The Design of Defined Contribution Plans

VIVEK BHATTACHARYA and GASTÓN ILLANES

Department of Economics, Northwestern University and NBER

April 14, 2022

ABSTRACT. Defined contribution (DC) plans are a major vehicle for retirement savings in the US, holding almost $10 trillion in assets under management. In recent years, the quality and availability of these plans has been the subject of active policy attention and of several major lawsuits. This paper studies how employers and plan providers (recordkeepers) design these plans. We argue that low plan quality and limited provision can come from two sources. First, employer willingness to pay may be misaligned with that of workers or of regulators. Second, the market for recordkeeping may be imperfectly competitive. We propose a model of plan design and estimate that while both frictions are at play, significant changes to plan quality require modifying employer preferences. Accordingly, we evaluate proposed policies and conclude that only direct quality regulation can lead to significant quality improvements. Recent proposals can increase plan provision but have negligible quality effects.

KEYWORDS. defined contribution plans, bargaining.

JEL CODES. C78, J32, L51.

1. Introduction

Defined contribution (DC) plans constitute one of the most important retirement savings vehicles in the United States. In 2019, over 100 million workers had DC plans, and total assets invested in them amounted to $9.1 trillion. In a DC plan, workers earn tax advantages for saving a portion of their salary into a pre-specified menu of investments, set in conjunction by the employer and a plan administrator, or "recordkeeper." Regulations place standards of care on employers to try to ensure that the plan contains high-quality investment options.

Nevertheless, many plans offered in the marketplace arguably do not meet this standard. Many limit workers' access to investment options that are typically considered higher-quality: for the firms in our sample, 40% of DC plans did not include an S&P 500 tracker, 26% did not include a target date fund, and 17% did not include any passively managed index fund at all. Fees for investment options are often high: the average asset-weighted expense ratio is 67 bp, and over 60% of all plans had an asset-weighted expense ratio higher than 57 bp, the average expense ratio among all mutual funds in the US (Simon, 2017). These patterns are even starker for small firms. Moreover, many employers do not even offer a DC plan, leaving their workers without access to a major vehicle for tax-advantaged retirement savings.

These facts have not gone unnoticed by industry participants: over the past decade, a string of lawsuits have alleged that large employers have violated their fiduciary responsibilities on many dimensions, including by offering high-fee options. Cases have even reached the Supreme Court (e.g., *Tibble v. Edison International* and *Hughes v. Northwestern University*), and litigation related to DC plans has been growing over time; over 300 such cases were filed in various courts since 2015 (Schembari, 2022).

This paper studies how employers design their defined contribution plans. We study how both market imperfections and misalignment in employers' willingness to pay for quality can distort the menu of investment options made available to workers, or cause employers to not offer a plan at all. We then analyze which policy levers can improve outcomes for workers, studying both the impact of policies designed to target plan provision and the impact of direct quality regulation. We place emphasis on evaluating policies that are under current debate or that are provisions of the SECURE Act, a law passed in December 2019 designed to improve the private retirement landscape.

Why would employers design a plan without high-quality investment options? There are

two main incentives for the provision of high-quality investment options: (i) as with any other non-wage benefit, it might lead to wage savings, advantages in hiring and retention, or increases in worker morale; and (ii) offering a low-quality plan exposes the firm to litigation risk. Under the Employee Retirement Income Security Act of 1974 (ERISA), employers are typically considered fiduciaries to their workers and are expected to choose a diverse menu and monitor investments prudently. In practice, however, there are several reasons that may lead a firm to choose lower-quality plans. The firm may simply lack the sophistication or the bandwidth to properly evaluate investment options. The quality of the plan may not be salient (at hiring or even later), limiting the incentives to provide higher quality. Even though ERISA-related lawsuits about fiduciary duty are growing, litigation risk remains low—especially for small firms, whom plaintiffs may not find worthwhile to take to court. Moreover, even if the firm internalized the workers' objectives entirely, workers themselves may not value high-quality investment options, perhaps again due to lack of sophistication. Under these rationalizations, employers' willingness to pay for quality may be misaligned —either with workers' (an agency problem) or with the preferences a regulator or an informed participant would have.

Additionally, compensation structures in this setting can lead to agency frictions that further misalign employers' willingness to pay. Recordkeepers primarily earn revenue from two sources: direct payments from the employer and indirect payments from the investments on the menu. Indirect payments are typically a share of the expenses paid by workers (called "revenue sharing") and are thus larger for higher-fee investment options (Pool et al., 2021). Employers that do not fully internalize their workers' utility, or who face internal budget constraints for non-wage benefits, have an incentive to shift plan design towards higher-fee investments in order to save on direct payments.

Finally, while industry observers have typically viewed the issue of low-quality plans as one generated by flawed employer decisions, the structure of the market for plan provision may also have imperfections. Recordkeepers may have market power, stemming from imperfect search by employers for potential plan providers, differentiation (due to either cost or quality), or simply bargaining power in negotiations with employers. Since recordkeepers also participate in discussions on plan design, market power may skew the design of plans away from ones that are beneficial for workers—or even preferred by the employer.

In this paper, we develop a model of the market for DC plan provision that incorporates both these classes of frictions. We allow for a variety of sources of market power for recordkeepers: employers search for recordkeepers that may administer their plan, and

they bargain with them over both the menu of funds that is made available to workers and the transfer that will be paid to the recordkeeper. We also account for the two sources of misalignment of willingness-to-pay. First, employers perceive a benefit from offering a plan that is a function of the plan's characteristics, including the number of investment options, the mean expense ratio, and whether the plan has certain investment vehicles; we make no assumptions on the efficiency of these preferences. Second, employers may place differential disutility on spending money in direct transfers to recordkeepers, in employer contributions to workers, and in the money their workers spend compensating recordkeepers through fees—which affects their willingness to substitute direct transfers paid by sponsors for expenses paid indirectly by workers. Finally, we account for a variety of different other elements that are appropriate for the setting, such as heterogeneous preferences for recordkeepers, heterogeneous costs across recordkeepers, and economies of scale in plan provision.

We use data on all DC plans in 2016 with at least 100 participants from Form 5500, collected by the Department of Labor. We observe the investment menu and the total allocations in each asset by both the employer and the employees. We also observe all sources of revenue for the recordkeeper, including the direct transfers and indirect revenues in the form of revenue sharing. Observing plan choice together with the payments between parties is useful in identifying the magnitude of the frictions outlined above: broadly, the characteristics of the chosen plan are informative of the preferences of the employer, and patterns in the transfer help us identify the extent of recordkeeper market power.

We find that both misalignment of willingness-to-pay and market imperfections are at play. First, small employers value plans that have a S&P 500 tracker, target date funds, index funds in general, and lower expense ratios much less than large firms. This gap can come from several sources: different demands from workers, different competitive pressure from the labor market, or different exposure to fiduciary duty litigation would generate this pattern. Regardless of the source, it implies that smaller firms bargain with recordkeepers as if having a higher quality plan creates little marginal surplus. This is a main driver of the differences between small and large employers in the probability of having a plan and in the expenses their workers pay. In fact, when small employers' (less than 200 employees) preferences are aligned with large employers' (over 1000 employees), the probability of offering a DC plan ("participation") increases by nearly 10 pp and mean expenses decrease by around 10 bp.

Second, we estimate that all firms are reluctant to spend on direct transfers: small firms

3

view the marginal dollar spent on direct transfers as approximately $6 while large firms view it as approximately $3. This rejects the notion that workers trade off expenses paid by their employees and direct transfers to the recordkeeper equally, as would happen in a frictionless setting. Instead, firms have an incentive to compensate recordkeepers indirectly through revenue sharing rather than directly. This leads to plans where workers pay higher expenses. For small employers, eliminating this friction by forcing them to value direct transfers at $1 on the margin increases participation by 5 pp, and reduces expenses by about 4 bp. Aligning both elements of willingness-to-pay—the sensitivity to quality and the aversion to direct transfers—increases plan provision of small firms to levels beyond that of large firms and eliminates about two-thirds of the difference in expense ratios.

We also find that there is scope for the exertion of market power. Employers engage in limited search, only considering about 70% of the recordkeepers with whom similar firms contract. Eliminating search frictions would increase participation by between 5 and 6 pp for small plans, but we find an increase in average expenses paid (by about 3 bp). This is a direct implication of our previous findings: small employers do not value plan quality, so when they search more they do not get more of it. Employers also have limited bargaining power relative to recordkeepers. We estimate bargaining parameters ranging from 0.25 for small employers to 0.57 for large employers. This does not affect the extensive margin, as in our model a deal will be struck whenever there are gains from trade. However, eliminating recordkeeper market power has an effect on plan quality, reducing transfers that employers pay and thus decreasing expenses paid by small employers by about 5 bp. On net, however, we find that the incidence of market power lies primarily on transfer payments by employers and not on distorted plan quality.

We next analyze how policy can improve market outcomes. One class of policies encourages provision of plans but does not directly incentivize provision of high-quality plans. However, it is a priori plausible (and indeed possible in the model) that alleviating some of the frictions leading to limited plan provision also leads to higher quality: for instance, policies that increase competition could reduce the direct transfers employers pay and thus increase the relative willingness to pay for quality. One possibility is to subsidize sponsors for providing DC plans; in fact, the SECURE Act offers tax credits to (very) small businesses for offering these plans, up to $250 per participant for a limited time. Some states also offer start-up subsidies for employers. We find that offering firms who employ between 100 and 200 employees a $500 per worker subsidy by increases participation of small businesses by 6 pp, but we see no passthrough of this subsidy to higher plan quality.

Another widely discussed provision of the SECURE Act allows small businesses to band together to establish multiple-employer plans (MEPs). The theory behind this proposal is it can reduce costs by either harnessing economies of scale or increasing small employers' bargaining power to that of large employers. We estimate that this policy is ineffective, as the differences in outcomes between small and large employers are not driven by cost differences or by bargaining power: they are driven by differences in willingness to pay direct transfers and in valuations for plan quality. Thus, for MEPs to be effective, they would have to align small employers' demand for high-quality plans with large employers'. Since this demand is derived from the labor market and from exposure to fiduciary duty, we consider this to be unlikely.

These results suggest that if the goal is to improve plan quality significantly, it is not effective to take roundabout methods that simply incentive provision. We thus consider a variety of policies that target quality. We first consider "must-carry provisions," which mandate that plans must include certain types of investments. Overall, we find this to be a blunt tool: mandating S&P 500 trackers reduces expenses paid by workers in small employers by about 5 bp, mandating target date funds has no impact on overall expenses, and they reduce participation by by 3 and 7 pp respectively. We next consider expense ratio caps: mandating that unweighted expenses must be no more than 60 bp reduces this quantity to about 30 bp for small firms but reduces participation by 12 pp. Since such changes may be too stark, we finally consider targeted subsidies and penalties for offering plans with low average expense ratios. We find that these can be attractive alternatives to must-carry provisions. A targeted penalty of $100 per participant improves both provision and expenses relative to must-carry provisions on S&P and target date funds. Targeted subsidies of $500 per participant can lower expenses by 7 bp and increase participation by 1–2 bp, but are of course more costly.

Overall, we conclude that alleviating market imperfections and aligning willingness-to-pay affects plan provision significantly, and it also greatly increases plan quality. As a practical matter, however, employer's preferences for quality must be altered for this to happen—and this is difficult. Thus, policy intended to improve plan quality must regulate it directly, but care should be taken to ensure that extensive margin responses do not leave workers worse off.

Related Literature.   A large literature in household finance has focused on participant behavior in defined contribution plans, documenting a variety of behavioral frictions in par-

ticipation, contributions, and asset allocations. Benartzi and Thaler (2007) and Choi (2015) provide survey of the behavioral literature, and we discuss relevant papers in Section 2.2; Egan et al. (2021) study these patterns through the lens of a demand model. In contrast, there is much less work studying the supply side of this industry, although a small set of recent papers documents potential conflicts of interest: plan providers give preference to their own funds (Pool et al., 2016), they trade off direct fees from the sponsor with indirect fees from investments (Doellman and Sardarli, 2016; Badoer et al., 2020), and these indirect fees do affect whether funds are included in the plan (Pool et al., 2021). This paper accordingly builds off these observations to develop a model of the market for plan provision, respecting the possibility of behavioral frictions that may govern the underlying behavior of consumers, with the goal of studying the effects of regulation.

This paper fits into a literature on the industrial organization of financial markets, particularly using structural methods to study market structure and competition. Settings include car loans (Einav et al., 2012; Grunewald et al., 2019), credit cards (Nelson, 2020), insurance (Koijen and Yogo, 2016), mortgages (Allen et al., 2014, 2019; Robles-Garcia, 2020), municipal bonds (Brancaccio et al., 2020), pensions (Luco, 2019; Illanes, 2017; Illanes and Padi, 2021), advice (Bhattacharya et al., 2020; Egan, 2019; Guiso et al., 2022), and consumer and student loans (Bachas, 2019; Cuesta and Sepúlveda, 2019). Clark et al. (2021) provides an overview. More narrowly, we contribute to the smaller literature on market imperfections in business-to-business markets in finance (Koijen and Yogo, 2015; Di Maggio et al., 2021; Robles-Garcia, 2020).

This paper also contributes to the literature taking models of Nash bargaining, either between two parties or bilaterally between many interconnected pairs, to empirical work. Researchers have studied automobiles (Larsen, 2021), cable (Crawford and Yurukoglu, 2012; Crawford et al., 2018), medical devices (Grennan, 2013), financial markets (Robles-Garcia, 2020; Brancaccio and Kang, 2021), and hospital-insurer relationships (Gowrisankaran et al., 2015; Ho and Lee, 2017, 2019; Cuesta et al., 2019). Unlike most of these papers (Brancaccio and Kang (2021) is an exception), we study a setting in which quality is manipulable and a direct transfer between participants is a significant component of the payoffs. Understanding the interaction between these dimensions is a contribution of this paper.

2. Industry Background and Data

2.1. Setting and Data Sources

A defined contribution plan consists of a menu of investment options selected by the employer. Workers choose how much to contribute, and employers often match a portion of workers' contributions; the level of matching varies across employers. Typically, workers keep these funds in their individual accounts until they are spent down in retirement; withdrawals before age 59½ are subject to a 10% penalty except in certain life situations, underscoring its purpose as a retirement savings vehicle. All withdrawals are taxed as income. Importantly, unlike health insurance, employers are not required to provide DC plans—or any retirement benefit—at all.

Figure 1 illustrates the structure behind the administration of defined contribution plans. Employers (often called a plan "sponsor") decide on a recordkeeper (called a "trustee" in other sources) who administers both employer's and employees' contributions into the plan.[1] For this service, recordkeepers charge sponsors direct fees, usually on a per-participant basis. In some cases, employees also pay direct and management fees to recordkeepers. Moreover, the recordkeeper and employer also decide on the set of investment options available through the plans, including index funds, actively managed funds, common stock of the employer, etc. These investments are offered by "investment providers." It is common for recordkeepers to be vertically integrated into investment provision, and for them to offer their own funds. Investment providers pay recordkeepers fees for offering their funds, usually as a portion of assets under management, a practice called "revenue sharing." Finally, employees pay investment providers fees in the form of expense ratios.

Beyond these payments, both employers and recordkeepers face administrative costs when setting up a plan. These costs may include a fixed cost component, which would make it particularly expensive for small companies to set up a plan. Employers are also fiduciaries to their employees when designing a DC plan, and as such face the threat of legal action if plan quality is too low.

Currently, regulators have only instituted guidelines for certain features of the plan. For instance, plans must be designed to be "nondiscriminatory," i.e. benefit rank-and-file employees instead of just managers. Specific matching rules provide safe harbor provisions that exempt a plan from nondiscrimination testing, but no regulator mandates that plans

---

[1]In some cases, sponsors select multiple recordkeepers for employees to choose from. As this is rare, we describe the market ignoring this possibility.
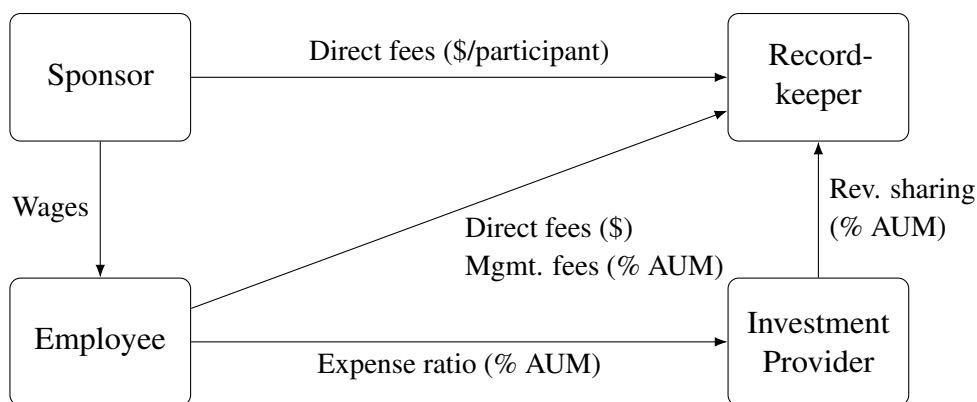
Figure 1: Illustration of DC plans, adapted from Collins et al. (2016).

follow such rules. Employers that automatically enroll their workers in the plan must choose
a default investment, and they are liable for choosing a high-quality one. The Department
of Labor issued a regulation in 2006 that designated certain types of funds—such as index
or target date retirement funds—as "qualified" default investment alternatives (QDIAs),
freeing employers from liability if they choose such funds but without mandating they do
so.[2] Moreover, there are no specific mandates for the dimensions of plan quality that are
often brought up in lawsuits—such as the types of funds included in the plan, the number of
options on the plan, or the expense ratios charged by the funds.

We leverage data from Form 5500, an annual filing all employers who offer benefits as a
part of compensation must file with the Department of Labor, to study this market. For all
firms offering employee benefits, we observe the number of participants in each benefit plan,
as well as their name, location, and NAICS code. Focusing on defined contribution plans, we
can identify all the entities involved in the plan and almost all the transfers outlined above.[3]
In particular, we observe revenue sharing outlays at the plan-asset level, contributions made
by employees and employers, and matching and vesting rules. Finally, we observe the menu
of investment options for all plans with 100 or more participants, including balances for
each investment option. While all this information is made public by the DOL, some of it
is not digitized, and we purchased a digitized version from Brightscope. We complement
this information with data from CRSP and Morningstar regarding asset fees, star ratings,

---

[2]See https://www.irs.gov/retirement-plans/operating-a-401k-plan and https://www.betterment.com/work/
resources/what-is-a-qualified-default-investment-alternative-qdia for more information.

[3]We do not have data on fees paid by the employer to the recordkeeper directly. However, our understanding is
that most plans tend not to have such fees, while direct fees from the sponsor to the recordkeeper and revenue
sharing are nearly universal.

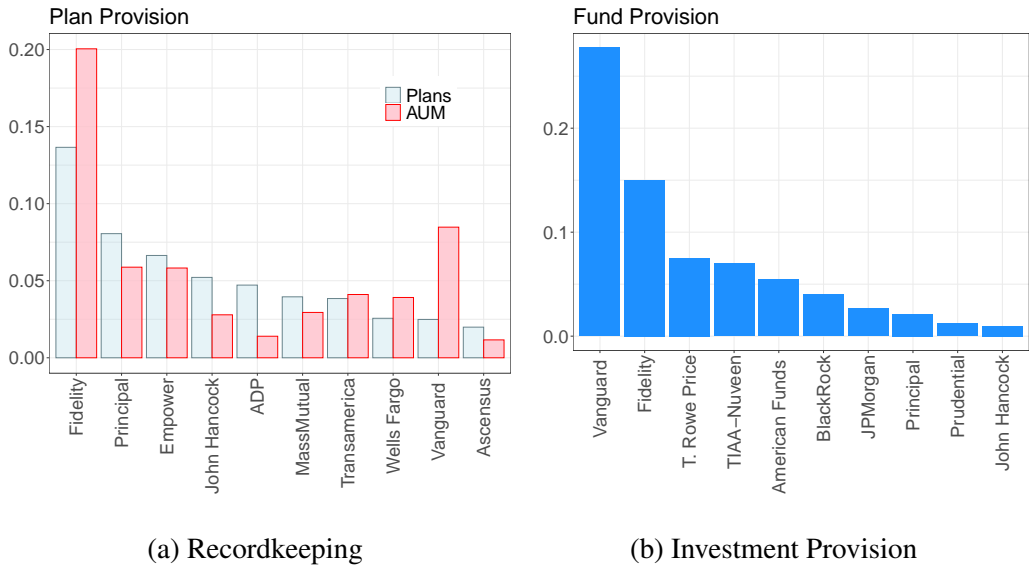(a) Recordkeeping   (b) Investment Provision

Figure 2: Market Shares

and investment styles. For the analysis that follows, we focus on plans sized between 100 and 5,000 participants, dropping sponsors that offer defined benefit (i.e., traditional pension) plans and a small handful that exhibit outsized fees. See Appendix A for further details.

2.2. Descriptive Facts

Figure 2 reports market shares for recordkeeping and investment provision. The market as a whole is not especially concentrated, but not all recordkeepers necessarily compete for the business of each sponsor. Comparing Panels (a) and (b) also highlights that vertical integration with investment providers is important, as Vanguard and Fidelity are market leaders in both plan provision (by AUM) and fund provision.

Table 1 reports summary statistics for our sample. Panel A illustrates the heterogeneity across sponsors and savings behavior in our sample. The average firm has 499 employees, with 167 and 515 as the first and third quartiles. Mean per-person balances are around $30,000. Mean employee and sponsor contributions are around $3,000 and $1,500, respectively. Some sponsors, however, have considerably larger balances and contributions. The empirical model allows for this heterogeneity to lead to difference in potential revenues (through revenue sharing agreements) and costs (through economies of scale).

We next document a number of descriptive facts to motivate elements of the model. First, the extensive margin is sizable: among firms that do not offer defined benefit plans, about

	N	Mean	SD	5%	25%	50%	75%	95%
A. General Plan Characteristics								
Total Participants	47835	499	678	120	167	262	515	1752
Balance / Person	47835	31436	33289	823	8037	21146	43666	96569
Employee Contribution / Person	47835	3071	2404	294	1279	2477	4223	8021
Employer Contribution / Person	47835	1535	1952	0	258	895	2017	5567
Has a DC Plan	60634	0.789	0.408	0	1	1	1	1
B. Plan Quality								
Has S&P Tracker	40792	0.589	0.492	0	0	1	1	1
Has Target Date Fund	40792	0.737	0.44	0	0	1	1	1
Has Index Fund	40792	0.831	0.375	0	1	1	1	1
Weighted Expense Ratio (bp)	40792	67.4	31.3	17.4	44.9	65.6	89.8	117.8
Expenses / Person	40792	193	185	7	53	137	279	570
C. Asset Allocation								
Share in Index Funds	40792	0.228	0.268	0	0.027	0.122	0.328	0.861
Weighted / Average Expenses	40792	0.93	0.172	0.621	0.855	0.951	1.014	1.154
D. Revenues and Payment								
Direct Transfer / Person	47835	52	114	0	2	15	60	212
Revenue Sharing / Person	47835	60	87	0	4	23	77	252
Direct Transfer / Total Revenue	47507	0.423	0.391	0	0.028	0.307	0.86	0.996

Table 1: Summary statistics. Panels A and D show summary statistics of all plans from Form 5500 for which we have relevant data. Panels B and C restrict to the sample for which we observe the plan menu.

21% also do not offer any DC plan. The extensive margin into plan provision is important, as such workers are only eligible for a much smaller tax-advantaged contribution through an IRA.

Second, plans have heterogeneous quality, and a fair number of them are arguably of low quality as measured by various metrics. Panel B shows that 41% lack an S&P index fund, 26% lack a target date retirement fund, and 17% lack any sort of index fund—investments that are normally thought of as high-quality. Asset-weighted expense ratios are not low: they average 67 bp, but over 25% of firms have one that exceeds 89 bp. As mentioned earlier, the average across all mutual funds was 57 bp in 2016. In the model, high expenses and the lack of such funds can come from sponsors undervaluing their benefits to workers or recordkeepers exerting market power to avoid including lower-fee funds on the menu.

Third, workers tend to invest across funds in the menu instead of simply allocating the majority of their assets to index funds or other lower-fee options. Panel C shows that only about 23% of assets on average are invested in index funds. We also tabulate the ratio between the asset-weighted average expense ratio and the equal-weighted average expense

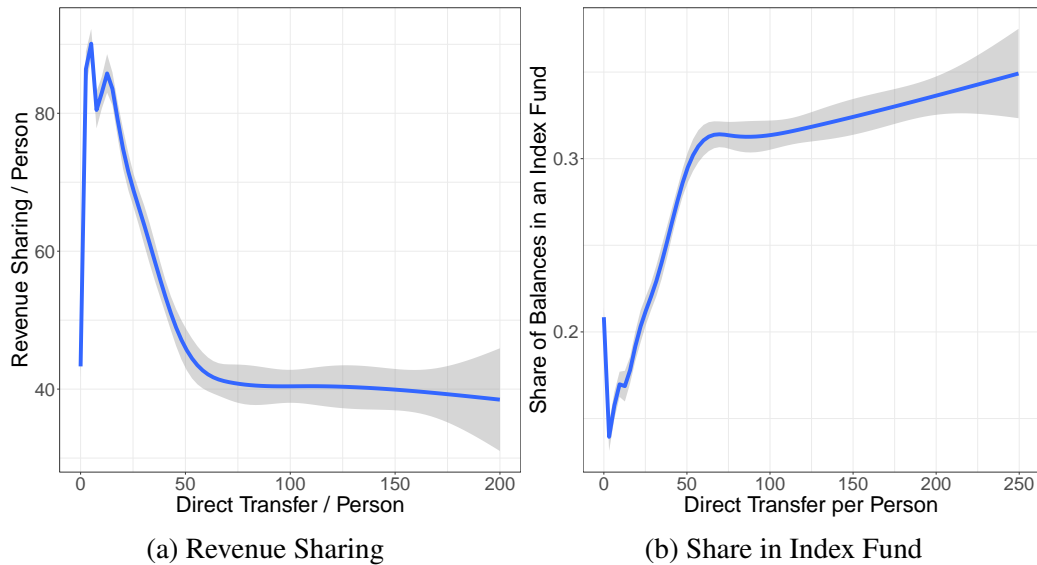(a) Revenue Sharing   (b) Share in Index Fund

Figure 3: Relationship between direct transfers and plan quality metrics

ratio of funds in the menu; this ratio is only slightly smaller than 1 (0.93), again suggesting that workers are not directing funds towards lower-fee options. (Figure D.1 in Appendix D provides more details on these quantities.)

These observations are in line with considerable evidence in this setting that decisions regarding contributions and asset allocations are made sub-optimally and subject to heuristics: examples include naive diversification by splitting assets roughly evenly across all options (Benartzi and Thaler, 2001; Huberman and Jiang, 2006; Iyengar and Kamenica, 2010), underdiversification by investing in only one fund or stock (Choi et al., 2005), ignoring assets in one account when deciding on the other (Choi et al., 2009), or limited sensitivity to fees when they are not salient (Kronlund et al., 2021). Moreover, contributions are also subject to behavioral frictions, which would affect the map between the plan menu and total expenses or revenue generated.[4] From a modeling perspective, we avoid conceptualizing asset allocations as the outcome of an optimal portfolio allocation problem, instead relying on flexible predictions of the outcomes necessary for estimation and counterfactuals.

Fourth, the two main sources of revenue mentioned above—revenue sharing agreements

---

[4]Participants are often reluctant to invest money even when deposits so would lead to employer contributions and could be immediately withdrawn (Choi et al., 2011). Conditional on contributing, participants often set contribution rates at "round" portions of income (5% or 10%), as documented by Benartzi and Thaler (2007). Increasing choice is not always always associated with greater takeup (Papke, 2004; Sethi-Iyengar et al., 2004; Huberman et al., 2007; Beshears et al., 2013).

and direct transfers—are both important to the recordkeeper. Panel D shows that the means per participant ($52 and $60, respectively) are of similar magnitudes, as are the quantiles. In fact, the share of total revenues to the recordkeeper due to direct transfers is on average 42%, but there is heterogeneity on this dimension as well, with significant mass near both 0 and 1. Moreover, Panel (a) in Figure 3 shows a trade-off between the two sources of revenue—with the exception of the case where transfers are especially small. The implication of this observation for modeling purposes is that, unlike other settings with bargaining, we cannot abstract away from this direct transfer; how we model the trade-off between these two revenue sources has implications for the nature of competition in this market. Additionally, the fact that quality is high when direct transfers are zero is an implication of our model; we return to this issue in Section 3. Panel (b) corroborates that plans with higher direct transfers have higher quality investment outcomes—consistent with recordkeepers being compensated either through direct transfers or through higher expenses.

Finally, plan size is an important driver of the heterogeneity documented above: both plan availability and quality are correlated with it.[5] Figure 4(a) shows that plan availability increases about 15 pp when moving from sponsors with 100 employees to those with 300. Figure 4(b) shows that average asset-weighted expenses move from 80 bp to 55 bp when moving from firms with 100 to 2,000 employees. The availability of various investments also generally increases with firm size, as show in Figure 4(c). In the empirical implementation of the model, we allow the magnitude of the frictions to depend on firm size to disentangle the sources of the gradients documented in Figure 4.

3. Model

In this section, we develop a model of the interactions between plan sponsors and record-keepers that lead to observed plan choice.

3.1. Payoffs to the Parties

A set $\mathcal{R}_s$ of recordkeepers is capable of providing a DC plan to sponsor s. The sponsor s chooses a subset $\tilde{\mathcal{R}}_s \subseteq \mathcal{R}_s$ with whom to negotiate the design of the plan. The sponsor s and a recordkeeper r can decide on a plan $p \in \mathcal{P}_{rs}$ and a dollar transfer T from s to r. If the

---

[5]For firms that offer a plan, we use the number of plan participants as the measure of firm size. For firms that do not, we use the number of participants in the health insurance plan.
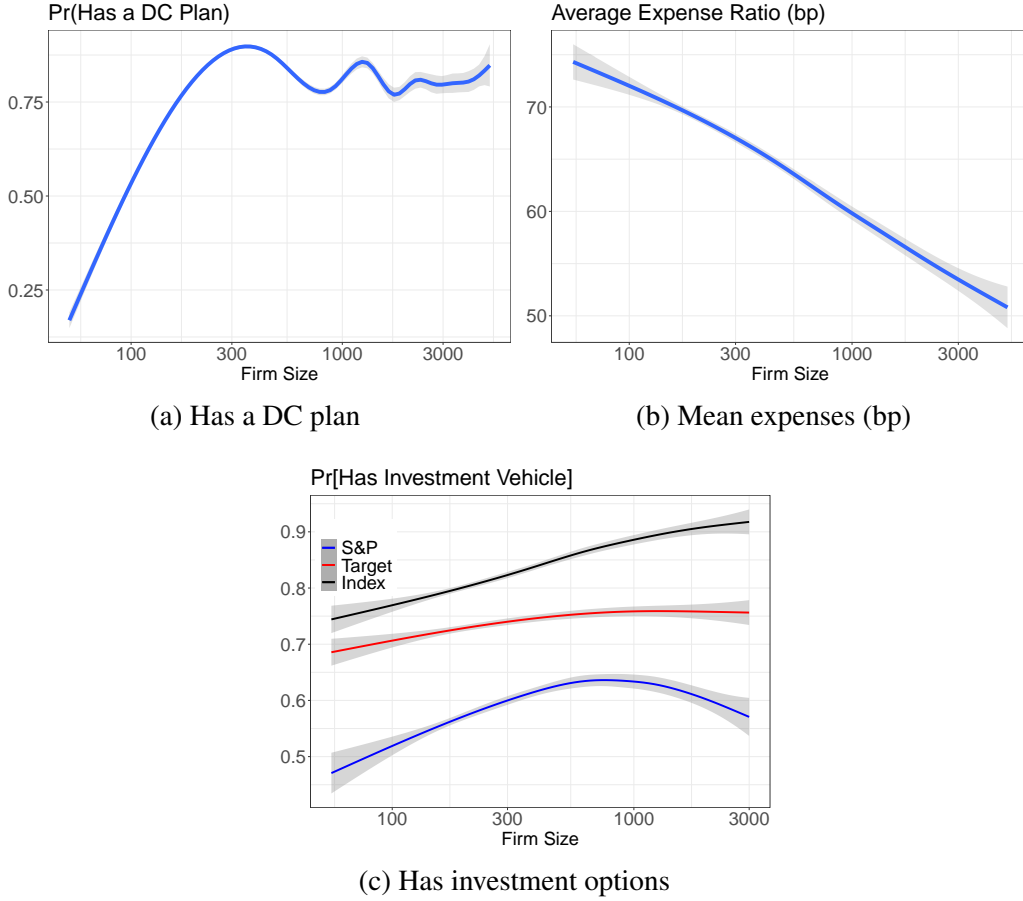
(a) Has a DC plan

(b) Mean expenses (bp)

(c) Has investment options

Figure 4: Relationship of plan provision and quality metrics with firm size

parties decide on a plan and transfer (p, T), the per-participant payoff to the sponsor s is

$$\pi_{rs}^{S}(p, T) \equiv B_{rs}(p) - C^{S} - T - \kappa(T). \tag{1}$$

In (1), $B_{rs}(p)$ is the per-participant net benefit to the sponsor from offering a plan p. We conceptualize this benefit as coming from four main sources. First, there is a compensating differential between wage and non-wage benefits, so that offering a higher quality plan allows the employer to also offer lower wages. Second, there may be additional benefits to employers to offering a better plan beyond pure wage savings—such a higher morale leading to a better work environment. Third, higher-quality plans reduce the probability that the sponsor is sued for breach of fiduciary duty, and sponsors may be differ in the likelihood of being sued and in legal costs. Costs of fiduciary insurance would fall into this category.

Finally, sponsors often offer matching, and these are deducted from benefits. Next, C^S is the per-person administrative cost faced by the sponsor.

The final two terms of (1) relate to the direct transfer T from the sponsor to the record-keeper. First, this transfer is a cost that directly impacts utility. Second, the function $\kappa(T)$ denotes an additional cost of paying for a defined contribution plan. We assume $\kappa(\cdot) \geq 0$, $\kappa'(\cdot) \geq 0$, and $\kappa''(\cdot) \geq 0$. A convex $\kappa(\cdot)$ corresponds to the case where the firm's marginal willingness to pay decreases with the transfer: the extreme case of $\kappa(\cdot)$ being especially large after some $\bar{T}$ would correspond to a firm unwilling to pay more than $\bar{T}$ for a DC plan. This cost can capture liquidity constraints, internal budgeting, managerial incentives, or any other frictions that would prevent a firm from wanting to pay especially large transfers.

The per-participant payoff to the recordkeeper r is

$$\pi_{rs}^{R}(p, T) \equiv \Pi_{sr}(p) - C_r^R + T. \tag{2}$$

The term $\Pi_{sr}(p)$ in (2) constitutes the per-participant revenues the recordkeeper earns. The primary component of this comes from revenue sharing agreements: the sum of the expected dollar amount invested by each of the sponsor's workers in fund f times q_{fr}, is the revenue sharing that r earns from fund f. If f is part of r's fund family, then q_{fr} is the expense ratio and otherwise it is the revenue sharing rate. Next, C_r^R is the per-participant administrative cost of the recordkeeper; in estimation, we let it depend on the number of participants to capture economies of scale. Finally, in the empirical implementation we allow for an idiosyncratic shock at the sponsor-recordkeeper level.

3.2. Plan Choice

As discussed in Section 2, sponsors decide on plans jointly with recordkeepers, and it is a priori unclear which party sets the terms of the plan. To allow for this, we consider that plans and transfers are set through Nash bargaining between the two parties. In particular, if a sponsor and recordkeeper are negotiating over a plan design, they solve

$$\underset{p \in \mathcal{P}_{rs}, T}{\arg\max} \left(\pi_{rs}^{S}(p, T) - \pi_{rs}^{Sd}\right)^{\eta} \cdot \left(\pi_{rs}^{R}(p, T)\right)^{1-\eta} \text{ s.t. } T \geq 0. \tag{3}$$

14

In (3), η is the bargaining power of the sponsor. The constraint $T \geq 0$ reflects the restriction that the recordkeeper cannot pay the sponsor for the right to provide the plan.[6] The recordkeeper does not face any capacity constraints, and its disagreement payoff is simply zero. The disagreement payoff π_{rs}^{Sd} to the sponsor, however, corresponds to extracting the maximum surplus possible from the second-best recordkeeper, subject to meeting its individual rationality constraint. That is, we let

$$\bar{\pi}_{rs} \equiv \max_{p,T} \pi_{rs}^{S}(p, T) \text{ s.t. } T \geq 0 \text{ and } \pi_{rs}^{R}(p, T) \geq 0, \tag{4}$$

and $\pi_{rs}^{Sd} \equiv \max_{r' \in \tilde{\mathcal{R}}_s, r' \neq r} \bar{\pi}_{r's}^{S}$.

Fixing a recordkepeer, (3) and (4) specify how the plan is chosen. The sponsor selects the recordkeeper that maximizes its own utility $\pi_{rs}^{S}(p^*, T^*)$, and the plan and transfer are given by this bargaining procedure. If for a choice of a recordkeeper r, π_{rs}^{d} is less than the maximum surplus $\pi_{rs}^{S}(p, T)$ the sponsor would earn with that recordkeeper, the sponsor s would not choose r. Thus, a consequence of this bargaining scheme is that sponsors contract with the recordkeeper that offers the highest $\bar{\pi}_{rs}$.

3.3. Properties of the Model

Broadly, the model captures market imperfection in two ways. First, the model captures search frictions naturally: sponsors only consider a limited subset $\tilde{\mathcal{R}}_s$ when deciding on the plan menu, and this could be a strict subset of all potential recordkeepers $\mathcal{R}_s$. Second, bargaining power η, together with the surplus generated by the second-best option, captures recordkeeper market power. The model also captures misalignment in willingness to pay through two channels. First, since we do not take a stance on the determinants of the benefit function $B_{rs}(p)$, the sponsor may not place the same emphasis on plan characteristics—e.g., having low-fee funds, or incorporating certain types of investment options—that informed workers would if they were selecting plans themselves. Second, $\kappa(\cdot)$ is an explicit agency friction: a sponsor has an incentive to reduce expenses (like the direct transfer) that are incident on itself, even if doing so comes at the cost of increasing expenses incident on workers.

These frictions interact with each other in the model: the bargaining procedure allows

---

[6] A natural concern is that the recordkeeper may offer "free" services in lieu of such payments that make the plan seem more attractive. In the empirical implementation, we let the mean utility for the recordkeeper depend on firm size, which would capture systematic differences in quality provision by size. However, our model does not allow us to introduce costly unobserved quality.

for market power to affect plan characteristics, and how much it does so depends on the misalignment in willingness to pay. To understand the economic forces embedded in this model, consider the case where we relax the $T \geq 0$ constraint and set $\kappa(\cdot) = 0$. Then, the model reduces to Nash bargaining with transferable utility, and the solution to (3) is to set the plan p to maximize joint sponsor-recordkeeper surplus $\pi_{rs}(p) \equiv \pi_{rs}^S(p, T) + \pi_{rs}^R(p, T)$, which does not depend on T by design. The transfer T is then set so that the sponsor captures a share η of the incremental surplus generated by the bargain:

$$T = \tilde{\pi}_{rs}^S(p^*) - \eta \left(\pi_{rs}(p^*) - \pi_{rs}^{Sd} \right), \tag{5}$$

where $\tilde{\pi}_{rs}^S(\cdot)$ denotes the utility of the sponsor ignoring the transfer. Here, the choice of the plan (and recordkeeper) incorporates agency frictions—captured in the benefit function. However, reducing the market power of the recordkeeper does not affect the plan choice: increasing η or increasing π_{rs}^{Sd} (which captures an increase in the level of competition in the market) only reduce the direct transfer without changing plan choice.

The constraint $T \geq 0$ and the cost function $\kappa(\cdot)$ allow bargaining and competition to directly affect plan quality. Consider an extreme case where $\kappa(T) = 0$ for $T \leq \bar{T}$ and prohibitively large otherwise, perhaps if the HR department of a firm has a budget to spend on retirement plans. If (5) is such that T would be negative, then the solution sets $T = 0$ and the plan to maximize $\left(\pi_{rs}^S(p, 0) - \pi_{rs}^{Sd} \right)^{\eta} \cdot \left(\pi_{rs}^R(p, 0) \right)^{1-\eta}$: since the recordkeeper cannot transfer surplus to the sponsor by lowering the transfer further—which would be efficient from the perspective of joint surplus—it does so by improving the plan quality (to the extent that the sponsor earns a higher benefit from higher-quality plans). Conversely, if the sponsor's bargaining power is so low (or competition is so weak that π_{rs}^{Sd} is low) that T would exceed $\bar{T}$, the sponsor would instead limit the transfer to $\bar{T}$ but be willing to accept plans of lower quality. Smoother functional forms for $\kappa(\cdot)$ embed analogous forces, and thus $\kappa(\cdot)$ is an important parameter that governs how competition and market power translates to plan quality.[7]

The elements can rationalize the descriptive facts outlined in Section 2.2. The model incorporates an extensive margin, as both the sponsor and at least one recordkeeper in the consideration set has to meet their individual rationality constraints for there to be a plan. Direct inspection of (3) shows that the recordkeeper view direct fees and revenue sharing as

---

[7]We should note that in this discussion, we refer to quality as measured by $B_{rs}(\cdot)$ —not necessarily higher quality as perceived by what one may expect to be beneficial to workers.

substitutes, setting up a tradeoff between them. High expenses or the lack of index funds could be explained by a low preference for them on the part of the sponsor ($B_{rs}(\cdot)$), that such funds bring especially high revenues to the recordkeeper and thus increase the Nash product in (3), or that $\kappa(\cdot)$ is large and the sponsor is averse to high transfers (and thus providing the recordkeeper revenues through revenue sharing increases the Nash product). The fact that surplus must be provided to the sponsor through higher-quality plans when $T = 0$ explains the low average revenue sharing (from lower average fees) documented in Figure 3.

Finally, we model plan design as a static game. Since in practice employers do not change their plan provider on a yearly basis, a natural concern is the possibility of inertia. Not accounting for inertia would cause us to miss potential dynamic pricing incentives due to lock-in or cause use to underestimate sensitivity to quality. We do not consider this concern to be first-order in our setting. First, survey evidence suggests that sponsors do actively evaluate their DC plans: a 2021 survey by Fidelity of more than 1,000 sponsors showed that 61% evaluate their providers and advisors annually, almost half of whom evaluate it more than once a year.[8] Moreover, these evaluations are not merely passive: 74% of sponsors made changes to plan menus in the past two years. Overall, these observations suggest that sponsors do not simply set a menu once and leave it be for many years; thus, we find the approximation of active choice to be sensible and not merely made for simplicity.

4. Identification and Estimation

4.1. Empirical Model

To flexibly estimate how model primitives change with firm size, we divide firms into groups as a function of the number of employees and estimate the model separately for each group. In particular, we bin together firms that hire 200 or fewer employees, between 201 and 500 employees, between 501 and 1,000 employees, and more than 1,000 employees (Groups 1–4, respectively). All parameters noted below are different for each group. We assume that there is no further dependence on firm size. We then parameterize

$$
\begin{aligned}
B_{sr}(p) &= \boldsymbol{\theta}_s \cdot \boldsymbol{Q}_s(p) + \delta_r + \epsilon_{rs}^S, \\
\Pi_{sr}(p) &= R_s(p) - C_r^R + \epsilon_{rs}^R
\end{aligned}
\tag{6}
$$

---

[8]See https://institutional.fidelity.com/app/item/RD_13569_26306/plan-sponsor-attitudes.html for the 2021 Plan Sponsor Attitudes Survey, conducted by Fidelty.

In the specification of the benefit function, we include a dependence on sponsor-plan properties $\boldsymbol{Q}_s(p)$ capturing plan quality: number of investment options, indicator variables for whether the plan has an S&P 500 tracker or a suite of target retirement date funds, number of index funds, unweighted average expense ratio, and per-person employer contributions. We allow the coefficient multiplying quality terms to be sponsor-specific and, for each dimension of quality q, parameterize $\theta_{qs} \sim N(\mu_q, \sigma_q^2)$. We also assume per-person costs to recordkeepers C_r^R are constant within bins of firm size: if there are economies of scales, this value will decrease over firm size groups. We also assume unobserved (net) preferences for recordkeepers δ_r are fixed within bins of firm size. We can conceptualize $\delta_r = \xi_r - C^S$, where ξ_r captures differences in quality or services provided by recordkeepers which may make them differentially attractive to sponsors, as well as the mean utility difference between having a plan and not having one. It also encapsulates the sponsor cost (C^S in the notation from Section 3.1); in the empirical model, disentangling the preferences for recordkeepers from costs to the sponsor will require an instrument that shifts one or the other, and doing so is not necessary for our counterfactuals of interest. Finally, we include an unobserved shock $\epsilon_{rs}^S(p) \sim \text{T1EV}(\sigma^S)$ to the benefit function, which we interpret as a "match value": a sponsor may prefer a particular recordkeeper because of a pre-existing relationship between the companies, or because of the possibility of getting more firm-specific services from them. The recordkeeper's revenue $\Pi(\cdot)$ consists of revenue from revenue-sharing agreements $R_s(p)$, per-person costs C_r^R, and a shock $\epsilon_{rs}^R \sim \text{T1EV}(\sigma^R)$. This shock also captures any unobserved prior relationship or synergies that may affect the revenues (or the service cost) for a particular pair.

The remaining elements of the model are the bargaining parameter η, the cost of transfers $\kappa(\cdot)$, and the search process that leads to $\tilde{\mathcal{R}}_s$. We parameterize $\kappa(T) = \kappa \cdot T^2$: the quadratic function allows for an increasing marginal cost of transfers, and firms with high κ are more willing to accept lower-quality plans in exchange for lower transfers. The quadratic form also allows for closed-form computation of the transfer conditional on a candidate plan, which is useful in estimation and counterfactuals.

Finally, we specify the search process as follows. We assume that $\mathcal{R}_s$, the set of all potential recordkeepers sponsor s might select, is the set of all recordkeepers serving firms in the same state-two digit NAICS combination. From this set, a sponsor s constructs the consideration set $\tilde{\mathcal{R}}_s$ by picking each of the top 19 recordkeepers with probability ρ. The sponsor can also select, with probability ρ, one recordkeeper at random from the fringe of remaining firms. When sponsor s draws recordkeeper r into the consideration set $\tilde{\mathcal{R}}_s$,

we assume that all plans offered by r to firms in the same state-two digit NAICS as s are made available to s. This is a reduced-form approach to consideration set formation, similar in spirit to Goeree (2008).[9] We adopt this approach in lieu of a model of optimal search partly for its parsimony but also because we find it appropriate for the setting: a model of optimal search requires a level of sophistication on the sponsor side that is at odds with them choosing plan quality in a potentially sub-optimal way. Since we think that allowing for this is of first-order importance, we are not comfortable imposing a high degree of sophistication in the search process.

Almost all elements of the parameterization are directly observed from any plan menu. However, per-person employer contributions and average per-person revenue sharing are only observed for selected plans; per-person dollar expenses (which are useful for interpretation of the counterfactuals) are also only observed for such plans. We treat the computation of these quantities for counterfactual plans—e.g., what revenue sharing would be if s selects a plan other than the one it chose—as a prediction problem. In particular, we use stochastic gradient boosting (Friedman, 2002) to predict these quantities from a large vector of plan and sponsor characteristics. Appendix B provides a description of the procedure.

We take this approach since it allows us to be agnostic about how workers decide on asset allocations, thus respecting the literature (and facts shown in Section 2.2) about behavioral biases in contribution and allocation decisions. Of course, the main assumption needed for this approach to be valid is that plans are not selected on the basis of unobserved sponsor-level variables that affect asset allocations.[10] Moreover, this approach limits counterfactuals to ones in which we expect the map from plans to asset allocations to be unchanged. We do not consider this to be a central concern in our counterfactuals of interest.

---

[9]Unlike Goeree (2008), we do not assume that consideration set probabilities are a function of observables, and as a result we do not leverage an exclusion restriction between the observables that enter into consideration set formation and the observables that shift sponsor utility. Instead, we leverage the fact that sponsors can choose between multiple plans offered by the same recordkeeper to identify consideration set probabilities. We discuss this further in Section 4.2.

[10]An example would be a sponsor that chooses to contract with Fidelity since it knows its workers have an affinity for Fidelity funds; we find this concern to be unlikely. In general, threats to this strategy seem to require some sophistication in asset allocation decisions by workers (a preference for certain types of funds or a propensity for certain investment decisions, which must be known to the sponsor), which we find to be broadly at odds with the behavioral literature on worker investments.

4.2. Identification

For intuition about identification, it is informative to consider the benchmark with $\kappa = 0$ and no constraint that transfers must be positive, and with $\rho = 1$. As discussed in Section 3.3, the parties would maximize joint surplus in this model. Like in standard discrete choice settings, plan choice would identify the sponsor's preferences for quality as well as the mean utility for a particular recordkeeper net of total costs ($\delta_r = \xi_r - C^S$). Ignoring (ϵ^S, ϵ^R) shocks, we could then compute the joint surplus $u^{(1)}$ of the chosen plan and that of the second-best option $u^{(2)}$. Since the recordkeeper earns $(1 - \eta)$ of the surplus, the equation that $R^{(1)} + T - C^R = (1 - \eta) \cdot (u^{(1)} - u^{(2)})$, where $R^{(1)}$ is the revenue-sharing by the chosen plan, would identify C^R and $(1 - \eta)$ given variation in the incremental surplus generated.[11]

Using variation in the incremental surplus to identify a bargaining parameter mirrors the argument of Grennan (2013). In our setting, this variation could be induced by differences in choice sets across sponsors. Firms in different industries and geographies have different sets of plans from which to choose, which is embedded into how we construct choice sets. Moreover, due to variation in firm characteristics (including policies such as match rates), firms would even face different revenue sharing levels and contributions for the same plan.

Simple departures from this baseline model help connect this intuition back to the setup in this paper. Consider adding back incomplete consideration, so that ρ could be less than 1 and must be identified. We make two observations. First, a sponsor still considers all plans offered by a recordkeeper; thus, variation in plan quality within the chosen recordkeeper can still identify the sensitivity to plan quality. Second, we have two sources of variation in incremental surplus—the joint surplus of the chosen plan and that of the second-best option. With $\rho = 1$, having two different sources of surplus is immaterial for identification: the slope of total revenues of the recordkeeper with respect to either is $1 - \eta$ in absolute value. However, when $\rho < 1$, the candidate for the second-best option among all potential options may not be in the actual (unobserved) consideration set. Thus, finding that total revenues to the recordkeeper are less sensitive to the quality of the second-best choice among all potential options than they are to the quality of the chosen option provides evidence of

---

[11]We make two comments about this discussion. First, identifying the scale of u from the discrete choice requires a numeraire; in our case, this is revenue sharing (which is observed and enters joint surplus with a coefficient of 1). Second, the logic in this paragraph can be extended to the case with ϵ shocks as well: if the observed component of the incremental surplus generated by the bargain is sufficiently large, one can show that the expected value of the terms due to ϵ in the equation is zero. Thus, a regression of transfers plus revenue sharing on incremental surplus—for sufficiently large incremental surplus—would return $1 - \eta$ as the slope and C^R as the constant.

limited search.

A third observation in the baseline model is that plan choice within a recordkeeper is not related to the quality of the second-best option, as the joint surplus is not affected by the second-best option. However, if variation in the outside option induces variation in the characteristics of a chosen plan conditional on the choice of recordkeeper, it must be the case that $\kappa > 0$. Moreover, the fact that the cost of transfers is convex means that the sensitivity of plan choice to the outside option changes with the level of the transfer. Loosely, if transfers are near 0, variation in the outside option primarily translates to variation in transfers in a manner dictated by η, much like the baseline model. Additionally, once the constraint on transfers is imposed again, the probability that the transfer is 0 is informative of η (and the cost of the recordkeeper) as well. As transfers increase, variation in the outside option also induces variation in the characteristics of the chosen plan, in a manner governed by a combination of both η and κ. With knowledge of η and κ, the choice of plan within recordkeeper is still informative of the sponsor's sensitivity to plan quality; instead of maximizing joint surplus, however, the parties maximize the Nash product.

The intuition in the above paragraph is incorporated into the estimation procedure through the choice of moments to match. All moments we target take the form $\mathbb{E}\left[X \cdot \mathbb{1}(s \text{ chooses } r)\right]$: the expectation of the product of some variable X and an indicator for firm s choosing recordkeeper r. The first set of moments are simply properties of the chosen plans and transfers: (i) an indicator for the transfer equalling 0, (ii) the value of the transfer, (iii) each quality dimension $Q_s(p)$ of the chosen plan, and (iv) revenue sharing.

The second set of moments captures correlations between transfers and quantities that would affect surplus of the chosen plan. These moments are (i) interactions between each quality dimension and T, (ii) interactions between revenue sharing and T, and (iii) the same interactions, replacing T with an indicator that $T = 0$. These moments capture the intuition that the correlation between elements of $u^{(1)}$ and transfers is informative of η. The final set of moments mirrors this second set but replaces characteristics of the chosen plan with characteristics of non-chosen plans in the choice set. These moments mirror the classic "rivals' characteristics" instruments of Berry et al. (1995) as shifters of price; as discussed above, here they shift plan choice and transfers, and how they do so informs ρ and κ.

4.3. Estimation

Estimation of this model can be thought of as joint estimation of a demand-side as in Berry et al. (2004) ("micro-BLP"), and a supply-side that allows for the optimal setting of transfers

as the outcome of a bargaining game between employers and recordkeepers. Our procedure involves inverting a share equation for the choice of recordkeepers to recover mean utilities for recordkeepers, and then using these recovered parameters to compute the micro-moments described in Section 4.2. This section provides an outline; see Appendix C for more details.

We begin with a guess of random coefficient parameters for θ_s, of the variance of recordkeeper and sponsor errors σ^R and σ^S, of disutilities from transfers κ, of recordkeeper costs C_r^R, and of search cost probabilities ρ. Call this vector of parameters Ω for notational convenience. Fix these values and consider a particular consideration set. The plan with the highest extractable surplus for each recordkeeper in the consideration set satisfies

$$
V_{rs}^*(\epsilon_{rs}^R; \theta, \sigma^R, \kappa, C_r^R) \equiv \max_{p \in \mathcal{P}_{rs}} \boldsymbol{\theta}_s \cdot \boldsymbol{Q}_s(p) - \max\left[0, C_r^R - R_r(p) - \sigma^R \epsilon_{rs}^R\right]
$$
$$
- \kappa \cdot \max\left[0, C_r^R - R_r(p) - \sigma^R \epsilon_{rs}^R\right]^2. \quad (7)
$$

Note that (7) reflects the impact of two constraints: the recordkeeper's individual rationality constraint and the restriction that transfers cannot be negative. Note that within-recordkeeper plan choice does not depend on sponsor shocks ϵ^S, on unobserved preferences for record-keepers ξ_r, or on sponsor-level administrative costs, as these do not vary within recordkeeper. Because of the lack of a plan-sponsor-recordkeeper shock, our model for within-recordkeeper plan choice is akin to a pure characteristics demand system (Berry and Pakes, 2007). We make this modelling assumption to ensure that the expected utility obtained from a particular recordkeeper does not increase mechanically when the number of plans offered by the recordkeeper increases. This is particularly important in counterfactuals where we restrict the set of admissible plans. Finally, in (7) we impose the scale normalization that the utility of a dollar of transfer payments is equal to -1.

Given a consideration set, the probability sponsor s chooses recordkeeper r is given by

$$
\Pr\left[V_{rs}^*(\epsilon_{rs}^R; \theta, \sigma^R, \kappa, C_r^R) + \delta_r + \epsilon_{rs}^S \geq V_{r's}^*(\epsilon_{r's}^R; \theta, \sigma^R, \kappa, C_{r'}^R) + \delta_{r'} + \epsilon_{r's}^S \text{ for all } r'\right]
$$
$$
= \iiint \frac{\exp(\frac{1}{\sigma^S}(V_{rs}^*(e_r; \theta, \sigma^R, \kappa, C_r^R) + \delta_r))}{1 + \sum_{r' \in \tilde{\mathcal{R}}_s} \exp(\frac{1}{\sigma^S}(V_{r's}^*(e_{r'}; \theta_s, \sigma^R, \kappa, C_{r'}^R) + \delta'_r))} dF(\theta) dF(e) dF(\tilde{\mathcal{R}}_s), \quad (8)
$$

where δ_r. For the sake of clarity, (8) has a slight abuse of notation: the integrand is 0 whenever $r \notin \mathcal{R}_s$. In (8), we impose the location normalization that the mean utility of not having a plan is equal to 0. We compute the integral over the distribution of random coefficients $F(\theta)$ using monomial rules for Gaussian quadrature, the integral over the vector

22

of recordkeeper shocks e by simulation, and the integral over consideration sets using importance sampling.

Denote the left-hand side of (8) by $s_{rs}(\Omega, \delta)$. Averaging over firms yields the market-level shares $s_r(\Omega, \delta) = \frac{1}{S} \sum_s s_{rs}(\Omega, \delta)$. Since this equation satisfies the conditions outlined in Berry et al. (2013), given a guess of the parameters in Ω, we can invert the share equation and recover the vector of mean utilities δ. Having obtained mean utilities, we can compute the aforementioned moments. We estimate the parameters using two-step GMM.

5. Parameter Estimates

Table 2 reports estimates from the model outlined in Section 4. The model is estimated separately for each group of firms, binned by the number of participants. Each column provides a separate set of estimates.

Panel A shows estimates relating to the market imperfections embedded in the model. We find that search probabilities ρ are between 0.59 and 0.73, with larger firms searching slightly less than smaller firms. However, the differences across groups are small and not economically meaningful. Small firms, however, have significantly lower bargaining parameters η than large firms; we estimate the largest firms to have over twice the bargaining power of the smallest ones. Across all groups, bargaining parameters are significantly below 1: sponsors capture much of the incremental surplus generated by plans. While this does not lead to lower participation—if a sponsor bargains with a recordkeeper that offers a plan with positive surplus, they will come to an agreement—it leads to more expensive plans and lower plan quality.

Panel B shows estimates related to willingness to pay. All groups have a value of κ that is greater than zero. To interpret κ, note that the sponsor perceives the marginal cost of an additional dollar of transfers to be $1 + 2 \cdot \kappa \cdot T$ when transfers are at T. Thus, the estimate of $\kappa = 0.025$ for small firms implies that at transfers of \$100 per participant, small firms view an additional dollar of transfers as costing \$6. For large firms, this magnitude is \$2.94. This additional level of sensitivity to transfers gives firms incentives to compensate recordkeepers through higher-fee plans rather than through direct transfers—and this is particularly relevant for small firms.

Panel B next shows estimates of θ, which corresponds to sensitivity to plan quality; the first set of rows shows means while the second set shows standard deviations of the distribution of random coefficients. We find that firms are not especially sensitive to the

	Group 1 ≤ 200	Group 2 201–500	Group 3 501–1000	Group 4 > 1000
A. Market Imperfections Frictions				
ρ	0.73	0.71	0.59	0.65
	(0.004)	(0.006)	(0.003)	(0.004)
η	0.28	0.04	0.36	0.57
	(0.003)	(0.001)	(0.005)	(0.007)
B. Willingness to Pay (κ and θ)				
$\kappa \times 100$	2.49	2.59	1.19	0.97
	(0.019)	(0.019)	(0.018)	(0.025)
Means of θ				
# Options	-22.24	-29.65	-30.32	-50.23
	(0.45)	(0.44)	(0.53)	(0.71)
Has Target Date Fund	866.42	1196.46	811.12	1082.06
	(9.79)	(6.83)	(11.66)	(12.81)
# Index Funds	0.00	0.00	0.10	50.98
	(0.00)	(0.00)	(0.79)	(0.80)
Has S&P 500 Tracker	245.48	806.57	1027.80	1457.04
	(5.19)	(5.48)	(9.78)	(11.83)
Mean Expense Ratio (bp)	0.00	0.00	-7.06	-8.40
	(0.000)	(0.000)	(0.103)	(0.136)
Employer Contributions	-1.21	-1.16	-1.97	-2.34
	(0.006)	(0.007)	(0.009)	(0.012)
Standard Deviations of θ				
Constant	115.91	125.10	113.05	100.68
	(8.64)	(10.89)	(18.57)	(22.64)
# Options	3.27	6.53	7.35	8.65
	(0.11)	(0.16)	(0.20)	(0.31)
Has Target Date Fund	11.05	11.40	9.93	9.17
	(88.22)	(77.15)	(92.56)	(106.32)
# Index Funds	13.68	10.45	10.96	7.13
	(0.510)	(0.613)	(1.213)	(1.199)
Has S&P 500 Tracker	9.88	10.54	10.17	12.45
	(65.59)	(78.03)	(86.97)	(128.90)
Mean Expense Ratio (bp)	13.62	10.58	9.37	9.52
	(1.07)	(0.89)	(2.32)	(3.05)
Employer Contributions	0.05	0.09	0.10	0.10
	(0.077)	(0.076)	(0.129)	(0.152)
σ_S	848.74	799.64	617.52	674.01
	(9.79)	(6.83)	(11.66)	(12.81)
C. Recordkeeper Costs				
σ_R	163.13	71.84	133.48	105.13
	(0.00)	(0.00)	(0.79)	(0.80)
Average Cost (per person)	11.69	9.76	11.19	11.47

Table 2: Parameter estimates. Standard errors are in parentheses.

number of options on a menu, preferring smaller plans by about $20–$50 per option. They are also not especially sensitive to the number of index funds: large sponsors are willing to pay about $50 per index fund, while small employers do not value them. The distribution of preferences for options and the number of index funds is also not especially disperse. On the other hand, we estimate significantly larger willingness to pay for target date funds and S&P trackers. There is no clear pattern by firm size in the willingness to pay for target date funds (and means are about $1,000 per person), but we see a clear pattern that large firms are more willing to select S&P trackers than small ones—with a willingness to pay almost 6 times as large. Moreover, we estimate that large firms prefer low-expense plans: they are willing to pay about $8 per basis point reduction (per participant) in average expenses. For comparison, an reduction in expenses of 1 bp corresponds to yearly savings of $3.14 per participant at the average balance in the data. On the other hand, we estimate that small firms are insensitive to expenses on average. The distribution of preferences for expense ratios, however, is significantly wider than than for other plan characteristic: there is significant heterogeneity in the disutility of expenses.

As discussed in Section 4.2, within-recordkeeper choice of plans informs these parameters: roughly, small firms do not choose plans with low expenses even though they have access to them, and the additional transfer payments cannot explain this choice. The lack of sensitivity of the sponsor to plan quality could come from a variety of sources: limited sophistication of the benefits department of small firms, a lower probability of legal recourse from choosing bad plans, or a lack of sensitivity of workers to plan quality.

The final line of θ reports sensitivity to employer contributions. We find estimates between -1.2 and -2.3, with larger firms being more averse to this expense. If $\kappa = 0$, this coefficient outlines how much the employer trades off a marginal dollar of transfers with a marginal dollar of contributions. It is not a priori clear that this number should be -1: contributions are usually tax-advantaged (a force towards lowering the magnitude of the coefficient), but they may also be less salient to workers than expenses in other benefits (a force towards increasing the magnitude). They may also be harder to target to particular groups of workers than wage increases. Given $\kappa > 0$, the marginal cost of another dollar on transfers is $1 + 2 \cdot \kappa \cdot T$. Compared to this value (and a rough benchmark of transfers of $100), we find that firms are actually less sensitive to contributions than to transfers. This finding is consistent with both an internal budgeting story for κ and the existence of benefits from offering contributions: firms avoid direct transfers and would prefer to spend the money on contributions.

Panel C shows parameters related to the recordkeeper's utility. We estimate per-participant costs of about $9–11, averaged across all recordkeepers. Table D.1 in Appendix D provides estimates at the recordkeeper level. Importantly, we do not find any evidence of economies of scale for the recordkeeper, as costs do not decrease with the group. In fact, the scale of the recordkeeper's cost shock seems significantly larger than the estimated costs, suggesting that idiosyncratic match values are important for choice.

6. Quantifying the Frictions in Plan Design

The parameter estimates identify both market imperfections and misaligned willingness to pay as present in this market. How much do these forces contribute to limited participation and to observed plan quality?

Figure 5 shows how outcomes would change if each friction were alleviated. To simulate these effects, we change model parameters to alleviate frictions one by one and solve for counterfactual equilibria. In doing so, we keep all choice sets $\mathcal{R}_s$ equal to the ones used in estimation, re-draw consideration sets $\tilde{\mathcal{R}}_s$, and re-compute the solution to the Nash bargaining problem. We show results for firms in Group 1 (≤ 200 employees, "small" firms) in light gray and Group 3 (501–1,000 employees, "medium" firms) in dark gray. Each group's bar graphs have as a base value the average value of the outcome in that group in the baseline. The dotted horizontal line corresponds to the outcome for Group 4 (i.e., those with more than 1,000 employees, "large" firms).

Alleviating Market Imperfections. The first set of columns in Figure 5 shows the impact of alleviating market imperfections: we evaluate setting $\rho = 1$, $\eta = 1$, and then both. Setting $\rho = 1$ eliminates the impact of limited search, while setting $\eta = 1$ eliminates recordkeeper market power. Eliminating search frictions affects plan quality through two channels. First, increasing the number of options in the choice set can match sponsors with recordkeepers that offer plans whose quality is better suited to their preferences. Second, eliminating search frictions can increase the utility from the outside option, lowering direct transfers. Since $\kappa > 0$, this reduction in the transfer reduces the incentive to distort plan design towards plans with high revenue sharing and away from what maximizes sponsor utility. Eliminating recordkeeper market power also affects plan quality through this second channel, as it also reduces transfers.

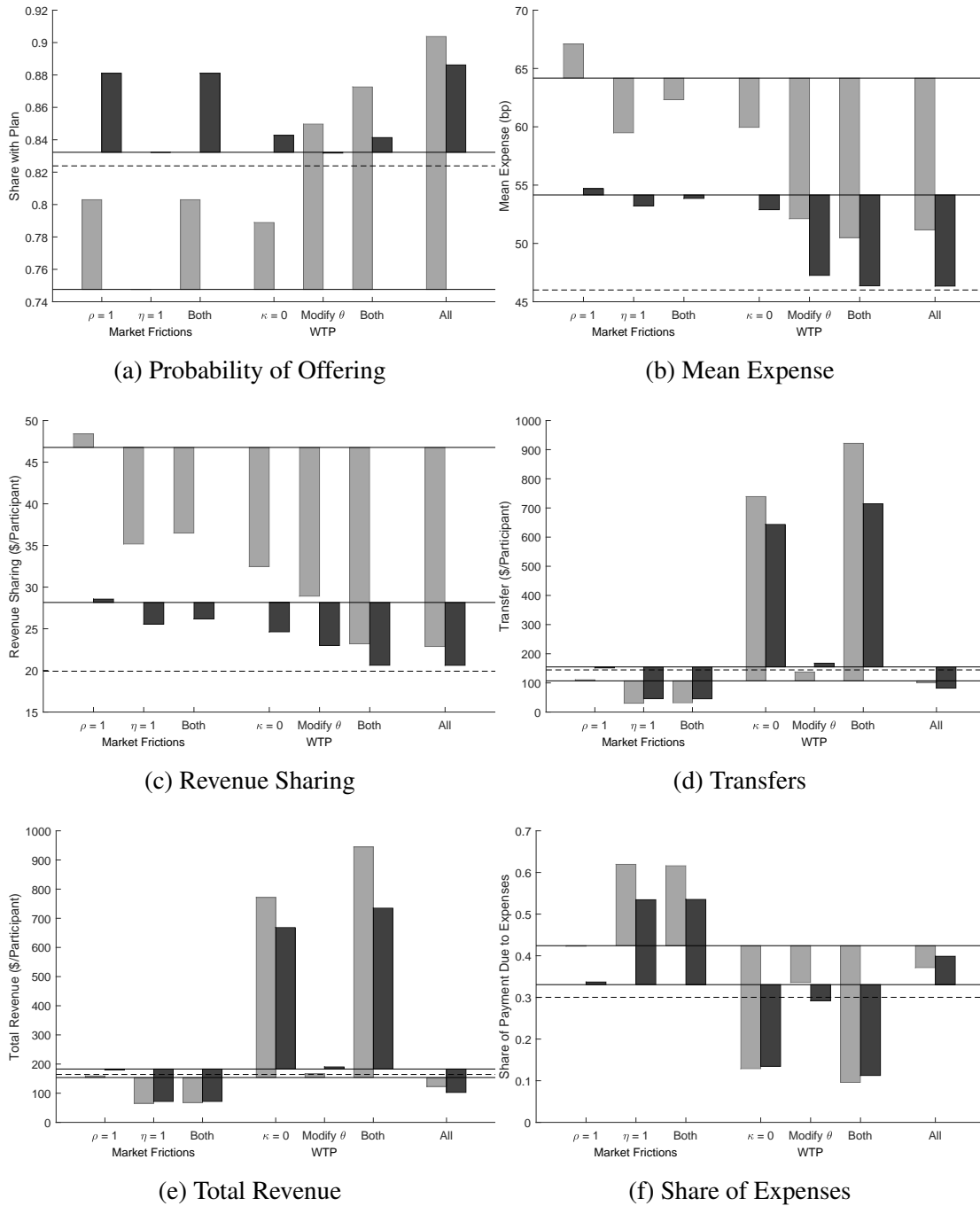Note that neither of these channels necessarily increases plan quality as measured by

(a) Probability of Offering   (b) Mean Expense

(c) Revenue Sharing   (d) Transfers

(e) Total Revenue   (f) Share of Expenses

Figure 5: Effects of alleviating various market frictions, using data from small firms (≤ 200 participants) in light gray and medium firms (Group 3, 501–1000 participants) in dark gray. The bars originate from solid horizontal lines that denotes the mean value of the metric for small or medium firms. The dotted line denotes the mean value for large firms (> 1000 participants).

what an informed observer would deem beneficial to workers; they only increase plan quality as measured by sponsors' willingness to pay. If these two objects are misaligned, alleviating market imperfections could even have adverse effects on worker welfare.

The first column in Figure 5 sets $\rho = 1$. Panel (a) shows that this increases the rate of offering a plan by around 5 pp for both small and medium firms, consistent with an increase in the likelihood of finding a profitable plan for both parties. The remaining panels show metrics conditional on offering a plan. Panel (b) shows that even though sponsors search more, the average unweighted expense ratio of chosen plans increases slightly (by 2.9 bp), and Panel (c) shows a slight increase in revenue sharing, consistent with the idea that expense ratios are correlated with revenue sharing. These results highlight that expanding the choice set does not necessarily improve outcomes for workers, as the choices are still made to maximize the weighted product of the sponsor's utility and recordkeeper profits— and we have already estimated that sponsors in Group 1 do not put high utility on metrics that an informed observer would consider beneficial to workers. Thus, while there are more profitable contracts available and more plan provision, the larger choice sets does not translate to higher quality. Results are similar but smaller for Group 3: these firms experience higher disutility from expenses on average, so the increase in mean expenses conditional on offering a plan is also smaller (about 0.5 bp).

The second column removes recordkeeper market power by setting $\eta = 1$. In this model, market power does not lead to quantity constraints: as long as there is some transfer and plan that is profitable for both the recordkeeper and the sponsor, an agreement will be reached. Accordingly, we see in Panel (a) that the probability of offering a plan is not affected. However, eliminating market power does improve plan quality: Panel (b) shows that mean expenses drop by 4.7 bp for small firms. Panels (c) and (d) shows that both sources of revenue for the recordkeeper drop: revenue sharing drops by about \$12/participant, and transfers drop by a much larger \$77/person. Together, the total revenue to the recordkeeper falls by almost 60%, as shown in Panel (e). We see similar but smaller results for medium firms: expense ratios drop by about 0.9 bp, for instance. This is partly because recordkeepers exert slightly less market power over them and partly because κ is lower, so the feedback from a reduction in transfers to plan quality is lower.

The third column shows the effects of setting both $\rho = 1$ and $\eta = 1$. In our model, this represents the limit of what competition can do to market outcomes. The net effect of moving to this limit on plan quality is small: expenses fall slightly. That is, while transfers and revenue to the recordkeeper fall, this does not translate to significantly better plans for

28

workers; outcomes are nowhere close to the plan quality offered by large employers, for example. Panel (f) provides a metric to summarize these observations. We report the share of total expenses (transfer payments plus dollar payment through expense ratios) that are due to expense ratios; this metric provides the relative incidence of policies on sponsors versus workers. This is on average 0.42 in the baseline for small firms, 0.33 for medium firms, and 0.30 for large firms: larger firms have relatively higher transfers compared to revenues from expense ratios. Alleviating all market imperfections actually increases the relative burden on workers, to 0.61 and 0.54 for Groups 1 and 3 respectively. This suggests that the incidence of market power lies primarily on the sponsor: given sponsor preferences (θ and κ), recordkeepers exercise market power by increasing transfers rather than by depressing plan quality. As transfers are higher than they would be if the market were competitive, an important margin of distortion is the fact that participation is significantly lower than it would be in a competitive market.

Aligning Willingness to Pay.   We have shown that, conditional on sponsor preferences, market power distorts the extensive margin of plan provision but does not significantly distort quality. A natural question to ask is to what extent do sponsor preferences cause distortion. The second set of columns in Figure 5 answers this question alleviating the misalignment in willingness-to-pay for plans. One direct source of misalignment is that $\kappa > 0$, which induces two distortions. The first is on the intensive margin: a positive κ implies that the sponsor is willing to trade off plan features that go into its benefit function if doing so reduces the transfer, especially if the transfer is already high. The second is on the extensive margin. Each recordkeeper has a minimum transfer that would make it willing to serve a particular sponsor; reducing κ makes it more likely that this minimum transfer would be feasible for the sponsor as well and thus increases plan provision.

In line with this intuition, we see that the share of firms with a plan increases if $\kappa = 0$, by 4.1 pp for small firms and 1.0 pp for medium ones. Moreover, we see drops in expense ratios of 6 bp and 2 bp, respectively. Recall that small firms do not value lower expense ratios; this result is due to a shift towards plans that carry target retirement date funds and that have S&P 500 trackers. Interestingly, Panels (c) and (d) show that this improvement in the intensive and extensive margins is paid for by a large increase in transfers; revenue sharing falls by about $16/participant ($4 for medium firms)—almost half the difference to large firms—but transfers increase four- to seven-fold. Clearly, the effect on transfers is much larger than the effect on expenses, and as a result recordkeepers greatly benefit. That

is, the fact that employers place a greater disutility on a dollar of transfers than on a dollar of revenue sharing paid out by their workers distorts both plan quality and coverage, but due to market power recordkeepers capture a large share of the additional surplus that is created when this friction is removed.

The second source of misalignment is that θ is such that sponsors may undervalue quality—either relative to workers or relative to what an expert would choose. Unlike for ρ, η, or κ, there is no natural benchmark for alleviating any frictions induced by a misaligned θ, as it is a priori unclear what the "right" preferences should be. Thus, we consider the experiment of moving θ to that for large firms (and move δ to this level as well to account for changes in the mean benefit of offering plans). Doing so has a large effect on plan provision for small firms, increasing it to beyond 0.9. It also has a larger effect on quality than alleviating market power or setting $\kappa = 0$: mean expenses for small firms drop by 12.1 bp, and those for medium firms drop by 7.0 bp. Revenue sharing drops significantly as well, particularly for small firms, and this drop is compensated with an increase in direct transfers. On net, total revenue to recordkeepers increases slightly. It is striking to find that differences in willingness to pay for plan quality between small and large employers are enough to rationalize the entire gap in outcomes. Of course, aligning preferences of small employers with preferences of large employers may not be feasible, so in Section 7.2 we discuss direct quality regulation.

The final column in the second set of columns in Figure 5 aligns both κ and θ. Here, the effects on quality compound: the extensive margin is slightly larger and mean expenses become similar to those paid by employees of large firms, although this comes at a cost of a large increase in transfers. Overall, fully aligning willingness to pay without addressing market power would benefit recordkepeers greatly as well.

Interaction Between Market Power and Willingness to Pay. The final column in Figure 5 shows the effect of both alleviating market imperfections and aligning willingness to pay. We find that relieving both sources of friction has an additional extensive margin effect. That is, even with aligned preferences there are many employers excluded from plan provision due to market imperfections. Moreover, when recordkeepers have market power, aligning willingness to pay directly increases transfers; when market power is addressed, this is no longer the case. This result highlights that any policy addressing misaligned preferences will be more effective at increasing plan provision if market power and imperfect search are addressed.

7. How Does Regulation Affect Plan Quality?

The results of Section 6 help understand the relative contribution of various frictions to the equilibrium we observe. Eliminating market frictions improves both provision and quality, but directly changing preferences has a much larger effect. This observation guides our analysis of counterfactual regulation. As discussed below, some of them can be viewed as attempt to address market imperfections indirectly while others attempt to target quality directly. We analyze such policies in Sections 7.1 and 7.2, respectively, and discuss the results together in Section 7.3. The takeaways from this section are in line with the lessons from Section 6: policies that directly promote or constrain plan quality are necessary to appreciably change it, and indirect measures are much less effective.

7.1. Targeting Market Imperfections

Targeting market imperfections through regulation broadly involves either increasing the level of competition, or indirectly increasing η or ρ. In this section, we consider two avenues to achieve these goals: (i) direct subsidies for plan provision and (ii) policies that let smaller firms "mimic" larger firms. Either avenue would reduce the transfers paid by sponsors and thus make them more willing to substitute to higher-quality plans.

Both approaches are included in the SECURE Act, signed into law in December 2019, which instituted a set of disparate changes to the private retirement landscape (Commito, 2020). One of its provisions provides tax credits of up to \$250/participant-year to subsidize the provision of DC plans—albeit the focus is on especially small plans (effectively fewer than 20 employees). Here, we consider the impact of expanding subsidies to larger firms, which would both directly incentivize provision and encourage competition. Another provision allows small employers to band together to form "multiple-employer plans" (MEPs) or pooled-employer plans, changing an earlier Department of Labor rule that drastically limited the ability of firms to form such plans.[12] One rationale for MEPs is to allow recordkeepers and small firms to realize economies of scale in plan provision; if these are large, it would directly increase plan provision and potentially also increase competition by increasing the set of recordkeepers willing to serve the sponsor. Another rationale is that these MEPs would give small firms more buying power to negotiate with plan providers,

---

[12]Previously, firms were only allowed to form them if they were tightly linked, e.g., in the same industry and geography ("nexus"). Moreover, such plans were subject to the "one bad apple" rule, which made all firms liable for a fiduciary violation from one of them. Industry participants claimed this sharply increased costs.
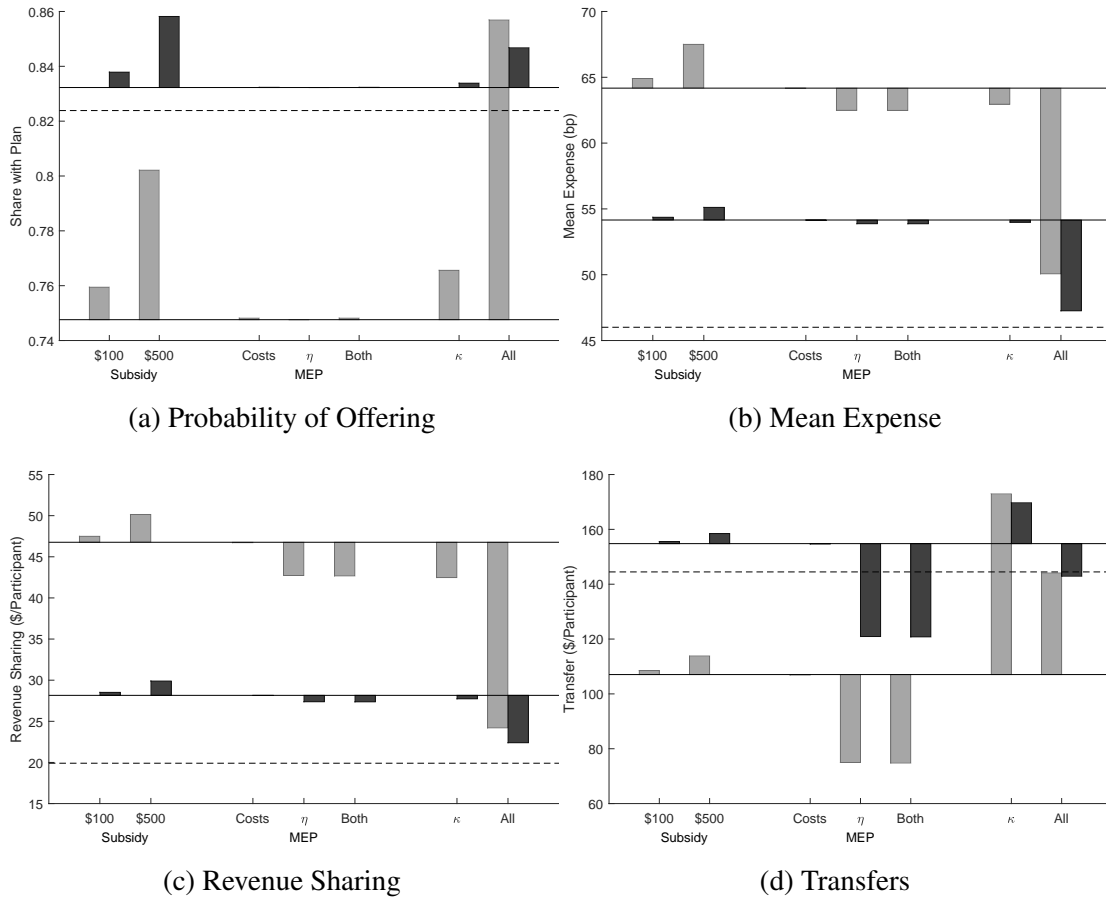
(a) Probability of Offering

(b) Mean Expense

(c) Revenue Sharing

(d) Transfers

Figure 6: Effects of changing parameters for small firms ($\leq$ 200 participants) in light gray and medium firms (Group 3, 501–1000 participants) in dark gray to those of large firms ($>$ 1000 participants). The final two columns show the effects of per-participant subsidies to plan providers for serving small firms. The solid horizontal lines denote the mean value of the metric for small and medium firms. The dotted line denotes the mean value for large firms.

providing one mechanism to increase η.[13]

Subsidizing Plan Provision.   Subsidizing sponsors that provide plans obviously increases coverage. However, such subsidies will affect equilibrium transfers—both directly and through the fact that the sponsor's disagreement point could change—and thus sponsors'

---

[13]For commentary from industry participants discussing this rationale, see "SECURE Act Alters 401(k) Compliance Landscape" (SHRM, 2020) at https://www.shrm.org/resourcesandtools/hr-topics/benefits/pages/secure-act-alters-401k-compliance-landscape.aspx or "How the SECURE Act is Changing Retirement" (QRPS, 2020) at http://www.qrps.com/Resources/Articles/ArticleID/1371/How-the-SECURE-Act-is-Changing-Retirement.

willingness to provide high-quality plans. We investigate the effects of subsidies around the levels proposed by the SECURE Act.[14]

The first set of columns for Figure 6 shows the effects of subsidies of $100 and $500 per participant. They have a direct impact on provision, with $500 subsidies increasing participation for small firms by about 6 pp, or roughly three-fourths of the distance to that of large firms. The increase for medium firms is 2.6 pp. However, subsidies do not improve plan quality, and we actually find slight increases in mean expenses, of 3.3 bp for small firms and 1.0 bp for large firms. Without changes to θ or a significant decrease in transfers paid, which we do not observe, the effective willingness to pay for quality does not change. The result suggests that without direct changes to the benefit function, it is difficult to improve plan quality significantly.

Effects of MEPs. Our general approach to modeling MEPs is to change certain parameters of small firms to those of large firms while keeping choice sets and other parameters fixed. Since the policy intends to exploit potential economies of scale, we start with the recordkeeper's cost parameters. The first column of Figure 6 shows that in our setting, however, the effects of such cost changes are negligible, mostly because the cost differences between large and small firms are also negligible (see Table 2).

A natural question is what happens if the sponsors' costs change as well. Our empirical strategy allows us to estimate δ, the mean net benefit of offering a plan with each record-keeper, but does not separate costs and benefits within δ. However, sponsor subsidies are isomorphic to cost reductions, so even a cost reduction of $100 per worker will not lead to large increases in participation or quality.

What if the effect of MEPs is to improve buying power? We can model this directly by increasing the bargaining power of small firms to those of large ones. Accordingly, this counterfactual is a less extreme version of the one presented in Section 6 to alleviate market power. We thus find similar, but smaller, results: transfers and revenue sharing decrease, but we find little impact on expenses: they decrease by 1.6 bp for small firms and are mostly unchanged for large firms. Putting them together, we see that the net effect of MEPs is almost identical to the effect of simply changing the bargaining parameter.

---

[14]We make these subsidies nominally incident on the sponsor to capture the setup of the SECURE Act. In a world with efficient bargaining ($\kappa = 0$ and no positive transfer constraint), nominal incidence of the subsidy would not matter for real outcomes. However, in our model, nominal incidence does matter; a subsidy incident on the recordkeeper is more likely to reduce transfers paid by the sponsor, for instance. We have checked that making the subsidy incident on the recordkeeper has qualitatively similar and quantitatively slightly larger effects on plan provision, with little impact on plan quality.

Thus, even though the policy motivation for MEPs comes from the size gradient documented in Figure 4, our analysis suggests that targeting the parameters most likely to be affected does not have an impact on plan quality or coverage. Which parameters actually generate the observed differences between small and large employers? To answer this question, the second set of columns in Figure 6 changes other parameters from their values for Groups 1 or 3 to those for Group 4. The first column changes κ; again, this is a less extreme version of the analysis in Section 6, and we see small increases in the share with a plan and a small improvement in plan quality, coupled with an increase in direct transfers. The next column repeats the exercise in Section 6 of changing θ: relative to all other parameters, this has by far the largest impact on quality. The final column changes all parameters of small firms and moves them to the counterparts for large ones. We see that outcomes are quite similar; the remaining differences are due to heterogeneity in choice sets and firm characteristics (which affect predicted mean expenses and revenue sharing).

Overall, we conclude that for MEPs to be effective at improving plan quality, they must cause small employers to adopt the willingness to pay for plan quality (θ) of large employers. One might envision policies that directly target θ. For instance, small firms might not value S&P indices or low expense ratios because of limited sophistication by their decision-makers, or because their workers do not value them and thus the firm does not accrue any labor market benefits from putting them on the plan; in such cases, informational policies may help. Alternatively, small firms may not face much litigation risk for lower-quality menus, in which case reforms may address legal recourse options available to individual workers. On net, however, we find it implausible that these policies could completely close the gap in willingness to pay between small and large employers, as significant labor market differences will always remain. Thus, we conclude that MEPs are unlikely to effectively close the outcome differences between small and large employers.

7.2. Direct Regulation of Quality

Targeting market imperfections provides an indirect method to regulating quality. A more direct approach would constrain the set of plans a sponsor can offer, or provide subsidies or penalties as a function of plan quality. As discussed in Section 2.2, no regulation currently imposes such restrictions. Nevertheless, this approach could be envisioned as a natural strengthening of the guidelines provided by the IRS or DOL in certain situations. For instance, regulators could impose "must-carry provisions," which mandate that all plans contain investments that would be appropriate as a QDIA, rather than simply offering safe
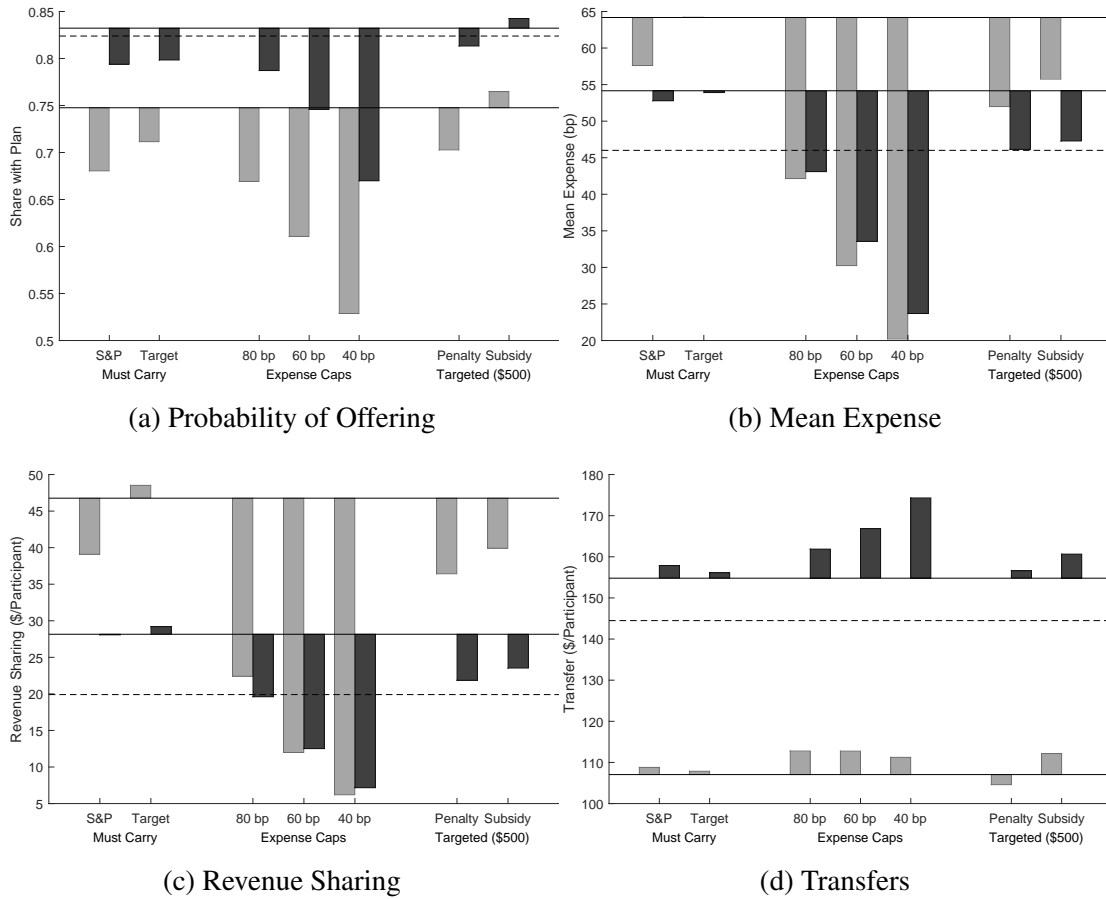
(a) Probability of Offering	(b) Mean Expense
(c) Revenue Sharing	(d) Transfers

Figure 7: Effects of must-carry provisions, expense caps, and targeted subsidies or penalties for small firms (≤ 200 participants) in light gray and medium firms (Group 3, 501–1000 participants) in dark gray to those of large firms (> 1000 participants). The solid horizontal lines denote the mean value of the metric for small and medium firms. The dotted line denotes the mean value for large firms.

harbor provisions for offering a QDIA. Regulation could also introduce "expense caps," prohibiting recordkeepers from plans with high average expense ratios.

In this subsection, we evaluate policies that directly regulate plan quality. The fundamental trade-off here is between the intensive and extensive margin. On the one hand, such policies mechanically improve plan quality conditional on selecting a plan—at least on the dimension that is regulated. On the other, imposing constraints on acceptable plans makes it less likely a sponsor offers a plan: a sponsor may have to compensate a recordkeeper more when plan quality is improved, and it may be unwilling to do so—both because it perceives a lower benefit from it and because $\kappa > 0$ and it prefers to avoid high transfers. A less drastic measure is to target subsidies to sponsors specifically for plans that meet certain criteria.

Figure 7 shows results from these policy experiments. We implement the must-carry provisions and expense caps by keeping parameters at the estimated levels for Group 1 or 3 but modifying the choice set of plans by eliminating any plan that does not meet the must-carry provision. We also consider targeted subsidy policies for expense caps; for this counterfactual, we retain the same choice set as in the data but add a fixed dollar amount (per participant) to the sponsor's payoff if the plan has average expenses less than 50 bp.

Must-Carry Provisions and Expense Caps. We first consider two simple and easily implementable must-carry provisions—mandating either an S&P index fund or a target date fund—modeled after strengthening QDIAs. We find that mandating an S&P index fund lowers the probability of offering a plan by 6.8 pp for small firms (3.9 pp for medium firms), and it also reduces expenses by 6.6 bp (and 1.3 bp for medium firms). Consistent with revenue sharing correlating with expense ratios, we see a drop of about $8/participant for revenue sharing for small firms, which is partially compensated by a small increase in transfers. The effects of mandating that all plans carry target date funds are even more muted: we estimate a drop in plan offerings of about 3.6 pp, and much smaller changes in other quantities. This does not mean the plan themselves do not change: by construction, all selected plans would carry either an S&P or a target date fund.

Why do these must-carry provisions have somewhat small effects on expenses, especially for medium firms? The problem is fundamentally one of targeting the wrong metric. Simply including an S&P index or a target date fund does not ensure that workers will invest heavily in these funds, nor does it ensure that higher-fee funds cannot be on the menu. Thus, we find that sponsors choose similar types of plans on dimensions such as expense ratios even though the plan now contains these mandated funds.

These results suggest that if the goal is to lower expenses, a policy should directly target them. Thus, we evaluate a expense ratio cap, which forces sponsors to only offer plans whose mean expense ratios are lower than a specified amount. In Figure 7, we consider successively stringent caps of 80 bp, 60 bp, and 40 bp. We focus on the (relatively mild) cap of 80 bp. Imposing such an expense cap reduces expenses to 42–43 bp for both small and medium firms. We see corresponding drops in revenue sharing, to about $20-22/person, with an increase in transfers about half as large. Essentially, this cap eliminates the right tail of plans (with especially high expense ratios), causing them to either adjust their choices or cease to provide a plan altogether. Indeed, we see a reduction in plan provision (7.8 pp and 4.5 pp for Groups 1 and 3, respectively). More stringent expense caps only amplify these

effects: an expense cap of 40 bp reduces mean expenses for small firms to 20 bp, but only 53% of small firms continue to offer a plan. This observation is relevant for the industry: many small firms must be able to offer low-quality plans in order to provide them at all. Of course, this extensive margin effect interacts with search imperfections: coupling expense caps with mechanisms for facilitating search would mitigate some of the extensive margin losses.

Targeted Penalties and Subsidies. A perhaps better-targeted version of a must-carry provision is to subsidize firms for offering certain kinds of plans (e.g., those with low expenses)—or to penalize them for not doing so. We consider policies that subsidize plans with expense ratios lower than 50 bp (or penalize the opposite) and call these reforms targeted penalties and subsidies. The final set of columns in Figure 7 shows the effects of $500/participant targeted penalties and subsidies. A penalty generally has a qualitatively similar effect on outcomes as an outright cap. Quantitatively, the effect of the penalty is about half of that of the effect of an expense cap of 80 bp. Of course, the penalty need not be this large, and Figure 8 (discussed in Section 7.3 below in more detail) shows that certain values of targeted penalties can dominate must-carry provisions on plan quality and expense. A $500 subsidy, on the other hand, has slight smaller effects on mean expenses and revenue sharing, but it leads to a small increase in plan provision (of about 1–2 pp). The government must balance these benefits with the cost of the transfer itself.

7.3. Discussion

Figure 8 summarizes the results of the policy counterfactuals for small firms (Group 1). For each policy, it plots in black the mean expense ratio of chosen plans against the share of firms who carry a DC plan. The dashed vertical and horizontal lines indicate the baseline values for these firms. Since mean expense ratios are plotted in descending order, any policy in the top right quadrant is preferred on these metrics, while those in the top left or bottom right are not unambiguously comparable. A variety of counterfactual policies considered in the previous two subsections are plotted with black crossmarks, and outcomes with subsidies (targeted or not) are plotted as dotted and dashed lines connecting various values of subsidies.

As shown in the above section, must-carry provisions improve expenses at the cost of reductions in plan provision. Targeted penalties, plotted as dashed black lines, can deliver higher coverage for any given level of plan quality than must-carry provisions, as do expense
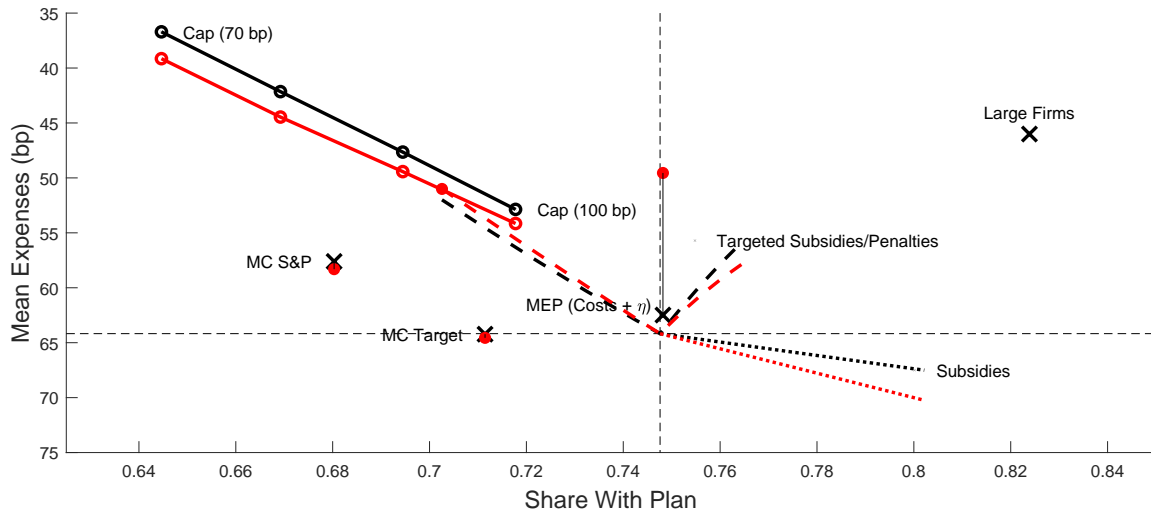
Figure 8: Summary of policy counterfactuals. The dotted axes indicate the values for Group 1 firms. The black crossmarks show values for must-carry policies (MC) and multi-employer plans (MEP). The bold dotted line shows targeted penalties or subsidies (from a $500 penalty to a $500 subsidy) for plan with mean expenses under 50 bp. The dotted bold lines show the effects of untargeted subsidies (from $0 to $500). The red circles and lines show the associated "effective" expense ratios of each of the policies if transfers are passed through entirely to workers.

ratio caps. However, neither of these policies is immune to a drop in coverage. Targeted subsidies are the only policy that improves both provision and expenses—but this of course comes at a cost to the government, and even a $500/person subsidy fails to improve quality and provision to the level seen in large firms. Untargeted subsidies, plotted as dotted black lines, increase plan provision but increase expenses slightly. The takeaway from the analysis in this paper is that the main barrier to small firms offering high-quality plans is not market power or aversion to transfers but rather limited sensitivity to plan quality. In the absence of direct methods to increase this sensitivity, the policy options to address quality must either constrain plan choice (and thus reduce provision) or subsidize it directly (which would be expensive).

One limitation of our analysis is that we may not capture all impact of these policies on workers, as one may expect changes in the quality of DC plans to pass through to wages. Without data on wages, we are unable to estimate this elasticity, and despite the fact that the substitution between wages and non-wage benefits is a classic question in labor economics (Woodbury, 1983), we are unaware of any such estimates in the literature in the context of DC plans. Moreover, in our setting we find it a priori unlikely that wages would be affected significantly by changes in plan quality. Workers tend to be unaware of the quality of their

38

defined contribution plan at the time of selecting jobs, other than matching, and even upon joining there still seem to be significant behavior frictions to participating in the DC plan (Choi, 2015). Additionally, we have estimated that κ is large. If a firm were aware that it could save on wages by improving plan quality, we might expect it to try to at least partially alleviate whatever internal budgeting constraint prevents it from paying for quality.

A more plausible caveat to the results is that we have thus far taken a stark approach to the incidence of counterfactual policies. We view this as an appropriate benchmark, as the employee pays expense ratios and the sponsor directly pays transfers. However, one could ask if the conclusions in this analysis change if the sponsor does not fully bear the predicted changes in transfers. To investigate this question, we consider the opposite benchmark in which the sponsor passes through all transfers to workers.[15] We implement this calculation by computing the basis-point equivalent change in transfers: if a policy causes expenses to decrease by 25 bp and transfers to increase by $10/participant relative to the baseline rate, and mean balances are $10,000/participant, we say that effective expenses only decrease by 15 bp.

Figure 8 also illustrates, in either red circles or red lines, effective expense ratios assuming full pass-through of transfers. For must-carry provisions, the quantitative impacts are small. Pass-through of transfers affects MEPs more strongly: we had identified that if MEPs increase buying power (η), they would reduce transfers but not expenses. However, these transfer savings would translate to reductions of about 15 bp in effective expense ratios if they were fully passed-through to workers. Finally, pass-through mostly does not affect outcomes for targeted subsidies and penalties, and slightly increases expenses under caps and subsidies. Overall, our general finding does not change: plan quality distortions come from misalignment in willingness to pay for plan quality, so directly targeting quality is the most effective policy, although it comes with an extensive margin trade-off.

8. Conclusion

This paper develops a model of the bargaining between employers and recordkeepers that lead to the design of defined contribution plans, the largest stock of retirement savings in the US. While employers are charged with keeping their workers' best interests in mind when

---

[15]A sponsor could in principle adjust wages to reflect changes in transfers, although we find such fine tuning of wages practically difficult and unlikely, as discussed above. However, since it is possible for plans to contain direct fees participants must pay recordkeepers, adjusting such fees (or introducing them if they do not exist) may be one way to pass through transfers.

designing these plans, there is significant policy concern over the prevalence of low-quality plans. Using data on plan menus and payments, we show that both market imperfections and a misalignment in willingness to pay (which incorporates agency frictions between the employer and the worker) contribute to lowering plan quality. However, alleviating market imperfections has its limits: although it can increase plan provision for small firms to near the level of large ones, it does not have nearly as large an effect on plan quality. Instead, the quality of plans is determined largely by the employer's willingness to pay for quality, which could be derived from its own sophisitication, its worker's preferences, or the labor market power it has over its workers.

These observations guide our analysis of a variety of proposed policies. In line with the previous results, we find that policies that target competition—by reducing costs, increasing small employers' bargaining power, or providing subsidies—are of limited benefit in terms of plan quality. We conclude that regulation must directly target quality itself, either by constraining the design of plans or by subsidizing or penalizing certain plan features, if it is to be effective at improving quality.

We view this paper as broadly contributing to our understanding of the market for non-wage benefits, which has largely focused on health insurance so far. In general, employers offer a variety of non-wage benefits, which are provided by other companies. This paper highlights that market frictions are present in other settings as well, and regulation can have real and economically significant impacts on workers. We view this area as a fruitful avenue of research in the future.

References

ALLEN, J., R. CLARK, AND J.-F. HOUDE (2014): "The Effect of Mergers in Search Markets: Evidence from the Canadian Mortgage Industry," *American Economic Review*, 104, 3365–96.

——— (2019): "Search Frictions and Market Power in Negotiated-Price Markets," *Journal of Political Economy*, 127, 1550–1598.

BACHAS, N. (2019): "The Impact of Risk-Based Pricing and Refinancing on the Student Loan Market," Working Paper.

BADOER, D. C., C. P. COSTELLO, AND C. M. JAMES (2020): "I Can See Clearly Now: The Impact of Disclosure Requirements on 401(k) Fees," *Journal of Financial Economics*, 136, 471–489.

BENARTZI, S. AND R. THALER (2007): "Heuristics and Biases in Retirement Savings Behavior," *Journal of Economic Perspectives*, 21, 81–104.

BENARTZI, S. AND R. H. THALER (2001): "Naive Diversification Strategies in Defined Contribution Saving Plans," *American Economic Review*, 91, 79–98.

BERRY, S., A. GANDHI, AND P. HAILE (2013): "Connected Substitutes and Invertibility of Demand," *Econometrica*, 81, 2087–2111.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 841–890.

——— (2004): "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 112, 68–105.

BERRY, S. AND A. PAKES (2007): "The Pure Characteristics Demand Model," *International Economic Review*, 48, 1193–1225.

BESHEARS, J., J. J. CHOI, D. LAIBSON, AND B. C. MADRIAN (2013): "Simplification and Saving," *Journal of Economic Behavior & Organization*, 95, 130–145.

BHATTACHARYA, V., G. ILLANES, AND M. PADI (2020): "Fiduciary Duty and the Market for Financial Advice," Tech. rep., Northwestern.

BRANCACCIO, G. AND K. KANG (2021): "Search Frictions and Product Design in the Municipal Bond Market," Tech. rep., NYU.

BRANCACCIO, G., D. LI, AND N. SCHÜROFF (2020): "Learning by Trading: The Case of the US Market for Municipal Bonds," Working Paper.

CHOI, J. J. (2015): "Contributions to Defined Contribution Pension Plans," *Annual Review of Financial Economics*, 7, 161–178.

CHOI, J. J., D. LAIBSON, AND B. C. MADRIAN (2009): "Mental Accounting in Portfolio Choice: Evidence from a Flypaper Effect," *American Economic Review*, 99, 2085–95.

——— (2011): "$100 Bills on the Sidewalk: Suboptimal Investment in 401(k) Plans," *Review of Economics and Statistics*, 93, 748–763.

CHOI, J. J., D. I. LAIBSON, AND B. C. MADRIAN (2005): "Are Empowerment and Education Enough? Underdiversification in 401(k) Plans," *Brookings papers on economic activity*, 2005, 151–213.

CLARK, R., J.-F. HOUDE, AND J. KASTL (2021): "The Industrial Organization of Financial Markets," in *Handbook of Industrial Organization*, Elsevier, vol. 5, 427–520.

COLLINS, S., S. HOLDEN, J. DUVALL, AND E. B. CHISM (2016): "The Economics of Providing 401(k) Plans: Services, Fees, and Expenses, 2015," *ICI Research Perspectives*, 22.

COMMITO, T. F. (2020): "SECURE Act Enhances Small Employer 401(k) Plans," *Journal of Financial Service Professionals*, 74, 12–15.

CRAWFORD, G. S., R. S. LEE, M. D. WHINSTON, AND A. YURUKOGLU (2018): "The Welfare Effects of Vertical Integration in Multichannel Television Markets," *Econometrica*, 86, 891–954.

CRAWFORD, G. S. AND A. YURUKOGLU (2012): "The Welfare Effects of Bundling in Multichannel Television Markets," *American Economic Review*, 102, 643–85.

CUESTA, I. AND A. SEPÚLVEDA (2019): "Price Regulation in Credit Markets: A Trade-off between Consumer Protection and Credit Access," Working Paper.

CUESTA, J. I., C. NOTON, AND B. VATTER (2019): "Vertical Integration Between Hospitals and Insurers," Tech. rep., Stanford University.

DI MAGGIO, M., M. L. EGAN, AND F. FRANZONI (2021): "The Value of Intermediation in the Stock Market," *Journal of Financial Economics*, Forthcoming.

DOELLMAN, T. W. AND S. H. SARDARLI (2016): "Investment Fees, Net Returns, and Conflict of Interest in 401(k) Plans," *Journal of Financial Research*, 39, 5–33.

EGAN, M. (2019): "Brokers vs. Retail Investors: Conflicting Interests and Dominated Products," *Journal of Finance*, 74, 1217–60.

EGAN, M. L., A. MACKAY, AND H. YANG (2021): "What Drives Variation in Investor Portfolios? Evidence from Retirement Plans," Tech. rep., National Bureau of Economic Research.

EINAV, L., M. JENKINS, AND J. LEVIN (2012): "Contract Pricing in Consumer Credit Markets," *Econometrica*, 80, 1387–1432.

FRIEDMAN, J. H. (2002): "Stochastic Gradient Boosting," *Computational Statistics & Data Analysis*, 4, 367–78.

GOEREE, M. S. (2008): "Limited Information and Advertising in the U.S. Personal Computer Industry," *Econometrica*, 76, 1017–1074.

GOWRISANKARAN, G., A. NEVO, AND R. TOWN (2015): "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry," *American Economic Review*, 105, 172–203.

GRENNAN, M. (2013): "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices," *American Economic Review*, 103, 145–77.

GRUNEWALD, A., J. LANNING, D. LOW, AND T. SALZ (2019): "Auto Dealer Loan Intermediation: Consumer Behavior and Competitive Effects," Working Paper.

GUISO, L., A. POZZI, A. TSOY, L. GAMBACORTA, AND P. E. MISTRULLI (2022): "The cost of steering in financial markets: Evidence from the mortgage market," *Journal of Financial Economics*, 143, 1209–1226.

H2O.AI. (2020): "h2o: R Interface for H2O. R package version 3.30.0.6." .

HO, K. AND R. S. LEE (2017): "Insurer Competition in Health Care Markets," *Econometrica*, 85, 379–417.

——— (2019): "Equilibrium Provider Networks: Bargaining and Exclusion in Health Care Markets," *American Economic Review*, 109, 473–522.

HUBERMAN, G., S. S. IYENGAR, AND W. JIANG (2007): "Defined Contribution Pension Plans: Determinants of Participation and Contributions Rates," *Journal of Financial Services Research*, 31, 1–32.

HUBERMAN, G. AND W. JIANG (2006): "Offering Versus Choice in 401(k) Plans: Equity Exposure and Number of Funds," *Journal of Finance*, 61, 763–801.

ILLANES, G. (2017): "Switching Costs in Pension Plan Choice," Working Paper.

ILLANES, G. AND M. PADI (2021): "Retirement Policy and Annuity Market Equilibria: Evidence from Chile," Tech. rep., National Bureau of Economic Research.

IYENGAR, S. S. AND E. KAMENICA (2010): "Choice Proliferation, Simplicity Seeking, and Asset Allocation," *Journal of Public Economics*, 94, 530–539.

KOIJEN, R. S. J. AND M. YOGO (2015): "The Cost of Financial Frictions for Life Insurers," *American Economic Review*, 105, 445–75.

——— (2016): "Shadow Insurance," *Econometrica*, 84, 1265–1287.

KRONLUND, M., V. K. POOL, C. SIALM, AND I. STEFANESCU (2021): "Out of Sight No More? The Effect of Fee Disclosures on 401(k) Investment Allocations," *Journal of Financial Economics*.

LARSEN, B. J. (2021): "The Efficiency of Real-World Bargaining: Evidence from Wholesale Used-Auto Auctions," *Review of Economic Studies*, 88, 851–882.

LUCO, F. (2019): "Switching Costs and Competition in Retirement Investment," *American Economic Journal: Microeconomics*, 11, 26–54.

NELSON, S. (2020): "Private Information and Price Regulation in the US Credit Card Market," Working Paper.

PAPKE, L. E. (2004): "Individual Financial Decisions in Retirement Saving Plans: The Role of Participant-Direction," *Journal of Public Economics*, 88, 39–61.

POOL, V. K., C. SIALM, AND I. STEFANESCU (2016): "It Pays To Set the Menu: Mutual Fund Investment Options in 401(k) Plans," *Journal of Finance*, 71, 1779–1812.

——— (2021): "Mutual Fund Revenue Sharing in 401(k) Plans," Tech. rep., Vanderbilt.

ROBLES-GARCIA, C. (2020): "Competition and Incentives in Mortgage Markets: The Role of Brokers," Working Paper.

SCHEMBARI, J. (2022): "The Supreme Court Rules for Participants in Hughes v. Northwestern," .

SETHI-IYENGAR, S., G. HUBERMAN, AND W. JIANG (2004): "How Much Choice is Too Much? Contributions to 401(k) Retirement Plans," in *Pension Design and Structure: New Lessons from Behavioral Finance*, ed. by O. S. Mitchell and S. P. Utkus, Oxford University Press, chap. 5, 83–95.

SIMON, J. (2017): "Morningstar reports record low in fund fees," .

STROUD, A. (1971): *Approximate Calculation of Multiple Integrals*, London: Prentice Hall.

WOODBURY, S. A. (1983): "Substitution Between Wage and Nonwage Benefits," *American Economic Review*, 73, 166–182.

ONLINE APPENDICES

A. Data Appendix

Appendix A.1 describes the data sources, and Appendices A.2 and A.3 describes the cleaning procedure use to extract information from the sources and arrive at the final dataset for estimation.

A.1. Data Sources

All firms that offer non-wage benefits are required to file a version of Form 5500 annually with the Department of Labor to comply with the requirements of ERISA. Plans with fewer than 100 participants are required to file a shorter form, Form 5500-SF, which contains a subset of the information in Form 5500. Our dataset comes from these public filings, but we import it from two sources. First, the DOL provides machine-readable data for a large subset of attachments in Form 5500 for direct download.[16] However, some information is submitted to the DOL by the firms in the form of PDF files without any common structure; these pdf files are also public and can be downloaded from the DOL's EFAST filing system.[17] Our second data source consists of these pdf files; we purchase data from BrightScope that extracts and cleans plan menus and asset allocations for plans from these pdf files.

The body of Form 5500 contains the employer's identification number (EIN), plan number, industry codes, main location, and participant counts. It also includes pension and welfare benefit codes, which allow us to identify firms that offer health insurance without offering a retirement plan or ones that offer defined benefit plans. Form 5500-SF also contains all this information.

Schedule C of Form 5500 (but not Form 5500-SF) contains information about service providers and the fees that they collect. Items 1 and 3 of Part 1 of the Schedule C contain information about which service providers received "indirect" compensation from the fund. The exact amount of the compensation is included in a format that is not easily machine-readable, and in Appendix A.3 we outline how we extract it. Item 2 lists service providers that received direct compensation for services, including payment amounts. This is the data source for what we refer to as transfers in the model. We use the payment to the

---

[16]This data can be accessed at https://www.dol.gov/agencies/ebsa/employers-and-advisers/plan-administration-and-compliance/reporting-and-filing/form-5500.

[17]See https://www.efast.dol.gov/welcome.html.

party designated as the main recordkeeper (see Appendix A.2) as our measure of T, but descriptive results are similar if we use total payment to all service providers as the measure.

Finally, Schedule H of Form 5500 contains information the plan's general financial positions—including total balances, contributions (by employers and employees), and expenses. The machine-readable data from the DOL contains this information in aggregate at the plan level. In addition, firms are also required to report information about the plan in Line 4(i) of this schedule. However, responses to Line 4(i) are in pdf files, and this is the information we obtain from BrightScope (who processes these pdf files into a machine-readable format). BrightScope records investment menus, asset allocations, and contributions by investment. It also adds tickers for publicly traded mutual funds and includes expense ratios and other fees of these funds, along with a description of investment styles.

Our sample consists of all firms who file a version of Form 5500 to declare non-wage benefits. We exclude firms that offer defined benefit plans (even if they offer defined contribution plans), as we do not wish to model the interaction between the two types of plans. We focus on plans sized between 100 and 5,000 participants for the main analysis in the paper; for firms that do not offer a defined contribution plan, we take the number of participants in the benefit that they do offer (usually health insurance) as the plan size. We drop plans that report matching outlays higher than $35,000 per participant, revenue sharing outlays higher than $1,000 per participant, expense ratio outlays higher than $3,000 per participant, or direct transfers higher than $10 million. This corresponds to 102 observations. We also drop 55 plans reporting negative values for one of these outlays, as these are likely data entry mistakes.

A.2. Identifying the Recordkeeper

In the model, we consider bargaining between the sponsor and the recordkeeper—treating the recordkeeper as the "main" entity involved in the negotiation. To identify the main provider, we use BrightScope data on service providers; this is based on Schedule C of Form 5500, which lists the services provided by each firm for the plan (based on a set of pre-specified service codes). If a firm is listed as providing a trustee service, we designate it the main provider. If no such firm is listed, we choose the firm listed as the recordkeeper. We then go down a pre-specified list of roles (e.g., "investment management services" and then "investment advisory services") until we find a firm that fulfills one of these roles. We have checked that variations of this order do not appreciably change the designated main

providers, since the vast majority of plans have one firm designed as either a trustee or a recordkeeper. This choice is relevant for the estimation procedure since (i) we allow costs and mean utilities to depend on the identity of the main provider, and (ii) we treat a firm as capturing the full expense ratio from funds with which it is vertically integrated.

A.3. Extracting Revenue Sharing Agreements

The primary source of revenue sharing agreements is Line 3 of Schedule C, which includes information on indirect compensation, including the type of payment, the payer and the payee, and some information about the formula for payment. There is no consistent reporting format for the formula, so we developed a series of regular expressions to extract as much information from the explanations as possible. We are left with either (i) a single number corresponding to the revenue sharing rate for that particular payer, (ii) a set of rates corresponding to multiple payers, or (iii) some indication that the revenue sharing agreement cannot be extracted from the data in Line 3.[18]

Since revenue sharing is given in rates, we have to merge these back to the menu to compute the total amount of revenue sharing. The payer listed in Line 3 can be a fund provider or a specific fund. In the case in which the payer is a fund provider, we assume that all funds provided by that provider have the same revenue sharing rate. In the case where the payer is a fund, we attempt to match back to the plan menu. Since the fund name is listed differently in Line 3 than in BrightScope, we first match the fund name to a ticker.[19] We then match the ticker to a row in the plan menu by merging on ticker if possible, and using minimum distance matching if not. If we are unable to match on tickers, we match funds in the plan menu to Line 3 based purely on the fund provider, averaging revenue sharing rates across funds in Line 3 if needed.

If we are unable to match a fund to a revenue sharing agreement from Line 3 using the procedure above, we use the listed 12b-1 fee as the revenue sharing rate.

---

[18]Approximately 58% of all strings contain a single numeric (percentage) value that can be extracted. Another 9% have multiple percentages—usually corresponding to multiple payers in the associated fields. We find that about 1% of string contain a dollar amount (preceded by a dollar sign) rather than a rate. About 22% have no digits at all, usually corresponding to a generic formula (such as "percentage of assets under management") or explicitly saying something like "see attachment." Overall, we are able to parse 91% of strings with the set of regular expressions derived.

[19]To do so at scale, we set up three Google Custom Search Engines to search the Morningstar, Financial Times, and US News websites. We programmatically input the strings specified in the payer field into the CSEs and accept a match if at least one returns a match and the different CSEs do not disagree.

B. Predictive Model

This section discusses implementation details for predicting revenue from expense ratios, revenue from revenue sharing agreements, and expenses due to matching outlays. Our objective is to construct a model for each of these outcomes that predicts their counterfactual values under alternative plan designs. This entails two steps: first, distilling features from each plan menu into a set of variables that can serve as inputs into a model; and second, selecting a suitable model. We discuss each of these issues in turn.

Recall that a firm's plan consists of a series of investment vehicles, each with their own features. We convert these heterogenous investment menus into a series of variables that describe the menu appropriately. The variables can be characterized as either counts, fractions, and minima/average/maxima.

The count variables included in the model are number of investment options, number of mutual funds, number of target retirement date funds, number of index funds, number of S&P 500 tracker funds, number of funds offered by each fund provider, number of bond alternatives by bond investment style (examples), number of bond alternatives by bond credit quality, number of bond alternatives by bond interest rate sensitivity and number of funds by decile of the expense ratio distribution (unweighted) within Brightscope and Morningstar fund categories. As for minima/average (unweighted)/maxima, we calculate these objects for the following variables: expense ratio, 12b-1 fee, management fee, front end load and deferred load. Finally, we also the fraction of investment vehicles by value of Morningstar rating (for example, fraction of options that are rated as 5 stars by Morningstar).

Beyond features of the investment menu, we also incorporate features of the plan itself and firm characteristics. Features of the plan include matching rates and amounts, vesting rules, total balances, total participants, active participants, number of participants with a positive balance, retired participants, participant contributions, identity of the main provider, whether the plan is collectively bargained, the date it was first opened, and participation rates. As for firm characteristics, we incorporate NAICS code, state, region, and FIPS code. We also construct interactions between region and 2-digit NAICS code, and augment the dataset with marginal tax, wage, and rent data at the county level from NBER.

We build a predictive model for revenues from expense ratios per plan participant, revenues from revenue sharing agreements per plan participant, and outlays from matching per plan participant using stochastic gradient boosting (Friedman, 2002).[20] We use 70% of

---

[20]This model was implemented in R via the h2o package (H2O.ai., 2020).

	Training	Testing
Revenue from Expense Ratios - RMSE	14.31	35.21
Revenue from Expense Ratios - MAE	9.43	20.31
Revenue from Revenue Sharing - RMSE	12.73	29.31
Revenue from Revenue Sharing - MAE	8.07	16.73
Matching Expenditures - RMSE	479.22	1,314.50
Matching Expenditures - MAE	324.78	818.58

Table B.1: In-sample and out-of-sample fit for predictive models

(a) Expenses   (b) Revenue Sharing   (c) Matching
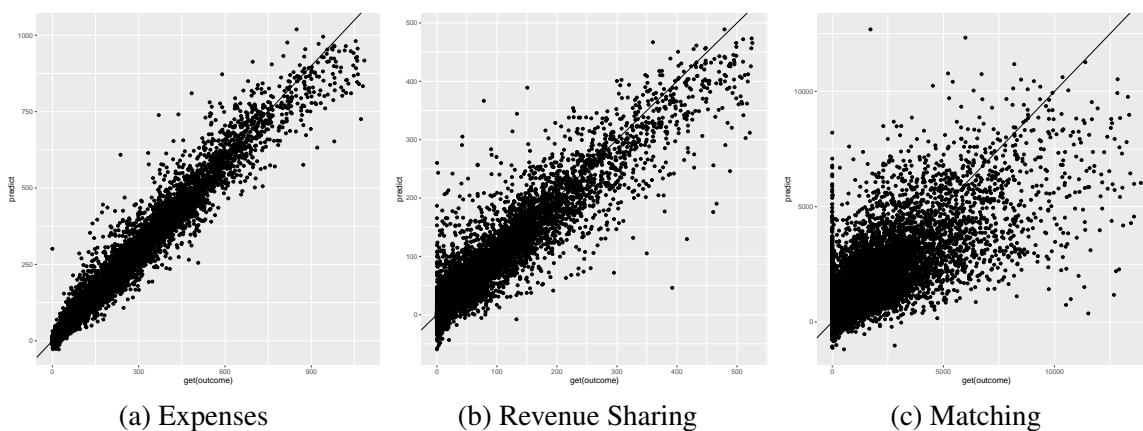
Figure B.1: Model Fit, Full Sample

the data for training the model, and the remainder for testing.

Table B.1 provides measures of root mean squared error (RMSE) and mean absolute error (MAE) for both samples. These objects allow us to assess the quality of the predictions. The models deliver low error values in the testing dataset, regardless of the error metric. For expense ratio revenues, RMSE is 35 dollars per person in the testing dataset, while MAE is 20 dollars. As for revenues from revenue sharing, these values are 29 dollars per person and 17 dollars per person, respectively. Finally, errors for matching expenditures are higher, but still reasonable. RMSE is 1314 in the testing dataset, and MAE is 818. Overall, these values reassure us that the procedure is working well. Additionally, Figure B.1 reports fit for each outcome using the testing sample.

To dig deeper into the determinants of these outcomes, Figure B.2 shows variable importance plots. For all three outcomes, the main predictor is balance per person. For expenses, the next predictors in terms of importance are mean expenses (an unweighted average of expense ratios) and contributions per participant. This is consistent with expense

ratios mechanically falling with balances, due to share classes, and with asset allocations that on aggregate are similar to a $\frac{1}{N}$ rule. For revenue sharing, the next most important predictors are the identity of the main provider, the number of investment vehicles without a star rating from Morningstar, the number of funds offered by provider 50 (who is particularly likely to use revenue sharing agreements), and the mean 12b-1 fee. We believe that these determinants are also reasonable. Since main providers differ in their propensity to use revenue sharing, particularly because some of them are vertically integrated, the identity of the main provider mattering makes sense. Unrated funds tend to be newer, smaller, and more expensive, so higher revenue sharing makes sense. Finally, matching outlays are predicted by NAICS and state, as one would expect when matching varies as a function of the labor market. Labor pool (NAICS-state combinations) also matter, as do wages. Overall, the main predictors in these models seem sensible.

C. Estimation

C.1. Integration

We solve the integrals in (8) as follows. The integral over θ is approximated using monomial rules for Gaussian quadrature, calculated by Stroud (1971).[21] We integrate over e and $\tilde{\mathcal{R}}_s$ by simulation. In particular, we draw $D = 1808$ consideration set probabilities ρ uniformly from the grid $\{0.1, 0.2, ..., 0.9, 0.99\}$,[22] and for each draw ρ_d, draw whether each possible recordkeeper from $\mathcal{R}_s$ is in the consideration set from a Bernoulli distribution with probability ρ_d. This gives us D draws of consideration sets, $\tilde{\mathcal{R}}_{s,d}$, and a value for the probability of observing the draw, ω_d. We hold these sets fixed throughout the estimation procedure, and for each guess of parameter values, solve for choice probabilities using importance sampling. The probability sponsor s chooses recordkeeper r is

$$\sum_{d=1}^{D} \frac{\omega_d(\rho)}{\omega_d} 1[r \in \tilde{\mathcal{R}}_{s,d}] \iint \frac{\exp(\frac{1}{\sigma^S}(V_{rs}^*(\theta, \sigma^R, \kappa, C_r^R) + \delta_r))}{1 + \sum_{r' \in \tilde{\mathcal{R}}_{s,d}} \exp(\frac{1}{\sigma^S}(V_{r's}^*(\theta_s, \sigma^R, \kappa, C_{r'}^R) + \delta_r'))} dF(\theta) dF(e)$$
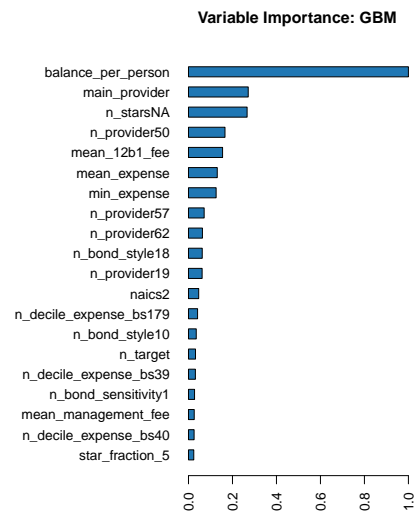
where $\omega_d(\rho)$ is the probability that consideration set $\tilde{\mathcal{R}}_{s,d}$ is drawn if the search cost is ρ.

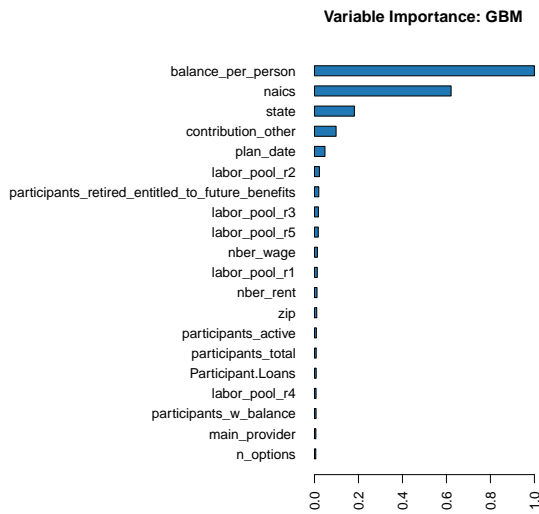[21]We use rule en_r2_07_2 in https://people.math.sc.edu/Burkardt/m_src/stroud/stroud.html.
[22]We use a multiple of the number of quadrature nodes (452).
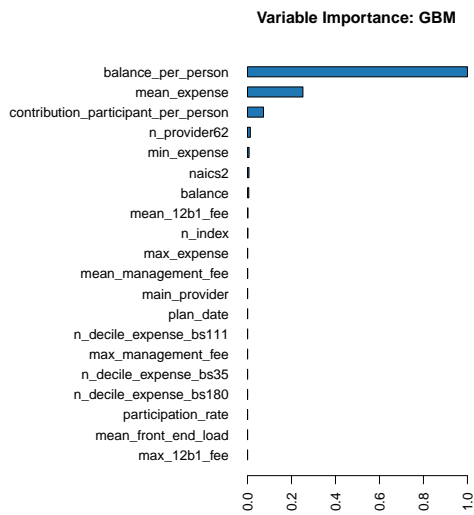
(a) Expenses

(b) Revenue Sharing

(c) Matching

Figure B.2: Relative Importance of Inputs into the Predictive Model

C.2. Moments

Recall that all moments take the form $\mathbb{E}[X \cdot \mathbb{1}[s \text{ chooses } r]]$, where X are different features of the chosen plan. This expectation is equal to

$$E[X_p^*] = \frac{1}{S} \sum_{s,d} \frac{\omega_d(\rho)}{\omega_d} \sum_{r \in \tilde{\mathcal{R}}_{sd}} \sum_{p \in \mathcal{P}_{rs}} X_{rsp} \cdot \Pr[s \text{ chooses } r \text{ and } p \text{ in draw } d]. \qquad (9)$$

Let $\pi_s^d \equiv \max_{r' \in \mathcal{R}_{rs} \backslash r} V_{r's}^*(\theta, \sigma^R, \kappa, C_{r'}^R) + \delta_{r'} + \sigma^S \epsilon_{r's}^S$ denote sponsor s's disagreement payoff, and define $l_{rs} \equiv \sigma^S \epsilon_{rs}^S - \pi_s^d$. When sponsor s chooses recordkeeper r and plan p, it must be the case that $V_{rs}^*(\theta, \sigma^R, \kappa, C_r^R) + \delta_r + l_{rs} \geq 0$ and

$$p = \underset{p \in \mathcal{P}_{rs}}{\arg\max} \; \underset{T_p \geq 0}{\max} (\boldsymbol{\theta}_s \cdot \boldsymbol{Q}_s(p) + \delta_r + l_{rs} - T_p - \kappa \cdot T_p^2)^\eta \cdot (R_r(p) - C_r^R + \sigma^R \epsilon_{rs}^R + T_p)^{(1-\eta)}$$

$$\equiv \underset{p \in \mathcal{P}_{rs}}{\arg\max} \; NP(l_{rs}, \epsilon_{rs}^R, \theta_s; \delta, \kappa, C_r^R, \eta).$$

With this notation, the probability in (9) can be expressed as

$$\Pr[s \text{ chooses } r \text{ and } p \text{ in draw } d]$$

$$= \Pr[s \text{ chooses } p \text{ in draw } d | s \text{ chooses } r \text{ in draw } d] \Pr[s \text{ chooses } r \text{ in draw } d]$$

$$= \Pr[p = \underset{p \in \mathcal{P}_{rs}}{\arg\max} \; NP(l_{rs}, \epsilon_{rs}^R, \theta_s; \delta, \kappa, C_r^R, \eta) | l_{rs} \geq -V_{rs}^*(\theta_s, \sigma^R, \kappa, C_r^R) - \delta_r]]$$

$$\times \; s_{rs}(\epsilon_{rs}^R, \theta_s; \sigma^R, \kappa, C^R, \sigma^S, p, \delta)$$

$$= \int \int \int_{l_{rs} \geq -V_{rs}^*(\theta, \sigma^R, \kappa, C_r^R) - \delta_r} 1[p = \underset{p \in \mathcal{P}_{rs}}{\arg\max} \; NP(l, e, \theta; \delta, \kappa, C_r^R, \eta)]$$

$$\times \; s_{rs}(e, \theta; \sigma^R, \kappa, C^R, \sigma^S, p, \delta) dF(l) dF(\theta) dF(e). \qquad (10)$$

To solve the integral in (10), we need the distribution of $l_{rs} \equiv \sigma^S \epsilon_{rs}^S - \pi_s^d$ conditional on being greater than $-V_{rs}^*(\theta, \sigma^R, \kappa, C_r^R) - \delta_r$. We begin by deriving the distribution of the disagreement payoff $\pi_s^d \equiv \max_{r' \in \mathcal{R}_{rs} \backslash r} V_{r's}^*(\theta, \sigma^R, \kappa, C_{r'}^R) + \delta_r' + \sigma^S \epsilon_{r's}^S$. Note $V_{r's}^*(\theta, \sigma^R, \kappa, C_{r'}^R) + \delta_r + \sigma^S \epsilon_{rs}^S$ distributes Gumbel with location $V_{r's}^*(\theta, \sigma^R, \kappa, C_{r'}^R) + \delta_r$ and scale σ^S. Since the maximum of independent Gumbel draws with the same scale parameter is also Gumbel, π_d distributes Gumbel with location

$$\sigma^S \log(1 + \sum_{r' \in \mathcal{R}_{rs} \backslash r} \exp(\frac{1}{\sigma^S}(V_{r's}^*(\theta, \sigma^R, \kappa, C_r^{R\prime}) + \delta_r')))$$

and scale σ^S. Since the difference between independent Gumbels with the same scale parameter is Logistic, $\sigma^S \epsilon^S_{rs} - \pi^d_s$ distributes Logistic with location

$$\mu \equiv -\sigma^S \log \left(1 + \sum_{r' \in \mathcal{R}_{rs} \backslash r} \exp(\frac{1}{\sigma^S}(V^*_{r's}(\theta, \sigma^R, \kappa, C^{R'}_r) + \delta'_r)) \right)$$

and scale σ^S. This is the unconditional distribution of $\sigma^S \epsilon^S_{rs} - \pi^d_s$. To derive the conditional distribution, note that for $s > a$,

$$\Pr[\sigma^S \epsilon^S_{rs} - \pi^d_s \leq s | \sigma^S \epsilon^S_{rs} - \pi^d_s \geq a] = \frac{\Pr[a \leq \sigma^S \epsilon^S_{rs} - \pi^d_s \leq s]}{\Pr[\sigma^S \epsilon^S_{rs} - \pi^d_s \geq a]}$$

$$= \left[\frac{1}{1 + \exp(-\frac{s-\mu}{\sigma^S})} - \frac{1}{1 + \exp(-\frac{a-\mu}{\sigma^S})} \right] \cdot \left[1 - \frac{1}{1 + \exp(-\frac{a-\mu}{\sigma^S})} \right]^{-1} = \frac{1 - \exp(-\frac{a-\mu}{\sigma^S})}{1 + \exp(-\frac{s-\mu}{\sigma^S})}.$$

so if $y = \Pr[\sigma^S \epsilon^S_{rs} - \pi^d_s \leq s | \sigma^S \epsilon^S_{rs} - \pi^d_s \geq a]$, $s = \mu + \sigma^S \left[\log(y + \exp(\frac{a-\mu}{\sigma^S})) + \log(1-y) \right]$. To integrate numerically over this conditional distribution, we obtain D uniform $(0,1)$ draws as y and transform them following this last expression, with $a = -V^*_{rs}(\theta, \sigma^R, \kappa, C^R_r) - \delta_r$.

Finally, we estimate the model on a random subsample of 7,500 firms for each of the four groups.

D. Additional Tables and Figures

Figure D.1 provides more details about the asset allocation. Panel (a) shows the share of assets in an index fund across all plans. The section of the histogram in blue illustrates plans that do not have an index fund at all (in which case the share is mechanically zero), and the remainder of the histogram shows the distribution for plans with an index fund. Panel (b) shows a scatter of the weighted expense ratio against the unweighted expense ratio for all plans, along with a locally linear best fit line. The best fit line is very close to the 45° line (but slightly lower), suggesting that workers do not seem to be weighing low expense ratio funds more in their allocation.

Table D.1 reports estimated costs by recordkeeper. We find some variation across record-keepers and limited evidence of economies of scale even recordkeeper-by-recordkeeper.
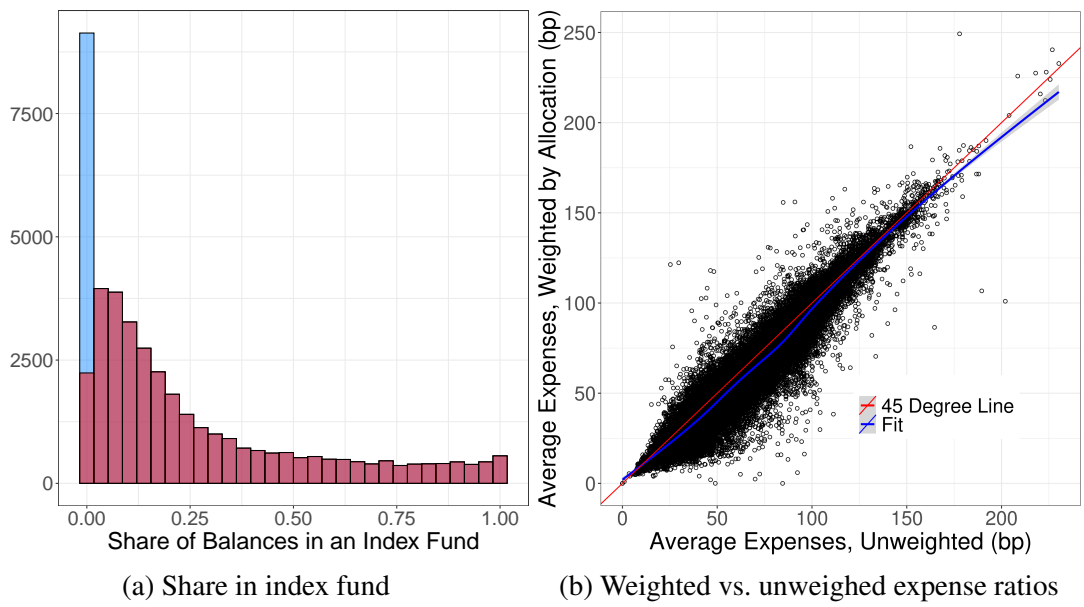
(a) Share in index fund

(b) Weighted vs. unweighed expense ratios

Figure D.1: Information about asset allocation

	Group 1 ≤ 200	Group 2 201–500	Group 3 501–1000	Group 4 > 1000
ADP	28.14	3.51	36.45	57.17
	(0.278)	(0.116)	(0.522)	(0.687)
Alerus Financial	12.04	16.83	9.33	9.26
	(0.527)	(0.750)	(1.001)	(0.995)
American Funds (Capital Group)	3.63	7.22	11.65	9.58
	(0.465)	(0.796)	(0.813)	(0.787)
Ascensus, Inc.	8.71	11.93	9.35	9.74
	(0.585)	(0.880)	(1.136)	(1.007)
Bank of America Merrill Lynch	10.20	9.95	8.72	9.77
	(0.572)	(1.483)	(1.005)	(1.178)
Charles Schwab	6.97	11.56	11.09	8.58
	(0.538)	(0.947)	(0.880)	(0.978)
Empower	9.50	8.79	8.62	8.26
	(0.679)	(0.920)	(1.083)	(1.114)
Fidelity	13.47	13.96	12.55	7.59
	(0.581)	(1.049)	(1.019)	(1.243)
John Hancock	12.28	9.65	10.43	10.24
	(0.550)	(1.011)	(1.138)	(0.937)
MassMutual	10.38	6.21	8.53	12.25
	(0.558)	(1.025)	(0.963)	(0.851)
Nationwide	10.65	11.81	9.20	10.15
	(0.539)	(0.779)	(1.127)	(1.042)
Newport Group Securities, Inc.	11.98	5.03	9.50	7.87
	(0.554)	(0.911)	(0.890)	(0.880)
Principal Financial Group	4.89	9.12	9.85	9.29
	(0.615)	(1.110)	(1.103)	(0.991)
T. Rowe Price	12.97	12.19	9.92	8.11
	(0.565)	(0.881)	(0.855)	(0.860)
TIAA-Nuveen	12.90	16.25	11.24	9.13
	(0.676)	(0.750)	(0.907)	(0.960)
Transamerica	13.76	8.34	9.82	9.47
	(0.477)	(0.950)	(1.057)	(0.861)
Vanguard Group	12.92	11.16	8.33	9.38
	(0.593)	(0.805)	(0.898)	(0.793)
Voya Financial (ING)	12.29	12.35	8.34	10.62
	(1.339)	(2.267)	(2.138)	(1.689)
Wells Fargo	12.90	13.70	8.90	8.51
	(0.461)	(0.693)	(0.824)	(0.806)
Other	13.46	5.41	12.32	7.28
	(0.486)	(0.914)	(0.897)	(0.848)

Table D.1: Estimated recordkeeper costs, in dollars per worker. Standard errors are in parentheses.