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).

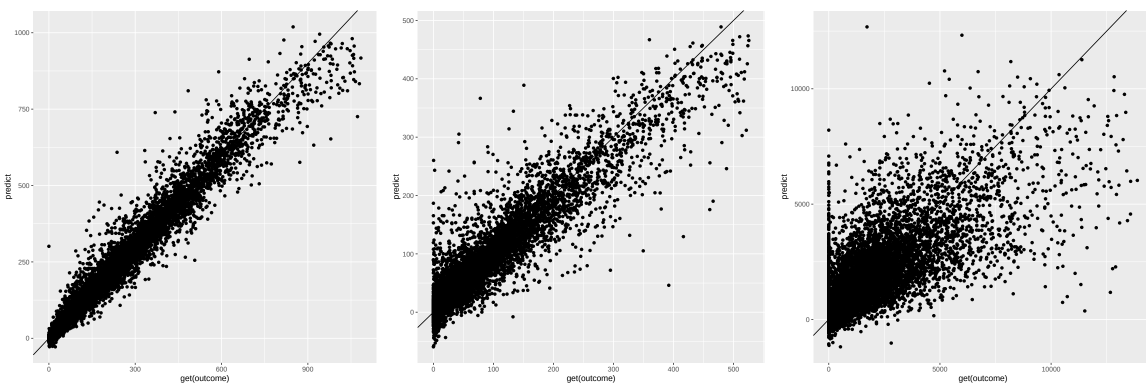
Beyond features of the investment menu, we also incorporate features of the plan itself and firm characteristics. Features of the plan include matching rates and amounts, vesting rules, total balances, total participants, active participants, number of participants with a positive balance, retired participants, 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

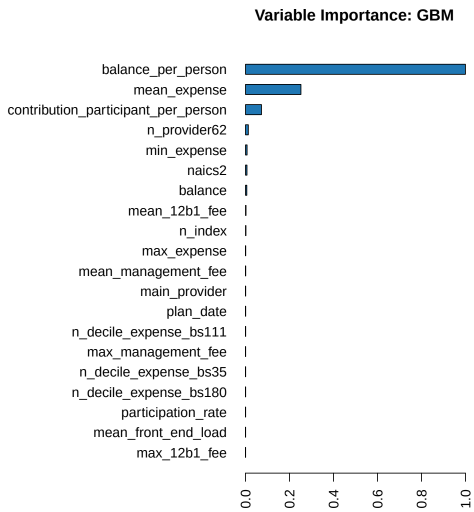
C.1. Integration

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

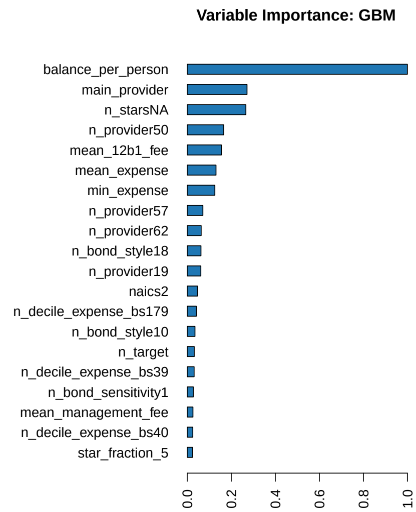
$$\sum_{d=1}^{D} \frac{\omega_d(\rho)}{\omega_d} 1[r \in \tilde{\mathcal{R}}_{s,d}] \iint \frac{\exp(\frac{1}{\sigma^S}(V^*_{rs}(\theta, \sigma^R, \kappa, C^R_r) + \delta_r))}{1 + \sum_{r' \in \tilde{\mathcal{R}}_{s,d}} \exp(\frac{1}{\sigma^S}(V^*_{r's}(\theta_s, \sigma^R, \kappa, C^R_{r'}) + \delta'_r))} dF(\theta) dF(e)$$

where $\omega_d(\rho)$ is the probability that consideration set $\tilde{\mathcal{R}}_{s,d}$ is drawn if the search cost is ρ.

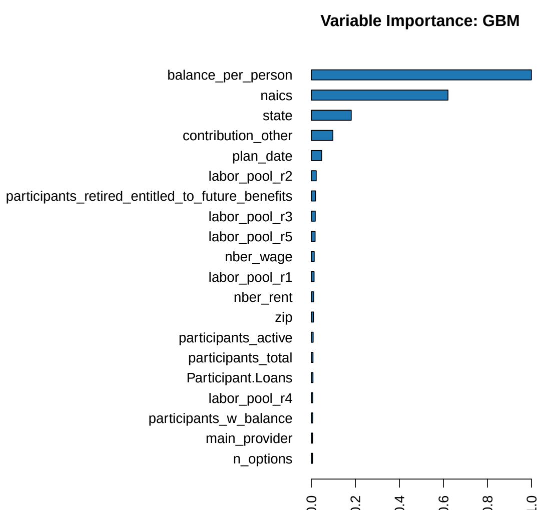
[21] We use rule en_r2_07_2 in https://people.math.sc.edu/Burkardt/m_src/stroud/stroud.html.
[22] We use a multiple of the number of quadrature nodes (452).



(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

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

Let $\pi_s^d \equiv \max_{r'\in\mathcal{R}_{rs}\setminus r} V_{r's}^*(\theta, \sigma^R, \kappa, C_{r'}^R) + \delta_{r'} + \sigma^S\epsilon_{r's}^S$ 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

$$\begin{aligned} p &= \arg\max_{p\in\mathcal{P}_{rs}} \max_{T_p\geq 0}(\boldsymbol{\theta}_s \cdot \boldsymbol{Q}_s(p) + \delta_r + l_{rs} - T_p - \kappa\cdot T_p^2)^{\eta}\cdot(R_r(p) - C_r^R + \sigma^R\epsilon_{rs}^R + T_p)^{(1-\eta)} \\ &\equiv \arg\max_{p\in\mathcal{P}_{rs}} NP(l_{rs}, \epsilon_{rs}^R, \theta_s; \delta, \kappa, C_r^R, \eta). \end{aligned}$$

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

$$\begin{aligned} &\Pr[s \text{ chooses } r \text{ and } p \text{ in draw } d] \\ &\quad = \Pr[s \text{ chooses } p \text{ in draw } d | s \text{ chooses } r \text{ in draw } d]\Pr[s \text{ chooses } r \text{ in draw } d] \\ &\quad = \Pr[p = \arg\max_{p\in\mathcal{P}_{rs}} NP(l_{rs}, \epsilon_{rs}^R, \theta_s; \delta, \kappa, C_r^R, \eta) | l_{rs} \geq -V_{rs}^*(\theta_s, \sigma^R, \kappa, C_r^R) - \delta_r]] \\ &\qquad\qquad \times s_{rs}(\epsilon_{rs}^R, \theta_s; \sigma^R, \kappa, C^R, \sigma^S, p, \delta) \\ &\quad = \int\int\int_{l_{rs}\geq -V_{rs}^*(\theta,\sigma^R,\kappa,C_r^R)-\delta_r} 1[p = \arg\max_{p\in\mathcal{P}_{rs}} NP(l, e, \theta; \delta, \kappa, C_r^R, \eta)] \\ &\qquad\qquad \times s_{rs}(e, \theta; \sigma^R, \kappa, C^R, \sigma^S, p, \delta) dF(l)dF(\theta)dF(e). \end{aligned} \tag{10}$$

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 $\pi_s^d \equiv \max_{r'\in\mathcal{R}_{rs}\setminus r} V_{r's}^*(\theta, \sigma^R, \kappa, C_{r'}^R) + \delta_r' + \sigma^S\epsilon_{r's}^S$. 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

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

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

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

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 $s > a$,

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

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

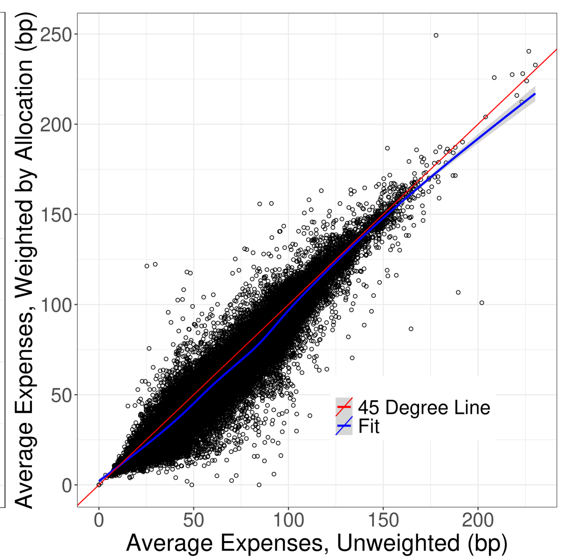
so if $y = \Pr[\sigma^S \epsilon^S_{rs} - \pi^d_s \leq s | \sigma^S \epsilon^S_{rs} - \pi^d_s \geq a]$, $s = \mu + \sigma^S \left[\log(y + \exp(\frac{a-\mu}{\sigma^S})) + \log(1-y) \right]$. To integrate numerically over this conditional distribution, we obtain D uniform (0,1) draws as y and transform 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 recordkeepers 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.