepers) to assist in designing their 401(k) menus, with different recordkeepers often having pre-existing relationships with specific mutual funds providers.[10] Consequently, sponsors served by different recordkeepers might encounter distinct sets of investment options to choose from, leading to differences in the resulting menus and expenses.[11]

From a supply-side perspective, several factors might deter funds from lowering their fees. For example, if plan sponsors are not sensitive to expenses or do not consider all available options, funds may have limited incentives to compete for being included in plan menus. Furthermore, conditional on plan inclusion, funds might lack the incentives to lower their fees to attract plan assets if plan investors do not respond to fees or there is no close substitute fund available in the menu.

In retirement investing, a non-negligible portion of investors is unresponsive to fees, often because they do not make active investment decisions. Instead, they allow their contributions to be automatically directed into their plan default option, typically a TDF. Moreover, while most sponsors aim to provide investors with a broad range of

[7]Table 1 and Figure 7 show that the plan-level (gross of fees) performance tends to be lower for more expensive plans. Similar patterns have been found in the context of active investing (Gil-Bazo and Ruiz-Verdu (2009)) contrasting what frictionless models with rational investors predict (Berk and Green (2004)).

[8]I provide more details of some recent lawsuits in Appendix D.

[9]Another factor that could lower sponsors' sensitivity to fees is liability risk. For example, sponsors of smaller plans may experience lower liability risk because employees may not consider pursuing legal action worthwhile. This diminished liability risk can disincentivize sponsors from investing resources in finding cheaper options.

[10]Plan sponsors typically outsource administrative tasks such as maintaining employees' account balances to a plan provider, also called recordkeeper, which in most cases is an investment provider. For example, Fidelity and Vanguard are among the largest recordkeepers.

[11]In fact, plan-level data indicates that recordkeepers' networks of funds are far from perfectly overlapping (Figures 8, 9, 10), implying that a considerable portion of the options accessible to one plan sponsor are likely unavailable to another sponsor who has a different recordkeeper.

investment categories,[12] they typically offer only one fund per category (Figure 11). Although this plan structure helps manage risk, it could weaken competition between funds, as price competition is more intense when products are more alike.[13]

To shed light on the factors contributing to the design of high-cost plans and to quantify the effects of plan design policies on investors' welfare, this paper develops and estimates an equilibrium model of retirement plan design, portfolio choice and fee competition between investment providers. The model features a two-layer demand system where, in the first layer, sponsors design their retirement plan and, in the second layer, plan investors form their retirement portfolio from the options available in their menu.

Plan sponsors compose their menu by selecting investment options from the pool available in their recordkeeper network. They evaluate investment funds with a linear random utility that depends on funds' characteristics, including fees, past (gross of fees) returns and whether funds are affiliated with the recordkeeper. The latter accommodates for the presence of agency frictions whereby sponsors favour recordkeepers' funds. Inclusion decisions are made independently for each investment category. First, sponsors decide whether to include or not a particular investment category. Once this decision is made, they proceed to choose the options within that category that offer the highest utilities. I assume that the number of options included is drawn randomly with decaying probability to capture the possibility that sponsors incur some costs when adding more than one option per category.[14]

In the second layer, plan investors form their optimal portfolio by allocating their contributions among the options available in their menu. To capture inertia, I allow for the possibility that some investors do not make an active investment decision. Instead, they default their asset allocation into a TDF or, if unavailable, a Balanced fund.

I estimate sponsors' and investors' preferences using comprehensive plan-level data. This information is collected from 401(k) plan menus and asset allocations as reported

[12] Securities and mutual funds are typically grouped into investment categories such as Large Cap, Small Cap, etc. Funds that belong to the same investment category tend to offer similar risk-return profiles to investors and, thus, should be more substitutable to each other. Figure 12 shows that sponsors tend to include around 14/15 different investment categories regardless of their size.

[13] e.g., we would expect fee competition to be more intense if two funds belong to the same investment category. Consistent with this Bertrand-type of intuition, plan-level data suggests that fees are inversely correlated with the number of options offered within a specific category (Table 2).

[14] Adding options may be costly for several reasons. First, sponsors fiduciary duty is not limited to the design of the retirement menu but also requires them to monitor the included options and provide plan investors with up-to-date information about their performance. Second, sponsors' asset base (i.e., the total contributions) is limited, and with too many options, it could be challenging to meet the minimum investment requirements required by investment funds. In estimation, I calibrate this probability to match the observed distribution of the number of funds included within each category (Figure 11). In Appendix C, I offer a simple microfoundation for the optimal choice of the number of options to be included within each category building on the Stigler (1961) simultaneous search model.

by plan sponsors on form F5500 to the Department of Labor (DOL).[15] Specifically, I compute the probability of a given investment fund being included in a plan from the observed plan menus. This is done by determining the fraction of plans including that particular fund. The observed variation in these inclusion probabilities enables me to identify and estimate plan sponsors' preferences. Similarly, the observed variation in plan-level asset allocations allows me to identify plan investors' preferences. To account for the possible endogeneity of investment fees, I exploit the granularity of the data to control for unobserved demand shocks along several dimensions, including investment category fixed effects, sponsors fixed effects, funds' brand fixed effects and a fixed effect for passive funds. I then instrument funds' fees with funds' turnover ratios, which measure trading-related transaction costs, typically passed on to investors through higher fees. The identifying assumption is that funds' turnover ratios affect sponsors' and investors' demand only through fees after controlling for those dimensions of unobserved heterogeneity.

Model estimates indicate that plan sponsors are less sensitive to fees than plan investors. This is particularly evident when comparing them to investors actively forming their retirement portfolios, who are nearly three times more elastic to fees than sponsors. This misalignment in sponsors' and investors' elasticity to fees suggests that sponsors may not adequately internalize their employees' preferences when designing retirement plan menus. The estimates also indicate that sponsors have a strong preference for including funds affiliated with their recordkeeper, consistent with recent empirical evidence underscoring the importance of agency frictions in this market (Pool, Sialm and Stefanescu (2016)). Quantitatively, being affiliated with the plan recordkeeper increases the inclusion probability by 0.47 percentage points, a magnitude four times higher than what a 10 basis points reduction in funds' fees would lead to.

After estimating sponsors' and investors' demands, I use the Nash-Bertrand equilibrium conditions to recover funds' marginal costs and price-cost margins. The median fund, spanning all fund types, charges a margin of 14 basis points. Given the median expense ratio, this translates to a markup of approximately 19%. Looking at passive funds and TDFs specifically, their estimated marginal costs indicate that these funds are more efficient than others. However, they don't fully transfer these efficiency gains to investors. This is especially evident for TDFs, where the estimated median markup stands at about 44%—over double the median markup observed across all funds. This finding is consistent with TDFs deriving their pricing power from targeting mostly inactive investors who are likely not responsive to fees. As previously mentioned, in retirement investing, the fraction of inactive investors is non-negligible. Model estimates corroborate this, indicating that, over my sample period, one out of four plan investors did not make an active investment decision. The estimated share of inactive

[15]See Section 3 for more details on the data.

investors has been increasing over time from around 15% in 2010 to more than 40% as of 2019.

In the last part of the paper, with the estimated demand and supply parameters, I explore a series of counterfactuals to assess the effects of plan design policies on plan investors' welfare. In the first counterfactual, I eliminate agency frictions whereby plan sponsors favour the inclusion of funds affiliated with their recordkeeper. This policy is ineffective because assuming sponsors do not value funds' affiliation does not prevent them from including expensive funds that are not affiliated with their recordkeeper. In other words, eliminating sponsors' preference for affiliated funds does not make them more responsive to fees.

In the second set of counterfactuals, I consider policies that mandate the inclusion of a low-cost equity index fund tracking the S&P 500 and the inclusion of a low-cost TDF. Mandating the inclusion of low-cost equity S&P 500 funds leads to an increase in investors' surplus of about 2%. The increase is modest because investors value lower fees, but they also want to diversify across all available assets, dampening the incentive to substitute toward the low-cost index fund. Moreover, this policy only benefits active investors, leaving inactive investors' surplus unchanged. Mandating the inclusion of low-cost TDFs increases investors' surplus by about 11%, a magnitude more than five times larger than the previous policy. This policy is more effective because it also benefits inactive investors who will reallocate their entire contribution to the low-cost TDF.

Lastly, I consider a policy that caps funds' expense ratios at 50 basis points. Under this policy, sponsors can only offer funds with expenses below this cap. This policy further improves investors' outcomes, leading to an increase in their surplus of about 14%. This policy is effective because it affects the entire menu of options by limiting the inclusion of the most inefficient funds. It benefits inactive investors because it replaces the most inefficient TDFs, and it benefits active investors by reducing the costs of all options available in the plan menu.

The rest of the paper proceeds as follows. Section 2 describes how this paper fits the literature. Section 3 describes the data. Section 4 sets up the demand side of the model. Section 5 focuses on the supply side and characterizes the equilibrium fees. Section 6 estimates the demand side of model. Section 7 turns to the supply side and recovers funds' price-cost margins. Section 8 presents the results of policy counterfactuals and Section 9 concludes.

2 Contributions to the Literature

This paper contributes to the literature that studies retirement investing and the design of retirement plans. A large part of this literature has focused on the demand side by studying 401(k) enrollment decisions with a particular focus on the role of automatic enrollment into default options (Madrian and Shea (2001), Beshears, Choi, Laibson and Madrian (2009), Carroll, Choi, Laibson, Madrian and Metrick (2009), Choi (2015)), behavioural biases in retirement investing (Benartzi and Thaler (2001), Huberman and Jiang (2006), Benartzi and Thaler (2007), Tang, Mitchell, Mottola and Utkus (2010)) and, more recently, the implications of automatic enrollment on saving behaviour over the life-cycle (Choukhmane (2021), Duarte, Fonseca, Goodman and Parker (2022)).

Another part of the literature has looked at the supply side and has mainly focused on empirically examining the role of agency frictions between plan providers (i.e., recordkeepers) and plan sponsors (i.e., employers). Pool, Sialm and Stefanescu (2016) show that plan providers vertically integrated into fund provision tend to favour affiliated funds, Badoer, Costello and James (2020) provide evidence of how plan providers trade-off direct fees from the sponsor with indirect fees paid by funds via revenue-sharing agreements and Pool, Sialm and Stefanescu (2022) show that funds paying revenue-sharing fees are more likely to be included in a plan and less likely to be deleted. Building on this evidence, Bhattacharya and Illannes (2022) take a more structural perspective and develop a model where revenue-sharing fees are the outcome of a bargaining game between sponsors and recordkeepers.

This paper lies in between these two strands of work. Its primary contribution is to develop and estimate an empirical model of plan design, retirement portfolio choice and fee competition between differentiated fund providers. Differently from Bhattacharya and Illannes (2022), my modelling of sponsors' plan design problem does not formalize agency frictions as a bargaining game between sponsors and recordkeepers. Instead, I take a simpler approach and model sponsors' plan design problem as a (multiple) discrete choice problem that builds on the workhorse discrete choice frameworks developed in Berry (1994) and Berry, Levinsohn and Pakes (1995). Nonetheless, I accommodate for agency frictions by allowing sponsors' preferences to depend on funds' characteristics that relate to the identity of the plan recordkeeper. For example, I allow sponsors to have a taste for whether or not a fund is affiliated with its recordkeeper. Moreover, by modelling sponsors' and investors' preferences separately, I can allow their sensitivity to expenses to be misaligned, consistent with agency frictions whereby sponsors favour expensive funds to reduce the direct fees paid to the recordkeeper.

Keeping the plan design problem simple enables me to model plan investors' portfolio decisions. After sponsors' menus have been designed, I assume plan investors with quadratic preferences (Markowitz (1952)) form their retirement portfolio from the options available in their menu. In a recent contribution, Egan, MacKay and Yang (2023) also model retirement portfolio choice as a mean-variance problem and, importantly, show how variation in expense ratios can identify investors' beliefs and risk aversion separately. I follow their methodology except for allowing investors to be inactive with some probability, in which case, investors' contributions are defaulted into some of the available TDFs. Variation in TDFs portfolio shares across plans allows me to identify the fraction of inactive investors together with the other investors' preference parameters.

I close the model by assuming that funds compete in a differentiated oligopoly by setting fees simultaneously a la Nash-Bertrand. I use the implied first-order conditions to recover funds' marginal costs and price cost margins. This links my paper to the empirical industrial organization literature on imperfect competition in the mutual fund industry. Most of this literature models investment decisions on the demand side as a standard discrete choice problem. This is done either because the focus is on competition between financially homogeneous products such as S&P index funds (Hortaçsu and Syverson (2004)), ESG and non-ESG funds that track the same underlying index (Baker, Egan and Sarkar (2022)) and variable annuities (Koijen and Yogo (2022)), or because investors are assumed to be risk neutral (Massa (2003), Roussanov, Ruan and Wei (2021)).

This paper contributes to this strand of literature by characterizing the Nash-equilibrium fees and margins when investors have a mean-variance demand rather than a discrete choice demand. The linear structure of investor demand allows me to go beyond the standard characterization of Nash-Bertrand price cost margins and to decompose equilibrium fees into three components. One of these components, which I refer to as 'Hotelling markdown' because it captures how much a monopolist should give up when it faces competitors that are closer in the characteristic space, is equivalent to the vector of Bonacich centralities of a network in which funds are the network nodes and funds' substitution patterns are (inversely) proportional to the network edges. A more central fund competes with similar products and must charge lower fees (e.g., has a higher Hotelling markdown).

This characterization of the equilibrium fees arises because the Bertrand game played by funds falls into a broader class of network games with quadratic payoff first studied in Ballester, Calvó-Armenagol and Zenou (2006). This literature on network games was born to study games with peer effects (Ballester, Calvó-Armenagol and Zenou (2006)) and has been recently applied to study oligopolistic competition (Ushchev and Zenou (2018), Pellegrino (2023), Loseto (2023)).[16] The key insight from this literature is that Nash equilibrium actions will generally depend on a player's network centrality. This insight carries over to the Bertrand network game I am considering. Still, my characterization is more general because the network is ex-ante unknown to players (e.g., funds do not know which plan will include them), and players' actions influence the resulting network structure (e.g., when setting fees, funds

[16] In Appendix E I describe and simulate a simple differentiated Bertrand-Network game.

influence their plan inclusion probability).[17]

3 Institutional Setting and Data Sources

In this section I describe what is a employer-sponsored retirement plan and overview its administrative structure. After that I describe the data sources and provide some summary statistics about my sample of retirement plan menus and the investment options available therein.

3.1 What is an employer-sponsored retirement plan?

Employer-sponsored retirement plans are savings vehicles designed to assist employees in accumulating wealth for their retirement years. Although there are various types of employer-sponsored retirement plans, the most common is the defined contribution (DC) plan. At its core, a defined contribution plan is a retirement plan in which an employee makes regular contributions. The final amount available upon retirement is not pre-determined but instead depends on the contributions made and on the returns obtained from the investment options available in the plan.

Under a DC plan, employees contributions represent a percentage of their salary which is subtracted from their paycheck before taxes, thereby reducing their current taxable income. The gains in the retirement account grow tax-deferred, implying taxes are not due until funds are withdrawn during retirement years. Withdrawals from a DC plan before a certain age (typically between 59 and 60) can result in penalties. After reaching retirement age, participants might withdraw distributions as lump sums, period payments or annuities. Importantly, all withdrawals at this point are subject to standard income tax.

Employers play a crucial role in the provision and design of these type of plans. First, many employers offer to match a portion of the employee's contributions. For example, an employer might contribute 50 cents for every dollar the employee contributes, up to a certain percentage of the employee's salary (typically around 6%). More importantly, employers are in charge of selecting and monitoring which investment options, typically mutual funds, are to be included in the plan. Once the employee decides what percentage of their income to contribute, they typically have the autonomy to allocate their contributions, together with the part matched by their

[17] Grice and Guecioueur (2023) also study fee competition between investment providers from a network perspective. Building on Grice (2023), they propose a model of competition between fund families where the competitive network is micro-founded from investors' imperfect consideration. My model instead features competition at the fund level and belongs to the class of network games studied in Loseto (2023) where the underlying network is determined by products' characteristics, thereby allowing for arbitrary substitution patterns between products. Moreover, in the model I am considering, the network structure is unknown to players because investment providers set fees before plan sponsors design their retirement menu.

employer, across the options available within their plan. In many cases, to help less financially savvy employees or those who do not make an active investment choice, plan sponsors include options, such as TDFs, known as Qualified Default Investment Alternative (QDIA) to be used as default option when an employee contributes to the plan without specifying how their contribution should be invested.

Figure 18 summarizes graphically the structure of a DC plan. Plan sponsors typically hire recordkeepers to assist in the design of their retirement menu and to perform administrative tasks such as maintaining account balances. Most recordkeepers, like for example Fidelity, are also providers of investment funds and it is not uncommon for retirement plans to include funds from the recordkeeper product line. Overall, a retirement plan consists in a set of assets, spanning a broad range of investment categories, selected by sponsors and recordkeepers. Under the Employee Retirement Income Security Act (ERISA), plan sponsors are fiduciaries to their employees and are subject to litigation risk if their retirement plan is not designed in the employees' best interest. The inclusion of high-cost investment options or the lack of low-cost options are among the main triggers of many recent lawsuits, as I describe in Appendix D.

3.2 Data

The primary data source for this study is collected from Form 5500. This form is annually filed by employers with the Department of Labor (DOL) in adherence to the Employee Retirement Income Security Act (ERISA) regulations. Within this form, Schedule H provides detailed information about retirement plan menus. Specifically, it contains data regarding the investment options offered within an employer's retirement plan and the plan-level (end-of-year) dollar allocation to each of these options.

Although Form 5500 filings are available for download from the DOL website, information about plan menus comes in a non-standardized format, stored in pdf images, that would need to be digitized manually. I acquired the digitized version of these filings for the years 2010 to 2019 from BrightScope Beacon who collects and digitizes these data directly from the publicly available DOL filings. Overall, the data covers more than 90% of total plan assets in each year from 2010 to 2019, auditing an average of 60,000 plans per-year.

I complement this data with additional information from the DOL Form 5500, including the number of plan participants and the identity of each plan recordkeeper. Also, I obtain data on funds' historical expense ratios from CRSP and merge those in the main dataset using funds' tickers. BrightScope also provides data on funds' expense ratios, but the historical expenses are only available starting from 2016. Before 2016, BrightScope reports the most recent funds' expense ratio which I replace with the one obtained from CRSP.

Table 3 provides some summary statistics for the plans in my sample. To limit the impact of outliers, I focus on plans whose number of participants is between 100 and 5000, representing roughly 95% of the whole sample. The average plan has close to 30 millions dollars in assets and around 475 participants. Both measures of plan size are right-skewed due to the presence of large plans, with the median plan having assets ranging around 10 millions and 250 plan participants.

Looking at the characteristics of the investment menu, the average plan offers 25 investment options across 15 different investment categories. For the average plan, one out of four options is affiliated with the plan recordkeeper. Moreover, most of the plans offer at least a TDF fund and a passive fund. Turning to plan expenses, the average expense ratio charged by the average plan is of about 63 basis points. This is more than 10 basis point higher than its asset-weighted counterpart, suggesting that plan investors tilt their allocations toward cheaper funds.

Table 4 reports some summary statistics for the funds offered in the sample of plan menus I observe. Excluding cash accounts and common stocks, there are 5600 distinct funds for which at least a ticker identifier is available.[18] In my sample of plans, the average fund manages 191 million dollars of retirement assets and has an average portfolio share of about 3%. The distribution of assets is right-skewed, with the median fund having 7 millions in assets. Interestingly, the share of plan assets allocated into a given fund varies significantly across plans, with the average standard deviation of the portfolio share for the average fund being as large as 3%.

The data also suggests that sponsors review their menu of investment options often. In Table 4, the penultimate row reports the average fraction of years a fund is included within a plan menu. On average, a fund is offered in only half of the years that I observe the plan menu, implying that sponsors regularly modify their menu offerings. The same does not appear to be true when looking at plans' recordkeepers, with more than 75% of plans having the same recordkeeper over the years (Table 3).

4 Demand

This section describes the two-layer demand side of the model. I start from the first layer where I describe sponsors' preferences and the plan design problem they face. After that, I derive funds' inclusion probabilities implied by the plan design demand model. Next, I turn to the second layer of demand where I first describe investors preferences and then derive both individual and aggregate demand systems.

4.1 Sponsors' plan design problem

Throughout the paper I will index plan sponsors by p and mutual funds by j. All vectors are in bold. The goal of sponsor p is to choose a set of mutual funds to include

[18] The same fund may have multiple tickers, one for each different share-class.

into its retirement plan. Typically sponsors hire recordkeepers to help in designing and managing their plan menu and, empirical evidence suggests that the set of funds sponsors consider to beging with is strongly influenced by the recordkeeper identity. From a modelling point of view, I do not model recordkeepers as separate agents but I allow for heterogeneity in sponsors' consideration sets. Formally, I assume that sponsor p considers with positive probability a subset $N_p \subset N$ of all mutual funds available. Empirically, I assume that fund j belongs to N_p if I observe at least a plan that includes j and shares the same recordkeeper as p. In other words, sponsors with the same recordkeeper have the same consideration set.

Sponsor p indirect utility from including fund j is given by

$$u_{jp} = V_{jp}(\boldsymbol{\theta}_p) + \varepsilon_{jp} \tag{1}$$

with

$$V_{jp}(\boldsymbol{\theta}_p) = \boldsymbol{w}'_{jp}\boldsymbol{\theta}_p + \zeta_j \tag{2}$$

where $\boldsymbol{w}_j$ is the vector of fund j's observed characteristics including its expense ratio, past returns and an indicator for whether j is a fund owned by p's recordkeeper; ζ_j captures characteristics, possibly correlated with fees, unobserved to the econometrician and, ε_{jp} is a random preference shock distributed as T1EV.

I assume that funds are classified into investment categories indexed by $g \in \{1, ..., G\}$ and model the menu design process as a two-stage decision problem. In the first stage, sponsors choose which category to offer in their plan and this decision is made independently across categories.[19] For each selected investment category, in the second stage, sponsors evaluate the options available making within category comparisons and selecting the options providing the highest indirect utility.

To complete the decision problem, I need to specify how the number of options to be included within each selected category is chosen. I do not attempt to model this choice as the outcome of a rational decision problem and instead assume that this number is drawn randomly from a geometric distribution with parameter q. This implies that the probability of n options being included in a given category is given by:[20]

$$q(n) \equiv q(1-q)^{n-1} \quad \text{for } n = 1, 2, ... \tag{3}$$

This modelling assumption is guided by the empirical observation that sponsors tend

[19] On average sponsors, regardless of their size, tend to include around 14/15 investment categories (Figure 12) out of a 25 categories in total.

[20] As I explain in Appendix C, where I offer a simple microfoundation for the optimal choice of the number of options included within an investment category, incorporating such decision in the full model would considerably complicate its estimation.

not to include more than one option per investment category. Figure 11 plots the empirical distribution of the number of options offered within investment category and shows how in 70% of instances sponsors only include one option per category. The probability then decays as the number of options included increases, consistent with the presence of some cost that sponsors incur when adding more than one option. In estimation, I calibrate q to match the observed empirical distribution allowing for heterogeneity at the recordkeeper-category level.

4.2 Funds' inclusion probabilities

In this section I derive funds' plan inclusion probabilities implied by sponsors' plan design problem described in the previous section. To this end, I will analyze sponsors' decision problem backward.

Consider sponsor p who has chosen to offer category g and needs to select n investment funds within g. At this stage, p ranks all the options according to (1) and selects the n options $\{j_1, ..., j_n\}$ such that

$$u_{j_1} > u_{j_2} > ... > u_{j_n}. \tag{4}$$

Fund j will be included in p's plan if and only if u_j is ranked among first nth highest utilities. Letting j_z be the option with the zth highest utility, the probability that j is included in plan p is given by

$$\phi_{jp}^{1:n} \equiv \sum_{z=1}^{n} \phi_{jp}^{z}$$

where ϕ_{jp}^{z} is the probability that j is ranked in the zth position i.e.,

$$\phi_{jp}^{z} \equiv \Pr\{j = j_z\}.$$

Under the assumption that preference random shocks are distributed as T1EV, an analytical expression for each ϕ_{jp}^{z} can be derived. For $z = 1$, ϕ_{jp}^{1} corresponds to the standard logit choice probability

$$\phi_{jp}^{1} = \frac{\exp(V_j(\boldsymbol{\theta}_p))}{\sum_{k \in g} \exp(V_k(\boldsymbol{\theta}_p))}. \tag{5}$$

For $z = 2$, ϕ_{jp}^{2} is the probability that j provides the 2nd highest utility which equals the sum of probabilities of all utility rankings where u_j is the 2nd largest utility. In a world in which there are only 4 options, say $\{j, k, l, m\}$,

$$\begin{aligned}\phi^2_{jp} = &\Pr\{u_{kp} > u_{jp} > u_{lp} > u_{mp}\} + \Pr\{u_{kp} > u_{jp} > u_{mp} > u_{lp}\}\\ &+ \Pr\{u_{lp} > u_{jp} > u_{kp} > u_{mp}\} + \Pr\{u_{lp} > u_{jp} > u_{mp} > u_{kp}\}\\ &+ \Pr\{u_{mp} > u_{jp} > u_{kp} > u_{lp}\} + \Pr\{u_{mp} > u_{jp} > u_{lp} > u_{kp}\}.\end{aligned}$$

In Appendix C, I show that the above six terms expression can be reduced to the following expression with only three addends

$$\begin{aligned}\phi^2_{jp} = &\frac{\exp(V_{kp})}{\exp(V_{kp}) + \exp(V_{jp}) + \exp(V_{lp}) + \exp(V_{mp})} \cdot \frac{\exp(V_{jp})}{\exp(V_{jp}) + \exp(V_{lp}) + \exp(V_{mp})}\\ &+ \frac{\exp(V_{lp})}{\exp(V_{kp}) + \exp(V_{jp}) + \exp(V_{lp}) + \exp(V_{mp})} \cdot \frac{\exp(V_{jp})}{\exp(V_{jp}) + \exp(V_{kp}) + \exp(V_{mp})}\\ &+ \frac{\exp(V_{mp})}{\exp(V_{kp}) + \exp(V_{jp}) + \exp(V_{lp}) + \exp(V_{mp})} \cdot \frac{\exp(V_{jp})}{\exp(V_{jp}) + \exp(V_{kp}) + \exp(V_{lp})}.\end{aligned}$$

This result is a consequence of the well-known independence of irrelevant alternatives (IIA) property of the logit model. IIA implies that how the 3rd and 4th options are ranked does not matter in determining the probability that j is ranked 2nd.

The above immediately generalizes to a case with an arbitrary number of options

$$\phi^2_{jp} = \sum_{j_1 \neq j} \frac{\exp(V_{j_1 p})}{\sum_k \exp(V_{kp})} \frac{\exp(V_{jp})}{\sum_{k \neq j_1} \exp(V_{kp})}.$$

and to the case in which we want to compute the probability of j being ranked in an arbitrary position zth

$$\phi^z_{jp} = \sum_{(j_1,\dots,j_{z-1}) \in g/\{j\}} \prod_{z'=1}^{z-1} \frac{\exp(V_{j_{z'}p})}{\sum_{z''=z'}^{N_{gp}} \exp(V_{j_{z''}p})} \cdot \frac{\exp(V_{jp})}{\sum_{z'''=z}^{N_{gp}} \exp(V_{z'''p})} \tag{6}$$

where $N_{gp} \subset N_p$ is the number of options in category g.

Moving backward in sponsor p's design problem, the probability that n options are chosen from investment category g is given by $q(1-q)^{n-1}$ so that, conditional on g being offered, the probability of j being included in p's plan is just

$$\sum_{n=1}^{\infty} q(1-q)^{n-1} \phi^{1:n}_j.$$

The choice of whether or not to include category g is assumed to depend on sponsors' expected utility from the highest ranked option, which under our T1EV assumption, is given by

$$\mathbb{E}[u_{j_1 p}] = \log\left(\sum_{k \in g} \exp(V_{kp}(\boldsymbol{\theta}_p))\right).$$

The probability that p decides to offer investment category g as part of its retirement plan equals

$$\lambda_{gp} = \frac{\exp(\mathbb{E}[u_{j_1 p}])}{1 + \exp(\mathbb{E}[u_{j_1 p}])}$$

which can be interpreted as the probabilistic outcome of a binary choice logit problem.

Combining all pieces together, it can be shown that the unconditional probability of fund j being included in sponsor p plan can be written as

$$\phi_{jp}(\boldsymbol{\theta}_p) = \lambda_{gp}(\boldsymbol{\theta}_p) \cdot \sum_{n=1}^{\infty} (1-q)^{n-1} \phi_{jp}^{n}(\boldsymbol{\theta}_p) \tag{7}$$

where I make explicit the dependence on the vector of preference parameters $\boldsymbol{\theta}_p$, $(1-q)^{n-1}$ is the probability that p includes a number of options greater or equal than n and ϕ_j^n is the probability that j is ranked in the nth position.[21]

Equation (7) extends the expression of the logit choice probabilities to the case in which decision makers can select more than a single option and where the number of option chosen is determined by the parameter q. Expression (7) collapses to the standard discrete choice logit probability if we assume sponsors can only include one fund within each investment category. This corresponds to setting $q = 1$, which implies that

$$\phi_{jp}(\boldsymbol{\theta}_p) = \lambda_{gp}(\boldsymbol{\theta}_p) \cdot \phi_{jp}^{1}(\boldsymbol{\theta}_p) = \frac{\exp(V_{jp}(\boldsymbol{\theta}_p))}{1 + \sum_{k \in g} \exp(V_{kp}(\boldsymbol{\theta}_p))}.$$

In this more general context, the decision of not including investment category g plays the role of the outside option in standard discrete choice models, with mean indirect utility normalized to 0.

Overall, equation (7) represents sponsor p individual demand for investment funds belonging to investment category g. Assuming sponsors preferences $\boldsymbol{\theta}_p$ follow some distribution F_θ, we can derive fund j's aggregate demand as

$$\phi_j = \int \phi_{jp}(\boldsymbol{\theta}_p) dF_\theta(\boldsymbol{\theta}_p). \tag{8}$$

The data counterpart to (8) corresponds to the share of plans that include fund j as part of their retirement menu. In Section 6 I estimate the distribution of sponsors preference parameters by matching (8) to these observed inclusion probabilities.

[21] I provide details for the derivation in Appendix C

4.3 Investors' retirement portfolio problem

Consider investor i who allocates its dollar contribution A across the investment funds available in plan p, indexed by $j \in \{1, ..., J_p\}$, and a cash account $j = 0$. In practice, not all plan investors make an active investment decision and many of them are automatically defaulted into one of the options available which often corresponds to a TDF fund or a balanced fund. I denote this default option $j = d$ and assume that with probability δ investor i does not make an active investment decision. In this case i's contribution will be allocated entirely to fund d. Conversely, with probability $(1-\delta)$ investor i makes an active investment decision and allocates A across all options available including d.

Conditional on making an active investment decision, investor i forms its retirement portfolio by choosing the vector of portfolio weights $\boldsymbol{a} \equiv (a_1, ..., a_{J_p})$ to maximize the following linear-quadratic problem:

$$\begin{aligned} \max_{\boldsymbol{a}} \quad & a_0 + \boldsymbol{a}'(X_{(1)}\boldsymbol{\beta} + \boldsymbol{\xi} - \boldsymbol{f}) - \frac{\gamma}{2}\boldsymbol{a}'\left(I + X_{(2)}X'_{(2)}\right)\boldsymbol{a} \\ \text{s.t. } & a_0 + \boldsymbol{a}'\mathbf{1} = 1 \end{aligned} \tag{9}$$

where a_0 is the fraction of wealth A allocated to the cash account, $X_{(1)}$ is a $J_p \times K_y$ matrix of funds characteristics entering the linear component of (9), $X_{(2)}$ is $J_p \times K_x$ a matrix of funds characteristics entering the quadratic component of (9), $\boldsymbol{f}$ is the vector of funds fees, β and γ are investors preference parameters and $\boldsymbol{\xi}$ is a vector of characteristics or demand shocks that are unobserved to the econometrician.

Problem (9) captures the idea that investors value assets along two margins, a 'perfect substitute' margin that pushes investors to just consume the asset with the highest linear component and a 'variety' margin that pushes investors to consume more than just one asset. In financial terms, problem (9) can be interpreted as a mean-variance portfolio problem where investors prefer assets with higher subjective expected returns, which depend on asset characteristics and investors preferences $(X_{(1)}\boldsymbol{\beta} + \boldsymbol{\xi})$, but at the same time dislike risk and would like to diversify across assets. How much they dislike risk depends on asset characteristics $X_{(2)}$ and the preference parameter γ. In standard mean-variance portfolio theory $X_{(2)}$ includes asset characteristics capturing measures of volatility such as the exposures to some aggregate risk factors.

4.4 Plan asset demand system

Under the previous assumptions, investor i's asset demand system is given by

$$\boldsymbol{a}_i(\boldsymbol{f}) = \begin{cases} \boldsymbol{e}_d & \text{if } i \text{ defaults} \\ \frac{1}{\gamma}(I + X_{(2)}X'_{(2)})^{-1}(X_{(1)}\boldsymbol{\beta} + \boldsymbol{\xi} - \boldsymbol{f}) & \text{o.w} \end{cases} \tag{10}$$

where $\boldsymbol{e}_d$ is a unit vector that takes value of 1 in its d element corresponding to the default option.

In the data I do not observe asset allocations at the individual level. Therefore, I need to aggregate individual investors demands to obtain the plan level demand system. Letting $\boldsymbol{s}_p$ be the J_p vector of allocations for plan p, A_p plan p total wealth and defining $\boldsymbol{\eta} \equiv (\boldsymbol{\beta}, \gamma, \delta)$ as the vector of demand parameters, we can sum demands across all investors in plan p to obtain:

$$\boldsymbol{s}_p(\boldsymbol{f};\boldsymbol{\eta}) \equiv \sum_{i \in I_p} \frac{A}{A_p} \boldsymbol{a}_i(\boldsymbol{f}) = \delta \boldsymbol{e}_d + \frac{1-\delta}{\gamma}(I + X_{(2)}X'_{(2)})^{-1}(X_{(1)}\boldsymbol{\beta} + \boldsymbol{\xi} - \boldsymbol{f}). \tag{11}$$

Empirically, I can use variation in the observed plan level allocations to estimate the demand system in (11). To this end, it is useful to multiply both sides of (11) by $(I + X_{(2)}X'_{(2)})$ to obtain a demand system where only own fees and own demand shock enter each equation :

$$\tilde{s}_p(\boldsymbol{f};\boldsymbol{\eta}) = \delta \tilde{\boldsymbol{e}}_d + X_{(1)}\tilde{\boldsymbol{\beta}} - \tilde{\gamma}\boldsymbol{f} + \tilde{\boldsymbol{\xi}}. \tag{12}$$

where $\tilde{\boldsymbol{s}}_p \equiv (I + X_{(2)}X'_{(2)})\boldsymbol{s}_p$, $\tilde{\boldsymbol{e}}_d \equiv (I + X_{(2)}X'_{(2)})\boldsymbol{e}_p$, $\tilde{\boldsymbol{\beta}} \equiv \boldsymbol{\beta}(1-\delta)/\gamma$ and $\tilde{\boldsymbol{\xi}} \equiv \boldsymbol{\xi}(1-\delta)/\gamma$. Because X is observed, I can estimate (12) via linear methods.

Before turning to the supply side, a couple of remarks are in order. First, so far I have assumed that individual investors' preferences are homogeneous. In this way, the parameters of the plan-level demand system are the same as the ones for the individual demand system. From an empirical perspective, the linear structure of the demand system prevents me from learning about unobserved heterogeneity in preference parameters.[22] Nevertheless, in Appendix C, I show how to interpret the parameters in (11) as weighted averages of the heterogeneous individual parameters. Second, I am assuming that the default option is the same for each individual. This is done only for expositional purposes and in the estimation and in Appendix C I allow for the presence of multiple default funds.

Lastly, in setting the supply side profit maximization problem, I will be working with a slightly rearranged version of equation (11). Specifically, I will rewrite the latter as

$$\boldsymbol{s}_p(\boldsymbol{f};\boldsymbol{\eta}) = \delta \boldsymbol{e}_d + \frac{(1-\delta)}{\gamma}(I - \mathcal{K})(\boldsymbol{\mu}_p - \boldsymbol{f}) \tag{13}$$

[22] One way to introduce heterogeneity would be interacting funds' characteristics, such as fees, with observable plans' characteristics. Egan, MacKay and Yang (2023) take this approach to uncover heterogeneity in investors' risk aversion.

where $\boldsymbol{\mu} \equiv X_{(1)}\boldsymbol{\beta} + \boldsymbol{\xi}$ and the matrix $\mathcal{K}$ is defined as

$$\mathcal{K} \equiv X_{(2)}(I + X'_{(2)}X_{(2)})^{-1}X'_{(2)}.$$

Rewriting aggregate asset demand as in (13) helps in visualizing own and cross substitution patterns across different assets. In particular, the fee elasticity between asset j and asset l is proportional to

$$\frac{\partial s_j}{\partial f_l} \propto \begin{cases} -(1-\kappa_{jj}) = -\left(1 - \boldsymbol{x}'_{(2)j}(I + X_{(2)}X'_{(2)})^{-1}\boldsymbol{x}_{(2)j}\right) & \text{if } j = l \\ \kappa_{jk} = \boldsymbol{x}'_{(2)j}(I + X_{(2)}X'_{(2)})^{-1}\boldsymbol{x}_{(2)k} & \text{if } j \neq l \end{cases} \tag{14}$$

which is always between $(-1, 1)$ if $j \neq k$ and between $(-1, 0)$ if $j = l$. It can be interpreted as a measure of how close/correlated asset j and l are in terms of their vectors of characteristics $\boldsymbol{x}_j$ and $\boldsymbol{x}_l$ respectively.[23]

5 Supply

I model supply as a differentiated Bertrand oligopoly where investment advisors set fees simultaneously before sponsors make their plan design decisions and plan investors form their portfolios. I assume that the same fund charges the same fee across different plans. In practice, the same fund can price discriminate by offering different share classes yet, in the context I am considering, 401(k) sponsors are almost always treated as institutional investors and almost all the observed variation in fees is across funds and not across share classes within the same fund. Moreover, many investment providers offer a specific share class for the retirement plan market.

Let P be the number of plan sponsors, A_p the dollar wealth of plan p and $\mathcal{S}_{jp} \subseteq 2^{N_p}$ be the set of all possible menus where p includes fund j. Fund j chooses its fee f_j to maximize the following expected dollar profit

$$\max_{f_j} \quad P \cdot (f_j - c_j) \cdot \int_p A_p \left(\sum_{S \in \mathcal{S}_{jp}} \phi_p(S, \boldsymbol{f}; \boldsymbol{\theta}_p) s_{jp}(\boldsymbol{f}; \boldsymbol{\eta}_p | S) \right) dF(A_p, \boldsymbol{\theta}_p, \boldsymbol{\eta}_p) \tag{15}$$

where $\phi_p(S, \boldsymbol{f}; \boldsymbol{\theta}_p)$ is the probability that sponsor p designs plan menu S and, conditional on menu S, $s_{jp}(\boldsymbol{f}; \boldsymbol{\eta}_p | S)$ is fund j's portfolio weight within plan p

$$s_{jp}(\boldsymbol{f}; \boldsymbol{\eta}_p | S) = \begin{cases} \frac{1-\delta_p}{\gamma_p}\left((1-\kappa^S_{jj})(\mu_{jp} - f_j) - \sum_{l \neq j, l \in S} \kappa^S_{jl}(\mu_{lp} - f_l)\right) & \text{if } j \neq d \\ \delta_p + \frac{1-\delta_p}{\gamma_p}\left((1-\kappa^S_{dd})(\mu_{jp} - f_j) - \sum_{l \neq j, l \in S} \kappa^S_{dl}(\mu_{lp} - f_l)\right) & \text{if } j = d \end{cases}$$

where I made explicit the dependence of the elements of $\mathcal{K}$ on the realized menu S and allow investor parameters to depend on p.

[23] I provide more details in Appendix C

Problem (15) is particularly complex to solve because it requires investment providers to internalize how a marginal increase in fees affects (i) investors portfolio allocation s_j conditional on a given menu S, (ii) the probability that plan menu S is chosen and (iii) trade-off these changes across all possible plan menus S and plan sponsors p. From a computational perspective, given the large number of funds available in the market, the number of possible plan menus in $\mathcal{S}_{jp}$ would be too large to make the computation of (15) and its derivatives feasible.[24] To overcome these difficulties, I will simplify funds' pricing problem in a way that allows me make the problem computationally tractable while at the same time preserving the two dimensions along which funds' compete in the retirement market, namely, competition for plan inclusion and competition for plan asset allocations.

To simplify problem (15) I will assume that funds' only consider the effect of a marginal change in fees on their own inclusion probabilities and not on competitors plan inclusion probabilities. Equivalently, funds' do not take into account that changing fees influences the probability that a particular menu is chosen but only consider how it affects the overall likelihood of being included. With this assumption, fund j fee setting problem can be rewritten more compactly as

$$\max_{f_j} P \cdot (f_j - c_j) \cdot \int A_p \phi_{jp}(\boldsymbol{f};\boldsymbol{\theta}_p) s_{jp}(\boldsymbol{f};\boldsymbol{\eta}_p) dF(A_p, \boldsymbol{\theta}_p, \boldsymbol{\eta}_p) \tag{16}$$

where $\phi_{jp}(\boldsymbol{f};\theta_p)$ is the probability that p includes j as defined in (7) and $s_j(\boldsymbol{f};\eta_p)$ is the expected portfolio allocation of fund j within plan p and is given by

$$s_{jp}(\boldsymbol{f};\boldsymbol{\eta}_p) = \begin{cases} \frac{1-\delta_p}{\gamma_p}\Big((1-\bar{\kappa}^p_{jj})(\mu_{jp}-f_j) - \sum_{l\neq j}\bar{\kappa}^p_{jl}(\mu_{lp}-f_l)\Big) & \text{if } j \neq d \\ \delta_p + \frac{1-\delta_p}{\gamma_p}\Big((1-\bar{\kappa}^p_{dd})(\mu_{jp}-f_j) - \sum_{l\neq j}\bar{\kappa}^p_{dl}(\mu_{lp}-f_l)\Big) & \text{if } j = d \end{cases} \tag{17}$$

with,

$$\bar{\kappa}^p_{jl} \equiv \begin{cases} \sum_{S\in\mathcal{S}_{jp}\cap\mathcal{S}_{lp}} \frac{\phi_p(S,\boldsymbol{f};\theta_p)}{\phi_{jp}(\boldsymbol{f};\theta_p)}\kappa^S_{jl} = \phi_{lp}(\boldsymbol{f},\theta_p)\cdot\mathbb{E}[\kappa^S_{jl}|j,l\in S] & \text{if } j\neq l \\ \sum_{S\in\mathcal{S}_{jp}} \frac{\phi_p(S,\boldsymbol{f};\theta_p)}{\phi_{jp}(\boldsymbol{f};\theta_p)}\kappa^S_{jj} = \mathbb{E}[\kappa^S_{jj}|j\in S] & \text{if } j = l \end{cases} \tag{18}$$

Problem (16) can be obtained from (15) after dividing and multiplying the term in the round brackets by ϕ_{jp} and exploiting the linear structure of s_{jp} to rewrite the expectation over S more compactly. The resulting term collapses to $s_{jp}(\boldsymbol{f};\boldsymbol{\eta}_p)$ as

[24] Goeree (2008) faces a similar problem when estimating a discrete choice demand model with imperfect consideration. She gets around the computational burden by simulating consumers consideration sets. My case is more complex because investors' consideration sets, or equivalently plan menus, are the outcome of sponsors' plan design problem. Moreover, consideration probabilities in my case are not independent for products that belong to the same investment category. Lastly, in my case prices affect consideration probabilities which means that derivatives of consideration probabilities will enter firms' first order conditions. This will be true for both own consideration probabilities but also competitors consideration probabilities.

defined in (17), which represents the asset demand from plan p investors that fund j expects before sponsor p designs its plan menu. Asset characteristics affect this expected demand through the matrix $\bar{\mathcal{K}}^p$, whose (j, l) element, defined in (18), captures how much competitive pressure j expects from competitor l. The latter depends on how likely is fund l to be included in plan p and, conditional on that, on how close substitute asset j and asset l are.[25]

My restriction on funds' fee-setting behavior assumes that funds' do not internalize how fees affect the elements $\bar{\kappa}_{jl}$. A constructive way to impose this restriction could be assuming that funds' believe sponsors' will include at most one fund per investment category, thereby assigning positive probability only to plan menus containing funds from different categories. Formally, this would require funds' to have a biased belief $\hat{q} = 1$ about the parameter q governing the distribution of the number of funds included within each category. In practice, we know that most plans do not include more than one option per category (Figure 11), making this assumption perhaps not so unreasonable. Under this assumption, I show in Appendix C that $\bar{\kappa}_{jl}$ would not depend on f_j because plan inclusion decision are assumed to be independent across investment categories. Moreover, this assumption is one of the sufficient conditions that allows me to prove existence of a Bertrand-Nash equilibrium.[26]

5.1 Nash equilibrium fees

In this section I derive the Nash equilibrium fees implied by problem (16) assuming that funds' take $\bar{\mathcal{K}}$ as given. Denoting by $s_j(\boldsymbol{f})$ fund j' expected dollar asset allocation, the first order conditions with respect to f_j is given by the usual Bertrand pricing equation

$$s_j(\boldsymbol{f}) + (f_j - c_j) \cdot \frac{\partial s_j(\boldsymbol{f})}{\partial f_j} = 0. \tag{19}$$

The difference between the current setting and standard oligopolistic problems is that funds' are competing along two dimensions, namely, they compete for being included in a plan and, conditional on plan inclusion, they compete for plan investors' allocations. These two layers of competition are enclosed in $\partial s_j(\boldsymbol{f})/\partial f_j$ which is made

[25]This measure of substitutatibility is given by the second term in (18) $\mathbb{E}[\kappa^S_{jl}|j, l \in S; \boldsymbol{f}]$. This is an expectation because κ^S_{jl} depends on the whole menu S and not only on fund j and fund l characteristics. Formally, this can be seen from the definition of κ_{jl} in (14) where κ_{jl} depends on the characteristics of all competitors through the weighting matrix $(I + X'_{(2)}X_{(2)})^{-1}$.

[26]In Appendix C I show that when $\hat{q} = 1$, sponsors preferences are homogeneous and a particular dominance diagonal condition on the jacobian of the demand system is satisfied there exists a Bertrand-Nash equilibrium.

of the following two terms

$$\frac{\partial s_j(\boldsymbol{f})}{\partial f_j} = \underbrace{\int A_p \frac{\partial \phi_{jp}}{\partial f_j}(\boldsymbol{f};\boldsymbol{\theta}_p) s_{jp}(\boldsymbol{f};\boldsymbol{\eta}_p) dF(A_p,\boldsymbol{\theta}_p,\boldsymbol{\eta}_p)}_{\text{demand loss from marginal sponsor}}$$

$$\underbrace{-\int A_p(1-\delta_p)\gamma_p^{-1}\phi_{jp}(\boldsymbol{f};\boldsymbol{\theta}_p)(1-\bar{\kappa}^p_{jj}) dF(A_p,\boldsymbol{\theta}_p,\boldsymbol{\eta}_p)}_{\text{demand loss from marginal investor}} < 0. \tag{20}$$

Expression (20) captures the classic expected revenue loss from marginal consumers not willing to buy at higher price. In this context, the reduction in demand comes from two forces (i) the marginal sponsor not willing to include j in its plan and (ii) the marginal plan investors reducing its allocation to fund j. Equation (19) then tells us that, for given competitors fees, fund j will choose an f_j that equalizes the profit reduction from losing the marginal sponsors and investors to the profit gain from charging the inframarginal ones an higher fee.

The linear structure of investors asset demand allows me to go beyond this standard intuition and to offer a novel characterization of the Nash equilibrium fees that sheds light on the forces driving price competition in this market. To this end, I will define the following variables,

$$\bar{\phi}_j \equiv \int A_p(1-\delta_p)\gamma_p^{-1}\phi_{jp} dF_p; \qquad \tilde{\kappa}_{jl} \equiv \bar{\phi}_j^{-1}\int A_p(1-\delta_p)\gamma_p^{-1}\phi_{jp}\bar{\kappa}^p_{jl} dF_p$$

$$\bar{\mu}_j \equiv \bar{\phi}_j^{-1}\int A_p(1-\delta_p)\gamma_p^{-1}\phi_{jp}[I-\bar{\mathcal{K}}^p]_j' \boldsymbol{\mu}_p dF_p; \qquad \tilde{\boldsymbol{\mu}} \equiv (I-\tilde{\mathcal{K}})^{-1}\bar{\boldsymbol{\mu}}$$

$$\bar{\delta} \equiv \bar{\phi}_j^{-1}\int A_p\phi_{jp}\delta_p dF_p; \qquad \iota_j \equiv -\bar{\phi}_j^{-1}\int A_p\frac{\partial \phi_{jp}}{\partial f_j} s_{jp}(f_j-c_j) dF_p$$

where I suppressed all functions' arguments and $[I-\bar{\mathcal{K}}^p]_j$ denotes the jth row of the matrix $I-\bar{\mathcal{K}}^p$. Next, I rewrite (19) in vector form for all funds in terms of these variables to obtain

$$\bar{\delta}\boldsymbol{e}_d + (I-\tilde{\mathcal{K}})(\tilde{\boldsymbol{\mu}}-\boldsymbol{f}) - \boldsymbol{\iota} - (I-diag(\tilde{\mathcal{K}}))(\boldsymbol{f}-\boldsymbol{c}) = 0$$

This system of Bertrand first order conditions can be further rearranged as

$$\boldsymbol{f}^* = \underbrace{\frac{\tilde{\boldsymbol{\mu}}+\boldsymbol{c}}{2}}_{\text{monopolist fee}} - \underbrace{\boldsymbol{h}\left(\tilde{\mathcal{K}},\frac{\tilde{\boldsymbol{\mu}}-\boldsymbol{c}}{2}\right)}_{\text{Hotelling markdown}} - \underbrace{\left(I-\frac{diag(\tilde{\mathcal{K}})}{2}-\frac{\tilde{\mathcal{K}}}{2}\right)^{-1}\frac{\boldsymbol{\iota}}{2}}_{\text{plan inclusion markdown}} \tag{21}$$

where for simplicity I assumed that there is no default fund.[27]

Equation (21) decomposes the fees charged by funds in an interior Nash equilibrium into three terms. The first term represents the vector of fees funds would charge as

[27] All derivations are presented in Appendix C

monopolist.[28] From these fees there are two types of markdowns that need to be subtracted to account for the two dimension of competition driving pricing incentives in this market. The plan inclusion markdown in (21) captures the optimal reduction in fees required to increase the probability of being included in a plan. If funds knew that they will be included with certainty i.e., $\phi_{jp} \equiv 1$, the plan inclusion markdown would disappear because $\iota_j = 0$. The Hotelling markdown $\boldsymbol{h}(\cdot)$ instead captures the optimal reduction in fees required to compete against similar funds. The simple Hotelling (1929) model predicts that when two firms located on a line are closer to each other, they will charge lower margins. The same intuition carries over in this more general setting. Funds that are closer to their competitors have an higher $\boldsymbol{h}$ an must lower their fee.

The natural question at this point is, what does being closer to competitors mean in this context and how does that relate to $\boldsymbol{h}$? Loosely speaking a fund is closer to its competitors when its characteristics are less differentiated from competitors' characteristics. In practice, this measure of proximity/differentiation is embedded in the asset demand cross-substitution patterns through the matrix $\mathcal{K}$ defined in (14). Fund j and fund l are closer substitutes if their characteristics are closer as measured by κ_{jl}. The proximity of each fund to all other competitors is summarized by the vector $\boldsymbol{h}(\cdot)$ which is defined as

$$\boldsymbol{h}\left(\tilde{\mathcal{K}}, \frac{\tilde{\boldsymbol{\mu}} - \boldsymbol{c}}{2}\right) \equiv \left(I - \frac{(\tilde{\mathcal{K}} - \kappa_0 I)}{2(1-\kappa_0)}\right)^{-1} \frac{(\tilde{\mathcal{K}} - \kappa_0 I)}{2(1-\kappa_0)} \frac{\tilde{\boldsymbol{\mu}} - \tilde{\boldsymbol{c}}}{2} \tag{22}$$

where for expositional purposes I assumed that $\tilde{\kappa}_{jj} \equiv \kappa_0$. Expression (22) is a measure of proximity because it is equivalent to the Bonacich network centrality measure studied in (Bonacich (1987)).[29] This network centrality measure appears in (21) because the Bertrand fee-setting game belongs to a broader class of network games first studied in Ballester et al. (2006).[30] The main insight from this literature is that Nash equilibrium actions will generally depend on a player's network centrality. In this case, funds' equilibrium fees depend on how central a fund is in the competitive network. A more central fund faces more similar competitors and charges lower margins.

So far, I have assumed that a Nash equilibrium exists. Given the complexity of funds' pricing problem deriving general results about existence and uniqueness is not an easy task. Nonetheless, in Appendix C I provide sufficient conditions for existence and uniqueness of an interior Bertrand-Nash equilibrium for some particular cases. Specifically, I start by showing that when funds know with certainty which plan menus will include them (e.g., $\phi_{jp} = 1$ if p includes j) then the following dominance diagonal

[28] The price charged by a monopolist facing a linear demand $q(p) = a - p$ with marginal cost c, is $(a+c)/2$.

[29] For any zero-diagonal adjacency matrix A, positive scalar $\delta > 0$ and non-zero vector $\boldsymbol{u}$, the vector of Bonacich centralities is defined as $\boldsymbol{b}(A, \boldsymbol{u}) \equiv (I - \delta A)^{-1} \delta A \boldsymbol{u}$.

[30] More details are provided in Appendix E.

condition

$$(1-\tilde{\kappa}_{jj})(\tilde{\mu}_j - c_j) > \sum_{l\neq j} |\tilde{\kappa}_{jl}|(\tilde{\mu}_j - c_l) \quad \text{all } j,$$

ensures that the Bertrand-Nash equilibrium is interior, with $f_j^* \in (c_j, \tilde{\mu}_j)$ for all j, and unique. Moreover, I am also able to show that when sponsors' preference are homogeneous and funds believe that sponsors will include at most one fund per category (e.g., $\hat{q} = 1$) there exists a Nash-equilibrium even when funds' do not know with certainty which plan menus will include them.

6 Demand Identification and Estimation

In this section I describe how to identify and estimate the model. I start from estimating sponsors' preferences from variation in the observed plan inclusion probabilities and investors' preferences from variation in the observed plan-level portfolio allocations. Lastly, I turn to the supply side and recover funds' marginal costs and markups using the demand estimates together with the Nash-Bertrand equilibrium conditions.

6.1 Identification and estimation of sponsors' preferences

The plan design model developed in Section 4 provides us with an analytical expression for the expected probability that fund j is included in a retirement plan for a given distribution of sponsors' preference parameters $\boldsymbol{\theta}_p \sim F(\boldsymbol{\theta}_p; \bar{\boldsymbol{\theta}})$ which I assume to be parametrized by the vector $\bar{\boldsymbol{\theta}}$:

$$\phi_j(\bar{\boldsymbol{\theta}}) = \int \lambda_{gp}(\boldsymbol{\theta}_p) \cdot \sum_{n=1}^{\infty} (1-q)^{n-1} \phi_{jp}^n(\boldsymbol{\theta}_p) dF(\boldsymbol{\theta}_p; \bar{\boldsymbol{\theta}}), \tag{23}$$

where $\lambda_{gp}(\boldsymbol{\theta}_p)$ is the probability that plan p offers category g and $\phi_{jp}^n(\boldsymbol{\theta}_p)$ is the probability that j is the fund with the nth highest utility.

The data counterpart to equation (23) is the share of retirement plans that include fund j. The estimation strategy is then to find the vector of parameters $\bar{\boldsymbol{\theta}}$ that makes the model implied inclusion probabilities in (23) as close as possible to the observed ones. As $q \to 1$, expression (23) collapses to the standard random-coefficient logit formula for product market shares

$$\phi_j(\bar{\boldsymbol{\theta}}) = \int \frac{\exp(V_{jp}(\boldsymbol{\theta}_p))}{1 + \sum_{k \in g} \exp(V_{kp}(\boldsymbol{\theta}_p))} dF(\boldsymbol{\theta}_p; \bar{\boldsymbol{\theta}}),$$

considered in the workhorse demand models of Berry (1994) and Berry, Levinsohn and Pakes (1995).

In estimation, I compute the observed inclusion probability for each fund at the year-recordkeeper-category level, where I define an investment category as the interaction between the standard investment categories and an indicator for passive funds. In this way, an index fund and active fund from say Large-Cap-Growth would be classified into two different categories Large-Cap-Growth-Active and Large-Cap-Growth-Passive. I refer to a particular year-recordkeeper-category combination as a market, indexed by t, and denote the share of retirement plans in market t that offer fund j as $\hat{\phi}_{jt}$.

I allow for the possibility that funds' fees are correlated with sponsors' demand shocks ζ_{jt}. These shocks enter sponsors' mean utilities $V_{jpt}(\boldsymbol{\theta}_p) = \boldsymbol{w}'_{jpt}\boldsymbol{\theta}_p + \zeta_{jt}$ and are observed by market participants, including investment funds, but unobserved to the econometrician. If funds set fees after observing ζ_{jt} or have better information about these demand shocks, demand and supply simultaneity would bias preference parameters estimates. I account for this type of price endogeneity in two ways. First, I exploit the granularity of the data to absorb unobserved heterogeneity in demand along three dimensions: (i) time by including year fixed effects, (ii) product quality by including funds' brand fixed effects, and (iii) financial characteristics by including investment category fixed effects. Second, I instrument funds' fees with funds' turnover ratios. The latter capture trading-related costs that funds' incur when selling and buying securities which are pass on to investors through fees. The identifying assumption is that the variation in funds' turnover ratio not explained by time, brand, investment category and passive fixed effects enters sponsors' demand only through fees.

To estimate sponsors preferences I use a nested-fixed point algorithm similar to the one developed in Berry et al. (1995).[31] To start with, I assume that the distribution of sponsors preference parameters is normal with mean $\boldsymbol{\mu}_{\bar{\theta}}$ and variance $\Sigma_{\bar{\theta}}$ (e.g., $\bar{\boldsymbol{\theta}} = (\boldsymbol{\mu}_{\bar{\theta}}, \Sigma_{\bar{\theta}})$) and write sponsor p mean utility as

$$V_{jt}(\bar{\boldsymbol{\theta}}) = \bar{v}_{jt} + \boldsymbol{w}'_{jt}\Gamma_{\bar{\theta}}\boldsymbol{\nu}_p$$

where $\bar{v}_{jt} \equiv \boldsymbol{w}'_{jt}\boldsymbol{\mu}_{\bar{\theta}} + \zeta_{jt}$ is the homogeneous component of preferences, $\boldsymbol{\nu}_p \sim N(\mathbf{0}, I)$ are random tastes for funds' characteristics and $\Gamma_\theta\Gamma'_\theta = \Sigma_\theta$. The estimation algorithm starts with a guess of $\bar{\boldsymbol{\theta}}$ and then for each market t finds the vector $\bar{\boldsymbol{v}}_t(\bar{\boldsymbol{\theta}})$ that matches observed and model-implied inclusion probabilities:

$$\hat{\boldsymbol{\phi}}_t = \boldsymbol{\phi}_t(\bar{\boldsymbol{v}}_t(\bar{\boldsymbol{\theta}})).$$

and recover the demand residuals $\boldsymbol{\zeta}_t(\bar{\boldsymbol{\theta}})$. The last step exploits the orthogonality condition between ζ_{jt} and an appropriate vector of instruments $\boldsymbol{Z}_{jt}$, $\mathbb{E}[\zeta_{jt}|\boldsymbol{Z}_{jt}] = 0$ to

[31] See Appendix C for more details.

form the GMM norm

$$\boldsymbol{\zeta}(\bar{\boldsymbol{\theta}})'\boldsymbol{Z}\Omega(\bar{\boldsymbol{\theta}})\boldsymbol{Z}\boldsymbol{\zeta}(\bar{\boldsymbol{\theta}}). \tag{24}$$

The algorithm keeps searching over $\bar{\boldsymbol{\theta}}$ until (24) is minimized.

6.2 Estimates of sponsors' preferences

Table 5 presents the estimates of sponsors' preference parameters. The first column reports the estimates of the means of the preference distribution and the second column reports the corresponding standard deviations. These estimates minimize the GMM objective (24) following the estimation algorithm I discussed in the previous section.

In the estimation of sponsors' preferences I include four characteristics, two of them are continuous and two of them are binary. The two continuous characteristics are funds' expense ratios, measured in basis points (bp.), and funds' returns in the previous year gross of expenses, measured in percentage points (pp.). Because I absorb investment category fixed effects, returns are relative to the average return of funds within the same category.

The two binary characteristics are indicators for whether a fund is affiliated with the sponsor's recordkeeper and for whether a fund is a target date. Including an indicator for affiliated funds allows me to accomodate for the presence of agency frictions whereby sponsors favor the inclusion of funds belonging to their recordkeeper product line. The inclusion of an indicator for TDFs instead captures the possibility that sponsors have a preference for funds that rebalance plan investors allocation automatically as they age. These type of funds have been created specifically for retirement investing and, after the Pension Protection Act of 2006, qualify as default option for plan participants who do not make an active investment decision. Since then, TDFs' market share in the retirement market has been growing substantially and it is reasonable to think that sponsors may have a preference for such funds even if just to comply with current regulations and reduce liability risk.

The parameter estimates reported in the first column of Table 5 suggest that sponsors value more whether a fund is affiliated or a TDF rather than how cost-efficient such fund is or how it performed relative to its investment category. Preference coefficients for both affiliated and target indicators are large and significant. The preference coefficient on funds' expense ratios is negative and significant but its magnitude looks small if compared with the affiliated and target coefficients. On the other hand, plan sponsors do not seem to value funds' returns gross of fees, as the estimated coefficient is close to zero and non-significant. I also allow for heterogeneity around the mean of sponsors' sensitivity to fees and report the estimated standard deviation in the second column of Table 5. Although modest in magnitude, the estimated heterogeneity is significant at conventional significance levels.

To get a better understanding of the estimated magnitudes, I report the marginal effects of each characteristic on the inclusion probabilities in the third column of Table 5. The marginal effect is the unit change in the expected probability of being included in a plan for a unit increase in the corresponding characteristic. For instance, the first number in the third column tells us that, on average, a ten basis points increase in expenses reduces the plan inclusion probability by 0.12 percentage points. On the other hand, the marginal effect of being an affiliated fund or a TDF is almost four times larger and increases the inclusion probability by about 0.46 percentage points. These magnitudes are far from being negligible given that the median inclusion probability in the estimation sample is roughly 0.72%. The marginal effect of a one percent increase in funds' gross returns is instead substantially smaller which is not surprising given that its preference coefficient is non significant.

The bottom part of Table 5 presents some additional information including information about sponsors' elasticity to fees. With the model estimates, I compute the elasticity of inclusion probabilities to fees for each fund-market combination and report the median of this distribution in the bottom part of Table 5. The latter is around around 1.3, suggesting that sponsors' demand is not too elastic to funds' fees.

As mentioned before, I allow for the possibility that funds' fees are endogenous but treat other characteristics as predetermined and independent of demand shocks.[32] In estimation I instrument for fees using a third order polynomial of funds' turnover ratios. Overall I have six moments, three included characteristics and three instruments function of funds' turnover, to estimate five model parameters, four means and one standard deviation. The resulting GMM objective is of 6.69 and rejects the overidentifying restriction at the 1% level. This may be due to the fact that there does not seem to be too much heterogeneity in sponsors' preferences even though the model allows for it. The estimates for the homogeneous model are indeed similar (Table 11) and come with a GMM objective of 5.8 which, although not perfect, is not rejected at conventional significance levels.

The nature of the data allows me to assess the heterogeneity of the estimates more directly. I do so along two dimensions. First, I split the sample based on plan size as measured by the number of plan participants and find that small plans are less responsive to fees and tend to have a stronger preference for affiliated funds than large plans (Table 12). This is consistent with smaller sponsors having less bargaining power and being less willing to pay or search for cheaper investment options (Bhattacharya and Illannes (2022)). Second, I split the sample before and after 2014 and find that sponsors have become more elastic to fees over time (Table 13). This is consistent with sponsors, as well as plan investors, becoming more attentive to fees in response

[32]This assumption is often used in the empirical industrial organization literature where

to regulatory interventions mandating the disclosure of funds' fees and performance (Kronlund, Pool, Sialm and Stefanescu (2021)).

6.3 Identification and estimation of investors' preferences

The identification of investors' preferences follows closely the logic for the identification of sponsors' preferences. The main difference is that the identifying variation comes from variation in the observed portfolio allocations rather than variation in plan inclusion probabilities.

The estimation of plan investors' preferences is different and simpler than the estimation of sponsors' preferences because investors' asset demand functions are linear in preference parameters or some known function of those. To see this recall that the plan level demand system is given by

$$\boldsymbol{s}_p(\boldsymbol{f}_p;\boldsymbol{\eta}_p) = \delta \boldsymbol{e}_p + \frac{(1-\delta)}{\gamma}(I + X_{(2)p}X'_{(2)p})^{-1}(X_{(1)p}\boldsymbol{\beta} + \boldsymbol{\xi}_p - \boldsymbol{f}_p) \tag{25}$$

which, after multiplying both sides by $I + X_{(2)p}X'_{(2)p}$, becomes a system of estimating equations whose RHS only depends on own demand shocks and whose LHS is an observed linear transformation of the observed plan-level portfolio allocations

$$\tilde{\boldsymbol{s}}_p(\boldsymbol{f}_p;\boldsymbol{\eta}_p) = \delta \tilde{\boldsymbol{e}}_p + X_{(1)p}\tilde{\boldsymbol{\beta}} - \tilde{\gamma}\boldsymbol{f}_p + \boldsymbol{\xi}_p \tag{26}$$

with $\tilde{\boldsymbol{s}}_p \equiv (I + X_{(2)}X'_{(2)})\boldsymbol{s}_p$, $\tilde{\boldsymbol{e}}_d \equiv (I + X_{(2)}X'_{(2)})\boldsymbol{e}_d$, $\tilde{\boldsymbol{\beta}} \equiv \boldsymbol{\beta}(1-\delta)/\gamma$ and $\tilde{\boldsymbol{\xi}} \equiv \boldsymbol{\xi}(1-\delta)/\gamma$. Equation (26) can be estimated via linear regression methods.

As before, I allow for the possibility that funds' fees are correlated with plan investors' demand shocks $\boldsymbol{\xi}_p$. As before, if funds make their price-setting decision after observing $\boldsymbol{\xi}_p$, demand and supply simultaneity would bias preference parameters estimates. To account for this type of price endogeneity I follow the same approach I used for the identification of sponsors' preferences. First, I exploit the granularity of the data to absorb unobserved heterogeneity in demand by including time, funds' brand, passive and investment category fixed effects. In this case, because the estimation is at the fund-plan level I am also able to include plan/sponsors' fixed effects to absorb plan-level preference shocks. Second, I instrument funds' fees with funds' turnover ratio which captures trading costs that are pass on to investors through higher fees. Again, the identifying assumption is that variation in funds' turnover ratio not explained by time, brand, category, passive and plan fixed effects enters investors' demand only through fees.

I estimate investors' preferences applying linear IV to equation (26) under the assumption that funds' turnover ratios Z_j are mean-independent of investors' demand shocks, formally, I require that $\mathbb{E}[\xi_{jp}|Z_j] = 0$. To compute the LHS of (26) I need to specify which asset characteristics form the matrix $X_{(2)p}$. In classic portfolio theory $X_{(2)}$ would include characteristics capturing the correlation structure between assets. Perhaps the most natural way to proceed would be to construct $X_{(2)}$ after estimating funds' loadings onto some underlying risk factors from the time-series of funds' returns and then, for each plan, compute the variance-covariance matrix of the assets available therein.

Although common in practice, I do not follow this approach and instead use funds' classification into investment categories to construct $X_{(2)}$. The reason I do so is that assets' loadings on standard factors do not seem to explain the retirement portfolio allocations observed in the data, whereas investment category fixed effects do. Table 6 presents the R2 from regressing observed portfolio allocations on categories and assets' factors loadings. Factors alone explain close to 4% of the observed variation in portfolio shares whereas investment categories fixed effects alone explain more than three times that. More importantly, after absorbing categories fixed effects, factors' R2 drops substantially to 0.1% suggesting that factors explanatory power was just proxing for investment categories. In a world in which investors' allocation decisions depend on assets' factor structure we would expect factor loadings to have some power in explaining observed portfolio shares.

Based on this evidence, I use asset categories to model how plan investors interpret assets substitution patterns. I do so by creating a three level nesting structure of asset categories. The first level consists into three broad asset classes Equity, Allocation and Bond. Then I create a second level for each of these classes. For instance, funds belonging to the Equity class are further classified into Equity-Large, Equity-Mid, Equity-Small and Equity-International. In the third level, Equity-Large fund are further classified into Equity-Large-Blend, Equity-Large-Growth and Equity-Large-Value and similarly for other second level categories. I consider each of these levels as a separate asset characteristic corresponding to a column of the matrix $X_{(2)}$; for instance if j is an Equity-Large-Value fund then the jth row of $X_{(2)}$ is a vector $\boldsymbol{x}_{(2)j}$ that takes value 1 for the Equity, Equity-Large and Equity-Large-Value columns and 0 everywhere else. The outer product of this matrix of category indicators $X_{(2)}X'_{(2)}$ is then a three-block diagonal matrix whose element (j, l) element $x'_{(2)j}x_{(2)l}$ equals 3 if fund j and l belong to the same 3rd level category (e.g., both are Equity-Large-Value funds), equals 2 if they belong to the same 2nd level but to a different 3rd level category, equals 1 if they only belong to the same 1st level and equals 0 otherwise.[33]

[33] Another way to see this is thinking about $X_{(2)}X'_{(2)}$ as a quadratic interaction of fixed effects where each level of classification is a fixed effect. I provide an illustrative example in Appendix C.

6.4 Estimates of investors' preferences

Table 7 presents the estimates of investors' preference parameters based on the linear specification in equation (26). The first three columns present some OLS estimates whereas the fourth column reports the IV estimates from instrumenting funds' fees with funds' turnover ratios. The estimates reported correspond to the coefficients on the RHS of equation (26). Besides funds' expenses, I assume that past returns gross of fees and funds' affiliation enter the set of asset characteristics $X_{(1)}$ determining the linear component of plan investors' preferences $X_{(1)p}\boldsymbol{\beta}+\boldsymbol{\xi}_p-\boldsymbol{f}_p$. Investment categories also enter this linear component of investors' utility because I absorb category fixed effects in all specifications. If one interprets this linear component as investors' subjective expectations about assets' returns, the implicit assumption is that investors' subjective beliefs depend on funds' past returns relative to their corresponding investment category.

The OLS estimates broadly suggest that plan investors dislike fees, like returns and have a preference for funds' that are affiliated with their sponsor recordkeeper. A closer look at the magnitudes further reveals that plan investors care more about funds' returns than their sponsors because, in this case, the preference coefficients on returns and fees are comparable in magnitude.[34] This is consistent with sponsors designing their plan to merely comply with regulation and minimize liability risk. Current ERISA regulation indeed prescribes that sponsors are not liable for funds' market performance to the extent that their plan includes high quality options compared to the alternative available in the market. Because fees are known whereas performance is uncertain, it is not surprising that sponsors care more about fees rather than gross performance as they cannot be held accountable for the latter.

Plan investors seem to value funds' affiliation less than their sponsors. For the latter, funds' affiliation is a crucial driver of plan inclusion decisions whereas for plan investors the importance of funds' affiliation is more modest although not irrelevant. This is consistent with agency frictions mostly biting at the plan design stage where sponsors tend to favor affiliated funds when constructing their retirement plan (Pool, Sialm and Stefanescu (2016)). Agency frictions could potentially spillover to the investment stage if, for instance, recordkeepers also offered advising services to plan investors and were to push investors towards affiliated funds. Although I am not aware of any empirical evidence on mis-advising for the particular context I am considering, a theoretical literature in financial economics contemplates this possibility (Inderst and Ottaviani (2012a), Inderst and Ottaviani (2012b)).

The coefficient on fees increases moving from the first to the third columns of the OLS estimates. In particular, it doubles when I control for fund brand fixed effects.

[34] The estimated magnitudes in Table 5 suggest the same is not true for sponsors. Also, note that in Table 5 fees are measure in basis points. while in Table 7 both are expressed in percentage points.

This is consistent with fund brands potentially capturing a good amount of unobserved heterogeneity in investors' preferences for particular fund brands thereby making funds of the same investment provider (and within the same category) perceived as more substitutable to each other. The same happens to the coefficient on funds' gross returns which is not surprising as we expect investors to care about returns net of fees. Including plan (or equivalently sponsor) fixed effects does not affect too much fees and returns coefficients but triples the coefficient on funds' affiliation suggesting that the extent with which agency frictions matter might depend on sponsor-level unobservables such as sponsors bargaining power in negotiations with the recordkeeper (Bhattacharya and Illannes (2022)). Lastly, across all specifications, the model estimates that one out of four investors do not make an active investment decision.

I report the IV estimates in the fourth column of Table 7. All coefficients, except for the one on funds' fees, are largely unchanged. Conversely, the coefficient on funds' expenses is almost four times larger. The discrepancy between OLS and IV estimates is common in contexts where prices and quantities are determined simultaneously in equilibrium. OLS estimates often imply inelastic demand curves because the observed variation in quantity and prices is also due to shifts in demand. However, after instrumenting for prices, the resulting estimates recover demand curves that are much more elastic. This pattern emerges clearly when looking at the elasticity to fees of investors' portfolio allocations, which I report in the second part of Table 7. Plan investors' portfolio allocations are inelastic under OLS but become elastic after instrumenting funds' fees with funds' turnover ratios. The median elasticity is around 2.3 which is almost twice larger than sponsors' elasticity, indicating that sponsors may not be internalizing investors' preferences when designing their plan menu. The misalignment in elasticities increases even more if we consider investors who actively form their retirement portfolio. The median fee elasticity for active investors is around 3, nearly three times larger than sponsors' elasticity to fees.

To get a better sense of the magnitudes the last column of Table 7 reports the marginal change in investors' portfolio allocations implied by a unit change of the corresponding characteristic. The change in allocations is measured in percentage points per basis point increase in expenses or gross returns. A 10 basis points increase in fees reduces the corresponding portfolio allocation by nearly 0.9 percentage points whereas a 10 basis points increase in past gross return increases the corresponding portfolio allocation by almost 0.4 percentage points. Because I am absorbing category fixed effects, the latter should be interpreted has the effect of a 10 basis points increase in funds' performance relative to their corresponding investment category, consistent with plan investors chasing performance (Chevalier and Ellison (1997)). These magnitudes are not small if one considers that the average retirement portfolio allocation is of about 3%.

The model estimates provided in Table 7 pool together all cross-sections from 2010 to 2019, although the identifying variation remains cross-sectional because I always include year fixed effects. That being said, nothing prevents me from estimating investors' preferences separately for each observed cross-section of plan menus and assess how such estimates have changed over time.

Figure 13 plots the estimated median fee elasticity for each cross-section from 2010 to 2019. Two broad patterns emerge from the picture. First, investors seem to have become more sensitive to fees over time, with a sharp drop in the estimated elasticity from 2011 to 2013. Second, investors have become more inactive, with the estimated elasticity for active investors diverging from the elasticity of all investors. The first pattern could be a consequence of the regulatory push that required plan sponsors and investment providers to disclose investment fees to plan investors. Specifically, starting from the year 2012 the Department of Labor (DOL) required plan sponsors to disclose information about funds' expenses and performance directly to plan investors and recent empirical evidence suggests that investors have become significantly more attentive to fees as a consequence of that (Kronlund, Pool, Sialm and Stefanescu (2021)).

The second pattern is likely a symptom of the growth in the demand and supply of TDFs following the 2006 Pension Protection Act which identified TDFs as one of the qualified default investment alternatives for retirement plans. Since then, TDFs have become a constant component of retirement plan menus with more than 80% of sponsors offering at least one TDF in their plan as of 2019 (Figure 15). At the same time, plan investors have been increasing their TDFs holdings, with the average portfolio share of TDFs across plans growing three-folds from approximately 10% in 2010 to more than 30% as of 2019 (Figure 16). This could be either due to the presence of more inactive investors whose contributions are automatically defaulted into some of the available TDFs or could be active investors choosing to allocate more of their contributions toward TDFs. My model gives credit to the former and attributes this increase in TDFs' portfolio share to the presence of more inactive investors. Indeed, the estimated fraction of inactive investors increases from roughly 15% in 2010 to 40% in 2019 (Figure 14). This large increase in the share of investors that do not actively form their portfolio explains why the estimated fee elasticity for active investors diverges from the fee elasticity of all investors.

7 Price-cost Margins and Fee Decomposition

In this section I combine the estimates of sponsors' and investors' preferences together with the Nash-Bertrand first order conditions derived in (20) to recover funds' price-cost margins. After that, I exploit the characterization of equilibrium fees derived in equation (21) to decompose the observed variation in fees into the monopolist margin, the Hotelling markdown and the plan inclusion markdown.

7.1 Recovering funds' price-cost margins

To begin with, I rewrite funds' profit maximization problem making explicit its dependence on the various dimensions of variation I have in the data. I index time periods (i.e., years) by t and recordkeepers by r. Next, I denote by R_{jt} the set of recordkeepers that include fund j in their network of funds and by P_{rt} the set of retirement plans administered by recordkeeper r. I assume that funds set fees simultaneously in each period before sponsors form their retirement menu but knowing which recordkeeper networks they belong to. I rewrite problem (15) as follows:

$$\max_{(f_{jt})_t} \quad \sum_t \sum_{r \in R_{jt}} P_{rt} \cdot (f_{jt} - c_{jt}) \cdot \int \phi^r_{jpt}(\boldsymbol{f}; \boldsymbol{\theta}_p) s^r_{jpt}(\boldsymbol{f}; \boldsymbol{\eta}_p) A_p dF(A_p, \boldsymbol{\theta}_p, \boldsymbol{\eta}_p) \tag{27}$$

where I made explicit the dependence of the inclusion probabilities and portfolio shares on the identity of the plan recordkeeper. This dependence is a consequence of the fact that different recordkeepers have different networks of funds.

The first order conditions associated with problem (27) imply the following price-cost margins for fund j:

$$f_{jt} - c_{jt} = \frac{\sum_{r \in R_{jt}} P_{rt} \cdot \int \phi^r_{jpt} s^r_{jpt} A_p dF_p}{\left(\sum_{r \in R_{jt}} P_{rt} \cdot \int \frac{\partial \phi^r_{jpt}}{\partial f_{jt}} s^r_{jpt} A_p dF_p\right) + \left(\sum_{r \in R_{jt}} P_{rt} \cdot \int \phi^r_{jpt} (1 - \delta_p) \gamma_p^{-1} (1 - \bar{\kappa}^{r,p,t}_{jj}) A_p dF_p\right)} \tag{28}$$

where the two addends in the denominator represent the revenue loss from the marginal sponsor and the revenue loss from the marginal investor respectively.

By plugging in (28) the estimated distributions of sponsors' and investors' preference parameters one obtains the margins charged by funds in an interior Nash-Bertrand equilibrium. I report the estimated margins and marginal costs in Table 8. The first set of columns presents the estimates for the full sample of funds, whereas the second and third sets of columns focus on passive funds and TDFs respectively.[35] Starting from the sample of all funds the estimates suggest that the median fund charges a margin of about 14 basis points which relative to a median expense ratio of 74 basis points implies a markup around 19%. The average fund is as efficient as the median fund and charges roughly the same markups suggesting that the overall distribution of costs is symmetric.

Perhaps not surprisingly things change when looking at passive funds. The median passive fund has a marginal cost more than three times lower than the median fund among all funds and charges a margin that is around 7 basis points. Interestingly, although the absolute margin for the median passive fund is twice smaller than the margin for the median fund among all funds, in relative terms, it charges a markup of

[35] I provide more details on the derivation in Appendix C

about 33% which is nearly twice bigger. This is suggestive of the fact that, although passive funds are typically perceived as more homogeneous products, they still enjoy substantial market power and do not pass all their cost efficiency down to investors (Hortaçsu and Syverson (2004)). Compared to all funds taken together, the distribution of costs for passive funds is more skewed, with the average fund bearing a marginal cost of about 21 basis points. However, the absolute margin charged by the average fund is similar to the median suggesting that the skewness is mostly driven by the cost structure.

The estimated margins and costs for Target Date Funds are reported in the last set of columns of Table 8. Starting from the fees, we can see that TDFs tend to be more expensive than passive funds but cheaper than all funds taken together, with the median and average TDFs charging an expense ratio of about 32 and 37 basis points respectively. On the other hand, the estimated cost structure for TDFs is remarkably similar to the one for passive funds, suggesting that TDFs, although being cost-efficient investment vehicles, charge higher margins to investors. In fact, the margin for the median TDFs is of about 14 basis points, with an implied markup of about 44%. The fact that TDFs charge higher margins is not surprising given that they target investors that are likely to be inactive and not too responsive to fees. Recent empirical evidence suggests that TDFs charge excessive fees because they are structured as funds of funds and, as such, their expenses reflect multiple layers of fees. Moreover, the vast majority of a TDFs' holdings are funds that belong to same fund family as the TDF itself and, some TDFs tend not to include the cheapest share classes of such funds (Brown and Davies (2021)).

7.2 Decomposition of equilibrium fees

In this section I decompose the observed fees exploiting the decomposition derived in equation (21). Specifically, I use the estimated preference parameters and the estimated marginal costs to decompose the observed fees into (1) monopolistic fee, (2) Hotelling markdown and (3) plan inclusion markdown. I perform this decomposition for each year of the sample from 2010 to 2019 and obtain the following decomposition for each fund j in each year t

$$f_{jt} = \frac{\tilde{\mu}_{jt} + c_{jt}}{2} - h_{jt} - \bar{\iota}_{jt} \tag{29}$$

where $\bar{\iota}_{jt}$ is the jth component of the vector

$$\left(I - \frac{diag(\tilde{\mathcal{K}}_t)}{2} - \frac{\tilde{\mathcal{K}}_t}{2}\right)^{-1} \frac{\boldsymbol{\iota_t}}{2}. \tag{30}$$

To get a sense of the magnitudes, Table 9 shows the average of each of the three components across all funds and cross-sections. The first column reports the average observed fee which is about 66 basis points.[36] The last three columns instead show the averages for each of the three components respectively. A monopolist, on average would charge a fee of about 120 basis points but, because of competition, it needs to give up about 44% of such fee. On average, the two competitive markdowns reduce the monopolist fee by nearly 54 basis points. The contribution of each of those is similar, suggesting that competition for entering investors' choice sets and competition in terms of product characteristics are equally important. On average, the Hotelling and inclusion markdowns each erode more than 20% of the fee a monopolist would be able to charge to its consumers.

Figure 17 repeats the same exercise for each cross-section from 2010 to 2019. The black solid line represents the average fee and shows its declining trend from around 80 basis points in 2010 to nearly 50 basis points in 2019. The decomposition sheds light on the sources of this decline. The blue bars suggest that the decline in fees is not a consequence of changes in investors willingness to pay, as captured by $\tilde{\mu}$, nor of changes in technological primitives as captured by funds' marginal costs c. The monopolist fee $(\tilde{\mu} + c)/2$ indeed has been roughly stable over time fluctuating between 100 and 130 basis points. On the other hand, the two markdowns seem to be the driver of such declining trend in fees. In absolute terms, they went from reducing the monopolist fee by about 29 basis points in 2010 to nearly 78 basis points in 2019. In relative terms, they accounted for a 27% reduction of the monopolist fee in 2010 which has more than doubled over time, accounting for a 59% reduction in 2019. Overall, these patterns are consistent with both sponsors and investors becoming more sensitive to funds' expenses.

8 Counterfactuals

In this section I evaluate the effects of three policy counterfactuals regulating the design of retirement plans. First, I consider the elimination of agency frictions whereby sponsors favor funds affiliated with their plan recordkeeper. Second, I consider the effects of a policy that mandates the inclusion of low-cost options such as low-cost S&P 500 index funds trackers or low-cost TDFs. Lastly, I consider a policy that caps funds' expenses.

For all counterfactuals, I quantify how the policy in question impacts plan investors welfare and plan expenses relative to the status quo. The latter corresponds to the welfare and expenses computed for the observed plan menus, plan expenses and portfolio

[36] This fee is slightly lower than the overall average fee because, to reduce the computational burden, I performed the decomposition only including the 200 largest funds in each recordkeeper network of funds. On average the 200 largest funds accounted for more than 80% of the AUM managed by the recordkeeper.

allocations. The main takeaway is that mandating the inclusion of low-cost default options and imposing expense ratio caps are the most effective policies. Assuming that sponsors do not value funds' affiliation does not improve investors' outcomes because nothing prevents them to include expensive options that are not affiliated. Similarly, mandating the inclusion of low-cost index funds has only a modest effect on welfare and expenses. The reason is that sponsors will still be including expensive funds and investors will still be investing in those either because they are inactive or for diversification purposes. Mandating the inclusion of low-cost TDFs improves outcomes because inactive investors benefit only from having access to cheaper default options.

To measure investors surplus I rely on the quadratic specification of investors' portfolio problem defined in (9). At the optimal portfolio allocation, the surplus for active investor i in plan p can be written as,[37]

$$IS_i \equiv \frac{1}{2}\sum_{j=1}^{J_p} a_{ji}(\boldsymbol{f};\boldsymbol{\eta}_p)(\mu_{jp} - f_j)$$

where recall that $\mu_{jp} = X'_{(1)}\beta + \xi_{jp}$. Investors' surplus is the sum of the areas below the demand curves of each asset. Because preferences are quadratic and, in turn, the demand for each asset is linear, this area corresponds to a rectangular triangle with height given by $(\mu_{jp} - f_j)$ and base given by given by a_j. Integrating over all active investors we obtain the average surplus for a plan p active investor:

$$IS_p^{\text{active}} = \frac{1}{2}\sum_{j=1}^{J_p} s_{jp}^{\text{active}}(\boldsymbol{f};\boldsymbol{\eta}_p)(\mu_{jp} - f_j)$$

where

$$s_{jp}^{\text{active}} \equiv \sum_{i\in I_{p,\text{active}}} \frac{A}{(1-\delta)A_p} a_{ji}.$$

To obtain a complete welfare measure I need to incorporate the surplus of inactive investors. Because, in the model I do not specify any preference for these investors, I define their surplus as

$$IS_p^{\text{inactive}} = \frac{1}{2}(\mu_d - f_d)$$

where d is plan p's default option. The surplus for a plan p investor is given by

$$IS_p = \delta \cdot IS_p^{\text{inactive}} + (1-\delta)\cdot IS_p^{\text{active}}$$

[37] See Appendix C for a derivation.

and, the overall surplus is then

$$\int IS_p(\boldsymbol{f};\boldsymbol{\eta}_p)dF(\boldsymbol{\eta}_p). \tag{31}$$

Each counterfactual amounts to (1) imposing the policy, (2) solve for funds' counterfactual equilibrium fees, (3) simulate sponsors' plan menus under the counterfactual policy, (4) compute investors' counterfactual portfolios and compute their surplus from equation (31) . Because investors have preferences over their portfolio allocations, their surplus is measured in units of (subjective) excess returns net of fees (e.g., $\mu_{jp} - f_j$). For active investors, I recover μ_{jp} from their estimated preference parameters by computing $\mu_{jp} = X'_{(1),j,p}\hat{\boldsymbol{\beta}} + \hat{\xi}_{jp}$, where $\hat{\xi}_{jp}$ are obtained from the residuals of the linear regression in (12) multiplied by the estimated $\tilde{\gamma}$.[38] For inactive investors, I use the average annual (excess) return to measure μ_d.

8.1 Eliminating preference for affiliated funds

The first counterfactual I consider restricts sponsors' preferences by forcing them not to value funds' affiliation when making plan design decisions. Under this restriction, holding every other characteristic constant, an affiliated fund and a non-affiliated fund will have the same likelihood of being included in a given plan. A practical way to implement such policy would be issuing penalties for sponsors that are found to be favoring affiliated funds even though they exhibit worse performance than otherwise similar alternatives (Pool, Sialm and Stefanescu (2016)).

The second row of Table 10 presents the results from this counterfactual exercise and shows that such policy would be ineffective, leaving investor surplus and plan expenses almost unchanged. In fact, plan expenses decrease by 1 basis point and investor surplus decreases by 3 basis points.

This policy is ineffective because removing sponsors' preference for affiliated funds does not prevent them from including expensive funds that are not affiliated with their recordkeeper. For this reason, counterfactual plan expenses remain unchanged as well as sponsors' sensitivity to funds' fees.

Investors' surplus is also almost unaffected. The slight decrease in their surplus could be either because investors have a small preference from affiliated funds or because affiliated funds are not systematically worse than other alternatives. For example, in many cases TDFs are affiliated with the plan recordkeeper and, we know that their expenses are typically below average (Table 8) and are particularly valued by inactive investors.

[38] Figure 19 shows the estimated distribution of the average μ_{jp} across plans by asset class.

8.2 Mandating the inclusion of low-cost options

The second set of counterfactuals studies the effect of policies mandating the inclusion of low-cost investment options. I consider both the inclusion of low-cost index funds that track the S&P 500 and the inclusion of low-cost TDFs. Both policies improve investors' outcome and lead to a reduction of the average plan expenses relative to the status quo.

Mandating the inclusion of at least one low-cost index fund increases investors' surplus by 2% and decreases the average plan expense by 10% relative to the status quo (Figure 20). In magnitudes, investors surplus increases by 5 basis points and the average expense decreases by 5 basis points as shown in the third row of Table 10. To add perspective, the third column of Table 10 computes the dollar savings implied by the policy change on an account balance of $50,000. The policy leads to a $25 savings per year. In the last column of Table 10 I instead consider the dollar savings over 40 years for an household receiving an annual income of $70,000 and contributing 10% to its 401(k) every year.[39] Such household would save approximately $12,000 in fees after the implementation of this policy.

While one might have expected significant impact from such a policy, its actual effects are relatively modest. This is noteworthy because I selected low-cost funds renowned for being both affordable. Each comes with an expense ratio well under 10 basis points, complemented by a 5-star Morningstar rating. My selection includes the Vanguard 500 Index fund (VFIAX) and ETF (VOO), Fidelity 500 Index Fund (FXAIX), Schwab S&P 500 Index Fund (SWPPX), Blackrock iShare Core S&P 500 ETF (IVV), and SPDR S&P 500 ETF (SPY).

Despite this, there are a couple of key reasons why this policy has had only a modest effect on investors' welfare. Firstly, although investors value lower fees, they also want to diversify across all assets available. As such, they will substitute toward the low-cost index only up to the point that does not hurt their diversification needs. This in practice requires maintaining part of the holdings into more expensive funds. Secondly, a significant segment of investors remain inactive in their investment approach. Instead of actively selecting funds, they default their contributions to the available TDF. The addition of a low-cost index fund does not alter the behavior of this group. Furthermore, it's worth noting that nearly half of sponsors already incorporate these type of funds in their existing offerings, implying that the potential welfare gains only come from the half the sponsors not offering those type of funds as part of their plan menu.

Investors' outcome improves if, instead of mandating low-cost index funds, the policy mandates the inclusion of low-cost TDF. In this case, investor surplus increases by 11% and the average plan expense decreases by 23%. In magnitudes, investors

[39] For such computation I assume an annual return of 6%.

surplus increases by 27 basis points whereas fees decrease by 12 basis points. An investor with a \$50,000 retirement account balance would save \$60 per year as a consequence of such policy, an amount more than twice larger than the savings under a policy mandating the inclusion of low-cost S&P 500 trackers. Similarly, an household with a \$70,000 income who contributes 10% to its 401(k) account would be saving \$28,832 in fees over a 40 years period, again an amount twice larger than the savings under a policy mandating the inclusion of low-cost S&P 500 trackers.

Why is it more effective to mandate the inclusion of low-cost TDFs than simply focusing on low-cost index funds? The primary reason lies in the benefit distribution: low-cost TDFs serve both active and, especially, inactive investors because they are used as qualified default options. In contrast, mandating low-cost index funds predominantly benefits active investors, leaving inactive ones with no benefit in terms of reduced fees. Additionally, while TDFs often carry higher fees compared to index funds, the most affordable TDFs have expense ratios closely aligned with the ones charged by low-cost index funds. Take, for instance, the Fidelity Freedom Index series and the Vanguard Target Retirement series — both offer TDFs with expense ratios under 10 basis points.

8.3 Capping funds' expenses

The last counterfactual I consider studies the effect of a 50 basis point expense ratio cap. Under this policy sponsors are allowed to include in their menu only funds with an expense ratio below 50 basis points. As a consequence, all funds whose marginal cost is higher than 50 basis point will exit the market. The latter are in most cases active funds, as more than 3/4 of passive funds and TDFs have an expense ratio below 50 basis points (Table 8).

This policy increases investor surplus by 14% and decreases the average plan expenses by 30%, corresponding to an increase in surplus of about 33 basis points and a decrease in plan expenses of about 15 basis points. A plan investor with a balance account of \$50,000 is expected to save \$75 per-year whereas an investor contributing 10% of its \$70,000 income is expected to save about \$36,000 in fees over 40 years.

Perhaps not surprisingly, this policy is the most effective among the ones I considered thus far. On the one hand, it benefits inactive investors by eliminating the right-tail of expensive TDFs. On the other, it benefits active investors by ensuring that all investment options available are not excessively expensive.

Before concluding, one remark is in order. My analysis so far abstracted away from extensive margin considerations and implicitly assumed that sponsors would be always willing to provide a retirement plan to their employees. In practice, plan provision could be affected by these type of policies. For example, under a 50 basis point expense ratio cap, it is likely that recordkeepers would lose revenues from revenue-sharing fees unless sponsors themselves compensate such loss by increasing their direct payments to their recordkeepers (Bhattacharya and Illannes (2022)). Some sponsors may be unwilling or might not have the resources to bear such costs and, consequently, might decide not to offer a retirement plan to their workers in the first place.

A plausible solution to minimize the extensive margin repercussions of these type of policies would be implementing such policies while at the same time subsidizing plan sponsors to incentivize plan provision. In practice, these type of subsidies have been already introduced in the 2019 SECURE Act to push small business to offer a retirement plan to their employees.

9 Conclusions

This paper proposes an equilibrium model of retirement plan design, portfolio choice and fee competition between investment providers to uncover the factors contributing to the design of low-quality and high-cost employer-sponsored retirement plans and quantify the welfare effects of policies regulating plan design.

The model features a two-layer demand system where, in the first layer, sponsors design their retirement plan and, in the second layer, plan investors form their retirement portfolio from the options available in their menu. On the supply side, investment funds compete by setting fees simultaneously while accounting for the two layers of demand. Funds compete for being included in a plan menu and for the plan assets.

I estimate the model using comprehensive data on retirement plan menus. Model estimates suggest that plan sponsors are less responsive to funds' fees than plan investors and value other fund characteristics, such as funds' affiliation with the plan recordkeeper. I use the estimated demand parameters to recover funds' price-cost margins and marginal costs from the implied Nash equilibrium conditions. Funds enjoy significant pricing power. This is particularly evident for Target-Date-Funds (TDFs), who, although almost as cost-efficient as index funds, charge double the margins, with an implied median markup of about 44%.

In the last part of the paper, I consider four policy counterfactuals that regulate the design of plan menus and quantify their effect on plan investors' welfare. The first counterfactual shuts down sponsors' preferences for funds' affiliation. The second set of counterfactuals considers mandating the inclusion of low-cost options. The last counterfactual instead imposes a 50 basis point cap on funds' expense ratios. Among those, requiring the inclusion of low-cost TDFs and capping expense ratios are the most effective in improving investors' outcomes. Specifically, mandating the inclusion of low-cost TDFs increases investors' surplus by 11%, whereas a 50 basis points expense ratio cap leads to a 14% increase. Both policies also significantly reduce average plan expenses by 23% and 30%, respectively.

An important caveat of my analysis is that it abstracts from extensive margin considerations. In particular, it assumes that these types of regulations do not affect sponsors' incentives to offer a retirement plan in the first place. In practice, imposing expense ratio caps might reduce plan provision (Bhattacharya and Illannes (2022)). A practical solution would be pairing these policies with plan provision subsidies. Quantifying the optimal subsidy scheme and the effect of this combination of policies on plan investors' welfare is an important direction for future research.

References

Badoer, D., Costello, C. and James, C. (2020), 'I can see clearly now: The impact of disclosure requirements on 401(k) fees', *Journal of Financial Economics* 136, 471–489.

Baker, M., Egan, M. and Sarkar, S. (2022), 'How do investors value esg?', NBER working paper.

Ballester, C., Calvó-Armenagol, A. and Zenou, Y. (2006), 'Who's who in networks. wanted: The key player', *Econometrica* 74(5), 1403–1417.

Beggs, S., Cardell, S. and Hausman, J. (1981), 'Assessing the potential demand for electric cars', *Journal of Econometrics* 16, 1–19.

Benartzi, s. and Thaler, R. (2001), 'Naive diversification strategies in defined contribution saving plans.', *American Economic Review* 91(1), 2997–3053.

Benartzi, s. and Thaler, R. (2007), 'Heuristics and biases in retirement savings behavior.', *Journal of Economic Perspectives* 21(3), 81–104.

Berk, J. and Green, R. (2004), 'Mutual fund flows and performance in rational markets', *Journal of Political Economy* 112(6), 1269–1295.

Berry, S. (1994), 'Estimating discrete-choice models of product differentiation', *The RAND Journal of Economics* 25(2), 242–262.

Berry, S., Levinsohn, J. and Pakes, A. (1995), 'Automobile prices in market equilibrium', *Econometrica* 63(4), 841–890.

Beshears, J., Choi, J., Laibson, D. and Madrian, B. (2009), 'The importance of default options for retirement saving outcomes: Evidence from the united states', *The Importance of Default Options for Retirement Saving Outcomes: Evidence from the United States.* University of Chicago Press.

Bhattacharya, V. and Illannes, G. (2022), 'The design of defined contributions plans', NBER working paper.

Bonacich, P. (1987), 'Power and centrality: A family of measures', *American Journal of Sociology* 92(5), 1170–1182.

Brown, D. and Davies, S. (2021), 'Off target: On the underperformance of target-date funds', SSRN working paper.

Carroll, G., Choi, J., Laibson, D., Madrian, B. and Metrick, A. (2009), 'Optimal defaults and active decisions', *The Quarterly Journal of Economics* 124(4), 1639–1674.

Chevalier, J. and Ellison, G. (1997), ‘Risk taking by mutual funds as a response to incentives’, *Journal of Political Economy* 105(6), 1167–1200.

Choi, J. (2015), ‘Contributions to defined contribution pension plans’, *Annual Reviews of Financial Economics* 7, 161–178.

Choukhmane, T. (2021), ‘Default options and retirement saving dynamics’, working paper.

Duarte, V., Fonseca, J., Goodman, A. and Parker, J. (2022), ‘Simple allocation rules and optimal portfolio choice over the lifecycle’, NBER working paper.

Egan, M., MacKay, A. and Yang, H. (2023), ‘What drives variation in investor portfolios? estimating the roles of beliefs and risk preferences.’, NBER working paper.

Gil-Bazo, J. and Ruiz-Verdu, P. (2009), ‘The relation between price and performance in the mutual fund industry’, *The Journal of Finance* 64(5), 2153–2183.

Goeree, M. (2008), ‘Limited information and advertising in the u.s. personal computer industry’, *Econometrica* 16, 1017–1074.

Grice, R. (2023), ‘Chain-linked markets’, working paper.

Grice, R. and Guecioueur, A. (2023), ‘Mutual fund market structure and company fee competition: Theory and evidence’, SSRN working paper.

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

Hotelling, H. (1929), ‘Stability in competition’, *The Economic Journal* 39(153), 41–57.

Huberman, G. and Jiang, W. (2006), ‘Offering versus choice in 401(k) plans: Equity exposure and number of funds’, *Journal of Finance* 61(2), 763–801.

Inderst, R. and Ottaviani, M. (2012a), ‘Competition through commissions and kickbacks’, *American Economic Review* 102(2), 780–809.

Inderst, R. and Ottaviani, M. (2012b), ‘How (not) to pay for advice?’, *Journal of Financial Economics* 105, 393–411.

Jackson, M. (2008), *Social and Economic Networks*, Princeton University Press.

Koijen, R. and Yogo, M. (2022), ‘The fragility of market risk insurance’, *The Journal of Finance* 77(2), 815–862.

Kronlund, M., Pool, V., Sialm, C. and Stefanescu, I. (2021), 'Out of sight no more? the effect of fee disclosures on 401(k) investment allocations', *Journal of Financial Economics* 141, 644–668.

Loseto, M. (2023), 'Network games of imperfect competition: An empirical framework', SSRN working paper.

Madrian, B. and Shea, D. (2001), 'The power of suggestion: Inertia in 401(k) participation and savings behavior', *The Quarterly Journal of Economics* 116(4), 1149–1187.

Markowitz, H. (1952), 'Portfolio selection', *Journal of Finance* 7(1), 77–91.

Massa, M. (2003), 'How do family strategies affect fund performance? when performance-maximization is not the only game in town', *Journal of Financial Economics* 67, 249–304.

Mellman, G. and Sanzenbacher, G. (2018), '401(k) lawsuits: What are the causes and consequences?', Center for Retirement Research, Boston College(18-8).

Pellegrino, B. (2023), 'Product differentiation and oligopoly: A network approach', SSRN working paper.

Pool, V., Sialm, C. and Stefanescu, I. (2016), 'It pays to set the menu: Mutual fund investment options in 401(k) plans', *The Journal of Finance* 521(4), 1179–1812.

Pool, V., Sialm, C. and Stefanescu, I. (2022), 'Mutual fund revenue sharing in 401k plans', NBER working paper.

Roussanov, N., Ruan, H. and Wei, Y. (2021), 'Marketing mutual funds', *Review of Financial Studies* 34, 3045–3094.

Stigler, G. (1961), 'The economics of information', *Journal of Political Economy* 69(3), 213–225.

Tang, N., Mitchell, O., Mottola, G. and Utkus, S. (2010), 'The efficiency of sponsor and participant portfolio choices in 401(k) plans', *Journal of Public Economics* 94(2), 1073–1085.

Ushchev, P. and Zenou, Y. (2018), 'Price competition in product variety networks', *Games and Economic Behavior* 110(2), 226–247.

A Figures

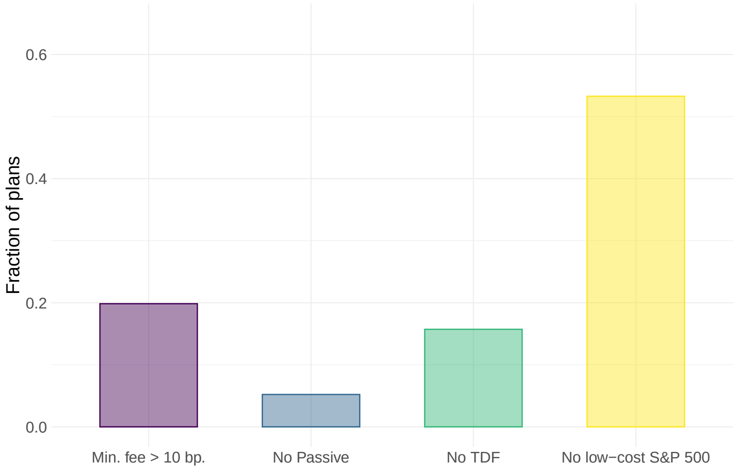

Figure 1: Plan quality for the year 2019. Low-cost S&P 500 Index funds include any of the Vanguard 500 Index (VFIAX/VOO), Fidelity 500 Index Fund (FXAIX), Schwab S&P 500 Index Fund (SWPPX), iShares Core S&P 500 ETF (IVV) and SPDR S&P 500 ETF Trust (SPY/SPLG).

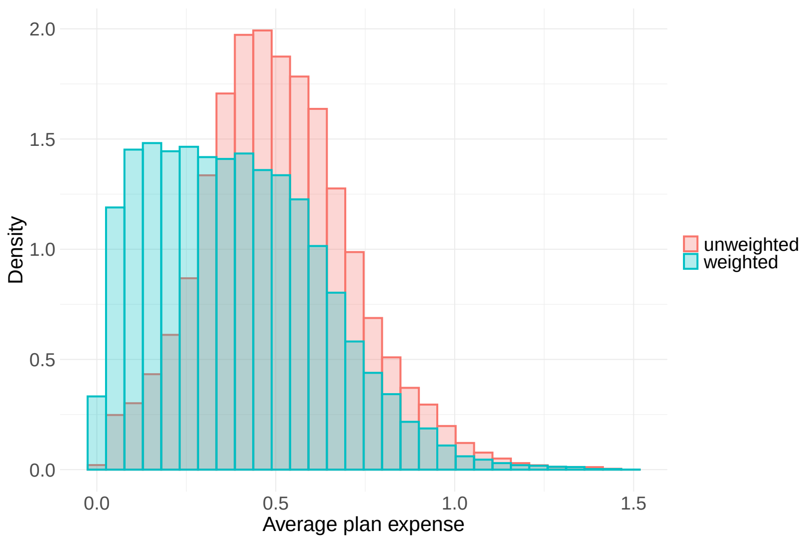

Figure 2: Distribution of average plan expense ratio across plans for the year 2019. Asset-weighted plan-level expenses are in blue. Expenses are measured in percentage points.

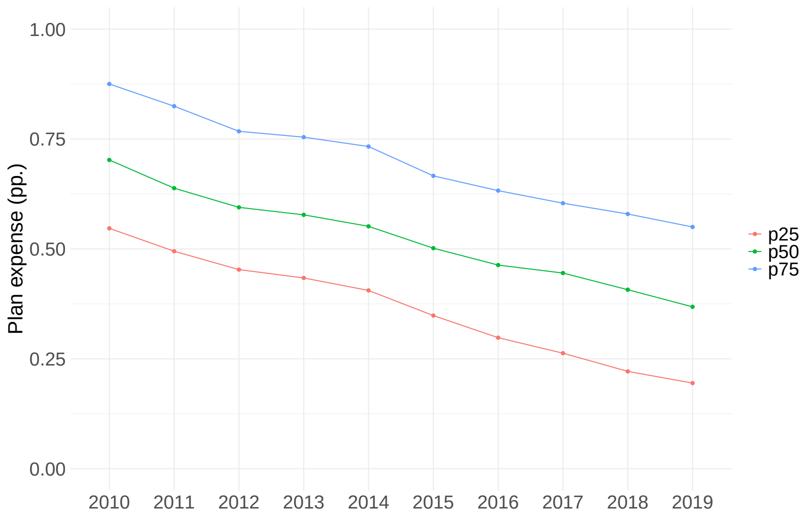

Figure 3: Distribution of average asset-weighted plan expense over time. Expenses are measured in percentage points

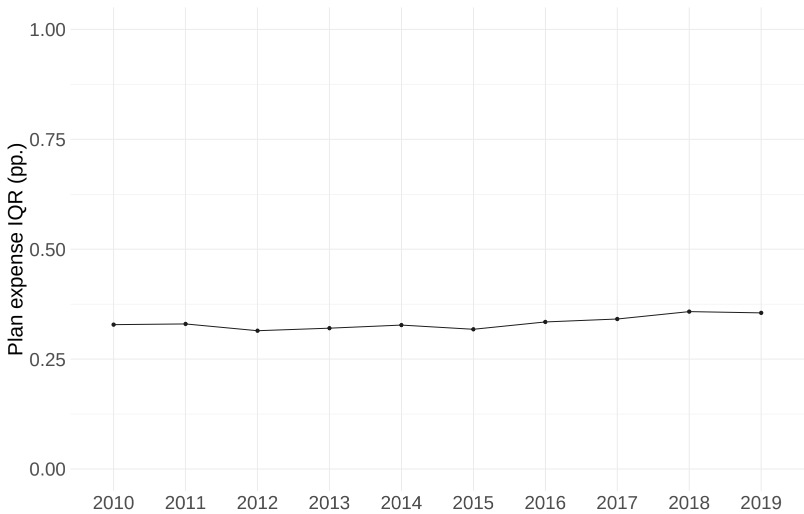

Figure 4: Interquartile range (IQR) of the average asset-weighted plan expense. Expenses are measured in percentage points.

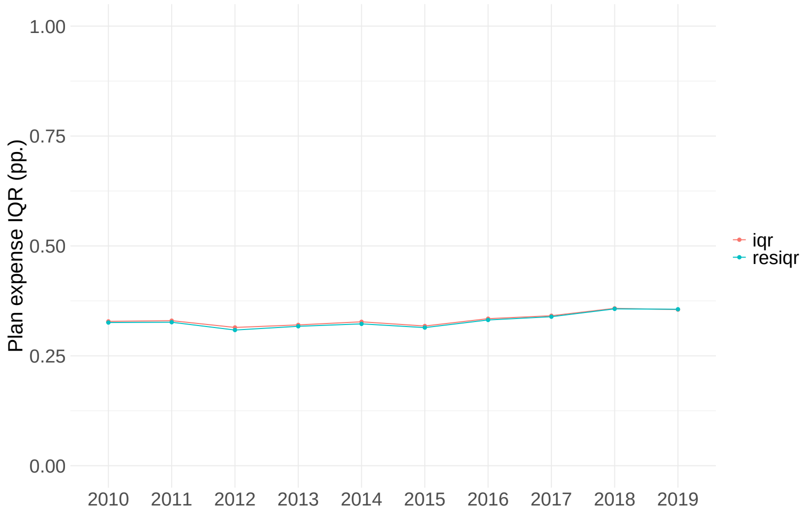

Figure 5: Interquartile range (IQR) of the average asset-weighted plan expense after partialling out plan size measured by total assets. Expense are measured in percentage points.

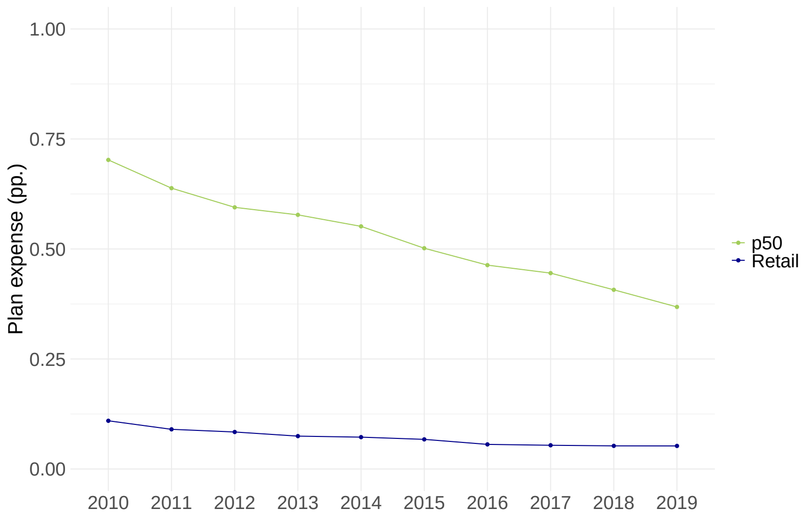

Figure 6: Median asset-weighted plan expense (green). Average expense ratio for a portfolio of Vanguard retail index funds (blue). Expenses are measured in percentage points.

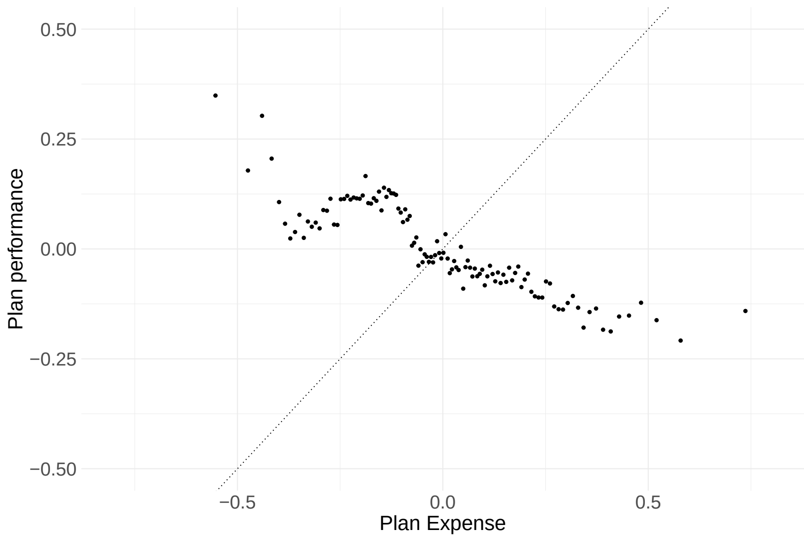

Figure 7: Bin-scatter of plan performance and average plan expenses corresponding to the specification in the second column of Table 1. Expenses and performance are demeaned with year fixed effects. Expenses and performance are measured in percentage points.

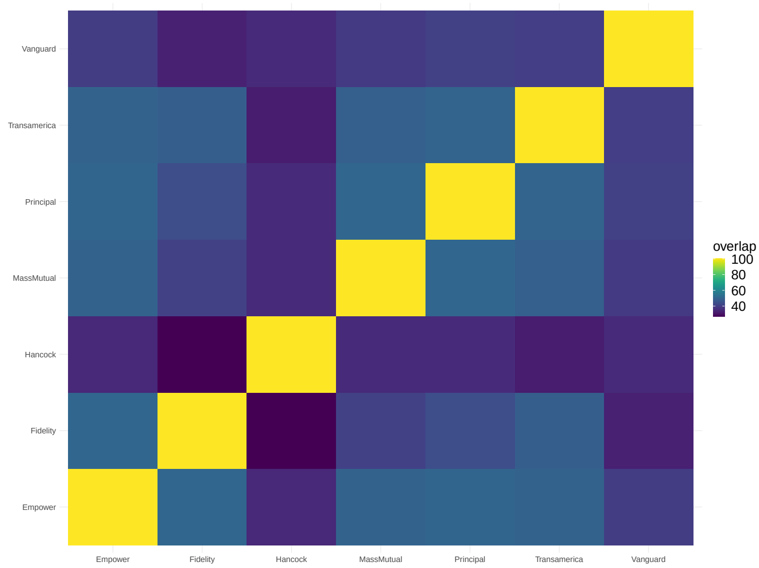

Figure 8: Average overlap in the network of funds of some of the largest recordkeepers for the year 2016. A fund belongs to a recordkeeper's network if it is offered in a plan managed by that same recordkeeper.

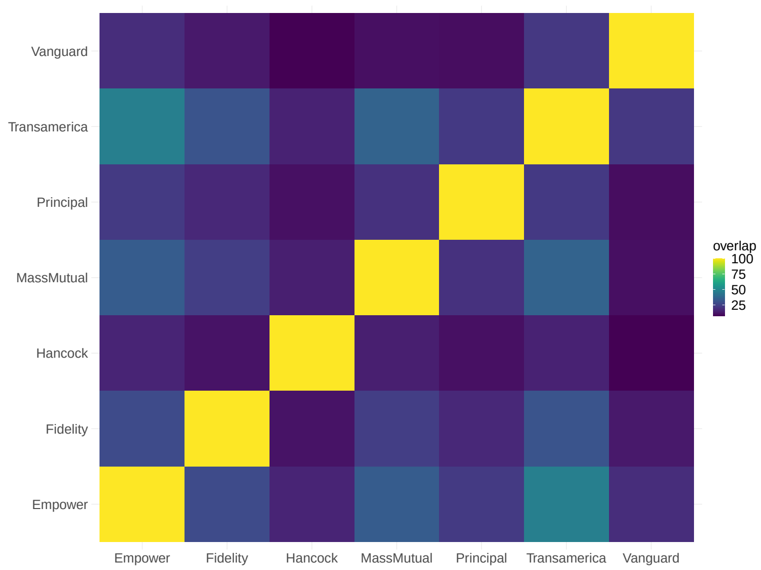

Figure 9: Asset-weighted average overlap in the network of funds of some of the largest recordkeepers for the year 2016. A fund belongs to a recordkeeper's network if it is offered in a plan managed by that recordkeeper.

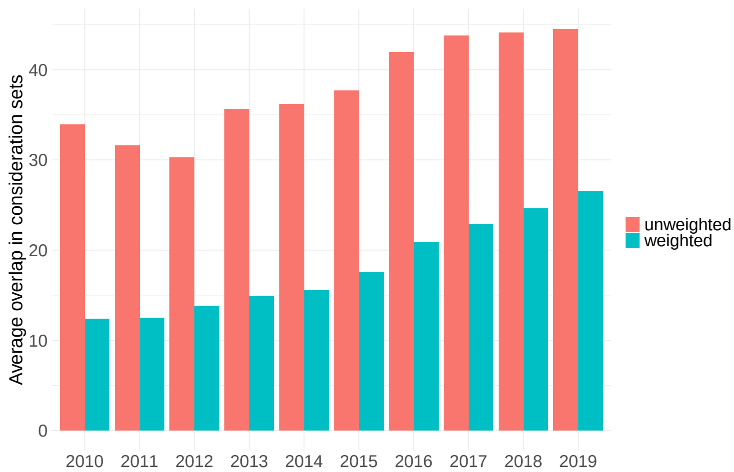

Figure 10: Average overlap in recordkeepers' network of funds. A fund belongs to a recordkeeper's network if it is offered in a plan managed by that same recordkeeper. The red bars represent the average fraction of funds that belong to the network of any two of the 10 largest recordkeepers. The turquoise bars represent the asset-weighted overlap.

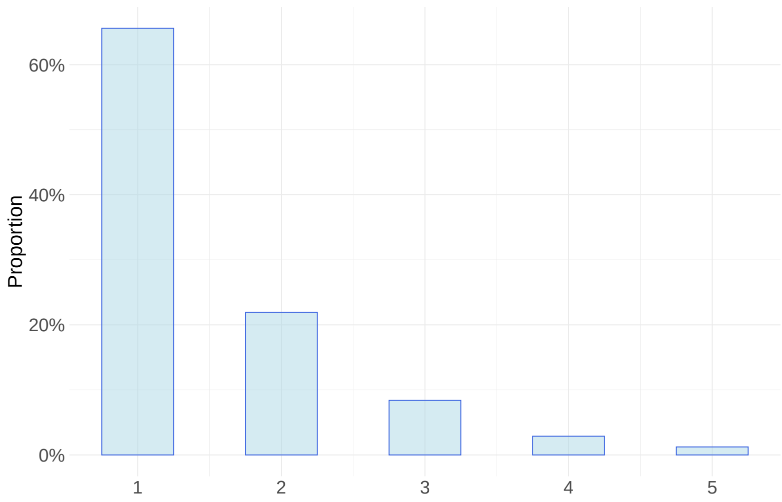

Figure 11: Distribution of number of options offered within investment category

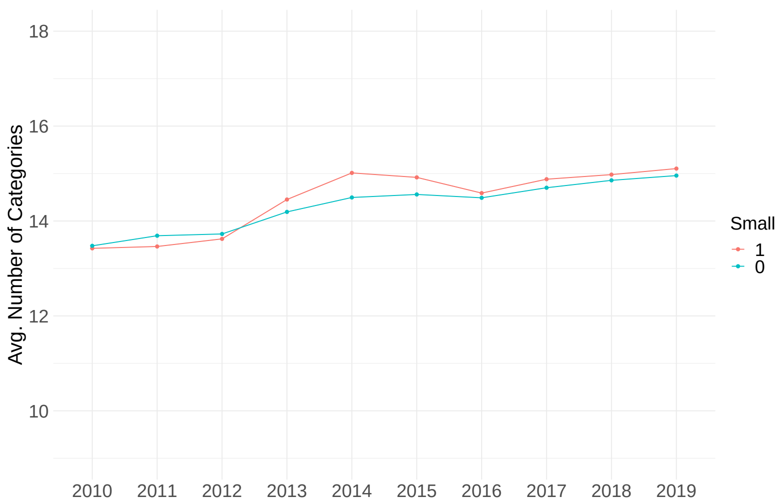

Figure 12: Average number of investment categories offered in a plan menu. The red line is for sponsors with number of participants below the median. The light blue line is for sponsors with number of participants above the median.

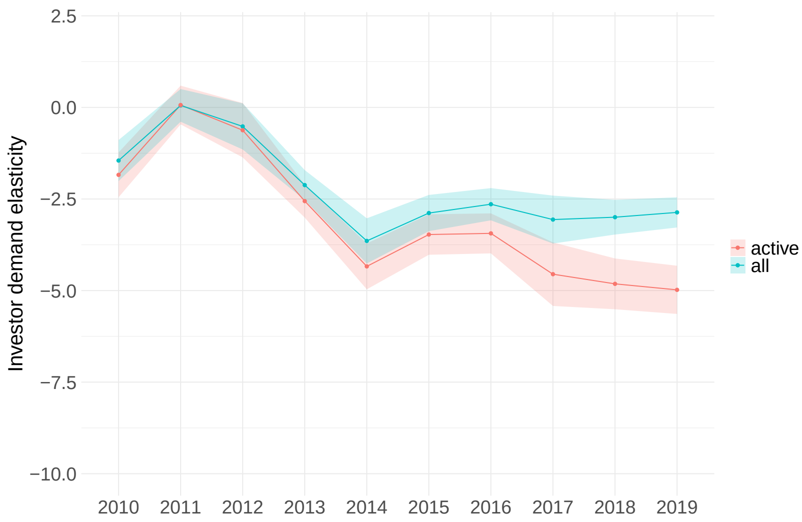

Figure 13: Cross-sectional estimates of plan investors median portfolio elasticity to funds' fees. Red line is for active investors. Blue line if for both active and inactive investors.

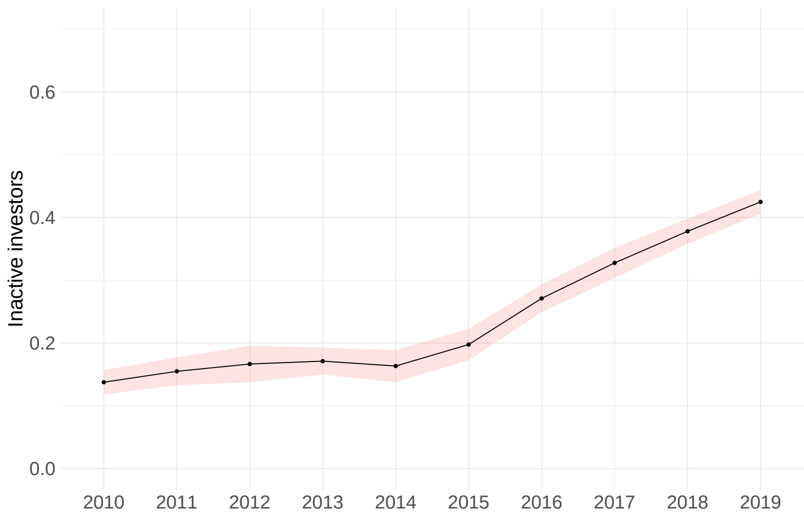

Figure 14: Cross-sectional estimates of the fraction of inactive investors.

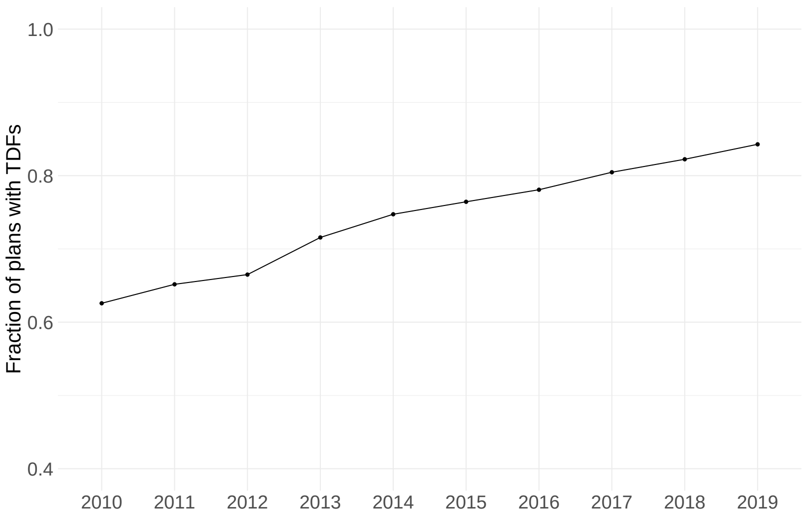

Figure 15: Share of retirement plans that offer at least one Target-Date-Fund (TDF).

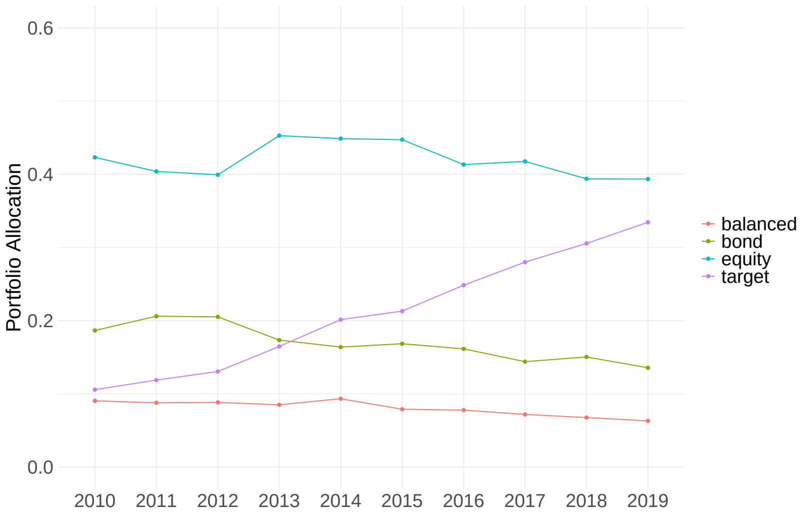

Figure 16: Average portfolio share across plan menus by asset class. Equity includes both US and International Equity funds. Balanced includes aggressive, moderate and conservative allocation funds that are not Target Date Funds (TDFs).

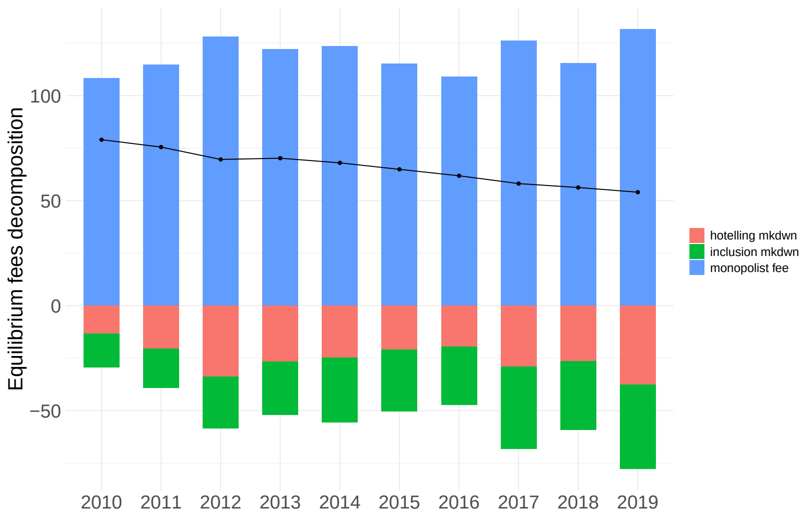

Figure 17: Cross-sectional decomposition of the average expense ratio into monopolist fee, hotelling markdown and plan inclusion markdown as defined in equation (21). Magnitudes are in basis points.

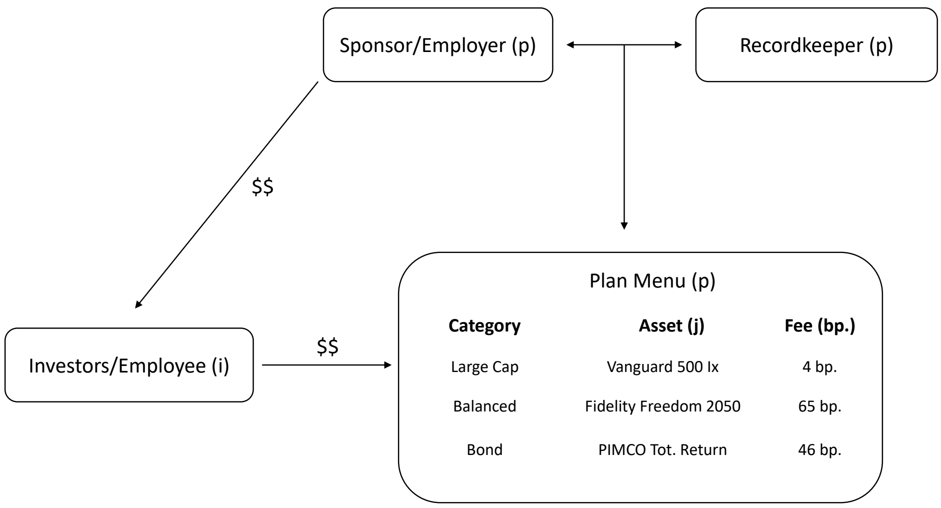

Figure 18: Administrative structure of a defined-contribution employer-sponsored retirement plan.

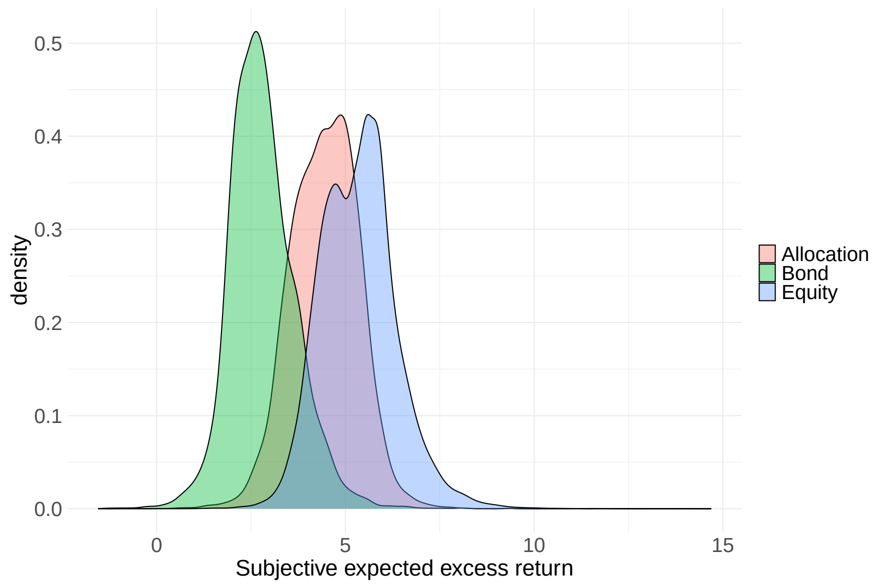

Figure 19: Estimated distribution of the average (across plans) subjective expected returns by asset class. Plan p subjective expected return for fund j in year t is estimated as $\mu_{jpt} = X'_{(1)}\hat{\boldsymbol{\beta}} + \hat{\xi}_{jpt}$. Magnitudes are in percentage points.

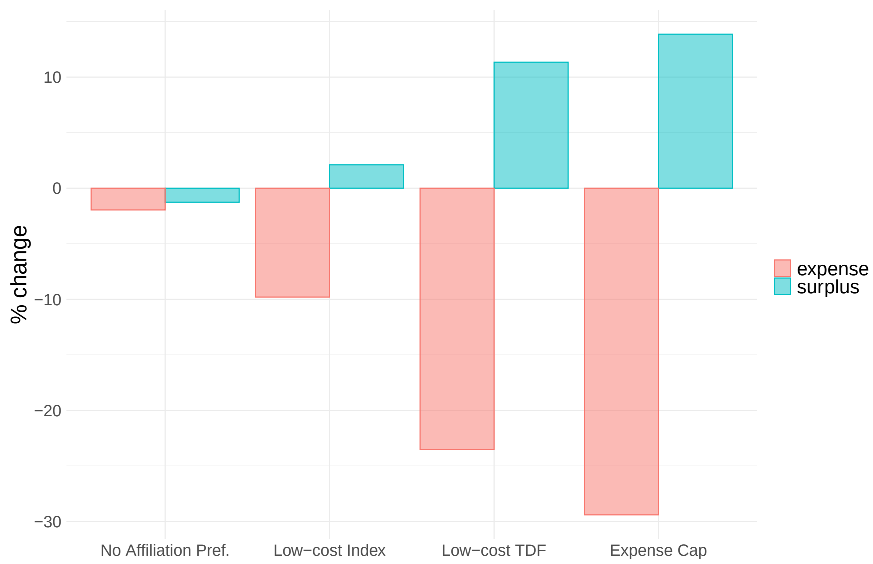

Figure 20: Percentage change in investor surplus and average plan expense for different plan design policies relative to the status quo. Magnitudes are in percentage points.

B Tables

	Plan performance	Plan performance	Plan performance	Plan performance
Plan expenses	-0.76	-0.38	-0.44	-0.06
	(0.01)	(0.01)	(0.01)	(0.01)
Weighted	Y	Y	N	N
Year FE	N	Y	N	Y
R2	0.02	0.19	0.01	0.10

Table 1: Fund performance is the difference between fund return and the average category return. Plan performance is the average performance (possibly asset weighted) of all funds in the plan. Returns are gross of fees. Returns and fees are in percentage points.

	Exp. Ratio	Exp. Ratio	Exp. Ratio
log(# opt. in cat)	-0.047	-0.077	-0.051
	(0.001)	(0.001)	(0.002)
R2	0.07	0.27	0.62
Year FE	Y	Y	Y
Sponsor FE	N	Y	Y
Fund brand FE	N	N	Y

Table 2: Dependent variable is funds' expense ratio. Independent variable is the number of funds within an investment category. Expense ratios are in percentage points.

	Sponsors preference parameters		
	Mean	S.D.	Marginal Effect (pp.)
Expense Ratio (bp.)	-0.015	0.001	-0.012
	(0.002)	(0.000)	
Affiliated (dummy)	0.439	-	0.469
	(0.044)		
Target (dummy)	0.421	-	0.450
	(0.088)		
Gross returns (pp.)	0.001	-	0.001
	(0.001)		
Average fee elasticity		-1.33	
Median fee elasticity		-1.23	
q (Calibrated)		0.70	
GMM objective (df)		6.69 (1)	

Table 5: Two-step GMM estimates of plan sponsor preferences. Robust standard errors are reported in parentheses. Year, category, passive and fund brand fixed effects are included. For the marginal effects, inclusion probabilities are in percentage points.

	Projected R2	R2
factors	0.037	0.038
categories	0.144	0.145
factors | categories	0.001	0.146
categories | provider FE	0.128	0.171
categories | provider FE + sponsor FE	0.116	0.223

Table 6: Dependent variable is plan-level portfolio allocations. All specification include year fixed effects. Factors are 3 Fama-French plus Momentum and 3 bond factors.

Plan investors preference parameters					
	OLS			IV	Marginal effect (pp.)
Expense Ratio ($\tilde{\gamma}$)	-0.006	-0.012	-0.014	-0.052	-0.089
	(0.000)	(0.000)	(0.000)	(0.003)	
Affiliated ($\tilde{\beta}_1$)	0.002	0.005	0.016	0.014	0.022
	(0.000)	(0.000)	(0.000)	(0.001)	
Gross returns ($\tilde{\beta}_2$)	0.007	0.015	0.017	0.021	0.036
	(0.001)	(0.001)	(0.001)	(0.001)	
Fraction inactive (δ)	0.222	0.246	0.299	0.254	-
	(0.004)	(0.004)	(0.006)	(0.007)	
Median fee elasticity	-0.256	-0.493	-0.573	-2.268	-
Median fee elasticity (active investors)	-0.329	-0.653	-0.817	-3.042	-
Fstat	-	-	-	84.20	-
R2	0.23	0.24	0.39	0.82	-
Fund brand FE	N	Y	Y	Y	-
Sponsor FE	N	N	Y	Y	-

Table 7: Estimates of plan investors preferences. All specifications include year, category and passive fixed effects. Expense ratios and gross returns are in percentage points. R2 for IV column is first stage. Marginal effect are for portfolio allocations in percentage points for a basis point increase in expenses or gross returns.

	All Funds			Passive Funds			Target Date Funds		
	Fee	MC	PCM	Fee	MC	PCM	Fee	MC	PCM
p25	43	27	8	9	2	4	13	2	8
p50	74	58	14	21	12	7	32	19	14
p75	101	87	19	43	34	11	52	39	15
Mean	73	59	14	29	21	8	37	26	11

Table 8: Price cost margins and marginal costs implied by the Nash-Bertrand first order conditions. Magnitudes are in basis points.

Equilibrium decomposition of observed fees			
Fee	Monopolist fee	Hotelling markdown	Plan inclusion markdown
65.75	119.44	25.15	28.54

Table 9: Decomposition of fees following equation (21). All magnitudes are in basis points. The figures shown are averages across time and funds.

Variable	N	Mean	SD	p5	p25	p50	p75	p95
Total assets (mln.)	60798	27.848	59.148	1.781	5.553	11.155	25.107	107.637
N. of parcipants	60798	475.193	597.083	118	160.85	250.102	494.828	1723.868
N. of options	60798	24.725	14.461	12.959	17.877	21.55	26.911	47.857
N. of categories	60798	14.958	3.268	9.552	13	15	17.009	20
Share of affiliated (%)	60798	25.127	27.559	0	0.64	14.327	43.243	82.328
Includes Target (%)	60798	78.355	37.629	0	73.171	100	100	100
Includes Passive (%)	60798	92.175	23.674	17.346	100	100	100	100
Avg. expense (pp.)	60798	0.626	0.214	0.261	0.492	0.626	0.76	0.982
Avg. expense (weighted)	60798	0.509	0.243	0.106	0.346	0.506	0.661	0.932
Assets per participant ($)	60742	50501	60798	4838	17498	35054	64670	143731
N. of years	60798	5.332	2.901	1	3	5	8	10
N. of recordkeepers	60798	1.197	0.474	1	1	1	1	2

Table 3: Plan level summary statistics for the years 2010 to 2019. Each variable is first averaged within plan across years and then tabulated across plans. The variable 'N' is the number of plans and the variable 'N. of years' is the number of years a plan is observed in the sample. Sample is for sponsors with number of participants between 100 and 5000.

Variable	N	Mean	SD	p5	p25	p50	p75	p95
Total Assets (mln.)	5615	190.736	1062.082	0.034	0.77	6.746	46.997	683.457
Portfolio share (avg.)	5615	2.896	3.821	0.119	0.887	1.795	3.279	10.039
Portfolio share (sd.)	5009	3.016	3.201	0.215	1.151	2.089	3.559	10.284
Fund-Plan turnover	5615	48.265	19.1	20.073	34.777	46.875	61.334	82.344
N. of share classes	5615	2.644	2.027	1	1	2	4	7

Table 4: Fund level summary statistics for the years 2010 to 2019. Each variable is first averaged (or summed in the case of 'total assets') within fund-year across plans, then within plan across years and tabulated across funds. The variable 'N' is the number of funds, excluding cash accounts and company stocks. Portfolio share (sd.) is the within fund-year standard deviation of the fund portfolio share across plans, which is then averaged within fund across years.

	Investor Surplus (bp.)	Average plan expense (bp.)	Fee savings/year ($)	Fee savings/40 years ($)
Status Quo	238	51	-	-
No Affiliation Preference	235	50	-	-
Low-cost Index Fund	243	46	25	11,897
Low-cost TDF	265	39	60	28,832
Expense cap (50 bp.)	271	36	75	36,192

Table 10: Investor surplus and average plan expense under different counterfactual policies. Magnitudes are in basis points. Savings are relative to the status quo. Fee savings per-year assumes a retirement account balance of $50,000. Fee savings over 40 years assumes an annual income of $70,000, contribution rate of 10% and an annual return of 6%. Expense ratio cap is at 60 basis points.

	Sponsors preference parameters			
	Homogeneous preferences		Heterogeneous q	
	Mean	Marg. Effect (pp.)	Mean	Marg. Effect (pp.)
Expense Ratio (bp.)	-0.017	-0.013	-0.021	-0.015
	(0.002)	-	(0.002)	-
Affiliated (dummy)	0.465	0.472	0.515	0.488
	(0.041)	-	(0.040)	-
Target (dummy)	0.402	0.428	0.378	0.416
	(0.082)	-	(0.088)	-
Gross returns (pp.)	0.001	0.001	0.004	0.002
	(0.001)	-	(0.001)	-
Average fee elasticity	-1.36		-1.74	
Median fee elasticity	-1.28		-1.62	
q (Calibrated)	0.70		0.66	
GMM objective (df)	5.76 (2)		5.59 (2)	

Table 11: Two-step GMM estimates of plan sponsor preferences. Robust standard errors are reported in parentheses. Year, category, passive and fund brand fixed effects are included. For the marginal effects, inclusion probabilities are in percentage points.

	Sponsors preference parameters			
	Small plans		Large plans	
	Mean	Marg. Effect (pp.)	Mean	Marg. Effect (pp.)
Expense Ratio (bp.)	-0.012	-0.009	-0.023	-0.016
	(0.003)	-	(0.003)	-
Affiliated (dummy)	0.660	0.524	0.345	0.381
	(0.052)	-	(0.048)	-
Target (dummy)	0.265	0.363	0.521	0.492
	(0.101)	-	(0.104)	-
Gross returns (pp.)	0.001	0.000	0.001	0.001
	(0.002)	-	(0.001)	-
Average fee elasticity	-1.20		-1.67	
Median fee elasticity	-1.11		-1.53	
q (Calibrated)	0.70		0.70	
GMM objective (df)	4.92 (2)		4.12 (2)	

Table 12: Two-step GMM estimates of plan sponsor preferences for plans with number of participants below the median (small) and above the median (large). Robust standard errors are reported in parentheses. Year, category, passive and fund brand fixed effects are included. For the marginal effects, inclusion probabilities are in percentage points.

	Sponsors preference parameters			
	Before 2014		After 2014	
	Mean	Marg. Effect (pp.)	Mean	Marg. Effect (pp.)
Expense Ratio (bp.)	-0.010	-0.007	-0.028	-0.019
	(0.003)	-	(0.004)	-
Affiliated (dummy)	0.455	0.461	0.483	0.482
	(0.060)	-	(0.057)	-
Target (dummy)	0.298	0.370	0.544	0.502
	(0.120)	-	(0.113)	-
Gross returns (pp.)	0.001	0.001	0.001	0.001
	(0.001)	-	(0.001)	-
Average fee elasticity	-1.14		-1.86	
Median fee elasticity	-1.02		-1.70	
q (Calibrated)	0.70		0.70	
GMM objective (df)	3.88 (2)		5.05 (2)	

Table 13: Two-step GMM estimates of plan sponsor preferences for the pre 2014 and post 2014 subsamples. Robust standard errors are reported in parentheses. Year, category, passive and fund brand fixed effects are included. For the marginal effects, inclusion probabilities are in percentage points.

	Sponsors preference parameters		
	Mean	S.D.	Marginal Effect (pp.)
Expense Ratio (bp.)	-0.014	0.001	-0.012
	(0.002)	(0.000)	
Affiliated (dummy)	0.444	-	0.474
	(0.042)		
Target (dummy)	0.425	-	0.451
	(0.088)		
Gross returns (pp.)	0.001	-	0.001
	(0.001)		
Average fee elasticity		-1.33	
Median fee elasticity		-1.23	
q (Calibrated)		0.70	
GMM objective (df)		6.70 (1)	

Table 14: Two-step GMM estimates of plan sponsor preferences. Robust standard errors are reported in parentheses. Year, category, passive ,fund brand and record-keeper fixed effects are included. For the marginal effects, inclusion probabilities are in percentage points.

C Model derivations

In this Appendix I provide more formal derivations of the results introduced in the main text.

Derivation of ranking probabilities. Consider the simple example in the main text where we have four options $\{j, k, l, m\}$ and we want to compute the probability that option j is ranked 2nd in terms of sponsors' indirect utilities. For simplicity I drop sponsor subscript p. The probability that j is ranked 2nd equals to the sum of all possible utility rankings in which u_j is the second highest:

$$\begin{aligned}\phi_j^2 &= \Pr\{u_k > u_j > u_l > u_m\} + \Pr\{u_k > u_j > u_m > u_l\} \\ &+ \Pr\{u_l > u_j > u_k > u_m\} + \Pr\{u_l > u_j > u_m > u_k\} \\ &+ \Pr\{u_m > u_j > u_k > u_l\} + \Pr\{u_m > u_j > u_l > u_k\}.\end{aligned} \tag{32}$$

Next, consider any of the six rankings above, say the first one and note that

$$\Pr\{u_k > u_j > u_l > u_m\} = \frac{\exp(V_k)}{\sum_{s\in\{j,k,l,m\}}\exp(V_s)} \cdot \frac{\exp(V_j)}{\sum_{s'\in\{j,l,m\}}\exp(V_{s'})} \cdot \frac{\exp(V_l)}{\sum_{s''\in\{l,m\}}\exp(V_{s''})} \tag{33}$$

Expression (33) can be derived analytically by integrating over the T1EV extreme value shocks and is known as ranked-ordered-logit (ROL),[40] which can be interpreted as a sequential multinomial logit decision problem.

Expression (33) applies analogously to all six terms in (32) and implies that ϕ_j^2 only depends on how the first two choices are ranked but not on the order of the 3rd and 4th choices. To see this consider the sum of the first two terms on the RHS of (32) and note that from (33) we can factor out the first two factors of each addend and that the sum of the last factors equals. Overall we obtain

$$\begin{aligned}&\Pr\{u_k > u_j > u_l > u_m\} + \Pr\{u_k > u_j > u_m > u_l\} = \\ &= \frac{\exp(V_k)}{\sum_{s\in\{j,k,l,m\}}\exp(V_s)} \cdot \frac{\exp(V_j)}{\sum_{s'\in\{j,l,m\}}\exp(V_{s'})}.\end{aligned}$$

Applying the same steps for all three lines in (32), the three term expression in the main text obtains.

Derivation of unconditional plan inclusion probability. Letting λ_{gp} the probability that p chooses to offer category g, $q(1-q)^{n-1}$ the probability that p includes n funds from category g and $\phi_j^{1:n}$ the probability that j is chosen conditional on n

[40] For more details see Beggs, Cardell and Hausman (1981).

options and g being offered, the unconditional probability that j ends up being in p's retirement plan can be rearranged as follows

$$
\begin{aligned}
\phi_{jp} &= \lambda_{gp} \cdot \left(\sum_{n=1}^{\infty} q(1-q)^{n-1} \phi_{jp}^{1:n} \right) \\
&= \lambda_{gp} \cdot \left(\sum_{n=1}^{\infty} q(1-q)^{n-1} \sum_{z=1}^{n} \phi_{jp}^{z} \right) \\
&= \lambda_{gp} \cdot \sum_{z=1}^{\infty} \sum_{n=z}^{\infty} q(1-q)^{n-1} \phi_{jp}^{z} \\
&= \lambda_{gp} \cdot \sum_{z=1}^{\infty} \phi_{jp}^{z} (1-q)^{z-1} \sum_{n=z}^{\infty} q(1-q)^{n-z} \\
&= \lambda_{gp} \cdot \sum_{z=1}^{\infty} \phi_{jp}^{z} (1-q)^{z-1}
\end{aligned}
$$

where the latter coincides with expression (7) provided in the main text.

Heterogeneous individual investors. Let us consider the case in which individual investors have heterogeneous preferences. Formally, let A_i, δ_i, $\boldsymbol{\beta}_i$ and γ_i be individual specific. Moreover, assume that there is a subset $D \subset J_p$ of default funds and denote by d_i investor i's default fund. Under these assumptions i's demand system is given by

$$
\boldsymbol{a}_i(\boldsymbol{f}) = \begin{cases} \boldsymbol{e}_{d_i} & \text{if } i \text{ defaults} \\ \frac{1}{\gamma_i}(I + X_{(2)}X_2')^{-1}(\boldsymbol{\mu}_i - \boldsymbol{f}) & \text{o.w} \end{cases} \tag{34}
$$

where to save on notation I defined $\boldsymbol{\mu}_i \equiv X_{(1)}\boldsymbol{\beta}_i + \boldsymbol{\xi}_i$. To obtain the aggregate demand system let us define the following weighted averages

$$
\begin{aligned}
\boldsymbol{\mu}_p &\equiv \sum_{i \in I_p} \frac{(1-\delta_i)\gamma_i^{-1}}{\sum_{i \in I_p}(1-\delta_i)\gamma_i^{-1}} \boldsymbol{\mu}_i \\
\gamma_p &\equiv \left(\sum_{i \in I_p} \frac{A_i}{\bar{A}} \frac{1-\delta_i}{\gamma_i} \right)^{-1} \\
\delta_{dp} &\equiv \left(\frac{\sum_{i \in I_{dp}} A_i}{\bar{A}} \right) \sum_{i \in I_{dp}} \frac{A_i}{\sum_{i \in I_{dp}} A_i} \delta_i
\end{aligned}
$$

where I_p is the set of plan p investors and I_{dp} is the set of investors that have d as default plan. Then taking the horizontal sum of the $\boldsymbol{a}_i$ it is easy to check that the plan level demand system is given by

$$
\boldsymbol{s}_p(\boldsymbol{f}; \eta_p) = \sum_{d \in D} \delta_{dp} \mathbf{e}_d + \frac{1}{\gamma_p}(I + X_{(2)}X_{(2)}')^{-1}(\boldsymbol{\mu}_p - \boldsymbol{f}).
$$

Overall, the estimated parameters from the aggregate demand system $\eta_p = (\delta_{dp}, \gamma_p, \boldsymbol{\mu}_p)$

are weighted averages of the underlying individual investors parameters.

Derivation of aggregate demand system in (13). To show that equations (11) and (13) are equivalent it is enough to note that

$$
\begin{aligned}
&(I - X_{(2)}(I + X'_{(2)}X_{(2)})^{-1}X'_{(2)})(I + X_{(2)}X'_{(2)}) = && (35)\\
&= I - X_{(2)}(I + X'_{(2)}X_{(2)})^{-1}X'_{(2)} + X_{(2)}X'_{(2)} - X_{(2)}(I + X'_{(2)}X_{(2)})^{-1}X'_{(2)}X_{(2)}X'_{(2)} && (36)\\
&= I + X_{(2)}(I - (I + X'_{(2)}X_{(2)})^{-1} - (I + X'_{(2)}X_{(2)})^{-1}X'_{(2)}X_{(2)})X'_{(2)} && (37)\\
&= I + X_{(2)}(I + X'_{(2)}X_{(2)})^{-1}(I + X'_{(2)}X_{(2)} - I - X'_{(2)}X_{(2)})X'_{(2)} = I && (38)
\end{aligned}
$$

Next, I show that $\kappa_{jj} \in (0,1)$ and $\kappa_{jl} \in (-1,1)$ for all j and l. To see this, consider note that by construction $\mathcal{K}_x$ and $I - \mathcal{K}_x$ are positive definite matrices. The let e_j be the jth unit vector and note the definition of positive definite matrix implies that

$$\kappa_{jj} = e'_j\mathcal{K}_x e_j > 0 \quad \text{and} \quad 1 - \kappa_{jj} = e'_j(I - \mathcal{K}_x)e_j > 0. \tag{39}$$

Next, take any (j,l) pair with $j \neq l$ and note that

$$\kappa_{jj} + \kappa_{ll} - 2\kappa_{jl} = (e_j - e_l)'\mathcal{K}_x(e_j - e_l) > 0 \tag{40}$$

where the first equality exploits the fact that $\mathcal{K}_x$ is symmetric. From (40) and the fact that $\kappa_{jj} < 1$ for all j, we can conclude that $\kappa_{jl} < 1$. To show that $\kappa_{jl} > -1$ it is enough to repeat the previous argument using $\kappa_{jj} + \kappa_{ll} + 2\kappa_{jl}$.

Derivation of expected κ_{jl} under biased beliefs about q. Suppose that fund j believes that sponsors will include at most one fund per investment category. This means that funds evaluate inclusion probabilities assuming that $q = 1$ and will assign positive probabilities only to menus S that include at most one fund per category.

Denoting by G_S the set of categories included in menu S, by j_g a generic option from category g and by g_j the investment category fund j belongs to, the probability that such menu $S \in \mathcal{S}_{jp}$ is chosen by sponsor p can be factored us

$$\phi_p(S) = \phi_{jp} \prod_{g \in G_S/g_j} \phi_{j_g p} \prod_{j_{g'} p}(1 - \lambda_{j_{g'}p}) \tag{41}$$

where the factorization is a consequence of the fact that inclusion decision are made independently across investment categories.

Next, consider fund j and fund l and note that

$$\bar{\kappa}_{jl} = \sum_{S \in \mathcal{S}_{jp}} \frac{\phi_p(S)}{\phi_{jp}} \kappa_{jl}^S \tag{42}$$

$$= \sum_{S \in \mathcal{S}_{jp}} \left(\prod_{g \in G_S/g_j} \phi_{j_g p} \prod_{j_{g'p}} (1 - \lambda_{j_{g'}p}) \right) \kappa_{jl}^S \tag{43}$$

which does not depend on f_j because inclusion probabilities of competitors funds belonging to different categories do not depend on f_j.

Derivation of equilibrium fees decomposition. Recall the system of Bertrand FOCs' derived in the main text

$$\bar{\delta}\boldsymbol{e}_d + (I - \tilde{\mathcal{K}})(\tilde{\boldsymbol{\mu}} - \boldsymbol{f}) - \boldsymbol{\iota} - (I - diag(\tilde{\mathcal{K}}))(\boldsymbol{f} - \boldsymbol{c}) = 0 \tag{44}$$

which can be rearranged as

$$2\left(I - \frac{diag(\tilde{\mathcal{K}})}{2} - \frac{\tilde{\mathcal{K}}}{2}\right)\boldsymbol{f} = (I - \tilde{\mathcal{K}})\tilde{\boldsymbol{\mu}} + (I - diag(\tilde{\mathcal{K}}))\boldsymbol{c} - \boldsymbol{\iota}. \tag{45}$$

Next define

$$\tilde{\boldsymbol{\iota}} \equiv \left(I - \frac{diag(\tilde{\mathcal{K}})}{2} - \frac{\tilde{\mathcal{K}}}{2}\right)^{-1} \frac{\boldsymbol{\iota}}{2} \tag{46}$$

and rewrite the system of FOCs as

$$\begin{aligned}
\boldsymbol{f} &= \frac{1}{2}\left(I - \frac{diag(\tilde{\mathcal{K}})}{2} - \frac{\tilde{\mathcal{K}}}{2}\right)^{-1} \left[(I - \tilde{\mathcal{K}})\tilde{\boldsymbol{\mu}} + (I - diag(\tilde{\mathcal{K}}))\boldsymbol{c}\right] - \tilde{\boldsymbol{\iota}} \\
&= \boldsymbol{c} + \frac{1}{2}\left(I - \frac{diag(\tilde{\mathcal{K}})}{2} - \frac{\tilde{\mathcal{K}}}{2}\right)^{-1} (I - \tilde{\mathcal{K}})(\tilde{\boldsymbol{\mu}} - \boldsymbol{c}) - \tilde{\boldsymbol{\iota}} \\
&= \boldsymbol{c} + \frac{1}{2}\left(I - diag(\tilde{\mathcal{K}}) - \frac{\tilde{\mathcal{K}} - diag(\tilde{\mathcal{K}})}{2}\right)^{-1} (I - \tilde{\mathcal{K}})(\tilde{\boldsymbol{\mu}} - \boldsymbol{c}) - \tilde{\boldsymbol{\iota}} \\
&= \frac{\tilde{\boldsymbol{\mu}} + \boldsymbol{c}}{2} - \left(I - diag(\tilde{\mathcal{K}}) - \frac{\tilde{\mathcal{K}} - diag(\tilde{\mathcal{K}})}{2}\right)^{-1} \frac{\tilde{\mathcal{K}} - diag(\tilde{\mathcal{K}})}{2} \frac{\tilde{\boldsymbol{\mu}} - \boldsymbol{c}}{2} - \tilde{\boldsymbol{\iota}} \\
&= \frac{\tilde{\boldsymbol{\mu}} + \boldsymbol{c}}{2} - (I - diag(\tilde{\mathcal{K}}))^{-1/2} \left(I - \frac{G(\tilde{\mathcal{K}})}{2}\right)^{-1} \frac{G(\tilde{\mathcal{K}})}{2} (I - diag(\tilde{\mathcal{K}}))^{1/2} \frac{\tilde{\boldsymbol{\mu}} - \boldsymbol{c}}{2} - \tilde{\boldsymbol{\iota}}
\end{aligned}$$

where

$$G(\tilde{\mathcal{K}}) \equiv (I - diag(\tilde{\mathcal{K}}))^{-1/2}(\tilde{\mathcal{K}} - diag(\tilde{\mathcal{K}}))(I - diag(\tilde{\mathcal{K}}))^{-1/2}. \tag{47}$$

Expression (22) in the main text obtains by setting $diag(\tilde{\mathcal{K}}) = k_0 I$. For the case in which there is a default fund the same steps apply after redefining $\tilde{\boldsymbol{\mu}} = \tilde{\boldsymbol{\mu}} + \bar{\delta}(I - \tilde{\mathcal{K}})^{-1}\boldsymbol{e}_d$.

Estimation algorithm. Before starting the estimation, I draw a vector $\boldsymbol{\nu}_s$ of random taste parameters for $s = 1, ..., S$ simulated sponsors from a normal $N(\mathbf{0}, I)$ and store it. Then the algorithm proceeds as follows.

Step 0: Guess Γ_θ

Step 1: For a given guess of the vector of the mean utility mean utility $\bar{\boldsymbol{v}}^{(k)}$. Compute the following variables for each fund j, market t

1.1 for each simulated sponsor s calculate the following objects

- the probability that j's category g is included by sponsor s

$$\lambda_{gst}(\Gamma_\theta, \bar{\boldsymbol{v}}_t^{(k)}) = \frac{\sum_{l \in g} \exp(\bar{v}_{lt}^{(k)} + \boldsymbol{w}'_{lt}\Gamma_\theta \boldsymbol{\nu}_s)}{1 + \sum_{l \in g} \exp(\bar{v}_{lt}^{(k)} + \boldsymbol{w}'_{lt}\Gamma_\theta \boldsymbol{\nu}_s)}$$

- for $n = \{1, 2, 3\}$ calculate the ranking probabilities

$$\phi_{jst}^n(\Gamma_\theta, \bar{\boldsymbol{v}}_t^{(k)}) = \sum_{(j_1, \ldots, j_{n-1}) \in g/\{j\}} \prod_{n'=1}^{n-1} \frac{\exp(\bar{v}_{j_{n'}t}^{(k)} + \boldsymbol{w}'_{j_{n'}t}\Gamma_\theta \boldsymbol{\nu}_s)}{\sum_{n''=n'}^{N_{gp}} \exp(\bar{v}_{j_{n''}t}^{(k)} + \boldsymbol{w}'_{j_{n''}t}\Gamma_\theta \boldsymbol{\nu}_s)}$$
$$\cdot \frac{\exp(\bar{v}_{jt}^{(k)} + \boldsymbol{w}'_{jt}\Gamma_\theta \boldsymbol{\nu}_s)}{\sum_{n'''=n}^{N_{gp}} \exp(\bar{v}_{j_{n'''}t}^{(k)} + \boldsymbol{w}'_{j_{n'''}t}\Gamma_\theta \boldsymbol{\nu}_s)}$$

- compute the probability that sponsor s includes fund j

$$\phi_{jst}^{1:3}(\Gamma_\theta, \bar{\boldsymbol{v}}_t^{(k)}) = \phi_{jst}^1(\Gamma_\theta, \bar{\boldsymbol{v}}_t^{(k)}) + (1 - q_t)\phi_{jst}^2(\Gamma_\theta, \bar{\boldsymbol{v}}_t^{(k)}) + (1 - q_t)^2\phi_{jst}^3(\Gamma_\theta, \bar{\boldsymbol{v}}_t^{(k)})$$

where q_t is calibrated to match the empirical distribution of the number of options within investment category in market t. The typical values for q_t are of 0.7 or higher so that $(1-q_t)^{n-1}$ decays quite fast in n. In simulations, I find that stopping at $n = 3$ works well in recovering the true parameters. Moreover the computational burden of calculating ϕ_{jst}^n for $n \geq 4$ is non negligible.

1.2 approximate the RHS of (23) with

$$\phi_{jt}^S(\Gamma_\theta, \bar{\boldsymbol{v}}_t^{(k)}) = \frac{1}{S} \sum_{s=1}^{S} \lambda_{gst}(\Gamma_\theta, \bar{\boldsymbol{v}}_t^{(k)}) \cdot \phi_{jst}^{1:3}(\Gamma_\theta, \bar{\boldsymbol{v}}_t^{(k)})$$

Step 2: For each t update the mean utility vector by computing $\bar{\boldsymbol{v}}_t^{(k+1)}$ as

$$\bar{\boldsymbol{v}}_t^{(k+1)} = \bar{\boldsymbol{v}}_t^{(k)} + \log(\hat{\boldsymbol{\phi}}_t(\Gamma_\theta, \bar{\boldsymbol{v}}_t^{(k)})) - \log(\boldsymbol{\phi}_t^S(\Gamma_\theta, \bar{\boldsymbol{v}}_t^{(k)})) \tag{48}$$

Step 3: For each t Repeat Step 1 and Step 2 until

$$||\bar{\boldsymbol{v}}_t^{(k+1)} - \bar{\boldsymbol{v}}_t^{(k)}||_\infty < \epsilon \tag{49}$$

for some tolerance level ϵ.

Step 4: Recover sponsors' preference parameters $\boldsymbol{\mu}_\theta$ that enter sponsors' utility linearly

$$\boldsymbol{\mu}_\theta(\Gamma_\theta) = (W'Z(Z'Z)^{-1}Z'W)^{-1}W'Z(Z'Z)^{-1}Z'\boldsymbol{v}(\Gamma_\theta) \tag{50}$$

where W is a matrix of funds' characteristics with number of rows equal to the total number of observations $\bar{N} = \sum_t N_t$ and number of columns equal to the number of characteristics and Z is including both excluded and included instruments.

Step 5: Recover demand residuals

$$\boldsymbol{\zeta}(\Gamma_\theta) = \boldsymbol{v}(\Gamma_\theta) - W\boldsymbol{\mu}_\theta \tag{51}$$

and compute the GMM norm

$$\boldsymbol{\zeta}(\Gamma_\theta)'Z\Omega(\Gamma_\theta)Z'\boldsymbol{\zeta}(\Gamma_\theta) \tag{52}$$

where $\Omega(\Gamma_\theta) = (Z'Z)^{-1}$ in the first GMM estimation step and then is updated to $\Omega(\Gamma_\theta) = Z'diag(\boldsymbol{\zeta}(\Gamma_\theta)^2)Z = \sum_{j,t} \zeta_{jt}^2(\Gamma_\theta)\boldsymbol{Z}_{jt}\boldsymbol{Z}_{jt}'$.

Interior equilibrium existence and uniqueness. Consider first the case in which funds know with certainty which plan menu will include them. Formally, this means that $\phi_{jp} = 1$ when p includes fund j and zero otherwise which further implies that the $\boldsymbol{\iota}$ in equation (22) equals zero. Given this, the system of funds' best replies becomes linear in $\boldsymbol{f}$:

$$(I - \tilde{\mathcal{K}})\boldsymbol{f} + (I - diag(\tilde{\mathcal{K}}))\boldsymbol{f} = (I - \tilde{\mathcal{K}})\tilde{\boldsymbol{\mu}} + (I - diag(\tilde{\mathcal{K}}))\boldsymbol{c} \tag{53}$$

because inclusion probabilities do not depend on fees anymore and with that also $\mathcal{K}$, $\tilde{\boldsymbol{\mu}}$ do not depend on fees. A well-defined solution is then guaranteed to exist as long as $(I - \tilde{\mathcal{K}})$ is invertible. Moreover, linearity would also imply that such solution is unique.

What we do not know is whether this solution is such that equilibrium fees are non-negative. In what follows I show that the following dominance-diagonal condition

$$(1 - \tilde{\kappa}_{jj})(\tilde{\mu}_j - c_j) > \sum_{k \neq j} |\tilde{\kappa}_{jk}|(\tilde{\mu}_k - c_k) \quad \text{all } j \tag{54}$$

implies that the system of best replies is a self-map over the interior of the set $\times_{j\in\{1,\dots,J\}}[c_j, \tilde{\mu}_j]$ which ensures that equilibrium fees are positive and above marginal costs. I start by defining the following linear operator $T : \mathbb{R}^J \to \mathbb{R}^J$ whose j-th component is fund j's best reply

$$T_j(\boldsymbol{f}) \equiv \frac{1}{2}\left[\tilde{\mu}_j - \sum_{k\neq j}\frac{\tilde{\kappa}_{jk}}{1-\tilde{\kappa}_{jj}}\mu_k + c_j + \sum_{k\neq j}\frac{\tilde{\kappa}_{jk}}{1-\tilde{\kappa}_{jj}}f_k\right]. \tag{55}$$

Assumption (54) implies that T is a self-map in the interior of $\times_{j\in\{1,\dots,J\}}[c_j, \tilde{\mu}_j]$. To see this take any $\boldsymbol{f} \in \times_{j\in\{1,\dots,J\}}[c_j, \tilde{\mu}_j]$ and note that

$$c_j < T_j(\boldsymbol{f}) < \tilde{\mu}_j \tag{56}$$

$$\Leftrightarrow \left|\sum_{k\neq j}\frac{\tilde{\kappa}_{jk}}{1-\tilde{\kappa}_{jj}}\frac{\tilde{\mu}_k - c_k}{\tilde{\mu}_j - c_j}\frac{\tilde{\mu}_k - f_k}{\tilde{\mu}_k - c_k}\right| < 1 \tag{57}$$

$$\Leftrightarrow \sum_{k\neq j}\frac{|\kappa_{jk}|}{1-\tilde{\kappa}_{jj}}\frac{\tilde{\mu}_k - c_k}{\tilde{\mu}_j - c_j}\frac{\tilde{\mu}_k - f_k}{\tilde{\mu}_k - c_k} < 1 \tag{58}$$

is always satisfied when (54) holds and $\boldsymbol{f} \in \times_{j\in\{1,\dots,J\}}[c_j, \tilde{\mu}_j]$. Because T maps a closed and bounded set into itself, Bower fixed point theorem implies that there exists an $\boldsymbol{f}^* \in \times_{j\in\{1,\dots,J_p\}}[c_j, \mu_j]$ such that $T(\boldsymbol{f}^*) = \boldsymbol{f}^*$. Moreover, from (58) we know that such fixed point is interior and unique, because T is linear.

We can also show that in this interior equilibrium each fund manages a positive amount of asset. To see this, consider fund j's first order condition evaluated at the optimum

$$s_j(\boldsymbol{f}^*) - (1-\tilde{\kappa}_{jj})(f_j^* - c_j) = 0 \tag{59}$$

which implies that

$$s_j(\boldsymbol{f}^*) = (1-\tilde{\kappa}_{jj})(f_j^* - c_j) > 0 \tag{60}$$

where the latter inequality holds because we just proved that $f_j^* > c_j$ for all j and $(1-\tilde{\kappa}_{jj}) > 0$ follows from the fact that $I - \mathcal{K}$ is positive definite and that $\tilde{\kappa}_{jj}$ is just an average of the κ_{jj} across plans:

$$\tilde{\kappa}_{jj} = \bar{\phi}_j^{-1}\int \mathbf{1}\{j \in S_p\}(1-\delta_p)\gamma_p^{-1}A_p\kappa_{jj}^{S_p}dF_p$$

with

$$\bar{\phi}_j = \int \mathbf{1}\{j \in S_p\}(1-\delta_p)\gamma_p^{-1}A_p dF_p. \tag{61}$$

Lastly I show that the equilibrium is stable. To see this define the following variable

$$\hat{f}_j \equiv \frac{f_j - c_j}{\tilde{\mu}_j - c_j} \tag{62}$$

and note that fund j best reply can be rewritten as

$$\hat{f}_j = \frac{1}{2}\left[1 - \sum_{k \neq j} \frac{\tilde{\kappa}_{jk}}{1 - \tilde{\kappa}_{jj}} \frac{\tilde{\mu}_k - c_k}{\tilde{\mu}_j - c_j}\left(1 - \frac{f_k - c_k}{\tilde{\mu}_k - c_k}\right)\right] \tag{63}$$

Defining the linear mapping on the above RHS as $\hat{T} : R^J \rightarrow R^J$, it can be shown that this mapping is a self-map into $[0,1]^J$ under assumption (54). Additionally, this mapping is a contraction in the L_∞ norm.

$$\begin{aligned}
||\hat{T}(\hat{\boldsymbol{f}}_1) - \hat{T}(\hat{\boldsymbol{f}}_0)||_\infty &= \max_j |\hat{T}_j(\hat{\boldsymbol{f}}_1) - \hat{T}_j(\hat{\boldsymbol{f}}_0)| && (64)\\
&= \max_j \left| \sum_{k \neq j} \frac{\kappa_{jk}}{1 - \kappa_{jj}} \frac{\mu_k - c_k}{\tilde{\mu}_j - c_j} (\hat{f}_{1k} - \hat{f}_{0k}) \right| && (65)\\
&\leq \max_j \sum_{k \neq j} \frac{|\kappa_{jk}|}{1 - \kappa_{jj}} \frac{\mu_k - c_k}{\mu_j - c_j} |\hat{f}_{1k} - \hat{f}_{0k}| && (66)\\
&< \max_k |f_{1k} - f_{0k}| && (67)
\end{aligned}$$

which ensures that the unique equilibrium is also stable.

Next, I consider the case in which sponsor preferences are homogeneous $\theta_p \equiv \theta$ and funds do not know with certainty whether or not they will be included in sponsors' retirement menus (e.g., $\phi_j \in (0,1)$). For simplicity, I also assume that there is only one recordkeeper or equivalently that all recordkeerpers have the same network of funds. Under this assumptions, I will show that a Nash-Bertrand equilibrium exists when funds believe that sponsors include at most one fund per category and the previous dominance diagonal condition holds.

To start with, note that fund j problem in any given period simplifies to

$$\max_{f_j} \quad P \cdot (f_j - c_j) \cdot \phi_j(\boldsymbol{f}_j; \boldsymbol{\theta}) \int s_{jp}(\boldsymbol{f}; \boldsymbol{\eta}_p) A_p dF(\boldsymbol{\eta}_p, A_p) \tag{68}$$

where fund j expected portfolio share is given by

$$\begin{aligned} s_{jp}(\boldsymbol{f};\eta_p) &= \sum_{S\in\mathcal{S}_j} \tilde{\gamma}_p \frac{\phi(S;\boldsymbol{\theta})}{\phi_j(\boldsymbol{f};\boldsymbol{\theta})} \left[(1-\kappa^S_{jj})(\mu_{jp}-f_j) - \sum_{l\neq j, l\in S} \kappa^S_{jl}(\mu_{lp}-f_l)\right] \\ &= \tilde{\gamma}_p(1-\bar{\kappa}_{jj})(\mu_{jp}-f_j) - \tilde{\gamma}_p \sum_{l\neq j}\sum_{S\in\mathcal{S}_j} \kappa^S_{jl}\mathbf{1}\{l\in S\}\frac{\phi(S;\boldsymbol{\theta})}{\phi_j(\boldsymbol{\theta})}(\mu_{lp}-f_l) \\ &= \tilde{\gamma}_p(1-\bar{\kappa}_{jj})(\mu_{jp}-f_j) - \tilde{\gamma}_p \sum_{l\neq j}\phi_l \mathbb{E}[\bar{\kappa}^S_{jl}|j,l\in S](\mu_{lp}-f_l) \\ &= \tilde{\gamma}[I-\bar{\mathcal{K}}]'_j(\boldsymbol{\mu}_p-\boldsymbol{f}) \end{aligned}$$

with

$$\bar{\kappa}_{jl} \equiv \begin{cases} \mathbb{E}[\bar{\kappa}^S_{jj}|j\in S] & \text{if } j=l \\ \phi_l(\boldsymbol{f};\boldsymbol{\theta})\mathbb{E}[\bar{\kappa}^S_{jl}|j,l\in S] & \text{if } j\neq l. \end{cases}$$

Overall, fund j pricing problem can be written more compactly as

$$\max_{f_j} \quad P\cdot(f_j-c_j)\cdot\phi_j(\boldsymbol{f}_j;\boldsymbol{\theta})\cdot[I-\bar{\mathcal{K}}]'_j(\boldsymbol{\mu}-\boldsymbol{f})\cdot\tilde{\gamma}\bar{A} \tag{69}$$

where for simplicity I assumed that investors preferences are homogeneous too.[41]

In what follows, I first prove a lemma that provides conditions ensuring that funds' problem is concave and then show that funds' objective satisfies such conditions under the previous assumptions. Equilibrium existence will then follow directly from Kakutani fixed point theorem.

Lemma 1 *Let $\pi(f)$ be a continuous and strictly concave function $\pi''(f)<0$ that is uniquely maximized at f^* and is such that $\pi(f^*)>0$. Let $\phi(f)$ a continuous and decreasing function such that*

$$\phi(f)\in(0,1) \tag{70}$$

$$\phi'(f) = -\phi(f)(1-\phi(f)) < 0 \tag{71}$$

then the function $\phi(f)\pi(f)$ is concave on a compact set $[\underline{f},\overline{f}]$ with unique interior maximizer $f^{**}\leq f^*$.

Proof: Because π is continuous and $\pi(f^*)>0$, we can construct $[\underline{f},\overline{f}]$ such that $\pi(f)>0$ for all $f\in[\underline{f},\overline{f}]$ and $f^*\in[\underline{f},\overline{f}]$.

Next suppose there exists a $f^{**}\in(\underline{f},\overline{f})$ such that

$$\phi'(f^{**})\pi(f^{**})+\phi(f^{**})\pi'(f^{**})=0 \tag{72}$$

[41] This assumption is irrelevant for the existence result.

which can be rearranged as

$$-\frac{\phi'(f^{**})}{\phi(f^{**})} = \frac{\pi'(f^{**})}{\pi(f^{**})} \tag{73}$$

Taking the second order condition

$$\phi''(f^{**})\pi(f^{**}) + 2\phi'(f^{*})\pi'(f^{*}) + \phi(f^{**})\pi''(f^{**}) < 0 \tag{74}$$

The last term of the above is negative by assumption. The sum of the first two terms is also negative:

$$\phi''(f^{**})\pi(f^{**}) + 2\phi'(f^{*})\pi'(f^{*}) < 0 \tag{75}$$

$$\Leftrightarrow \frac{\phi''}{\phi'} + 2\frac{\pi'}{\pi} > 0 \tag{76}$$

$$\Leftrightarrow \frac{\phi''}{\phi'} - 2\frac{\phi'}{\phi} > 0 \tag{77}$$

$$\Leftrightarrow -(1-2\phi) + 2(1-\phi) = 1 > 0 \tag{78}$$

where the latter equivalence uses the fact that $\phi'' = -\phi'(1-2\phi)$. This shows that if an interior f^{**} that satisfies the necessary FOC exists then it is always a maximum which implies that $\phi(f)\pi(f)$ is concave on $[\underline{f}, \overline{f}]$.

Lastly, we can show that such p^{**} indeed exists and is interior. To see this evaluate the FOC at $\underline{f}$ and note that under the above assumptions the following

$$\phi'(\underline{f})\pi(\underline{f}) + \phi(\underline{f})\pi'(\underline{f}) > 0 \tag{79}$$

holds by choosing $\underline{f}$ sufficiently low (for instance choosing $\underline{f}$ = marginal cost such that $\pi(\underline{f}) = 0$) and by noting that $\pi'(\underline{f}) > 0$ because $\underline{f} < f^{*}$. Then evaluate the FOC at $\overline{f}$ and note that

$$\phi'(\overline{f})\pi(\overline{f}) + \phi(\overline{f})\pi'(\overline{f}) < 0 \tag{80}$$

which implies, by continuity that there exists an interior f^{**} that satisfies the FOC. Moreover, it must be the case that $f^{**} < f^{*}$.

Overall, conditions all conditions of the lemma hold and we can be assured that fund j objective in (69) is concave in f_j which ensures that funds' best replies $f_j(f_{-j})$ are proper functions. Thus, all conditions of Kakutani fixed point theorem are satisfied and a Nash equilbrium exists (see MWG chp. 8 Proposition 8.D.3).

Funds' maximization problem in (69) satisfies the conditions in the previous lemma after defining $\pi(f_j) \equiv (f_j - c_j) \cdot [I - \bar{\mathcal{K}}]_j'(\boldsymbol{\mu} - \boldsymbol{f})$. First note that, because funds' believe that sponsors include at most one fund per category we have that

$$\phi_j(f_j) = \lambda_g \phi_j^1 = \frac{\exp(V_j(\boldsymbol{\theta}))}{1 + \sum_l \exp(V_l(\boldsymbol{\theta}))}$$

which is decreasing in f_j and such that

$$\phi_j' = -\theta_f \phi_j (1 - \phi_j) < 0.$$

Next, note that because funds' believe that $q = 1$, the matrix $\mathcal{K}$ does not depend on f_j as I showed in previous derivations. This in turn implies that $\pi(f_j)$ is quadratic in f_j and thus concave f_j if and only if

$$(1 - \bar{\kappa}_{jj}) > 0.$$

The latter holds because $\bar{\kappa}_{jj} = \mathbb{E}[\kappa_{jj}^S | j, l \in S]$ and $\kappa_{jj}^S \in (0, 1)$ for any S as I showed in previous derivations. Because π is globally concave in f_j it admits a unique maximum, f_j^*. Moreover we have $f_j^* > c_j$ whenever the previous dominance diagonal holds:

$$(1 - \bar{\kappa}_j)(\mu_j - c_j) > \sum_{l \neq j} |\bar{\kappa}_{jl}| (\mu_l - c_l) \tag{81}$$

and $f_l \in [c_l, \mu_l]$ for all l. To see this note that the above condition implies that

$$\begin{aligned}
(1 - \bar{\kappa}_j)(\mu_j - c_j) &> \sum_{l \neq j} |\bar{\kappa}_{jl}| (\mu_l - c_l) \\
&\geq \sum_{l \neq j} |\bar{\kappa}_{jl}| (\mu_l - f_l) \\
&> \sum_{l \neq j} \bar{\kappa}_{jl} (\mu_l - f_l)
\end{aligned}$$

where the last two inequalities holds whenever $f_l \in [c_l, \mu_l]$. But note that the previous inequality corresponds to fund j FOC (when maximizing π) evaluated at $f_j = c_j$

$$\begin{aligned}
\frac{\partial \pi(f_j)}{\partial f_j} &= (1 - \bar{\kappa}_{jj})(\mu_j - f_j) - \sum_{l \neq j} \bar{\kappa}_{jl} (\mu_l - f_l) - (1 - \bar{\kappa}_{jj})(f_j - c_j) \Big|_{f_j = c_j} \\
&= (1 - \bar{\kappa}_j)(\mu_j - c_j) - \sum_{l \neq j} \bar{\kappa}_{jl} (\mu_l - f_l) > 0.
\end{aligned}$$

Then it must also be the case that $\pi(f_j^*) > 0$, if not setting $f_j = c_j$ would lead to higher π which would be a contradiction.

Microfoundation of the distribution of the number of options. In what follows I offer a simple microfoundation for the distribution of the number of options sponsors include in any given category building on the Stigler (1961) simultaneous search model.

Sponsor p first commits to include n investment options in investment category g and then conditional on n, selects the n options providing her with the n highest utilities. I assume that sponsors choose n before observing their random utility shocks ε_{jp} but knowing the mean utility of each option $V_j(\boldsymbol{\theta}_p)$. From this perspective the utility sponsor p derives from option j is distributed as

$$u_{jp} \sim T1EV(V_j). \tag{82}$$

I assume that sponsors incur a cost $c(n)$ for including n options which is increasing, convex in n and is such that $c(1) < \mathbb{E}[u_{j_1 p}]$. The benefit from choosing n options is given by

$$\sum_{s=1}^{n} \mathbb{E}[u_{j_s p}] \tag{83}$$

where j_s is the option that provides the sth highest utility. Overall, sponsor p problem becomes

$$\max_{n \geq 1} \sum_{s=1}^{n} \mathbb{E}[u_{j_s p}] - c(n). \tag{84}$$

The are two main differences between this model and the Stigler (1961) model. First, in this case the agent commits to consume n options whereas in Stigler (1961) the agent commits to sample n options and among those to consume the one with the highest utility. Second, in this model after committing to n, sponsor p observes the realized utility of all options available but has chosen to only consume the n highest whereas in the Stigler (1961) model the agent observes the utility realizations of the searched options only, and among those selects the highest.

The solution to the above problem is given by the n^* such that the marginal benefit from chosing to include $n^* + 1$ options is lower than the change in the cost

$$\mathbb{E}[u_{j_{n^*+1}}] \leq c(n^* + 1) - c(n^*). \tag{85}$$

Heterogeneity in the cost of adding options c_p or in the benefits u_{jp} across sponsors would produce a different n_p^* for each sponsors. In the data such distribution of n^* corresponds to the one plotted in Figure 11 suggesting that most sponsors do not include more than one option and that the likelihood decreases geometrically in the number of options.

In estimation I do not attempt to estimate the distribution of costs c_p that matches the observed distribution in Figure 11 because it would complicate substantially the estimation of sponsor preferences. First, it would require estimating such distribution at each iteration of the estimation algorithm because the optimal number of options

n^* itself depends on sponsors preference parameters. Second, it would require finding a solution to problem (84) which is non-convex without further restrictions.

To keep estimation tractable I instead model sponsors' choice of n as a random draw from the empirical distribution which I parametrize as geometric with parameter q. In estimation I allow for such distribution to be heterogeneous at the recordkeeper-year-category level. In general, the distribution of the number of options included within each category looks similar to the one in Figure 11 for many cuts of the data I have considered.

Derivation of investor surplus. Consider active investor i in plan p. Investor i has preferences over its retirement portfolio allocation $\boldsymbol{a}_i$ given by

$$u_i(\boldsymbol{a}_i) = \boldsymbol{a}_i'(\boldsymbol{\mu} - \boldsymbol{f}) - \frac{\gamma}{2}\boldsymbol{a}_i'V\boldsymbol{a}_i$$

where

$$V \equiv I + X_{(2)}X_{(2)}'.$$

Investor i's demand implied by the previous problem is

$$\boldsymbol{a}_i(\boldsymbol{f}) = \frac{1}{\gamma}V^{-1}(\boldsymbol{\mu} - \boldsymbol{f}).$$

Next, combine the expression for $\boldsymbol{a}_i$ with the expression for $u_i(\boldsymbol{a}_i)$ as follows

$$\begin{aligned} u_i(\boldsymbol{a}_i) &= \boldsymbol{a}_i'(\boldsymbol{\mu} - \boldsymbol{f}) - \frac{\gamma}{2}\boldsymbol{a}_i'V\boldsymbol{a}_i \\ &= \boldsymbol{a}_i(\boldsymbol{f})'(\boldsymbol{\mu} - \boldsymbol{f}) - \frac{1}{2}\boldsymbol{a}_i(\boldsymbol{f})'(\boldsymbol{\mu} - \boldsymbol{f}) \\ &= \frac{1}{2}\boldsymbol{a}_i(\boldsymbol{f})'(\boldsymbol{\mu} - \boldsymbol{f}), \end{aligned}$$

which is the measure of investors' surplus provided in the main text.

Example of category-based correlation structure. Consider a plan menu with 8 assets classified in the following four investment categories 'Equity-Growth', 'Equity-Value', 'Bond-Government' and 'Bond-Corporate'. Also assume that there are two assets for each of the four categories.

The vector of characteristic for a given asset j, $\boldsymbol{x}_{(2)j}$, has six elements corresponding to the 1st level characteristics (Equity, Bond) and 2nd level characteristics (Equity-Growth, Equity-Value, Bond-Gov, Bond-Corp). If asset j is an Equity-Value fund its vector of characteristics is given by:

$$\boldsymbol{x}_{(2)j} = (1, 0, 0, 1, 0, 0)'. \tag{86}$$

With the assumption that there are two assets within each category the 8×8 outer-product matrix $X_{(2)}X'_{(2)}$ is given by:

$$X_{(2)}X'_{(2)} = \begin{bmatrix} 2 & 2 & 1 & 1 & 0 & 0 & 0 & 0 \\ 2 & 2 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 2 & 2 & 0 & 0 & 0 & 0 \\ 1 & 1 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 2 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 1 & 1 & 2 & 2 \end{bmatrix}$$

Funds' investment category classifications capture cross-substitution patterns between assets. In this example investors treat Equity and Bond assets as independent and within Equity, Growth and Value assets as less substitutable than the two Growth or the two Value assets.

D 401(k) Lawsuits

This appendix provides details about some recent 401(k) lawsuits. I start by providing some background on 401(k) regulations building on (Mellman and Sanzenbacher (2018)) and then offer few specific examples of recent lawsuits.

The design of 401(k) retirement plans is governed by the Employee Retirement Income Security Act of 1974 (ERISA), with the Department of Labor (DOL) in charge of updating and enforcing such regulation. The law specifies that plan sponsors (i.e., employers) have a fiduciary duty to their plan investors requiring them to design and administer the plan in the 'sole benefit' of plan participants.

While the regulation is clear about the role of plan sponsors as fiduciaries, it provides almost no guidance on how to fulfill such duty in practice. For example, not much is said about how plan fiduciaries should select the type and number of investment options or determine a reasonable level of fees. Instead of laying out specific regulations or guidance, the DOL's general approach to overseeing 401(k)s has been through its own enforcement actions or through privately initiated litigation. Overall, plan fiduciaries are often left to guess what practices comply with ERISA and may only become aware of an alleged violation from a DOL investigation or lawsuit.

Typically there are two reasons that trigger 401(k) lawsuits. First, the inclusion of inappropriate investment options and second, the inclusion of options charging excessive fees. The former was the most common cause after the Great Recession mainly as a consequence of the inclusion of poor-performing employers' own stock. However, this kind of lawsuit has become less common since a 2014 Supreme Court ruling in the case of *Dudenhoeffer v. Fifth Third Bancorp* indicating that plan fiduciaries will not be held liable for failure to predict the future performance of the employer's stock. Since then, most lawsuits involved allegations of excessive investment and administrative fees. In what follows, I describe a few recent examples.

Allen v. M&T Bank Corp (2016). The Plaintiff (Allen) alleges that the Defendant (M&T) breached their fiduciary duties by retaining their proprietary funds within the plan despite the availability of similar lower cost and better performing investment options. According to the plan's Form 5500 filed for 2010, of the 22 mutual fund investment options in the Plan, 8 were from proprietary M&T mutual funds, representing over 30% of all mutual fund investments. However, these proprietary mutual funds charged significantly higher fees than average for performance that most often trailed both the Fund benchmarks and the mutual fund averages.

The Plaintiff provides some specific examples. For instance, the Wilmington Large Cap Value Institutional lagged the performance of a more reasonably priced alternative, Vanguard Equity Income Fund Admiral Shares. The Wilmington fund charged an expense ratio of 1.17%, higher than the Large Cap Value average of 0.83% and the

0.21% fee charged by the Vanguard Equity Income Fund Admiral Shares. A similar observation is made for the Wilmington Funds Small Cap Growth Institutional Fund charging an expense ratio of 1.39% against the 0.40% fee charged by Vanguard Stragic Small Cap Equity Fund.

Creamer v. Starwood Hotels & Resorts Worldwide Inc (2016). The Plaintiff (Creamer) alleges that the Defendant (Starwood Hotels & Resorts Worldwide, Inc.) serially breached its fiduciary duties in the management, operation and administration of its employees' 401(k) plan. It failed to ensure that fees charged to participants were reasonable. It caused plan participants who invested in index funds to pay seven times more than a reasonable fee. Indeed, the Starwood Plan received from the BrightScope rating service a score of only 61. The top BrightScope rating for peer plans was 90. The Plaintiff highlights that this difference would require sixteen years of additional by Starwood employees to reach the same level of savings as peer plan participants. Starwood participants lost savings of $110,871 per participant.

The Plaintiff also provides some specific examples. For instance, the BlackRock LifePath Index funds (the plan TDF) just hold other BlackRock index funds. BlackRock Life Path 2050 Index Fund institutional shares have net operating expenses of 0.20%. The 2050 Index Fund is a fund that invests all of its assets in other BlackRock funds. 52% of the Life Path Index Fund was invested in the BlackRock Russell 1000 Index Fund. The Russell 1000 Index fund had net operating expenses of 0.08%. Thus, the fee paid by plan participants is 0.20% plus 0.08% for a total of 0.28%. In contrast, the Vanguard Institutional Index Fund Institutional Shares had a total expense ratio of only 0.04% so the plan has chosen funds with fees that are 700% more than the comparable Vanguard fund - a difference of 24 basis points

McCorvey v. Nordstrom, Inc. (2017). The Plaintiff (McCorvey) alleges that the Defendant (Nordstrom) failed to adequately and prudently manage the plan. It allowed unreasonable fees to be incurred by participants and failed to use lower cost investment vehicles. The annual operating fees charged for many of the plan's investment options were substantially higher than reasonable management and operating fees of comparable funds, both index and actively managed funds. These fees were up to 16 times higher than comparable index funds, and up to 2.7 times higher than comparable actively managed funds.

The Plaintiff highlights that the high fee funds in the Nordstrom plan could have been easily replaced by lower cost index funds, TDF, or actively managed funds. For example, the PIMCO Total Return charging 46 basis points in fees could have been replaced by the Vanguard High Dividend Yield Index Fund charging 15 basis points or the Vanguard Growth and Income Fund Admiral Shares (an active fund) charging 23 basis points. Similarly, the average expense ratio for the set of TDFs available in Nordstrom (42 basis points) could have been five times lower by replacing it with

Vanguard Institutional Target Date Funds whose average expense was around 9 basis points.

Pledger v. Reliance Trust (2015). The Plaintiff (Pledger) alleges that the Defendant (Reliance Trust) breached its fiduciary by providing to the plan investment options that contained unreasonable management fees when cheaper versions of the same investments were available to the plan, as were other high-quality, low-cost institutional alternatives. The Plaintiff also alleges that the plan recordkeeper (Insperity) and the Defendant engaged in self-dealing by offering higher-cost investments to the plan's participants, because Reliance selected those investments in order to pay a larger amount of revenue-sharing to the recordkeeper.

Schapker v. Waddell & Reed Fin. Inc (2017). The Plaintiff (Schapker) alleges that the Defendants (WR Financial) selected the investment opportunities made available to the plan participants. During the Class Period, more than 97% of the investment opportunities made available to the plan participants were established and managed by WR Financial or its affiliates. Only one unaffiliated investment option—out of dozens of funds offered each year—was ever offered to plan participants. Because nearly all the investment options the Defendants made available to plan participants were established and managed by WR Financial or its affiliates, the Defendants caused the plan to pay its own Sponsor, WR Financial.

Further the Plaintiff points out that the fees charged to plan participants for their investments were in excess of the fees typically charged by unaffiliated companies for comparable mutual funds and products, and the performance levels of the investment options within the plan were worse than the performance achieved by unaffiliated companies for comparable mutual funds and investment products. Defendants could have selected comparable investment products from unaffiliated companies that cost less and performed better than the proprietary branded investment products to which the Defendant limited the plan participants.

E Differentiated Bertrand Network Game

In this Appendix I describe a simple differentiated Bertrand Network game and show how firms Nash equilibrium prices and margins relate to firms' network centrality. The discussion is based on Loseto (2023). In what follows, I frame everything in terms of products or firms, instead of calling them investment funds. I will assume that consumers have quadratic preferences with a taste for variety which makes the consumers' problem mathematically equivalent to an investor mean-variance portfolio problem. The same type of preferences are considered in Pellegrino (2023) who instead instead studies Cournot competition.

Consider a market with $j \in \{1, ..., J\}$ products available. Each product j is characterized by a set of K attributes whose values are collected in the K dimensional real-valued vector $x_j = (x_{jk})_{k=1}^{K}$ where x_{jk} is measured in units of quantity consumed. Characteristic x_{jk} tells you how much of attribute k you would get if you consume one unit of product j.

I assume there is a representative consumer who takes product prices $p = (p_j)_{j=1}^{J}$ as given, and chooses how much to consume of each product available. I denote by $q = (q_j)_{j=1}^{J}$ their consumption vector and define their preference as

$$u(q, X) = q_0 + q'\mu - \frac{\gamma}{2} q' \left(I + XX'\right) q \tag{87}$$

where q_0 is an outside good, X the $J \times K$ matrix of products attributes, μ is a J-vector parameters determining the marginal utility that comes from the linear term in (87). Finally, γ captures consumer's taste for variety.

The representative consumer takes prices p as given and maximizes (87) subject to

$$q_0 + q'p \leq y \tag{88}$$

where y income. After substituting for the budget constraint in (88), the demand system is given by

$$\begin{aligned} q(p) &= \frac{1}{\gamma} \left(I + XX'\right)^{-1} (\mu - p) && (89) \\ &= \frac{1}{\gamma} \left(I - \Theta\right)^{-1} (\mu - p) && (90) \end{aligned}$$

which is always well defined because $(I + XX')$ is positive definite and therefore non-singular and $\Theta \equiv X(I + X'X)^{-1}X'$

Next consider assume that J single-product firms producing the J products with constant marginal costs. Firm j takes the vector of competitor prices p_{-j} as given and solves

$$\max_{p_j} \quad (p_j - c_j)q_j(p_j, p_{-j}) \tag{91}$$
$$\text{s.t.} \quad q_j(p_j, p_{-j}) = a_j - \frac{1}{\gamma}(1-\theta_{jj})p_j + \frac{1}{\gamma}\sum_{l\neq j}\theta_{jl}p_l \tag{92}$$

which is equivalent to

$$\max_{p_j} \left(a_j + \frac{c_j}{\gamma}(1-\theta_{jj})\right) p_j - \frac{1}{\gamma}\left(1-\theta_{jj}\right) p_j^2 + \frac{1}{\gamma}\sum_{l\neq j}\theta_{jl}p_j p_l. \tag{93}$$

The payoff function in equation (93) is analogous to the linear-quadratic utility functions considered in Ballester, Calvó-Armenagol and Zenou (2006) and, as such, defines a linear-quadratic network game in which each product is a node and the $J \times J$ matrix

$$A(\Theta) \equiv \Theta - diag(\Theta) \tag{94}$$

is the weighted and undirected adjacency matrix of the network.

Network game interpretation. The adjacency matrix defined in (94) shows that network connections and products' substitution patterns are isomorphic to each other. We know that an off-diagonal element θ_{jl} of the matrix Θ captures the degree of substitution between product j and product l because it is defined as the (j,l) element of the demand jacobian. From equation (94), we can interpret θ_{jl} as a network link between product j and product l and therefore, we can think of the product differentiation space as being a network whose nodes are the products and whose links tell us how close, or equivalently how substitutable, are any two products.

Framing the product differentiation space as a network enables us to learn how product differentiation affects equilibrium outcomes by studying the topological properties of the competitive network. Loseto (2023) shows that equilibrium Bertrand price-cost margins depend negatively on a product's Bonacich centrality, which, following Jackson (2008), is defined as

Definition 1 *Let (A, J) be a network with J nodes and adjacency matrix A. The J-vector of (weighted) Bonacich centralities* $\mathbf{b}(A, \delta, u)$ *is given by*

$$\mathbf{b}(A, \delta, u) \equiv (I - \delta A)^{-1}\delta A u = \sum_{k=1}^{\infty} \delta^k A^k u, \tag{95}$$

where $\delta > 0$ is a scalar and $u > 0$ is J-vector.

The j-th element of $\mathbf{b}(A, \delta, u)$ summarizes how central node j is in the network. This measure of centrality is widely used in social networks because it captures a node's importance in terms of how close/connected this node is to others and how close/connected the nodes it is connected to. According to the definition of Bonacich centrality, a node's importance is a weighted sum of the walks that emanate from it. Moreover, if $\delta \in (0,1)$, walks of shorter length are weighted more.[42]

In an interior Betrand-Nash equilibrium, firms' equilibrium price-cost margins can be decomposed as

$$p^* - c = \underbrace{\frac{\mu - c}{2}}_{\text{monopolist margin}} - \underbrace{\mathbf{b}\left(A(\Theta), \frac{1}{2(1-\theta)}, \frac{\mu - c}{2}\right)}_{\text{Bonacich centrality}}, \tag{96}$$

or equivalently, firms equilibrium fees can be decomposed as

$$p^* = \underbrace{\frac{\mu + c}{2}}_{\text{monopolist fee}} - \underbrace{\mathbf{b}\left(A(\Theta), \frac{1}{2(1-\theta)}, \frac{\mu - c}{2}\right)}_{\text{Bonacich centrality}}, \tag{97}$$

The key insight is that the more central a product is in the competitive network, the lower its equilibrium price-cost margins. What does this mean in practice? From Definition 1, we can see that the higher any of the entries of the j-th row of A, the more central node j is. In this setting, product j is more central the higher its substitutability with any other product (i.e., the higher the elements $(\theta_{jl})_{l \neq j}$ of the j-th row of Θ). Overall, the expression for the Bertrand price-cost margins in (97) tells us two things. First, a less central or, equivalently, more differentiated product will be able to charge higher markups. Second, a product's Bonacich centrality is a sufficient statistic to measure how product differentiation allows firms to price above marginal costs.

Next, I perform a simple simulation exercise to summarize and visualize how the Bertrand Network model works. Table 15 describes the parameters used in the simulation. There is a single market with $J = 30$ products and $K = 7$ characteristics whose values are drawn from a uniform distribution in between $[0,1]$. The demand intercept μ is the same across all products and set to 0.15 whereas marginal costs are heterogeneous across products and drawn from a $[.01, .03]$ uniform distribution.

Given this parameters, Figure 21 plots the underlying Bertrand network. Each product is a node and the edges capture the degree of substitution between any two products/nodes; the longer the edge the less substitute are the two products. The location of dots and edges is exogenous and entirely determined by the realization of the draws of product characteristics. Conversely, the size of the dots is endogenous and it is proportional to the equilibrium price-cost margins. The plot shows that nodes that are more peripheral tend to have larger dot sizes whereas dots that are more central are smaller. The intuition for this result is the following: peripheral products

[42]This interpretation is motivated by the fact that when A is a binary $\{0,1\}$ it k-th power A^k counts how many walks of length k are between any two nodes.

Parameter	Value
J	30
K	7
μ	0.15
c	U[0.01, 0.03]
x_{jk}	U[0, 1]

Table 15: Parameters for simulation of Bertrand network game

are more unique or equivalently less central and, per equation (97) will charge higher margins in equilibrium. On the other hand, more central nodes face more intense competition and must lower their margins.

Figure 22 instead visualizes the previous decomposition and plots the equilibrium price-cost margins on the y-axis against the Bonacich product centrality on the x-axis. It should be clear by now why the relationship is decreasing; higher centrality implies lower equilibrium markups. The noise around the downward sloping relationship is due to the fact that marginal costs are heterogeneous. By increasing the variance of the distribution of costs, Figure 22 would start looking noisier and the resulting relationship between centrality and margins might not look as clear. This highlights how empirically it is important to control for the unobserved costs in order to recover the downward sloping relationship. The same would be true if we were to introduce heterogeneity in the demand intercept μ.

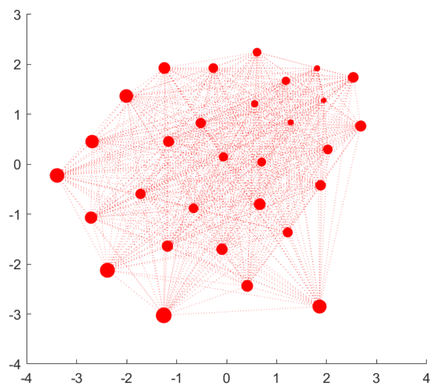

Figure 21: Simulated Network. Location is exogenous. Node size is proportional to markups.

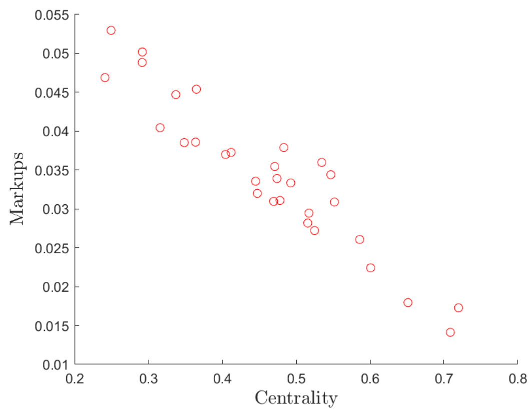

Figure 22: Simulated Network. Price-cost margins (y-axis) against Bonacich centrality (x-axis).