Surveying Counterfactuals: Improving 401(k) Matches Using Hypothetical Choices*

Guillermo Carranza Taha Choukhmane Fiona Greig
Cormac O'Dea Lawrence Schmidt

March 2026

Abstract

Evaluating reforms requires predicting how individuals would behave under counterfactual policies and mapping those behavioral responses into welfare-relevant criteria. We use survey responses to hypothetical scenarios linked with administrative 401(k) data to address both in the context of designing employer matching formulas. We find that (i) these survey responses can accurately predict saving responses in administrative data, (ii) saving is inelastic to the match rate, and (iii) non-elective contributions do not crowd out employee saving. These patterns imply that (iv) plans combining lower match rates with non-elective contributions generate higher savings and more equitable match distributions, and (v) many existing plans—including safe-harbor formulas—are dominated along both dimensions. These conclusions hold for any objective function that values either higher saving or lower inequality, or both.

*Carranza: Yale University,guillermo.carranzajordan@yale.edu; Choukhmane: MIT Sloan School of Management, tahac@mit.edu; Greig: Vanguard fiona_greig@vanguard.com; O'Dea: Yale University, cormac.odea@yale.edu; Schmidt: MIT Sloan School of Management, ldws@mit.edu. We are very grateful to Aaron Goodman, Kelly Hahn, Anna Madambda, and Ben Polak for illuminating comments and suggestions. We are very grateful to the Yale University Tobin Center for Economic Policy for co-funding this work. Any errors are our own.

1 Introduction

Every year, U.S. private employers contribute over $200 billion to their employees' retirement savings plans, mostly in the form of matching contributions. These incentives tie employee compensation to voluntary saving behavior and vary widely across employers, with more than 1,000 different matching formulas featuring different match rates, caps, and often multiple tiers.[1] A commonly cited rationale for these programs is to increase employee contributions for more than 100 million eligible workers while complying with federal regulations that limit the share of benefits accruing to highly paid workers. Given these twin goals of increasing employee savings and mitigating inequality, this paper asks a simple question: which matching formula should an employer choose to increase savings, reduce inequality, while containing costs? The answer to this question matters not only for employers seeking to align benefit design with these goals, but also for public policy. Over half of firms adopt one of small number of "safe harbor" formulas, which are exempt from non-discrimination testing. That share is growing, yet these widely adopted designs have never been evaluated against cost-equivalent alternatives.

Evaluating the welfare consequences of policy reforms requires solving two distinct problems: predicting how individuals would behave under counterfactual policies, and mapping those predicted behavioral distributions into welfare-relevant criteria. Existing approaches handle these problems jointly, each with distinct limitations. The sufficient statistics approach (see Chetty, 2009, for a review) derives welfare formulas that depend on estimable elasticities, but because these are typically local approximations around the status quo, they cannot deliver the quantitative comparisons needed to rank discrete alternatives such as two specific match formulas. The structural approach (see Low and Meghir, 2017, for a review) can handle large counterfactual changes, but its conclusions depend on functional form assumptions about preferences and optimization that are difficult to validate.

Our approach separates the two problems and makes weaker assumptions on each. For prediction, we elicit counterfactual responses directly via survey, requiring no parametric assumptions on the distribution of heterogeneity in preferences or constraints. Whatever drives individual saving behavior, whether shaped by frictionless optimization, or by (e.g.) present bias or liquidity constraints, is implicitly captured in the stated responses without the researcher needing to specify a model of it. We validate these predictions by comparing them to those that we see in the administrative data. For welfare evaluation, we show that progress is possible without committing to a specific social welfare function. The theoretical

[1] For instance, a single-tier plan might match 100% of employee contributions up to 5% of pay. A two-tier plan could match 100% of contributions up to 3% of pay, and then 50% of contributions on the next 2% of pay.

literature on retirement saving policy since Diamond (1977) and Feldstein (1985) has emphasized two objectives: increasing saving and reducing inequality in retirement resources. We do not take a stance on the relative weight a planner should place on each. Instead, we characterize the frontier of achievable equity-savings combinations, and identify plans that are dominated, in the sense that cost-equivalent alternatives exist that improve on both dimensions simultaneously. This delivers concrete guidance on the design of a $200 billion policy without requiring the strong modeling assumptions that such quantitative conclusions typically entail. The cost of not specifying a social welfare function is that we cannot identify a single optimal policy. But the dominated plans we identify, and the frontier alternatives we propose, are robust to the class of objective function that values both higher saving and lower inequality.

Concretely, we combine administrative data from 1.8 million Vanguard participants with a novel survey methodology that elicits responses to counterfactual matching designs. A distinctive feature of the survey is that it asks respondents about their likely behavior under match formulas that are rare or non-existent in the current landscape. We integrate these survey responses with a predictive model of saving behavior and show that, when calibrated with our survey responses, the model closely replicates observed saving choices. This provides validation for using the model, augmented with survey input, to predict responses under a wide range of counterfactual policies.

There is good reason to believe that current matching formulas could be improved on both equity and efficiency grounds. Using administrative data from nearly 1,000 retirement plans managed by Vanguard, we first show that most low-income workers do not contribute enough to receive the full match. Even among high-income workers, around 20% fall short of the level at which they would receive their full match. This pattern implies significant inequality in the distribution of matching resources, one that aligns with the pattern observed in nationally-representative tax data by Choukhmane et al. (2024). Second, we find that the majority of employer match dollars accrue to *inframarginal* savers—employees who already contribute beyond the maximum matched amount. This implies that most matching dollars are not targeted at creating incentives for workers to contribute more.

Next, we use a new survey of retirement contributions under hypothetical match formulas to study the potential of alternative formulas to improve outcomes along both dimensions. We have four main results. First, employee self-reported saving is inelastic with respect to the match rate: dollar-for-dollar match rates are much more costly than 25-cents-on-the-dollar match, and yet only generate modestly more additional saving. Second, non-elective contributions—employer contributions made regardless of employee saving—do not appear to crowd out employee contributions. This holds in both survey responses (where hypothet-

ical non-elective contributions do not reduce intended elected contributions) and administrative data (where actual increases in non-elective contributions do not reduce employee contributions). Third, bringing together the survey evidence and the predictive model, we characterize the frontier of the equity-efficiency trade off. In other words, we identify the set of contribution formulas where, for a given employer cost, reducing inequality necessarily requires lowering contributions, and increasing contributions requires more inequality. Formulas on the frontier tend to pair lower match rates with non-elective contributions. Fourth, we find that many existing retirement plans, including nearly all safe harbor plans, are strictly dominated in the sense that cost-neutral alternatives could increase saving without increasing inequality. We estimate that moving all plans in our sample from their current formulas to the dominating alternatives we identify would increase the average employee contributions induced by matches by 50%, without raising employer costs or exacerbating inequality. Alternatively, moving all plans to the efficiency frontier could reduce inequality in the distribution of employer contributions by 40%, again without reducing voluntary saving or increasing plan costs.

Our work relates to three strands of the literature. The first concerns the causal effects of saving incentives. Researchers have addressed this question using a variety of empirical strategies, including cross-sectional variation in match schedules (e.g., Engelhardt and Kumar (2007); Mitchell et al. (2007)), firm-level changes in match rates (e.g., Papke (1995); Kusko et al. (1996); Choi et al. (2002)), variation in tax incentives (e.g., Ramnath (2013); Chetty et al. (2014)) and experimental designs (e.g., Duflo et al. (2006)). A review by Choi (2015) concludes that while matching has a small positive effect on participation, its effect on average contribution rates is ambiguous.

We advance this literature by addressing a fundamental challenge: the difficulty of identifying responses to multidimensional and heterogeneous saving incentives that are often not observed in reality. Our approach estimates stated contribution preferences across a wide range of match formulas that extend well beyond those observed empirically. By combining a model with purpose-built survey data designed specifically for model calibration, we can predict contribution behavior under a large set of match schedules, thereby overcoming the limitations of existing variation in incentive structures.

Second, we contribute to the literature on the design of retirement saving institutions. Going back to Diamond (1977), the literature has identified two primary rationales for intervention in retirement savings: paternalism (addressing undersaving) and redistribution (both across people and states of the world). Beginning with Feldstein (1985), studies have used theoretical models to characterize optimal paternalistic retirement saving policies. These models typically assume that individuals undersave either because of behavioral biases (e.g.,

Beshears et al. (2025)) or because of the disincentives created by means-tested social safety nets (e.g., Hubbard et al. (1995)). In parallel, redistribution and insurance have been studied as an independent source of welfare gains in the design of these policies (e.g., Hosseini and Shourideh (2019), O'Dea (2018)). Recent work increasingly recognizes that these motives cannot be considered in isolation. For instance, Moser and Olea de Souza e Silva (2019) analyze optimal savings policies under joint concerns for retirement undersaving and income inequality, while Choukhmane and Palmer (2025) estimate a structural lifecycle model to quantify the welfare implications of alternative policy designs when policymakers value both objectives. While this literature provides important insights for the design of saving policies, it often relies on stylized settings that abstract from the complexities of actual retirement saving incentives. Real-world matching formulas involve multiple design parameters (i.e., match rates, caps, and often multiple tiers) and must account for wide heterogeneity in employee responses and employer financial resources, making analytical characterization often intractable. We contribute to this literature by providing the first systematic answer to which matching formulas actually achieve the two objectives identified in the theoretical literature, increasing savings and reducing inequality, while accounting for real-world constraints and both employee and employer heterogeneity.

Third, we contribute to the growing literature that combines administrative data, models of economic behavior, and custom surveys (Bachmann et al. (2022); Fuster and Zafar (2023); Stantcheva (2023); Caplin (2025)). These surveys bring rich variation to identify preferences, beliefs, and constraints that are generally missing from most administrative datasets but critical for model calibration (see Koşar and O'Dea (2023) for a review). Recent work has applied this approach to household consumption and saving behavior: Giglio et al. (2021) and Carranza and Goodman (2024) show that survey responses correlate with behavior in administrative data, while Ameriks et al. (2020) and Indarte et al. (2025) use surveys to calibrate key parameters in lifecycle consumption models. Similar to our approach, Colarieti et al. (2024) show that survey responses to hypothetical scenarios (namely marginal propensities to save and borrow) align quantitatively with empirical estimates in the literature. We contribute to this literature by showing how survey responses to hypothetical policies not only predict actual behavior in administrative data but also enable quantitative policy evaluation.

The remainder of the paper proceeds as follows. Section 2 describes the institutional setting and reports descriptive analysis of contributions using the administrative data. Section 3 introduces a simple model of retirement saving in the presence of employer matches. Section 4 presents our survey design and validation methodology. Section 5 brings together the model and survey data. Section 6 characterizes the efficiency frontier and evaluates existing

plans, and Section 7 concludes.

2 Matching Incentives: Institutional Context and Key Empirical Patterns

2.1 Employer Matches in the US

Defined contribution (DC) plans are the dominant vehicle for private retirement saving in the United States. Two-thirds of workers of US workers have access to employer-sponsored accounts (U.S. Bureau of Labor Statistics (2024)), and Americans hold over $13 trillion of assets within them.[2] A central feature of these plans is the employer match, which provides a direct financial incentive for employees to contribute.

Match formulas vary widely in their structure. Key elements include:

- the **match rate** (e.g., 50% or 100%) applied to employee contributions,
- the **cap** on contributions eligible for a match (e.g., up to 6% of pay),
- the **number of tiers**, with some plans offering a higher match rate on initial contributions and a lower rate thereafter,
- **non-elective contributions**, where employers contribute regardless of employee action.

Other plan design features also shape participation and savings behavior, including automatic enrollment, vesting schedules, eligibility rules, and withdrawal options. These plan design features may also influence whether the worker takes full advantage of the match and associated employer costs.

2.2 Non-Discrimination Testing and Safe Harbor Plans

While employers decide whether to offer a match and choose its specific formula, legislation and regulation play a pivotal role in guiding these choices. To retain tax-advantaged status, retirement plans must pass annual non-discrimination tests that ensure benefits, and in particular employer contributions, do not disproportionately accrue to highly compensated employees. Failure to pass these tests requires corrective action and may trigger

[2]This dollar value is an underestimate of the reach of employer-sponsored DC accounts; if IRAs are included the stock of wealth is worth $30 trillion, and most IRA balances represent rollovers from employer sponsored accounts.

excise taxes. Employers can bypass testing altogether by adopting standardized "safe harbor" plan designs, which are automatically deemed compliant. Through the structure of non-discrimination tests and the design of safe harbor provisions, public policy can effectively steer employers toward specific matching formulas. Reflecting this influence, Vanguard (2025) show that a third of Vanguard plans have adopted a safe-harbor design, while Arnoud et al. (2021) find that nearly half of plans in a nationally-representative sample adopt a match formula that satisfies the safe harbor employer matching requirements.

2.3 Administrative Data

Our primary administrative data come from records for Vanguard plan participants in compliance-testing plans, for which Vanguard manages non-discrimination testing on behalf of the plan sponsor. These records include W-2 earnings reported directly by plan sponsors, plan formula details (match rates, caps, and tiers), and annual employee and employer contributions, with employer contributions separated into matching and non-elective components. We also observe whether each plan has auto-enrollment and, if so, the date of adoption. Plan sponsors additionally provide basic demographic information: age, gender, and tenure within the plan.

We complement these annual records with a monthly panel of elected contribution rates, capturing the sequence of rates participants actively choose over time.[3] We use the contribution rate in effect at the time of survey participation to validate our model predictions.

2.4 Two facts about DC retirement saving behavior

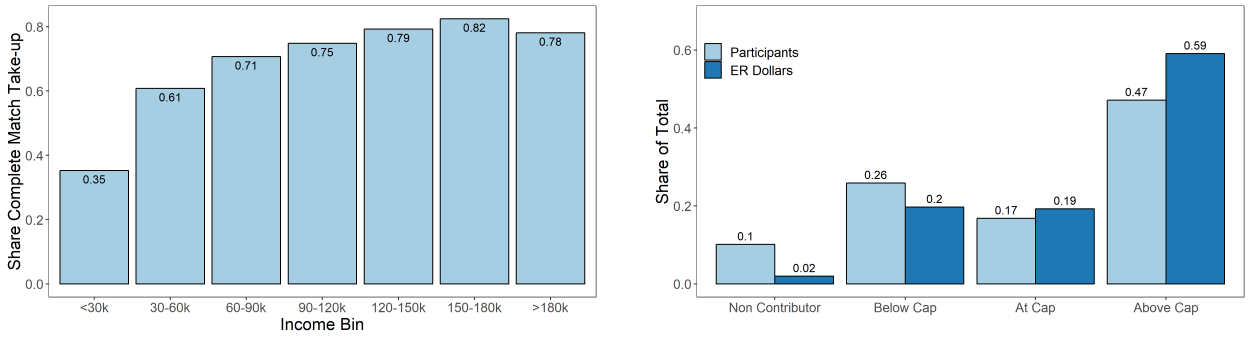
Before turning to our analysis of counterfactual policies, we summarize two empirical facts about defined contribution saving behavior in our data. These facts suggest that current match formulas may be poorly targeted: many workers do not fully benefit from the match, and most of the employer's match budget flows to workers whose saving the match is unlikely to have changed.

1. Take-up is incomplete, especially among low-income workers. Despite strong positive gradients in saving, many high earners do not contribute enough to receive the full employer match. Figure 1(a) shows that nearly 20% of workers earning $150,000 or more fall below the match threshold.

[3]We draw on a dataset of monthly contribution rates constructed by Aaron Goodman and used to measure income volatility among 401(k) participants in Goodman (2025).

2. Most match dollars go to inframarginal savers. Figure 1(b) shows that close to half of of employees contribute more than the match ceiling, and this group receives almost 60% of total match dollars. If the goal of matching is to increase saving at the margin, the heavy allocation of match dollars to inframarginal savers suggests that existing incentives may be poorly targeted.

Figure 1: Patterns of Match Take-Up and Dollar Flows



(a) Match Take Up by Income

(b) Distribution by Marginality

Notes: Panel (a) shows the proportion of workers in the administrative data who fully take up the match they are offered. Panel (b) divides the distribution of workers in groups of those who do not participate, those who participate but save less than match cap, those who save at the match cap, and save above the match cap. The figure also shows the proportion of match dollars going to employees in each of those four groups. Results are based on administrative population in the Vanguard data in 2023.

3 From Survey Responses to Counterfactual Predictions

Our aim in this paper is to evaluate both the positive and normative features of match formulas, including those not currently observed in practice. To do so, we present a simple model of retirement saving in the presence of employer matches. We show that this model, combined with a small set of variables collected through our survey, can be used to predict saving behavior across a wide range of match formulas.

Let $s(m, c)$ denote the contribution rate an agent chooses under a match rate of m with a cap at $c \in [0, 1]$, where c is expressed as a proportion of salary. We show that under three testable assumptions, knowledge of two objects for an individual: $s(0, .)$ and $s(m, 1)$ (saving under no match, and saving under an unlimited match) is sufficient to predict $s(m, c)$ for any cap c.

3.1 Assumptions

We impose three restrictions on saving behavior:

Assumption 1 (Incentives matter)

(a) If $c' > c$ then

$$s(m, c') \geq s(m, c)$$

Interpretation: Savings do not fall when incentives become more generous. This requires that increasing the cap does not reduce contributions. It is a weak monotonicity condition.

Assumption 2 (Irrelevance of non-binding caps) *If $s(m, c) < c$ and $c' > c$ then*

$$s(m, c') = s(m, c)$$

Interpretation: When a non-binding cap is increased, savings do not change. Employees who contribute below the cap are unaffected by further increases in the cap. This assumes agents respond only to the marginal incentive they face, not to incentives at higher contribution levels.

Assumption 3 (No wealth effects) *If $s(m, c) > c$ then*

$$s(m', c) = s(m, c) \quad \forall \quad m'$$

Interpretation: Savings do not change unless the marginal return to saving changes. When an employee contributes at or above the cap, they receive the full match subsidy on infra-marginal dollars. Changes in the match rate therefore generate pure income effects for these individuals. This assumption requires that such income effects do not affect total contributions. This is a strong but testable restriction.

These assumptions are mutually consistent and empirically testable.

3.2 Prediction Rule

Our aim is to predict saving under an arbitrary match rate $\hat{m}$ and cap c. The key insight is that the two observables $s(0, .)$ and $s(\hat{m}, 1)$ allow us to classify each individual into one of three mutually exclusive cases, each with a distinct prediction.

- **Case 1: High saver**: $s(0, .) > c$.

  Interpretation: This individual would save above c even with no match.

- From Assumption 3, since $s(0,.) > c$, we obtain $s(\hat{m}, c) = s(0,.)$.
- Relative to the no match case, an introduction of a match provides an income effect but does not change the marginal return to saving above c.

- **Case 2: Medium saver**: $s(0,.) < c$ and $s(\hat{m}, 1) > c$.

  Interpretation: This individual would save above c with an unlimited match but below c with no match.

  - Suppose $s(\hat{m}, c) < c$. By Assumption 2, $s(\hat{m}, c) = s(\hat{m}, 1) > c$, which is a contradiction. Therefore $s(\hat{m}, c) \geq c$.
  - Suppose $s(\hat{m}, c) > c$. By Assumption 3, $s(\hat{m}, c) = s(0,.) < c$, which is a contradiction. Therefore $s(\hat{m}, c) \leq c$.
  - Together, these yield $s(\hat{m}, c) = c$. The individual saves exactly at the cap.

- **Case 3: Low saver**: $s(0,.) < c$ and $s(\hat{m}, 1) < c$.

  Interpretation: This individual would save below c even with an unlimited match. The cap is therefore irrelevant.

  - Suppose $s(\hat{m}, c) \geq c$. From Assumption 1, $s(\hat{m}, 1) \geq s(\hat{m}, c) \geq c$, which contradicts the low saver condition. Therefore $s(\hat{m}, c) < c$.
  - By Assumption 2, since $s(\hat{m}, c) < c$, we have $s(\hat{m}, c) = s(\hat{m}, 1)$.

These three cases are exhaustive and mutually exclusive. They lead to the following closed-form prediction rule:

$$s(\hat{m}, c) = \max\{\, s(0,.),\ \min[\, s(\hat{m}, 1),\ c\,]\}\,. \tag{1}$$

The prediction rule has an intuitive interpretation. An individual's saving is the maximum of (a) what they would save with no match, and (b) the minimum of what they would save with an unlimited match and the cap itself. High savers are governed by (a), low savers by their unlimited-match saving in (b), and medium savers by the cap in (b).

Figure 2 summarizes the prediction rule. Panel (a) fixes the cap and classifies individuals. Those whose no-match saving exceeds the cap are the 'high savers' (by assumption 3, the match is not relevant to their behavior as it represents a pure income effect). Those whose unlimited-match saving exceeds the cap but whose no-match saving does not are the 'medium savers' (they bunch at the cap). Those whose unlimited-match saving falls below the cap are the 'low savers' (the cap does not bind). Panel (b) fixes the individual and varies the cap,

tracing out the three regimes as a function of c: flat at $s(0,.)$ when the cap is low, rising along the 45-degree line when the individual bunches at the cap, and flat at $s(\hat{m},1)$ when the cap is non-binding.

Appendix A.1 shows that the same logic extends to multi-tier plans, requiring only one additional observable (saving under an unlimited match) for each additional tier. Appendix A.2 provides a parametric model in which the three assumptions are satisfied.

4 Survey Data Collection and Validation

4.1 Survey Design and Implementation

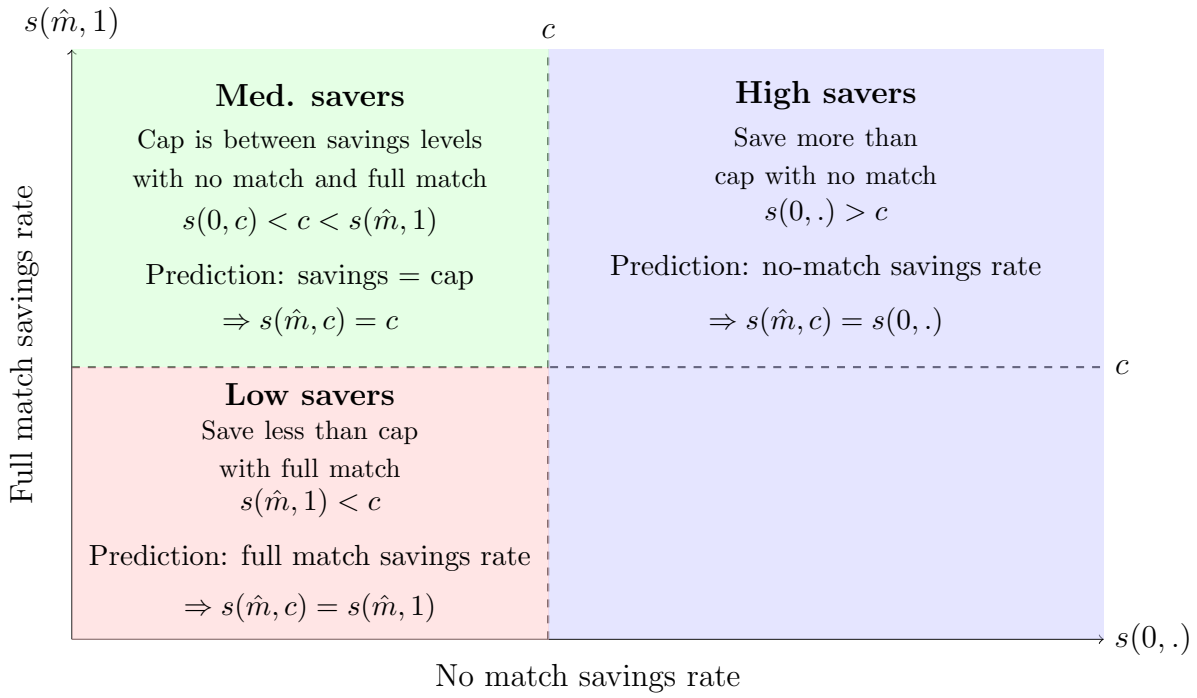
The survey was designed to serve two main purposes: (i) to evaluate saving responses under counterfactual match formulas, and (ii) to generate responses that could be used to calibrate our model.

Survey Design The core of the survey consisted of a series of hypothetical contribution vignettes. Respondents were asked to evaluate several retirement plan designs. Before viewing the vignettes, they were presented with a set of assumptions that would apply to all scenarios. Appendix Figure B.2(a) displays the introductory instructions. Participants were told to assume that employer contributions would be immediately vested, that employee contributions would receive the same tax treatment as in their current plan, and that, in scenarios where the match rate changed, all other aspects of their jobs would remain the same.

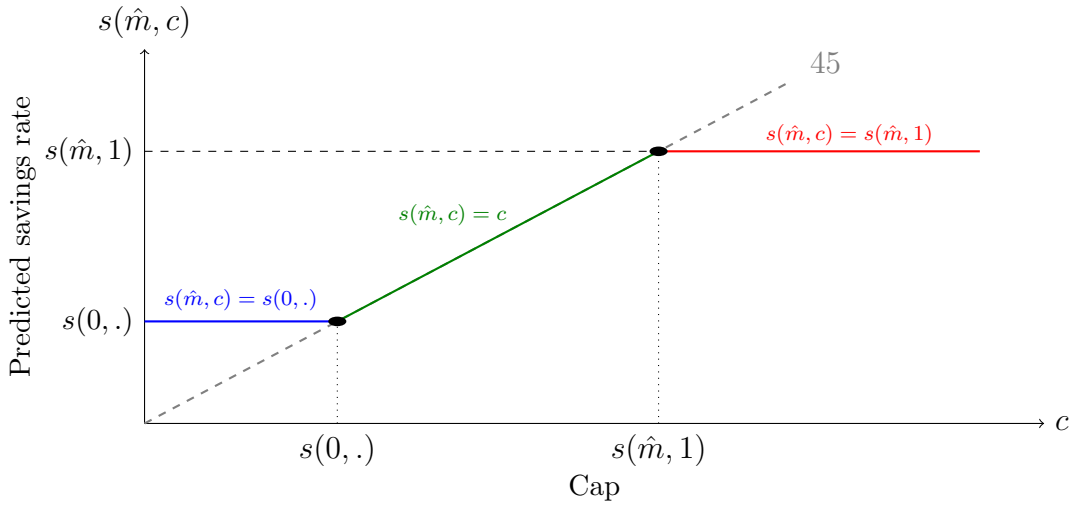
Following this, respondents were asked how much they would contribute under each hypothetical match schedule. Each vignette followed a standard format. Figure B.2(b) shows the interface for the "No Match" vignette. Respondents used a slider to choose their contribution rate as a percentage of income. As the slider moved, the display updated in real time to show the employee contribution, the employer contribution, and the total combined amount.

The match schedules tested in the survey included:

1. One vignette with no match.
2. Three vignettes with unlimited matching at rates of 25%, 50%, and 100%.
3. One vignette with a match rate up to 6%, randomly assigned a match rate of 25%, 50%, or 100%.



(a) Model Summary



(b) From data to prediction

Figure 2: Model illustration

Notes: This figure illustrates the prediction rule of saving under a match up to a cap. Panel (a) classifies individuals into three types based on the relationship between the cap and two observable quantities: saving under no match $s(0,.)$, and saving with an unlimited match rate of m denoted $s(\hat{m},1)$. High savers already contribute above the cap without a match, so the match does not affect their behavior. Medium savers would exceed the cap if the match applied to all contributions, so they bunch at the cap. Low savers contribute below the cap even with the match, so the cap does not bind. Panel (b) plots predicted saving $s(\hat{m},c)$ as a function of the cap c for a given individual. It shows three regimes: a flat region at $s(0,.)$ when the cap is below the individual's no-match saving, a 45-degree segment where the individual bunches at the cap, and a flat region at $s(\hat{m},1)$ when the cap is non-binding.

4. One vignette with no match but a universal employer contribution of 10% of pay.

The vignettes in items 1 and 2 provide the main inputs for predicting matching contributions under alternative plan designs, capturing $((s(0,.), s(0.25,1), s(0.5,1), s(1,1))$. The vignettes in item 3 is used to test the consistency of responses for respondents who actually have that match in reality. The vignette in item 4 allows us to evaluate the size of wealth effects.

The survey also included questions about current contribution behavior, as well as measures of behavioral traits such as present bias, risk aversion, self-control, and liquidity constraints.

Survey Implementation The survey was fielded in February 2025. The sample consisted of current plan participants for whom we had administrative data on contributions and income in 2023. The survey oversampled participants who had contributed less than the plan match cap in that year.

Table B.1 compares the respondent sample to the full administrative population across age, income, and tenure. The final sample closely matches the population in terms of income and tenure but includes a somewhat higher proportion of older participants. In our empirical analysis, we re-weight responses to match the joint distribution of age and income.

4.2 Survey Internal Validity

To test the internal validity of the survey responses, we compare survey responses to administrative data on measures for which both sources are available. First, we compare self-reported contribution rates to employer-sponsored retirement accounts with the actual contribution rates recorded in administrative data. Appendix Figure B.3 shows the comparison. Each point represents a percentile of the administrative contribution rate distribution, plotted against the corresponding percentile of the stated contribution rate. The median response falls close to the 45-degree line, indicating that participants generally report their current contributions accurately. There is some dispersion at higher contribution rates, where the 25th and 75th percentiles diverge from the 45-degree line, but the overall alignment is strong.

Second, we examine responses to the hypothetical "50 on 6" and "100 on 6" match scenarios for the subset of respondents who actually face those match formulas in reality. Appendix Figure B.4 plots each respondent's hypothetical response against their actual contribution rate. In both panels, the stated responses track the 45-degree line, though with greater dispersion than in Appendix Figure B.3, partly reflecting smaller sample sizes.

These exercises provide evidence that respondents understand and respond meaningfully to the hypothetical incentives presented in the survey.

As a final piece of internal validation, Appendix Figure B.5 examines internal consistency by comparing responses to a "no match" scenario elicited at the start of the survey with responses to the same scenario elicited at the end. The median response falls close to the 45-degree line, indicating that answers are stable across the survey and not sensitive to fatigue.

4.3 Survey Evidence on Saving Responses

Before using the survey responses to calibrate our model, we present initial evidence on the economic mechanisms reflected in the survey data.

Match *incentives* have a modest impact on saving Figure 3(a) illustrates how saving responds to different match rates offered without a cap. Its shows that the average increase in saving relative to the no-match baseline. Under an unlimited match, a 25% match increases saving by 1.9 pp; a 50% match yields 2.7 pps in extra saving, and a 100% match increases saving by approximately 4.8 pps.

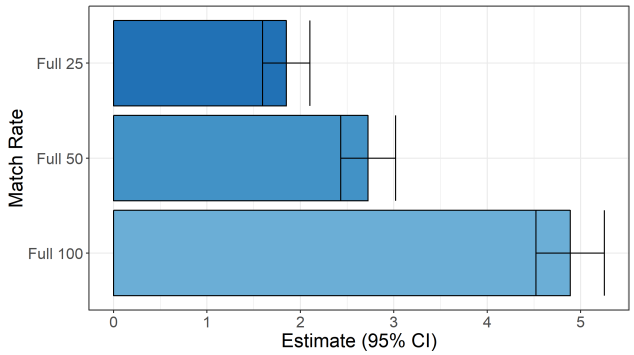
Low match rates provide better bang per buck than high match rates These patterns suggest that while saving is responsive to the presence of a match, the elasticity with respect to the match rate is modest. The table in Figure 3(b) illustrates this point. Column (1) shows the change in employee saving (relative to the no match case) from Figure 3(a). Column (2) shows the employer cost. Column (3) shows the cost per percentage point increase in employee saving. If the goal is to encourage saving, lower match rates deliver substantially greater value per dollar spent. Under this unlimited match it costs 1.5% of payroll to generate a one percentage point increase in saving using a 25% match rate. In contrast, it costs 2.9% of payroll to achieve the same effect using a 100% match.

Wealth effects from non-elective contributions are small Many firms offer both matching and non-elective contributions to employee 401(k) accounts. Non-elective contributions do not affect the marginal return to saving, but they do increase wealth, which could cause employees to reduce their own contributions. Our model assumes no such crowd-out. We present two pieces of evidence supporting this assumption.

First, we examine survey responses to a hypothetical scenario in which the employer contributes 10% of pay as a non-elective contribution. Appendix Figure B.6 plots stated contributions under this scenario against stated contributions under no match. The median

Figure 3: Survey Treatment Effects and Cost Effectiveness, by match rate

(a): Treatment Effects



Unlimited matching at different match rates

(b) Cost Effectiveness

	Δ Employee savings (p.p.)	Cost	Cost per p.p. Δ
25% Match	1.9	2.8	1.5
50% Match	2.7	6.0	2.2
100% Match	4.8	14.1	2.9

Notes: Panel (a) shows the estimated average treatment effect of introducing unlimited employer matching at different match rates, calculated as the average difference between saving under the match and saving under no match controlling for participant fixed-effects. Panel (b) reports the corresponding mean treatment effects, the cost of the policy as a share of payroll, and the cost per percentage point increase in employee saving.

response falls close to the 45-degree line, indicating that most respondents do not reduce their own saving in response to the non-elective contribution.

Second, we exploit discrete increases in non-elective contribution rates in our administrative data. We identify 20 plans that experienced sustained increases in their non-elective contribution rate during the period of observation: the 10 largest such plans with auto-enrollment and the 10 largest without. We pool across plans and estimate an event-study specification with participant fixed effects and demographic controls.[4] Figure 4 reports the results. The left panel confirms a sharp, persistent increase in NEC rates at the event date with no pre-trend. The right panel shows that employee contribution rates are flat both before and after the NEC increase, with point estimates close to zero and confidence intervals that rule out economically meaningful crowd-out. Appendix Figure B.7 shows similar results when restricting to new participants, defined as individuals observed only in their year of hire, confirming that the finding is not driven by inertia among incumbents whose plan changed mid-tenure.

Across both the survey and administrative evidence, we find no indication that employer non-elective contributions crowd out participants' voluntary saving.

[4]Figure B.10 in the appendix plots the raw time series of non-elective contributions and employee contribution rates for each plan. Despite large jumps in NEC rates, across plans, employee contributions remain on stable trends with no visible breaks.

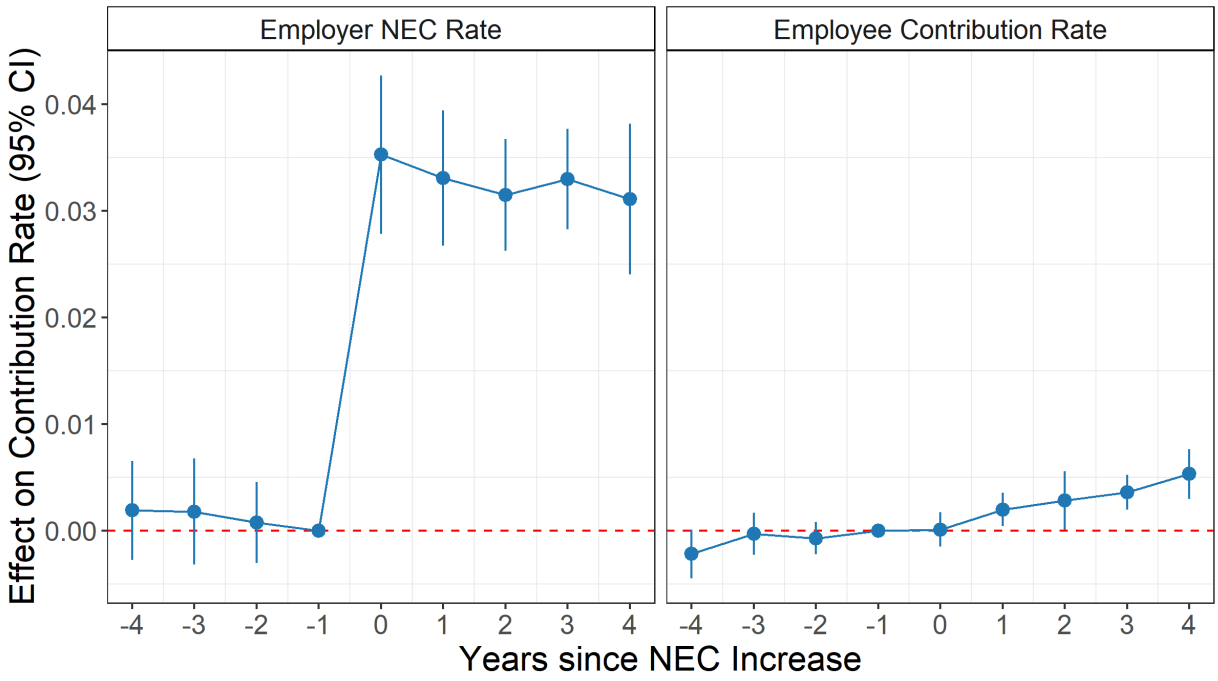


Figure 4: Change in Saving when Non-Elective is increased

Notes: The figure plots event-study estimates from a staggered difference-in-differences design. The sample is a panel covering 2010 to 2023 and includes the 10 largest auto-enrollment (AE) and 10 largest non-auto-enrollment (non-AE) plans that experienced a non-elective contribution (NEC) rate increase. The specification controls for plan and year fixed effects and quadratics in age, income, and tenure. Panel (a) plots the estimated effect on the NEC rate and Panel (b) plots the estimated effect on the elective contribution rate. Standard errors are clustered at the plan level. See Figure B.7 for a version of this graph which uses only new participants.

5 From Predictions to Policy Design: The Equity-Savings Frontier

Our survey data provide two key reference points for each respondent: how much they would save if the employer offered no match, and how much they would save if the employer offered an unlimited match at different rates. Equation 1 gives prediction rule.

5.1 Accounting for non-participants

We know from the administrative data that a substantial share of eligible employees do not participate in the retirement plan. Because our survey was not representative of non participants, without adjusting we might risk missing extensive margin responses. To account for this, we augment the survey population by introducing a synthetic group of non-participants.

Let π_0 denote the non-participation rate observed in the administrative data when no match is offered. We assign the synthetic non-participant group a weight of

$$w_{\text{NP}} = \frac{\pi_0}{1 - \pi_0} \cdot w_{\text{P}}$$

where w_{P} is the sum of weights of survey participants. This ensures that the combined population (survey participants plus synthetic non-participants) matches the observed non-participation rate under no match.

When a match is introduced, some of these non-participants may begin contributing. To estimate the magnitude of this response, we use our administrative data to estimate the effect of match introduction on participation. Specifically, we regress a participation indicator on match availability, controlling for plan fixed effects, year fixed effects, and demographic controls, exploiting within-plan variation in whether a match is offered. Table 1 reports the results. For plans with auto-enrollment, the estimated effect is 2.7 percentage points but is not statistically distinguishable from zero (column 1); we therefore assume no participation effect for these plans. For plans without auto-enrollment, the estimated participation effect is 5.1 percentage points for plans (column 2). Columns (3) and (4) allow the effect to vary by match rate and show that we cannot reject equality across rates, supporting the use of the pooled estimate.

When a match is introduced at rate m, we divide the synthetic non-participant population into two groups: those who remain non-participants, and new participants whose weight equals $0.051 \times (w_{\text{NP}} + w_{\text{P}})$ for non-auto-enrollment plans (and zero for auto-enrollment plans). For new participants, we assign predicted saving rates the average saving rate pf survey respondents with similar age and income characteristics, subject to one restriction: if the

predicted saving rate exceeds the match cap, we cap their contribution at that threshold. This reflects the assumption that marginal participants induced to contribute by the match are unlikely to save beyond the point where the highest match rate ends.

Table 1: Effect of Match on Participation

	Dependent variable: Participation			
	Pooled		By Match Rate	
	AE	Non-AE	AE	Non-AE
	(1)	(2)	(3)	(4)
Non-Zero Match	0.027	0.051**		
	(0.025)	(0.023)		
Match Rate 25%			0.031	0.023
			(0.027)	(0.040)
Match Rate 50%			0.044	0.043
			(0.028)	(0.031)
Match Rate 100%			0.012	0.065**
			(0.033)	(0.027)
Plan FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes	Yes
Observations	1,574,200	1,785,345	1,574,200	1,785,345
R^2	0.197	0.408	0.197	0.408

*p<0.1; **p<0.05; ***p<0.01. Standard errors clustered at the plan level.

Notes: This table reports estimates from regressions of a participation indicator on match availability, using within-plan variation. Columns (1) and (2) pool all non-zero match rates into a single indicator. Columns (3) and (4) allow the effect to vary by match rate. We cannot reject equality of the match rate coefficients in columns (3) and (4). All specifications include plan fixed effects, year fixed effects, and demographic controls (age, income, and tenure). Standard errors are clustered at the plan level. The regression uses all clusters, and the sample of participants is a 10% random sample of those in administrative population from 2010 to 2023.

5.2 Robustness

Our baseline results use the full sample of 1,634 survey respondents. As a robustness check, we also report results for two restricted samples. The first is a "rationalizable" sample of 1,088 respondents whose survey responses are consistent with the three behavioral assumptions underlying our prediction rule. The second excludes respondents who fail the test-retest

exercise. These are defined as those whose responses to the no-match vignette at the beginning and end of the survey differ by more than one percentage point, yielding a sample of 1,014 respondents. Our main findings are robust across all three samples (see Appendix C).

5.3 Validation

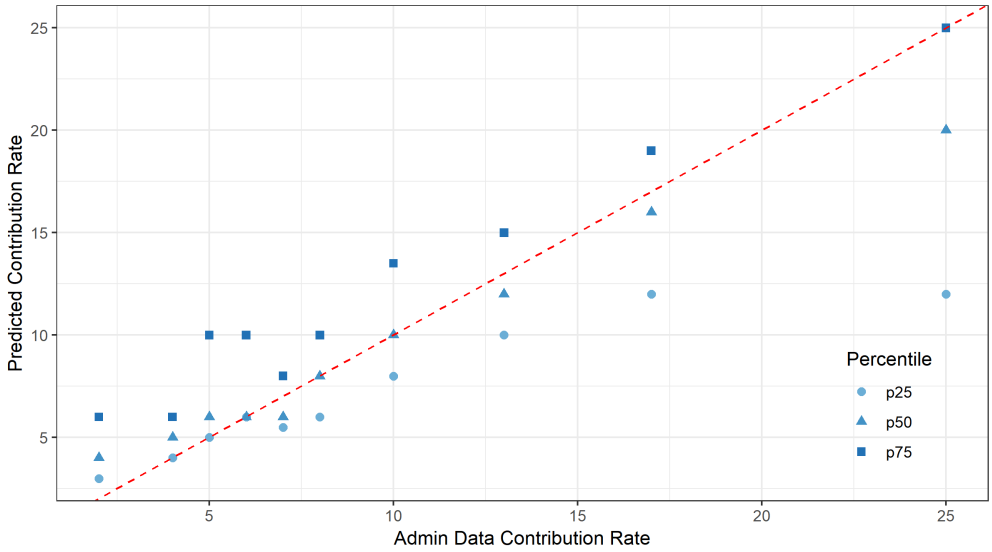
We compare the model's predictions to administrative records of actual contribution rates. The analysis focuses on the approximately 85% of the sample) whose real-world plans are simple one-tier or two-tier matches with match rates of 25%, 50%, or 100%, and so can be predicted.

Figure 5 plots the distribution of model predictions against the observed administrative contribution rates. Each point represents a bin of the administrative contribution rate, with the 25th, 50th, and 75th percentiles of the predicted rate plotted for each bin. Median predictions track the 45-degree line closely, indicating good alignment in central tendency. Dispersion is greater than in the survey validation, with noise increasing at very high contribution levels, but both the 25th and 75th prediction percentiles rise systematically with observed contribution rates.

As a more formal test, Table 2 reports regressions of administrative contribution rates on our predicted rates, with alternative sets of controls. Column (1) includes only a constant and the predicted rate; column (2) adds the two survey measures used in the prediction; column (3) includes only demographics; column (4) includes the predicted rate, survey measures, and demographics. The results show that administrative rates rise nearly one-for-one with predicted rates, and R^2 values are between 0.40 and 0.44. Most of the explained variance comes from the prediction algorithm itself: demographics alone explain only 8% of the variance (column 3), while adding the predicted rate raises the R^2 by a factor of more than 5 (column 4). The coefficient on the predicted rate remains large and statistically significant even after adding the survey inputs and demographic controls. We conclude that the model's predictions are strongly correlated with observed contribution rates, and that the predictive power is not simply due to the survey inputs or demographics alone, but reflects additional structure captured by the model.

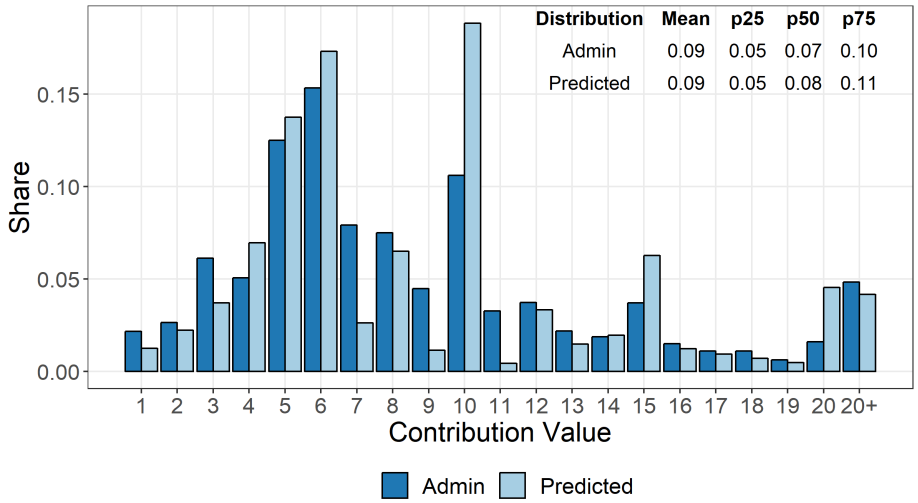
Finally, we assess the extent to which our predicted algorithm used on our sample of plan participants can replicate plan-level distributions of contributions. Our ability to match real-world contribution distributions depends on both the accuracy of the prediction algorithm and the representativeness of our sample data. We demonstrate that by appropriately re-weighting for observable demographics, we can capture the key features of contribution distributions under alternative formulas. Specifically, we divide both the survey sample and

Figure 5: Model Prediction of Current Administrative Plan Saving Decision



Notes: This figure compares respondents' saving as measured in the administrative data (on the horizontal axis), with the model-predicted contribution that they would do under their observed plan (on the vertical axis). Figure C.12 shows versions of this figure for each of our three survey samples.

Figure 6: Predicted Distributions



Notes: The figure compares the distribution of saving in the administrative data with model-predicted distributions for the 25 largest plans in our sample. Survey responses are reweighted to match administrative demographics within each plan and aggregated across plans using participant counts. Appendix Figure B.8 shows the distributional fit for the two largest individual plans. Appendix Figure C.13 shows this figure separately for each of our three survey samples.

the administrative population into mutually exclusive age-income strata, and weight each survey observation by the ratio of the administrative share to the sample share in its stratum.

Figure 6 shows the administrative distribution of contributions alongside our reweighted distribution of predicted rates. We capture the main features of the administrative distri-

bution, including the peak at the cap. The predicted distribution exhibits somewhat higher levels of bunching at round numbers (e.g., especially at 10%). This is a known feature of survey response data (Manski and Molinari (2010); Gideon et al. (2017)), and we note that a large share of the excess mass at 10% appears to draw from a deficit in the mass at 9% and 11%. We accurately match the mean, median, and the 25th and 75th percentiles of the distribution, as shown in the tables within each plot. Appendix Figure B.8 shows the same comparison separately for respondents facing a 50% match up to 6% and a 100% match up to 6%, with similarly close alignment. We conclude that our prediction algorithm, combined with appropriate reweighting, can reproduce the distribution of real-world contributions under alternative matching formulas.

Table 2: Model Predictive Power

	Dependent variable:			
	Contribution Rate			
	(1)	(2)	(3)	(4)
Predicted Rate	0.877***	1.076***		1.034***
	(0.054)	(0.164)		(0.160)
Full Match		0.174***		0.171***
		(0.034)		(0.033)
No Match		−0.321**		−0.315**
		(0.135)		(0.130)
Log Income			1.957***	0.981***
			(0.342)	(0.316)
Age			−0.468***	−0.338**
			(0.159)	(0.136)
Exp(Age,2)			0.006***	0.004***
			(0.002)	(0.001)
Tenure			0.008	−0.017
			(0.081)	(0.065)
Exp(Tenure,2)			0.001	0.001
			(0.002)	(0.002)
Observations	1,326	1,326	1,326	1,326
R^2	0.405	0.429	0.075	0.443
Note:				*p<0.1; **p<0.05; ***p<0.01

Notes: This table reports the results of linear regressions of the contribution rate observed in the administrative data on the model-predicted values. Column (1) includes on a constant and the predicted rate. Column (2) includes the answers to the questions which are used to estimate the model. Column (3) contains includes demographics. Column (4) contains the model prediction, the survey questions, and includes demographics. The regression sample is the 1,080 plan participants with one and two-tier match formulas for which we are able to predict contributions (85% of survey sample). Standard errors are heteroskedasticity-robust.

6 Are plan formulas on the equity efficiency frontier?

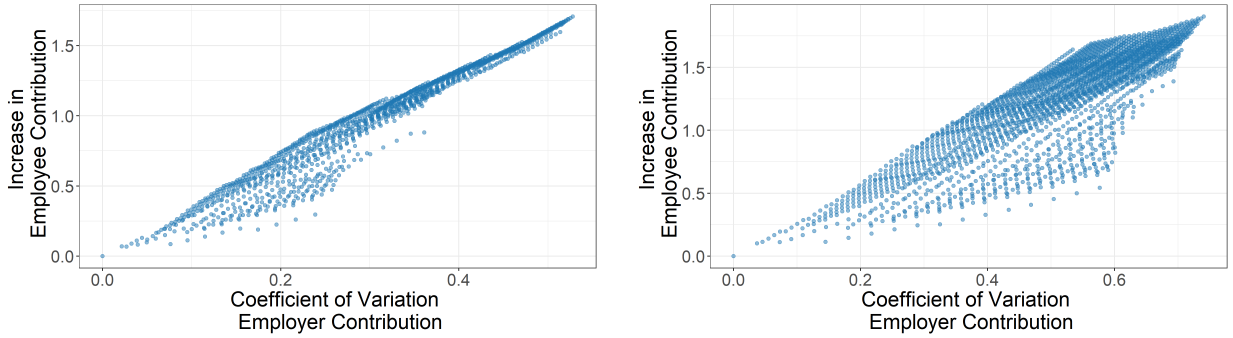
Our normative analysis begins with the assumption, commonly adopted in theoretical analyses of optimal retirement systems (starting with Diamond (1977)), that the social planner values both increased retirement saving and a more equal distribution of employer contributions. These two goals reflect the dual objectives of efficiency (encouraging greater saving) and equity (reducing inequality).

Using our survey data and model, we can predict employee saving behavior under any match formula and evaluate the equity and efficiency properties of all possible match formulas. We focus on one measure of efficiency: the increase in employee saving relative to a no-match scenario. For equity, our primary measure is the coefficient of variation in total employer contributions.

Figure 7 shows the equity and efficiency outcomes for the set of plans that meet three criteria: they cost the employer 3% of payroll, include either one or two tiers, and have a match rate that does not increase with employee contributions.

Each point in the figure represents a feasible plan, and the upper-left envelope traces out the equity-efficiency frontier: the set of plans that maximize employee saving for a given level of inequality, or equivalently, minimize inequality for a given level of saving. Two key patterns emerge. First, there is a clear equity-efficiency trade-off: along the frontier, holding employer cost constant, increasing employee saving requires accepting greater inequality in employer contributions. Second, many existing plans are strictly dominated — alternative designs exist that cost the same, generate higher employee saving, and yield more equitable distributions of employer contributions

Figure 7: Equity Efficiency Frontier



(a) Plans with auto-enrollment (b) Plans without auto-enrollment

Notes: These figures illustrate the equity and efficiency characteristics of the set of plans which cost 3% of payroll, are either single tier or two tier, and have a match rate that is non-increasing in employee contribution. Efficiency is measured using the increase in employee saving. Panel (a) shows plans with auto-enrollment and panel (b) shows plans without auto-enrollment.

6.1 Characterizing the Efficient Frontier

Figure 7 shows that many plans lie inside the frontier. Figure 8 turns to characterizing which plan features are associated with being on the frontier, and which are associated with being inside it. Each panel plots the same equity-efficiency space, but now colors each region by a plan characteristic, averaged over all plans in that region.

Moving to the frontier Panel (a) shows the average achievable match rate. Plans on or near the frontier tend to have lower match rates. This reflects the fact that high match rate concentrates benefits towards saving that would have been done anyway. Panel (b) shows the match cap, and panel (c) shows the non-elective contribution. Together, these two panels reveal the key trade-off *along* the frontier: holding the budget constant, moving toward greater efficiency (higher employee saving) is associated with a higher cap and a lower non-elective contribution, while moving toward greater equity is associated with a lower cap and a higher non-elective contribution. This is intuitive: a non-elective contribution is distributed to all employees regardless of their saving decisions, which compresses the distribution of employer contributions, while a higher cap rewards those who save more, increasing both total saving and inequality. The frontier thus traces out how the planner allocates a fixed budget between an unconditional transfer (the NEC) that promotes equity and a conditional incentive (the match with a higher cap) that promotes efficiency.

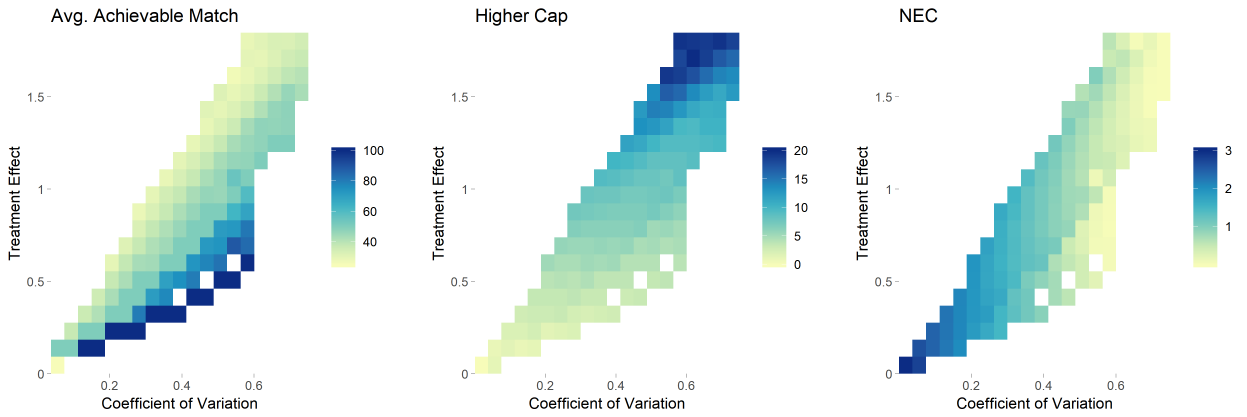
Choosing a point on the frontier. The choice of a formula on the frontier ultimately depends on the relative weight placed on efficiency and equity. A firm or policymaker who prioritizes aggregate saving will prefer plans toward the high-cap, low-NEC end of the frontier, while one who places greater weight on equalizing employer contributions will prefer plans toward the low-cap, high-NEC end. Since our dominance results hold for objective functions that values both dimensions, we remain agnostic about where on the frontier a given plan should sit. Beyond these normative considerations, firms may also weigh practical factors when selecting among frontier plans. One such consideration is simplicity: a single-tier match formula on the frontier may be easier to communicate to employees and administer than a two-tiers formula, even if the two have broadly similar efficiency and equity properties.

6.2 Moving to the frontier for the most common plans

We have shown that many plan formulas are dominated, in the sense that cost-neutral reforms could generate higher saving without increasing inequality in employer contributions. We now ask whether such dominated plans are actually observed in practice.

The second most common plan in our data offers a 50% match on contributions up to 6% of pay. Using our model, we identify all single and two tier plans that are equally costly to

Figure 8: Characterizing the Efficient Frontier



(a) Average Achievable Match (b) Top Cap (c) NEC

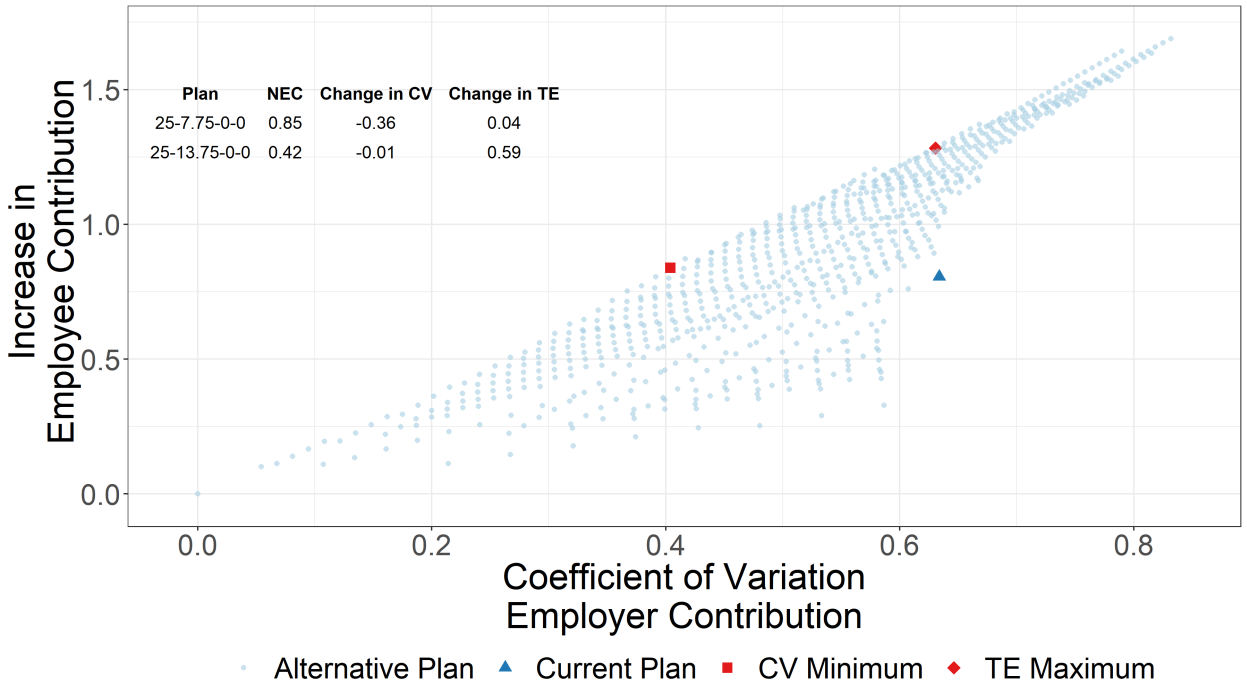
Notes: This figure characterizes which plan features are associated with different positions in the equity-efficiency space. Each panel plots the same set of plans as in Figure 7, with the horizontal axis showing the coefficient of variation in employer contributions and the vertical axis showing the increase in employee saving. Cells are colored by the average value of a plan characteristic among all plans in that region. Panel (a) shows the average achievable match rate: plans on or near the frontier tend to have lower match rates. Panel (b) shows the match cap: moving along the frontier toward greater efficiency is associated with a higher cap. Panel (c) shows the non-elective contribution (NEC): moving toward greater equity is associated with a higher NEC.

the firm and compute, for each, the resulting increase in saving (relative to no match) and the coefficient of variation in employer contributions. Figure 9 plots these outcomes. The blue triangle represents the 50% match up to 6% plan, which lies well inside the frontier.

We highlight two alternative plans. The red square shows the plan that minimizes inequality, subject to achieving at least as much saving as the current plan. The red diamond shows the plan that maximizes saving, subject to no increase in the coefficient of variation in employer contributions. The optimal plans have lower match rates and higher caps, and they balance the budget with non-elective contributions. The saving-maximizing tier increases employee contributions by almost 0.5 percentage points, an almost 60% increase in the treatment effect (relative to no match) induced by the formula. The inequality-minimizing formula reduces the coefficient of variation in employer contributions from 63% to 40%.

We can extend this analysis beyond the single most common plan. Table 3 reports the gains from moving the five most common plans in our data to their cost-equivalent, dominating alternatives, showing both the increase in employee saving and the reduction in inequality in employer contributions. Appendix Tables B.3 and B.4 present the corresponding results for the 25 most common plans assuming no auto-enrollment, and Tables B.5 and B.6 for the case with auto-enrollment, revealing a similar pattern. Most of these widely used plans lie well inside the efficiency–equity frontier.

Figure 9: Moving to the frontier for plans which have a 50% up to 6%



Notes: This figure shows the set of equity and efficiency characteristics for the set of plans which are as costly as the 50% up to 6% of pay. We show four alternative plans. We show plan which maximizes employee saving, subject to no increase in inequality. This is the red diamonds. We also show the plans which minimize inequality, subject to no decrease in employee saving. This is the red square. The characteristics of plans which achieve these frontier points, as well as the changes in equity and efficiency they generate are given in the legend.

Average Effects. Averaging across all plans, we estimate that replacing each existing formula with its identified dominating alternative would increase average employee contributions by 0.7 pps in plans without auto-enrollment and by 0.4 pps in plans with it. These increases occur without raising employer costs or increasing inequality. Alternatively, moving all plans to the frontier could reduce inequality in employer contributions, measured by the coefficient of variation by 0.3 in plans without auto-enrollment and by 0.1 in plans with auto-enrollment, again without reducing voluntary saving or increasing costs. Both reforms achieve these gains in part by introducing a non-elective contribution, which mechanically raises plan participation by 5 pp in plans with auto-enrollment and 25 pp in plans without it.

Heterogeneous Effects of the Saving-Maximizing Formulas. A natural concern is that the saving gains from moving all plans to the corresponding saving-maximizing formulas might be concentrated among a small number of very high savers, leaving most workers unaffected. In fact, we find the gains are broad-based. As expected, workers who report being financially unconstrained see the largest increase in total contributions, around 0.7 pp of salary, as they are most likely to have the capacity to save up to the higher cap. Yet even participants who report being financially constrained are predicted to contribute approximately 0.4 pp more. Finally, non-participants — presumably the most constrained group — see a substantial increase in total contributions of 0.6 pp, driven by the non-elective contribution component of the dominating formulas. The gains are also larger for workers who report currently saving too little (0.4 percentage points) than for those who report saving too much (0.2 percentage points), consistent with the formula delivering the largest saving gains to those who would presumably gains most from saving more.

Heterogeneous Effects of the Inequality-Minimizing Formulas. The reduction in inequality achieved by moving all plans to the corresponding inequality-minimizing formulas operates primarily through a redistribution of employer contributions from participants to non-participating workers. Among the former, average employer contributions decline by 0.5 pp of salary; among the latter, they increase by 1.4 pp. Importantly, within the group of participants, the decline in employer contributions is largest for workers who report having strong self-control, being financially unconstrained, and greater financial literacy, groups that are plausibly at lower risk of undersaving for retirement. Together, these patterns suggest that the redistribution implemented through the inequality-minimizing formula does not come at the expense of the workers most at risk of undersaving for retirement.

Table 3: The effect of moving to the frontier: five most common plans

Panel A: No Auto-Enrollment

	Minimize Inequality				Maximize Savings			
Current Plan	Plan	NEC	Baseline CV	Change CV	Plan	NEC	Baseline Savings	Change Savings
100-3-50-5	25-6.75-0-0	1.81	0.60	−0.35	50-1.25-25-20	0.82	6.51	0.99
50-6-0-0	25-7.75-0-0	0.85	0.63	−0.23	25-13.75-0-0	0.42	6.61	0.48
100-1-50-6	25-8-0-0	1.21	0.62	−0.27	25-17.5-0-0	0.68	6.66	0.64
100-5-0-0	25-7-0-0	2.48	0.61	−0.40	50-6-25-20	0.69	6.56	1.11
50-4-0-0	25-5-0-0	0.58	0.61	−0.23	25-8-0-0	0.22	6.30	0.38

Panel B: Auto-Enrollment

	Minimize Inequality				Maximize Savings			
Current Plan	Plan	NEC	Baseline CV	Change CV	Plan	NEC	Baseline Savings	Change Savings
100-3-50-5	50-2-25-6	1.91	0.27	−0.12	50-3.25-25-10	1.07	8.52	0.49
50-6-0-0	50-1-25-7	0.93	0.33	−0.10	50-3-25-9	0.18	8.61	0.29
100-1-50-6	50-1.75-25-7.25	1.20	0.31	−0.10	50-3-25-9.75	0.58	8.66	0.31
100-5-0-0	50-1-25-6.75	2.93	0.29	−0.16	50-2.5-25-15	1.80	8.57	0.72
50-4-0-0	50-1.25-25-4.25	0.57	0.29	−0.09	50-1.5-25-6	0.20	8.29	0.21

Notes: This table shows, for the five most common plans, the counterfactual plan which is equally costly and either minimizes inequality (subject to no fall in employee saving), or maximizes saving (subject to no increase in inequality). Panel A assumes no auto-enrollment; Panel B assumes auto-enrollment. See Appendix Tables B.3, B.4, B.5, and B.6 for further plans for the main sample. Tables C.8 and C.9 provide versions of this table the rationalizable and test-retest samples, respectively.

7 Conclusion

This paper proposes and implement a new method using surveys to improve policy design. In the context of workers responses to 401(k) matching formulas, we show that survey responses to hypothetical scenarios can accurately predict saving behavior. Using this approach, we estimate saving responses to a wide range of match schedules, including those not observed in reality, and trace out their implications for both aggregate saving and inequality.

Our findings highlight a central trade-off in plan design: raising aggregate saving often comes at the cost of greater inequality in employer contributions. Yet, many plans, including the most common ones and several safe-harbor designs, lie well inside the frontier. More saving, and lower inequality in employer contribution rates are attainable within existing budget constraints.

These conclusions rest on no specific objective function: a social planner who values both higher saving and lower inequality should prefer the plans we identify as dominating. More broadly, our results point to the value of counterfactual surveys as a general tool for policy design. This approach applies naturally to other settings, especially, when the policy space is large relative to what is observed, behavioral responses are heterogeneous and hard to predict from observational data alone, and policy objectives can be expressed through a small number of measurable criteria.

References

Ameriks, J., J. Briggs, A. Caplin, M. D. Shapiro, and C. Tonetti (2020). Long-term-care utility and late-in-life saving. *Journal of Political Economy 128*(6), 2375–2451.

Arnoud, A., T. Choukhmane, J. Colmenares, C. O'Dea, and A. Parvathaneni (2021). The Evolution of U.S. Firms' Retirement Plan Offerings. Evidence from a New Panel Data Set. Technical report, NBER.

Bachmann, R., G. Topa, and W. van der Klaauw (2022). *Handbook of economic expectations*. Elsevier.

Beshears, J., J. J. Choi, C. Clayton, C. Harris, D. Laibson, and B. C. Madrian (2025). Optimal illiquidity. *Journal of Financial Economics 165*, 103996.

Caplin, A. (2025). Data engineering for cognitive economics. *Journal of Economic Literature 63*(1), 164–196.

Carranza, G. and A. Goodman (2024). Retention or regressivity? the empirical effects of 401 (k) vesting schedules. *The Empirical Effects of 401*.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annu. Rev. Econ. 1*(1), 451–488.

Chetty, R., J. N. Friedman, S. Leth-Petersen, T. H. Nielsen, and T. Olsen (2014). Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from Denmark. *The Quarterly Journal of Economics 129*(3), 1141–1219.

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

Choi, J. J., D. Laibson, B. C. Madrian, and A. Metrick (2002). Defined contribution pensions: Plan rules, participant choices, and the path of least resistance. *Tax policy and the economy 16*, 67–113.

Choukhmane, T., J. Colmenares, C. O'Dea, J. L. Rothbaum, and L. D. Schmidt (2024, August). Who benefits from retirement saving incentives in the u.s.? evidence on gaps in retirement wealth accumulation by race and parental income. Working Paper 32843, National Bureau of Economic Research.

Choukhmane, T. and C. Palmer (2025). How Do Consumers Finance Increased Retirement Savings? Technical report.

Colarieti, R., P. Mei, and S. Stantcheva (2024). The how and why of household reactions to income shocks. Technical report, National Bureau of Economic Research.

Diamond, P. A. (1977). A framework for social security analysis. *Journal of public economics 8*(3), 275–298.

Duflo, E., W. Gale, J. Liebman, P. Orszag, and E. Saez (2006). Saving incentives for low-and middle-income families: Evidence from a field experiment with H&R Block. *The Quarterly Journal of Economics 121*(4), 1311–1346.

Engelhardt, G. V. and A. Kumar (2007). Employer matching and 401 (k) saving: Evidence from the health and retirement study. *Journal of Public Economics 91*(10), 1920–1943.

Feldstein, M. (1985). The optimal level of social security benefits. *The Quarterly Journal of Economics 100*(2), 303–320.

Fuster, A. and B. Zafar (2023). Survey experiments on economic expectations. In *Handbook of Economic Expectations*, pp. 107–130. Elsevier.

Gideon, M., B. Helppie-McFall, and J. W. Hsu (2017). Heaping at round numbers on financial questions: The role of satisficing. In *Survey research methods*, Volume 11, pp. 189.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus (2021). Five facts about beliefs and portfolios. *American Economic Review 111*(5), 1481–1522.

Goodman, A. (2025). Impatience or illiquidity? the case of early 401(k) withdrawals. *Available at SSRN 5205483*.

Hosseini, R. and A. Shourideh (2019). Retirement financing: An optimal reform approach. *Econometrica 87*(4), 1205–1265.

Hubbard, R. G., J. Skinner, and S. P. Zeldes (1995). Precautionary saving and social insurance. *Journal of political Economy 103*(2), 360–399.

Indarte, S., R. Kluender, U. Malmendier, and M. Stepner (2025). Consumption wedges: Measuring and diagnosing distortions.

Koşar, G. and C. O'Dea (2023). Expectations data in structural microeconomic models. In *Handbook of Economic Expectations*, pp. 647–675. Elsevier.

Kusko, A. L., J. M. Poterba, and D. W. Wilcox (1996). Employee decisions with respect to 401 (k) plans.

Low, H. and C. Meghir (2017). The use of structural models in econometrics. *Journal of Economic Perspectives 31*(2), 33–58.

Manski, C. F. and F. Molinari (2010). Rounding probabilistic expectations in surveys. *Journal of Business & Economic Statistics 28*(2), 219–231.

Mitchell, O. S., S. P. Utkus, and T. Yang (2007). Turning workers into savers? Incentives, liquidity, and choice in 401 (k) plan design. *National Tax Journal 60*(3), 469–489.

Moser, C. and P. Olea de Souza e Silva (2019). Optimal paternalistic savings policies. *Columbia Business School Research Paper* (17-51).

O'Dea, C. (2018). Insurance, efficiency and the design of public pensions. In *2018 meeting papers*, Volume 1037, pp. 373–383.

Papke, L. E. (1995). Participation in and contributions to 401 (k) pension plans: evidence from plan data. *Journal of Human Resources 30*(2), 311–326.

Ramnath, S. (2013). Taxpayers' responses to tax-based incentives for retirement savings: Evidence from the Saver's Credit notch. *Journal of Public Economics 101*, 77–93.

Stantcheva, S. (2023). How to run surveys: A guide to creating your own identifying variation and revealing the invisible. *Annual Review of Economics 15*(1), 205–234.

U.S. Bureau of Labor Statistics (2024, Apr). Access to and participation in defined contribution plans, march 2023. The economics daily, U.S. Department of Labor. 67

Vanguard (2025, June). How america saves 2025. Technical Report 24, Vanguard Workplace Solutions. 24th edition; data reflect savings behavior and plan trends through 2024.

A Model Appendix

A.1 Extension to Multi-Tier Formulas

The prediction rule in Equation (1) applies to single-tier match formulas. We now show that the same logic extends to formulas with multiple tiers.

Consider a two-tier plan that matches at rate m_1 on contributions up to c_1 and at rate m_2 on contributions between c_1 and c_2, where $m_1 > m_2$. An individual facing this plan encounters three distinct marginal returns to saving: $1 + m_1$ below c_1, $1 + m_2$ between c_1 and c_2, and 1 above c_2.

To predict saving under this plan, we require three observable quantities: saving under no match, $s(0, .)$; saving under an unlimited match at the lower rate, $s(m_2, 1)$; and saving under an unlimited match at the higher rate, $s(m_1, 1)$.

The prediction proceeds in two steps, working from the top tier down.

Step 1: Predict saving under the second tier alone. Consider a hypothetical single-tier plan that matches at rate m_2 up to c_2. Applying the prediction rule from Equation (1), we obtain:

$$s(m_2, c_2) = \max\{s(0, c_2),\ \min[s(m_2, 1),\ c_2]\}$$

Step 2: Predict saving under both tiers. Now consider the effect of introducing the first tier, which raises the match rate from m_2 to m_1 on contributions below c_1. For an individual saving at or above c_1 under the second-tier plan alone, this change in the match rate on inframarginal contributions generates a pure income effect. By Assumption 3, their saving does not change. For an individual saving below c_1 under the second-tier plan, the higher match rate m_1 applies at the margin, just as in the single-tier case.

Formally, replacing the no-match benchmark with the second-tier regime, we have:

$$s(m_1, c_1; m_2, c_2) = \max\{s(m_2, c_2),\ \min[s(m_1, 1),\ c_1]\}$$

The first argument of the max operator is now $s(m_2, c_2)$ rather than $s(0, .)$: an individual who would save above c_1 under the second tier alone is unaffected by the introduction of the first tier. The second argument is unchanged: an individual whose unlimited-match saving at rate m_1 falls below c_1 saves $s(m_1, 1)$, and one whose unlimited-match saving exceeds c_1 but whose second-tier saving falls below c_1 bunches at c_1.

Substituting the expression from Step 1, the full prediction rule for the two-tier plan is:

$$s(m_1, c_1; m_2, c_2) = \max\Big\{ \max\{s(0, c_2),\ \min[s(m_2, 1),\ c_2]\},\ \min[s(m_1, 1),\ c_1]\Big\}$$

General K-tier formulas. The same logic applies recursively to plans with K tiers, where tier k matches at rate m_k on contributions between c_{k-1} and c_k (with $c_0 = 0$ and $m_1 > m_2 > \cdots > m_K$). The prediction proceeds from tier K down to tier 1, with each step applying the single-tier prediction rule using the prediction from the previous step as the baseline. This requires $K + 1$ observable quantities: saving under no match and saving under an unlimited match at each of the K match rates.

A.2 A Parametric Model of Retirement Saving Choices

An individual i has income Y to allocate between current consumption (C) and saving (S), such that $Y = C + S$. Tomorrow's wealth (W) comprises saving S plus an employer match $M(S)$, which is a function of saving. Individuals have quasilinear utility:

$$\max_{S} \;\; \alpha_i \ln(Y - S) + \big(S + M(S)\big)$$

where α_i captures how much agent i values current consumption relative to future wealth. The quasilinear specification assumes that there are no wealth effects associated with the receipt of matches. We return to the assumption of no wealth effects below.

We consider match formulas that are weakly increasing and piecewise linear. For illustration, we use a simple one-tier match: the employer matches saving at rate m up to a cap S_{cap}, and provides no match beyond that point.[5]

$$M(S) = m \cdot \min(S, S_{\text{cap}}) \tag{2}$$

We will refer to this below as the 'Partial Match', and refer to the saving that i does under this as $S_i^{\text{Partial Match}}$. Our aim is to obtain an estimate ($\hat{S}_i^{\text{Partial Match}}$) of this.

Suppose we observe two pieces of data for individual i (these will be the the objects that we will collect in our survey). These are i) the saving that they would do under no match ($S_i^{\text{No Match}}$) and ii) the saving saving that they would do under an *unlimited* match at rate m ($S_i^{\text{Full Match}}$).

There are three distinct cases to consider in search of of our estimate ($\hat{S}_i^{\text{Partial Match}}$) . These are (suppressing the i subscript):

1. **Low Saver**: $S^{\text{Full Match}} < S_{\text{cap}}$. That is, even with an unlimited match, agent i would save less than the cap. Because the cap is not binding, their behavior under the partial match is the same as that under the full match:

   $$\hat{S}^{\text{Partial Match}} = S^{\text{Full Match}} \tag{3}$$

   Figure A.1a illustrates this case for an individual with 100 units of income and a 50% match up to 50 units.

2. **High Saver**: $S^{\text{No Match}} > S_{\text{cap}}$. That is, the agent would exceed the cap even without a match. Since the marginal incentive from matching under the partial match is irrelevant, the agent saves as they would without any match:

   $$\hat{S}^{\text{Partial Match}} = S^{\text{No Match}} \tag{4}$$

   Figure A.1b illustrates this cases.

3. **Medium Saver**: $S^{\text{Full Match}} > S_{\text{cap}}$ but $S^{\text{No Match}} < S_{\text{cap}}$. That is, the agent would exceed the cap under a full match, but fall short of it without any match. In this case

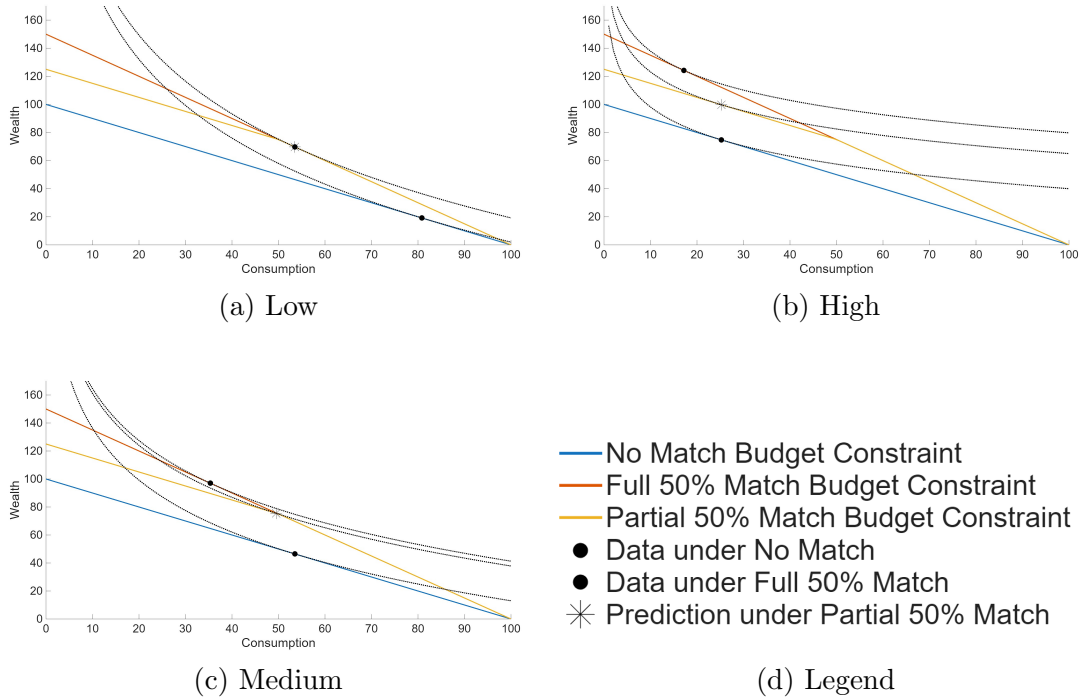
[5] The approach extends straightforwardly to multi-tier plans, as described in the appendix.

the agent optimally bunches at the cap:

$$\hat{S}^{\text{Partial Match}} = S_{\text{cap}} \tag{5}$$

Figure A.1c illustrates this case.

Figure A.1: Model Predictions for one-tier concave match formulas



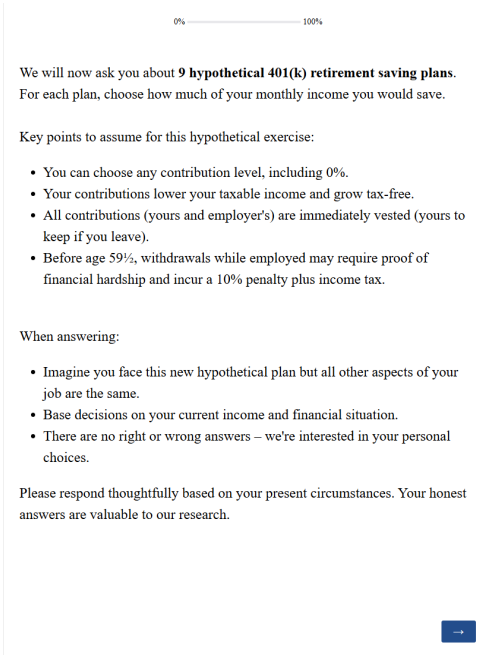
(a) Low (b) High

(c) Medium (d) Legend

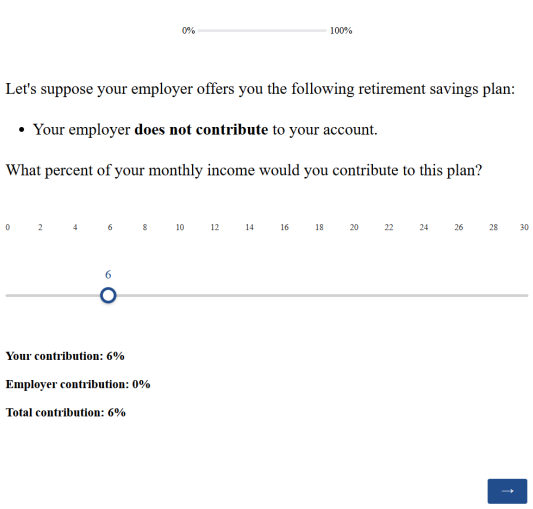
Notes: This figure illustrates the optimal period 1 consumption and period 2 wealth choices for three different agents under three budget constraints: a case where there is no matching, a case where there is unlimited matching at 50% and a case where there is matching at 50% up to a cap (S_{cap}). Panel (a) (a 'low' saver) shows the optimal consumption and wealth allocation for an agent who would save below (S_{cap}) under both no match and the unlimited match. Panel (b) (a 'high' saver) shows the optimal allocation for an agent who would save above (S_{cap}) under both no match and the unlimited match. Panel (c) (a 'medium' saver) shows the optimal allocation for an agent who would save below (S_{cap}) under no match and above (S_{cap}) the unlimited match.

This model, together with data on the saving individuals would do under each of the two rates of return that feature in the partial match (0% and m%), can be used to predict saving under any cap. We show in the appendix that this approach extends naturally to match formulas with multiple tiers. For example, to evaluate saving under a counterfactual formula that matches at 100% up to one level and 25% up to another, we would need data on saving under no match, an unlimited 25% match, and an unlimited 100% match. The logic holds because, with quasilinear utility, the marginal return to saving fully determines the agent's optimal choice — and at each segment, that marginal return is known.

B Additional Exhibits



(a) Introductory Instructions



(b) No Match Vignette

Figure B.2: Survey Interface for Hypothetical Contribution Scenarios

Notes: Panel (a) displays the introductory instructions presented before the nine hypothetical contribution vignettes. These instructions explain key assumptions that apply across all scenarios, including contribution flexibility, tax treatment, vesting, and withdrawal rules. Panel (b) shows the interface for the "No Match" vignette. Respondents use a slider to choose the share of monthly income they would contribute, with real-time updates displaying their own contribution, the employer contribution, and the total.

Table B.1: Demographic Characteristics of Survey Sample

Sample	Administrative Population	Survey Population	Final Sample
Count	1,859,488	340,653	1,634
Age			
18-35	26%	30%	7%
35-45	27%	27%	13%
45-55	24%	24%	29%
55-65	20%	15%	38%
65-75	4%	4%	13%
Income			
0-30k	10%	10%	9%
30-60k	24%	25%	22%
60-90k	24%	25%	24%
90-120k	15%	15%	16%
120-150k	10%	10%	13%
150-180k	6%	6%	7%
180k+	10%	10%	9%
Tenure			
0-3	21%	23%	17%
3-6	20%	22%	15%
6-9	14%	15%	11%
9-12	11%	11%	10%
12-15	7%	7%	8%
15+	27%	23%	40%

Notes: This table compares demographic characteristics across three groups: (i) the full administrative population of active plan participants, (ii) the subset of participants selected to receive the survey, and (iii) the final set of respondents who completed the survey. Percentages are calculated within each group for age, annual income, and job tenure categories. The final survey sample over-represents older participants and those with longer tenure compared to the administrative population, while the income distribution is broadly similar across groups. To account for these differences, all estimates in the empirical analysis are re-weighted to match the observable characteristics of the administrative population.

Table B.2: Effect of NEC Rate on Contribution Rates

	Dependent variable: Administrative Contribution Rate					
	All Participants			New Participants		
	(1) All	(2) AE only	(3) Non-AE	(4) All	(5) AE only	(6) Non-AE
NEC Rate	0.007	0.063***	0.0005	0.076***	0.089***	0.031***
	(0.005)	(0.002)	(0.0004)	(0.009)	(0.004)	(0.011)
Observations	1,322,047	1,039,592	282,455	156,070	128,360	27,710
R^2	0.142	0.155	0.115	0.221	0.234	0.178
Plan FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses.

*p<0.1; **p<0.05; ***p<0.01

Notes: This table reports estimates from linear regressions of participants' voluntary contribution rates on employer non-elective contribution rates, using the 10 largest auto-enrollment (AE) and 10 largest non-auto-enrollment (non-AE) plans that experienced a NEC rate increase from a 2010–2023 administrative panel. Panel A shows all participants. Panel B restricts to new participants, defined as individuals hired and first eligible to contribute in a given year. Columns (1) and (4) report the baseline specification with both plans with auto-enrollment, and plans without. Columns (2) and (5) show results for auto-enrollment plans only. Columns (3) and (6) show results for plans without auto-enrollment only. All regressions include fixed effects for plan, year, and demographic controls.

Table B.3: Moving to the Efficiency-Maximizing Frontier: 25 Most Common Plans (No Auto-Enrollment)

			Employee Saving			Inequality (CV)					
Current Plan	Frontier Plan	NEC	Base	New	Δ	Base	New	Δ	Δ Part.	ΔS≥6	ΔS≥12
100-3-50-5	50-1.25-25-20	0.82	6.51	7.50	0.99	0.60	0.59	−0.00	25	−8	6
50-6-0-0	25-13.75-0-0	0.42	6.61	7.09	0.48	0.63	0.63	−0.00	26	−7	7
100-1-50-6	25-17.5-0-0	0.68	6.66	7.30	0.64	0.62	0.61	−0.01	25	−7	8
100-5-0-0	50-6-25-20	0.69	6.56	7.67	1.11	0.61	0.60	−0.01	25	−5	12
50-4-0-0	25-8-0-0	0.22	6.30	6.68	0.38	0.61	0.59	−0.01	26	−2	−1
100-6-0-0	50-10-25-20	0.78	6.72	7.87	1.16	0.61	0.61	−0.00	25	−5	−6
50-7-0-0	25-17.5-0-0	0.55	6.76	7.30	0.54	0.65	0.65	−0.00	26	−7	8
100-4-0-0	50-1.5-25-20	0.79	6.35	7.51	1.16	0.60	0.60	−0.00	25	−7	6
100-1.25-0-0	25-4.75-0-0	0.13	5.95	6.28	0.33	0.58	0.55	−0.03	25	7	−0
50-8-0-0	50-1-25-19.75	0.53	6.90	7.48	0.58	0.67	0.66	−0.00	26	−7	−6
50-10-0-0	50-5.25-25-20	0.21	7.20	7.63	0.43	0.70	0.69	−0.01	26	0	−6
25-4-0-0	—	—	—	—	0.00	—	—	—	—	—	—
50-1-25-6	—	—	—	—	0.00	—	—	—	—	—	—
100-3-0-0	25-14-0-0	0.58	6.18	7.11	0.93	0.59	0.59	−0.00	25	−8	8
100-3-50-6	50-3.5-25-20	0.69	6.68	7.56	0.88	0.60	0.60	−0.01	25	−8	6
100-2-50-6	50-1-25-20	0.78	6.66	7.50	0.83	0.61	0.60	−0.00	25	−2	5
100-4-50-6	50-5.75-25-20	0.68	6.71	7.66	0.95	0.60	0.60	−0.00	25	−5	11
100-1-25-4	25-7-0-0	0.17	6.29	6.56	0.27	0.59	0.59	−0.00	25	12	−0
50-5-0-0	25-10.5-0-0	0.30	6.45	6.88	0.44	0.62	0.61	−0.00	26	−7	7
25-6-0-0	—	—	—	—	0.00	—	—	—	—	—	—
100-8-0-0	50-19-25-20	1.39	7.18	8.21	1.03	0.64	0.62	−0.02	25	−5	−12
100-3-50-7	50-5.5-25-20	0.62	6.85	7.64	0.79	0.61	0.61	−0.00	25	−6	−10
50-3-0-0	25-6-0-0	0.12	6.17	6.44	0.27	0.60	0.59	−0.01	26	12	−0
100-8.5-0-0	50-19-25-20	1.65	7.25	8.21	0.97	0.64	0.59	−0.05	25	−1	−10
100-2-50-4	25-13.75-0-0	0.57	6.34	7.09	0.75	0.59	0.59	−0.00	25	−8	7

Notes: This table reports the effect of moving each of the 25 most common plans to its cost-equivalent, efficiency-maximizing frontier alternative, assuming No Auto-Enrollment. Column (1) gives the current plan. Columns (2) and (3) give the frontier plan formula and its non-elective contribution rate. The "Employee Saving" columns report the baseline and new employee saving rates (in percent of pay) and the change between them. The "Inequality" columns report the baseline and new coefficient of variation in employer contributions and the change between them. The final three columns report changes in participation and in the share of employees saving above specified thresholds. Δ Part. is the predicted change in the participation rate (in percentage points). ΔS≥6 and ΔS≥12 report the change in the fraction of employees contributing at least 6% and 12% of pay, respectively (in percentage points). Plans marked with dashes are already on the frontier. Plans are ordered by the number of participants in the administrative data.

Table B.4: Moving to the Inequality-Minimizing Frontier: 25 Most Common Plans (No Auto-Enrollment)

			Employee Saving			Inequality (CV)					
Current Plan	Frontier Plan	NEC	Base	New	Δ	Base	New	Δ	Δ Part.	ΔS≥6	ΔS≥12
100-3-50-5	25-6.75-0-0	1.81	6.51	6.53	0.01	0.60	0.25	−0.35	25	−8	−4
50-6-0-0	25-7.75-0-0	0.85	6.61	6.65	0.03	0.63	0.40	−0.23	26	−7	−1
100-1-50-6	25-8-0-0	1.21	6.66	6.68	0.02	0.62	0.35	−0.27	25	−2	−1
100-5-0-0	25-7-0-0	2.48	6.56	6.56	0.01	0.61	0.21	−0.40	25	−3	−3
50-4-0-0	25-5-0-0	0.58	6.30	6.31	0.01	0.61	0.38	−0.23	26	−4	−13
100-6-0-0	25-8.75-0-0	2.92	6.72	6.73	0.01	0.61	0.22	−0.40	25	−3	−15
50-7-0-0	25-9-0-0	0.98	6.76	6.76	0.01	0.65	0.41	−0.24	26	−7	7
100-4-0-0	25-5.5-0-0	1.99	6.35	6.37	0.02	0.60	0.20	−0.39	25	−7	−4
100-1.25-0-0	25-1.75-0-0	0.61	5.95	5.98	0.02	0.58	0.21	−0.38	25	−0	0
50-8-0-0	50-1-25-9.5	0.98	6.90	6.90	0.00	0.67	0.43	−0.24	26	−2	−7
50-10-0-0	25-15.75-0-0	1.28	7.20	7.20	0.01	0.70	0.47	−0.23	26	−2	−8
25-4-0-0	—	—	—	—	0.00	—	—	—	—	—	—
50-1-25-6	25-6.75-0-0	0.11	6.52	6.53	0.00	0.63	0.61	−0.02	26	−1	−0
100-3-0-0	25-4-0-0	1.51	6.18	6.20	0.03	0.59	0.20	−0.39	25	−4	−1
100-3-50-6	25-8-0-0	1.95	6.68	6.68	0.00	0.60	0.27	−0.34	25	−8	−5
100-2-50-6	25-8-0-0	1.58	6.66	6.68	0.02	0.61	0.30	−0.30	25	−2	−5
100-4-50-6	25-8.75-0-0	2.24	6.71	6.73	0.02	0.60	0.26	−0.35	25	−7	7
100-1-25-4	25-5-0-0	0.43	6.29	6.31	0.03	0.59	0.43	−0.16	25	8	−0
50-5-0-0	25-6.25-0-0	0.73	6.45	6.46	0.01	0.62	0.39	−0.23	26	−7	−0
25-6-0-0	—	—	—	—	0.00	—	—	—	—	—	—
100-8-0-0	25-15.5-0-0	3.68	7.18	7.19	0.01	0.64	0.26	−0.38	25	−0	−13
100-3-50-7	25-10-0-0	2.02	6.85	6.87	0.02	0.61	0.30	−0.31	25	−8	−6
50-3-0-0	25-3.75-0-0	0.44	6.17	6.17	0.00	0.60	0.37	−0.22	26	−1	0
100-8.5-0-0	25-16.75-0-0	3.91	7.25	7.26	0.01	0.64	0.26	−0.39	25	−0	−13
100-2-50-4	25-5.5-0-0	1.28	6.34	6.37	0.03	0.59	0.27	−0.32	25	−3	−1

Notes: This table reports the effect of moving each of the 25 most common plans to its cost-equivalent, inequality-minimizing frontier alternative, assuming No Auto-Enrollment. Column (1) gives the current plan. Columns (2) and (3) give the frontier plan formula and its non-elective contribution rate. The "Employee Saving" columns report the baseline and new employee saving rates (in percent of pay) and the change between them. The "Inequality" columns report the baseline and new coefficient of variation in employer contributions and the change between them. The final three columns report changes in participation and in the share of employees saving above specified thresholds. Δ Part. is the predicted change in the participation rate (in percentage points). ΔS≥6 and ΔS≥12 report the change in the fraction of employees contributing at least 6% and 12% of pay, respectively (in percentage points). All changes are in percentage points. Plans marked with dashes are already on the frontier. Plans are ordered by the number of participants in the administrative data.

Table B.5: Moving to the Efficiency-Maximizing Frontier: 25 Most Common Plans (Auto-Enrollment)

			Employee Saving			Inequality (CV)					
Current Plan	Frontier Plan	NEC	Base	New	Δ	Base	New	Δ	Δ Part.	$\Delta S{\geq}6$	$\Delta S{\geq}12$
100-3-50-5	50-3.25-25-10	1.07	8.52	9.01	0.49	0.27	0.27	−0.00	5	−11	6
50-6-0-0	50-3-25-9	0.18	8.61	8.90	0.29	0.33	0.33	−0.00	6	−9	12
100-1-50-6	50-3-25-9.75	0.58	8.66	8.98	0.31	0.31	0.30	−0.00	5	−3	11
100-5-0-0	50-2.5-25-15	1.80	8.57	9.29	0.72	0.29	0.28	−0.00	6	−4	9
50-4-0-0	50-1.5-25-6	0.20	8.29	8.51	0.21	0.29	0.29	−0.00	6	−8	0
100-6-0-0	50-3.25-25-20	2.24	8.73	9.60	0.87	0.30	0.30	−0.00	6	−1	−21
50-7-0-0	50-1.75-25-11.75	0.55	8.75	9.05	0.30	0.36	0.36	−0.00	6	−9	10
100-4-0-0	50-3.5-25-10	1.04	8.35	9.02	0.66	0.27	0.27	−0.00	6	−6	6
100-1.25-0-0	50-1.25-25-3.25	0.15	7.99	8.18	0.19	0.24	0.24	−0.01	6	0	0
50-8-0-0	50-3-25-13.75	0.44	8.91	9.23	0.32	0.38	0.38	−0.00	6	−3	−12
50-10-0-0	50-3.5-25-20	0.63	9.23	9.61	0.38	0.43	0.43	−0.00	6	−3	−12
50-1-25-6	—	—	—	—	0.00	—	—	—	—	—	—
25-4-0-0	—	—	—	—	0.00	—	—	—	—	—	—
100-3-0-0	50-1.75-25-8	0.82	8.18	8.77	0.59	0.26	0.26	−0.00	6	−9	0
100-3-50-6	50-2-25-13	1.51	8.69	9.14	0.45	0.28	0.28	−0.00	5	−11	11
100-2-50-6	50-4-25-10	0.80	8.67	9.04	0.36	0.29	0.29	−0.00	5	1	6
100-4-50-6	50-2-25-15	1.85	8.73	9.27	0.55	0.29	0.29	−0.00	5	−5	9
100-1-25-4	50-1.25-25-5	0.22	8.27	8.39	0.12	0.26	0.26	−0.00	5	12	0
50-5-0-0	50-1.5-25-7.75	0.33	8.43	8.73	0.29	0.31	0.31	−0.00	6	−8	0
25-6-0-0	—	—	—	—	0.00	—	—	—	—	—	—
100-8-0-0	50-11-25-20	2.36	9.26	10.04	0.79	0.34	0.34	−0.01	6	−1	−12
100-3-50-7	50-2.5-25-15	1.61	8.87	9.29	0.42	0.30	0.30	−0.00	5	−11	−10
50-3-0-0	50-1-25-4.5	0.16	8.17	8.32	0.15	0.27	0.26	−0.01	6	12	0
100-8.5-0-0	50-13-25-20	2.51	9.32	10.12	0.80	0.35	0.35	−0.00	6	−1	−13
100-2-50-4	50-1.75-25-8	0.79	8.35	8.77	0.42	0.26	0.26	−0.00	5	−3	0

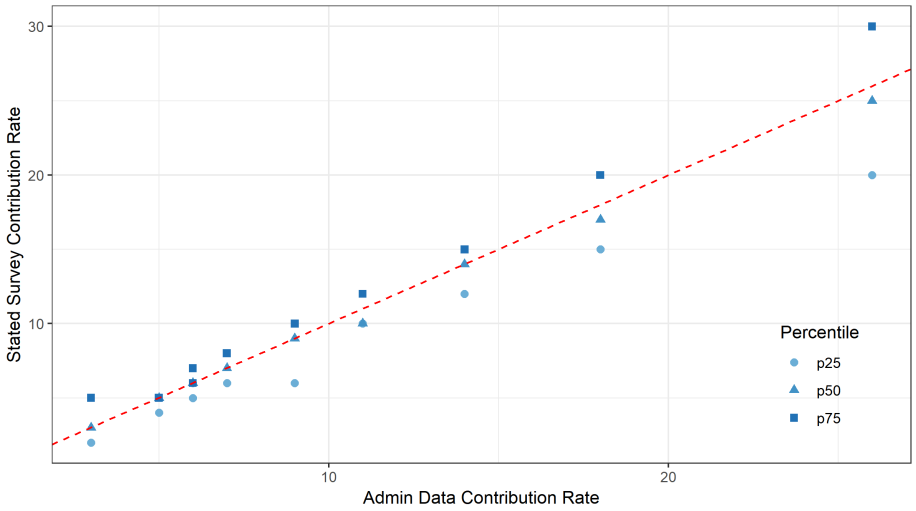
Notes: This table reports the effect of moving each of the 25 most common plans to its cost-equivalent, efficiency-maximizing frontier alternative, assuming Auto-Enrollment. Column (1) gives the current plan. Columns (2) and (3) give the frontier plan formula and its non-elective contribution rate. The "Employee Saving" columns report the baseline and new employee saving rates (in percent of pay) and the change between them. The "Inequality" columns report the baseline and new coefficient of variation in employer contributions and the change between them. The final three columns report changes in participation and in the share of employees saving above specified thresholds. Δ Part. is the predicted change in the participation rate (in percentage points). $\Delta S{\geq}6$ and $\Delta S{\geq}12$ report the change in the fraction of employees contributing at least 6% and 12% of pay, respectively (in percentage points). Plans marked with dashes are already on the frontier. Plans are ordered by the number of participants in the administrative data.

Table B.6: Moving to the Inequality-Minimizing Frontier: 25 Most Common Plans (Auto-Enrollment)

			Employee Saving			Inequality (CV)					
Current Plan	Frontier Plan	NEC	Base	New	Δ	Base	New	Δ	Δ Part.	ΔS≥6	ΔS≥12
100-3-50-5	50-2-25-6	1.91	8.52	8.52	0.00	0.27	0.15	−0.12	5	−5	−5
50-6-0-0	50-1-25-7	0.93	8.61	8.62	0.01	0.33	0.23	−0.10	6	−2	−0
100-1-50-6	50-1.75-25-7.25	1.20	8.66	8.67	0.00	0.31	0.21	−0.10	5	−2	0
100-5-0-0	50-1-25-6.75	2.93	8.57	8.58	0.01	0.29	0.13	−0.16	6	−4	−2
50-4-0-0	50-1.25-25-4.25	0.57	8.29	8.30	0.01	0.29	0.20	−0.09	6	−7	−17
100-6-0-0	50-1.75-25-7.75	3.39	8.73	8.74	0.00	0.30	0.13	−0.17	6	3	−16
50-7-0-0	50-1.25-25-8	1.05	8.75	8.75	0.00	0.36	0.24	−0.12	6	−2	1
100-4-0-0	50-1.25-25-4.75	2.34	8.35	8.36	0.01	0.27	0.11	−0.16	6	−3	−5
100-1.25-0-0	25-2-0-0	0.72	7.99	8.01	0.02	0.24	0.11	−0.13	6	−0	0
50-8-0-0	50-1.25-25-9.75	1.14	8.91	8.91	0.00	0.38	0.27	−0.11	6	−3	−9
50-10-0-0	50-2.75-25-14	1.03	9.23	9.24	0.01	0.43	0.33	−0.10	6	−3	−12
50-1-25-6	—	—	—	—	0.00	—	—	—	—	—	—
25-4-0-0	—	—	—	—	0.00	—	—	—	—	—	—
100-3-0-0	50-1.25-25-3.25	1.76	8.18	8.18	0.00	0.26	0.10	−0.16	6	−7	−1
100-3-50-6	50-1.5-25-7.5	2.17	8.69	8.69	0.00	0.28	0.16	−0.12	5	−6	0
100-2-50-6	50-1-25-7.5	1.82	8.67	8.68	0.00	0.29	0.18	−0.11	5	−2	−5
100-4-50-6	50-1.75-25-7.75	2.53	8.73	8.74	0.01	0.29	0.15	−0.13	5	−5	12
100-1-25-4	50-1.25-25-4	0.42	8.27	8.28	0.00	0.26	0.21	−0.05	5	0	0
50-5-0-0	50-1.5-25-5.5	0.68	8.43	8.44	0.00	0.31	0.21	−0.10	6	−2	0
25-6-0-0	—	—	—	—	0.00	—	—	—	—	—	—
100-8-0-0	50-2.75-25-14.5	4.04	9.26	9.26	0.00	0.34	0.19	−0.16	6	3	−11
100-3-50-7	50-3-25-8.75	1.98	8.87	8.87	0.00	0.30	0.19	−0.11	5	−2	−9
50-3-0-0	50-1-25-3.25	0.42	8.17	8.18	0.00	0.27	0.19	−0.08	6	0	0
100-8.5-0-0	50-3.5-25-15	4.17	9.32	9.33	0.00	0.35	0.19	−0.16	6	2	−11
100-2-50-4	50-1-25-4.75	1.49	8.35	8.35	0.01	0.26	0.14	−0.12	5	−2	−1

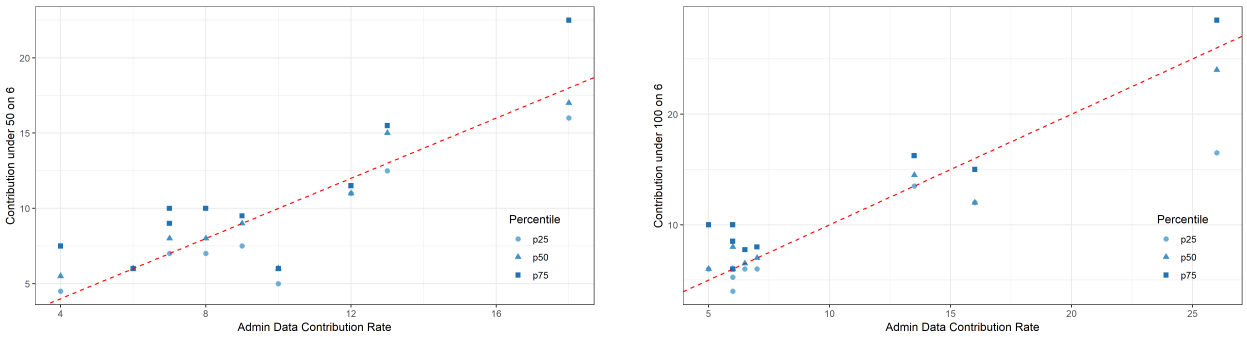
Notes: This table reports the effect of moving each of the 25 most common plans to its cost-equivalent, inequality-minimizing frontier alternative, assuming Auto-Enrollment. Column (1) gives the current plan. Columns (2) and (3) give the frontier plan formula and its non-elective contribution rate. The "Employee Saving" columns report the baseline and new employee saving rates (in percent of pay) and the change between them. The "Inequality" columns report the baseline and new coefficient of variation in employer contributions and the change between them. The final three columns report changes in participation and in the share of employees saving above specified thresholds. Δ Part. is the predicted change in the participation rate (in percentage points). ΔS≥6 and ΔS≥12 report the change in the fraction of employees contributing at least 6% and 12% of pay, respectively (in percentage points). Plans marked with dashes are already on the frontier. Plans are ordered by the number of participants in the administrative data.

Figure B.3: Comparison of Administrative and Stated Contribution Rates



Notes: This figure compares self-reported contribution rates from the survey with actual contribution rates from administrative data. Each point represents a bin of the administrative contribution rate distribution, with the 25th, 50th, and 75th percentiles of the stated rate plotted for each bin. The dashed line is the 45-degree line.

Figure B.4: Comparison of Hypothetical and Actual Contribution Rates for Aligned Plans

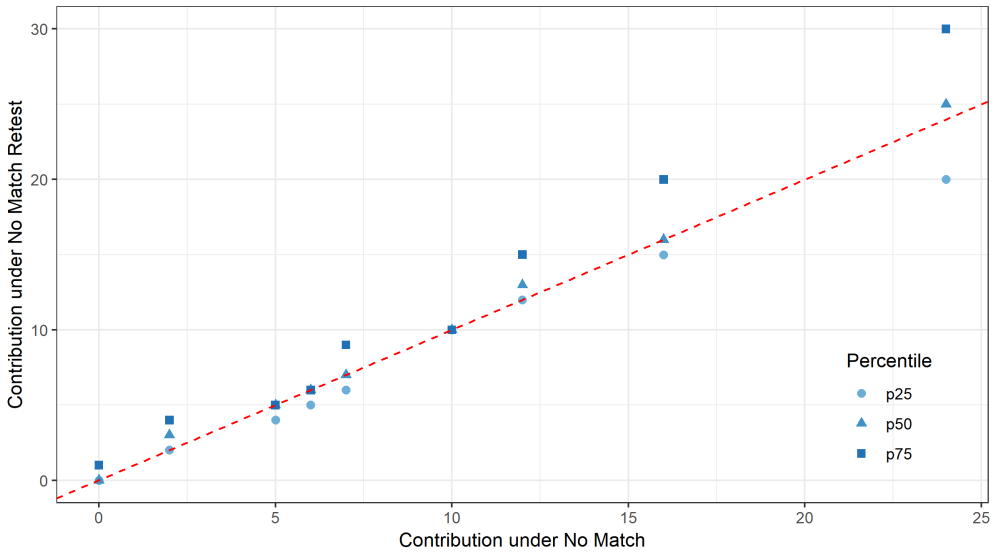


(a) 50% up to 6% plan

(b) 100% up to 6% plan

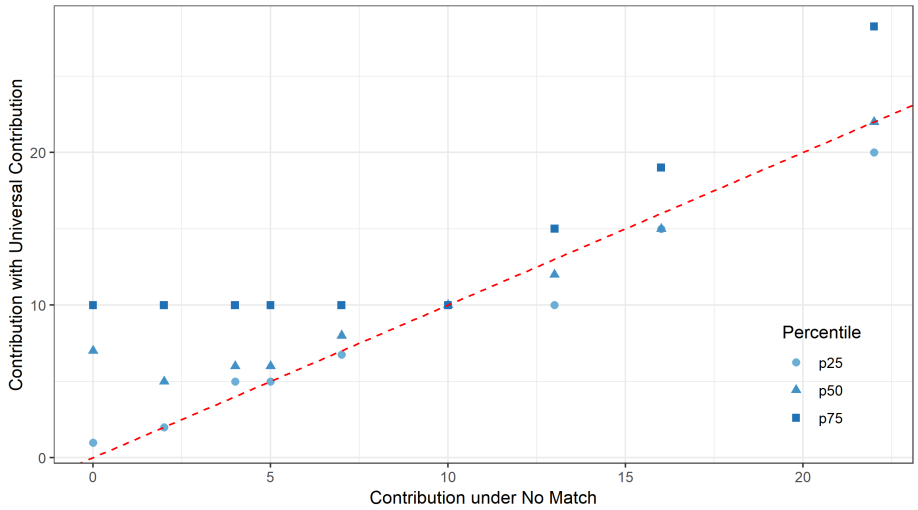
Notes: This figure compares hypothetical contribution rates from the survey with actual contribution rates from administrative data, for the subset of respondents whose actual plan matches the hypothetical scenario. Panel (a) shows respondents facing a 50% match up to 6% of pay, and panel (b) shows respondents facing a 100% match up to 6% of pay. Each point represents a bin of the administrative contribution rate, with the 25th, 50th, and 75th percentiles of the hypothetical stated rate plotted for each bin. The dashed line is the 45-degree line.

Figure B.5: Test-Retest Responses to Saving Under No Match



Notes: This figure compares responses to a "no match" hypothetical scenario asked at the beginning of the survey with responses to the same scenario asked at the end. Each point represents a bin of the initial response, with the 25th, 50th, and 75th percentiles of the retest response plotted for each bin. The dashed line is the 45-degree line. Points near this line indicate that responses are stable across the survey.

Figure B.6: Wealth Effects Survey



Notes: This figure compares responses to a "no match" hypothetical scenario with responses to a "universal contribution" scenario, in which the employer contributes a fixed amount regardless of employee contributions. Each point represents a bin of the initial response, with the 25th, 50th, and 75th percentiles of the universal contribution response plotted for each bin. The dashed line is the 45-degree line.

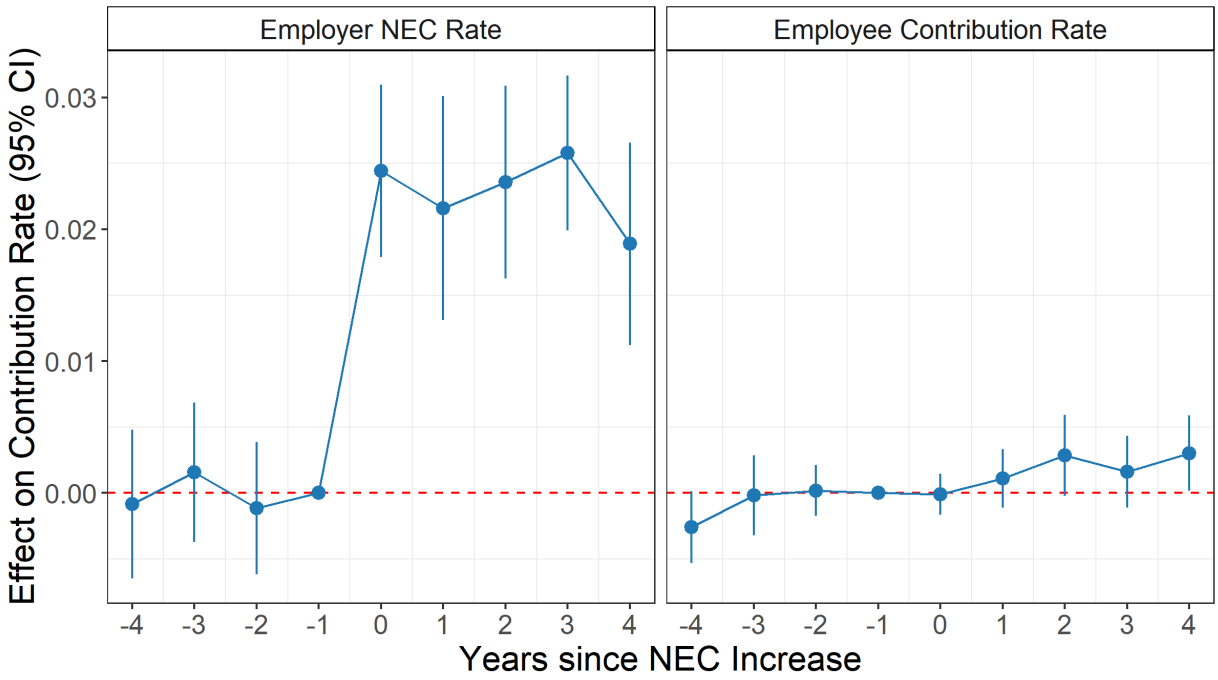
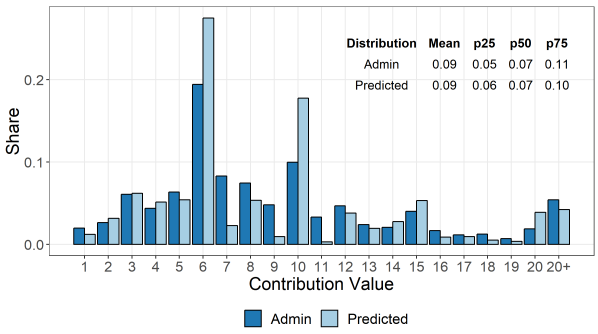


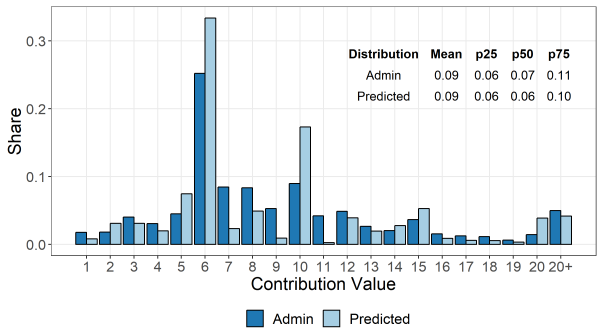
Figure B.7: Change in Saving when Non-Elective is increased: New Participants Only

Notes: This figure reports event-study estimates from a staggered difference-in-differences design using a panel spanning 2010 to 2023, restricted to new participants, defined as individuals hired and first eligible to contribute in a given year. The sample comprises the 10 largest auto-enrollment (AE) and 10 largest non-auto-enrollment (non-AE) plans that experienced a non-elective contribution (NEC) rate increase. All specifications include plan and year fixed effects and control for quadratics in age, income, and tenure. Panel (a) reports the estimated effect on the NEC rate; Panel (b) reports the estimated effect on the elective contribution rate. Standard errors are clustered at the plan level. See Figure 4 for the analogous estimates using all participants.

Figure B.8: Predicted Distributions under Two Match Formulas



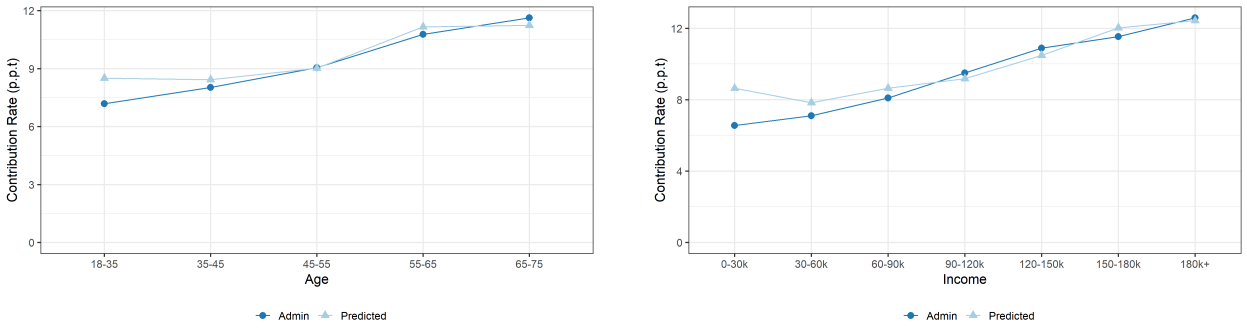
(a) Prediction for 50% up to 6% Plan



(b) Prediction for 100% up to 6% Plan

Notes: These figures compare, for two specific plans, the distribution of saving in the administrative data with model-predicted distributions. Survey responses are reweighted to match administrative demographics within each plan. Panel (a) shows the distributional fit for a plan with a 50% match up to 6% of salary; Panel (b) shows the fit for a plan with a 100% match up to 6% of salary. See Figure 6 for the aggregate distributions across the 25 largest plans in our sample.

Figure B.9: Predicted vs. Administrative Distributions by Demographics



(a) By age (b) By income

Notes: This figure compares average contribution rates in the administrative data with model-predicted average contribution rates across the 25 largest plans in our sample, disaggregated by age group (Panel a) and income group (Panel b). Survey responses are reweighted to match administrative demographics within each plan and aggregated across plans using participant counts. See Figure 6 for the aggregate distributional fit.

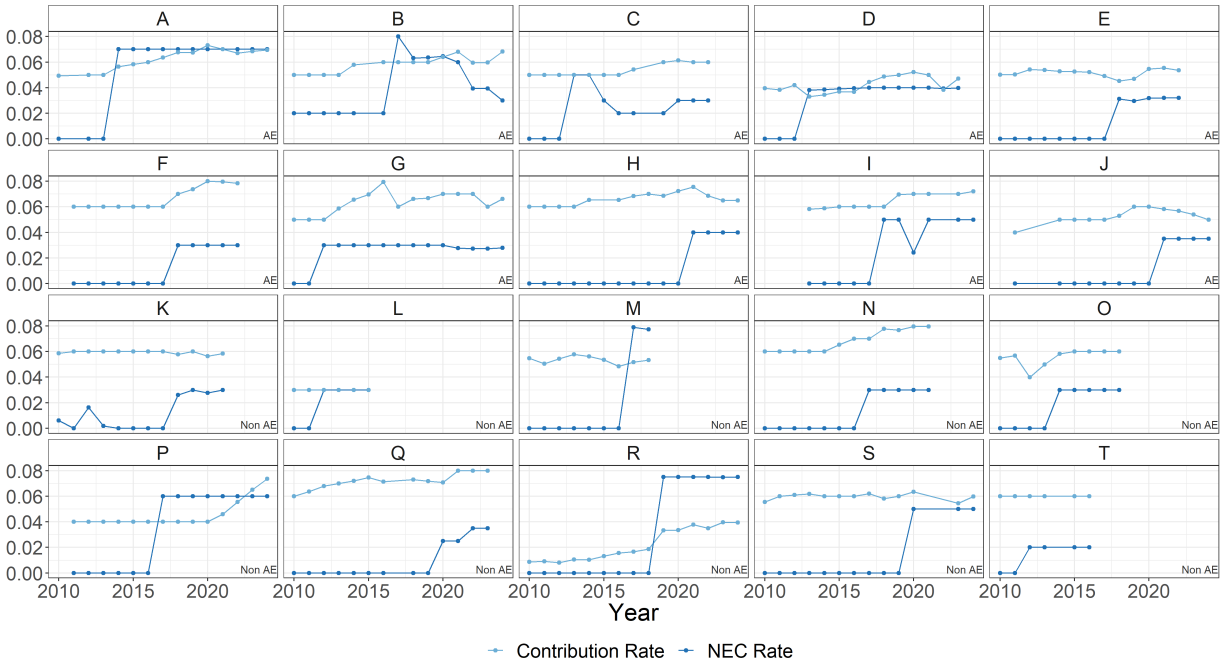
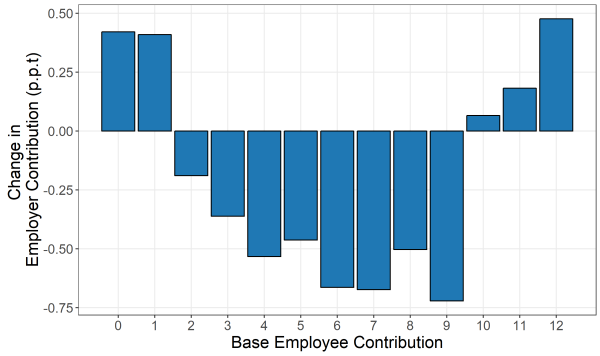


Figure B.10: Evolution of non-elective and contribution rates for 20 large plans

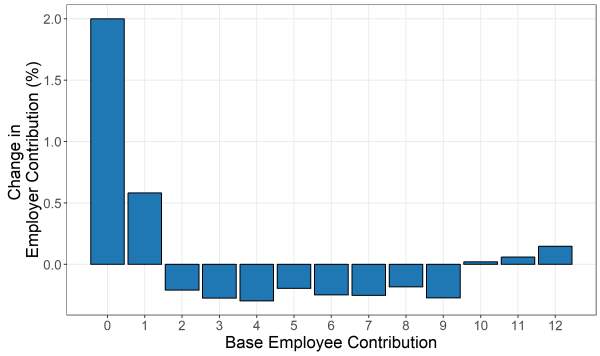
Notes: The figure displays the time series of non-elective contribution (NEC) rates and employee contribution rates for the 20 plans used in the event-study regression: the 10 largest auto-enrollment (AE) and 10 largest non-auto-enrollment (non-AE) plans that experienced a NEC rate increase. Each panel corresponds to a single plan, labeled by an anonymized identifier, with the plan type (AE or Non-AE) indicated in the bottom corner.

Figure B.11: Change in Employer Contributions, by Base Contribution Level

Efficiency Maximizing

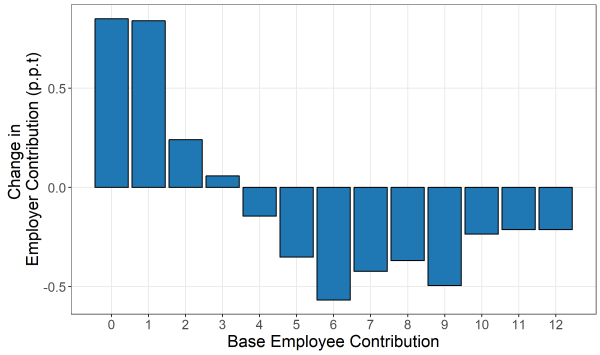


(a) Percentage points

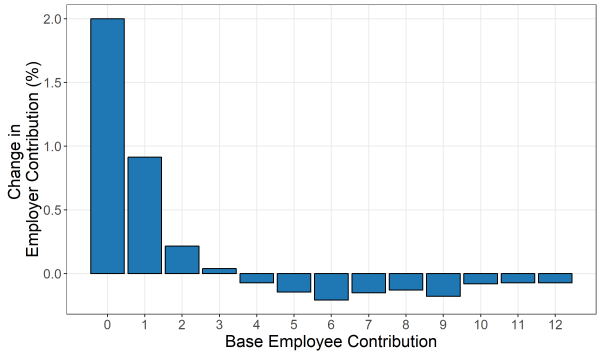


(b) Percent

Inequality Minimizing



(c) Percentage points



(d) Percent

Notes: This figure shows the change in employer contributions from moving to the efficiency-maximizing and inequality-minimizing frontier plans, plotted against the base contribution level under the current plan. Panels (a) and (c) shows the change in percentage points of pay. Panels (b) and (d) shows the change as a symmetric percent change, defined as the change divided by the average of the old and new values.

C Robustness To Different Selection

Our prediction rule relies on three behavioral assumptions. Because we can observe multiple survey responses for each individual, we can test whether each respondent's answers are consistent with these assumptions. We define a "rationalizable" sample consisting of respondents whose survey responses do not violate any of the testable implications of the model. We allow for proportionate measurement error of 20% when conducting these tests.

We implement four tests, corresponding to the three model assumptions. Not all tests apply to all respondents; some require the individual's survey responses to place them in a particular region of the contribution distribution.

Test 1: Incentives matter (cap). Assumption 1 implies that saving under an unlimited match should be at least as high as saving under a match with a cap. Each respondent answers both a vignette with an unlimited match at their randomly assigned match rate m and a vignette with the same match rate capped at 6% of pay. We require:

$$s(m, 1) - s(m, 0.06) \geq -\tau \cdot s(m, 0.06)$$

where $\tau = 0.20$. This test applies to all respondents.

Test 2: Incentives matter (match rate). Assumptions 1 and 3 also imply that saving under a match should be at least as high as saving under no match. We compare the unlimited match response at rate m with the no-match response:

$$s(m, 1) - s(0, .) \geq -\tau \cdot s(0, .)$$

This test applies to all respondents.

Test 3: Irrelevance of non-binding caps. Assumption 2 implies that if an individual saves below the cap under a capped match, raising the cap should not change their behavior. For respondents who report $s(m, 0.06) < 0.06$, the cap is non-binding, and their response under the capped match should equal their response under the unlimited match. We require:

$$\frac{|s(m, 1) - s(m, 0.06)|}{\frac{1}{2}[s(m, 1) + s(m, 0.06)]} \leq \tau$$

This test applies only to respondents with $s(m, 0.06) < 0.06$.

Test 4: No wealth effects. Assumption 3 implies that for individuals saving above the cap, changes in the match rate on inframarginal contributions should not affect saving. For respondents who report $s(m, 0.06) > 0.06$, the match is inframarginal and their saving under the capped match should equal their saving under no match. We require:

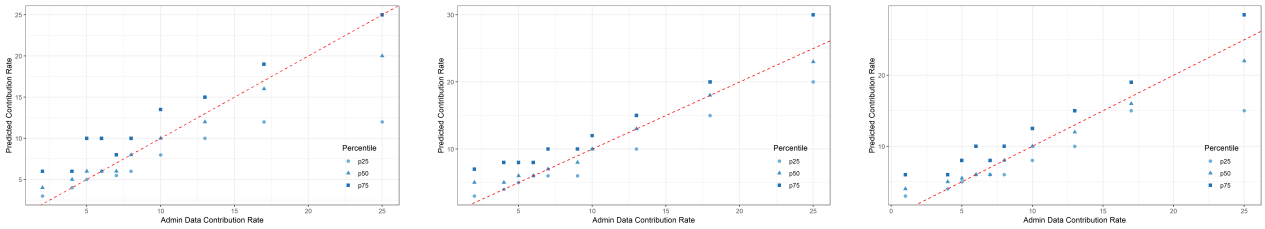
$$\frac{|s(m, 0.06) - s(0, .)|}{\frac{1}{2}[s(m, 0.06) + s(0, .)]} \leq \tau$$

This test applies only to respondents with $s(m, 0.06) > 0.06$.

Our second alternative sample uses a different exclusion criterion: rather than testing consistency with the model's assumptions, it excludes respondents who fail the test-retest, defined as those whose responses to the no-match vignette at the beginning and end of the survey differ by more than one percentage point. This yields a sample of 1,014 respondents. The remaining appendix tables report our main results under both the rationalizable and test-retest samples.

The exhibits below repeat the exhibits from our results section for each of the three samples.

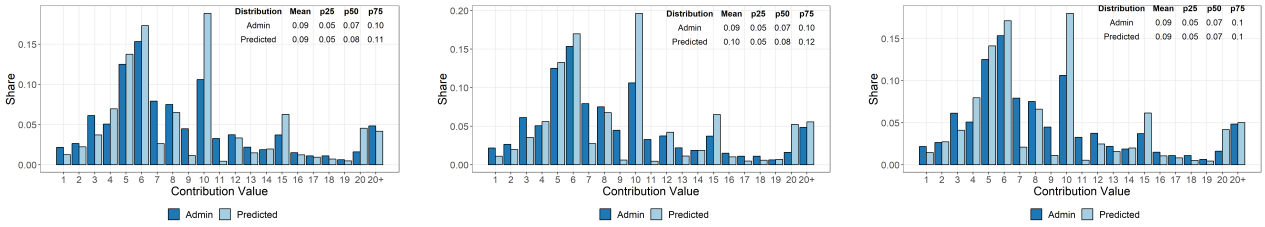
Figure C.12: Model Prediction of Current Administrative Plan Saving Decision



(a) Full sample (b) Rationalizable Sample (c) Passes Test Retest

Notes: This figure compares respondents' saving as measured in the administrative data (on the horizontal axis), with the model-predicted contribution that they would do under their observed plan (on the vertical axis).

Figure C.13: Predicted Distributions



(a) Full sample (b) Rationalizable Sample (c) Passes Test Retest

Notes: These figures compares, for distributions of saving in the administrative data, with model predicted distributions under each respondent's plan. We weight survey responses to match demographics in the administrative data.

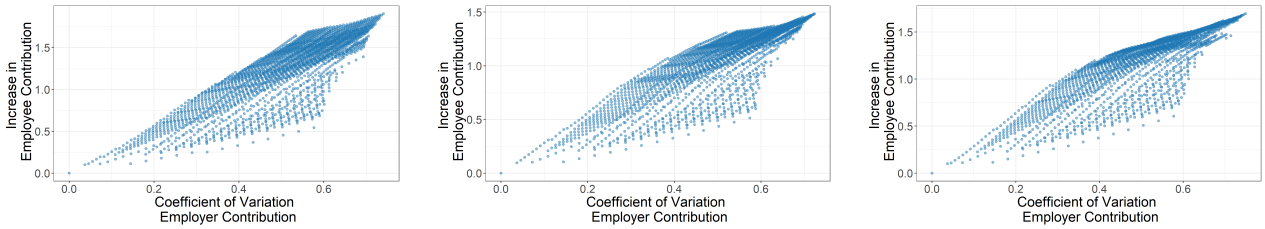
Table C.7: Model Predictive Power

	Dependent variable: Contribution Rate											
	Full Sample				Rationalizable Sample				Passes Test Retest			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Predicted Rate	0.877***	1.076***		1.034***	0.975***	1.168***		1.136***	0.938***	0.993***		0.982***
	(0.054)	(0.164)		(0.160)	(0.065)	(0.174)		(0.174)	(0.061)	(0.169)		(0.170)
Full Match		0.174***		0.171***		0.034		0.035		0.154***		0.151***
		(0.034)		(0.033)		(0.035)		(0.036)		(0.037)		(0.037)
No Match		−0.321**		−0.315**		−0.212*		−0.206*		−0.173		−0.187
		(0.135)		(0.130)		(0.116)		(0.114)		(0.131)		(0.133)
Log Income			1.957***	0.981***			1.991***	0.773*			2.149***	0.824**
			(0.342)	(0.316)			(0.425)	(0.404)			(0.411)	(0.378)
Age			−0.468***	−0.338**			−0.245	−0.149			−0.213	−0.070
			(0.159)	(0.136)			(0.198)	(0.164)			(0.172)	(0.148)
Exp(Age,2)			0.006***	0.004***			0.004*	0.002			0.003**	0.001
			(0.002)	(0.001)			(0.002)	(0.002)			(0.002)	(0.001)
Tenure			0.008	−0.017			−0.130	−0.113			−0.031	−0.089
			(0.081)	(0.065)			(0.106)	(0.085)			(0.097)	(0.076)
Exp(Tenure,2)			0.001	0.001			0.004	0.003			0.001	0.002
			(0.002)	(0.002)			(0.003)	(0.002)			(0.003)	(0.002)
Observations	1,326	1,326	1,326	1,326	873	873	873	873	1,014	1,014	1,014	1,014
R^2	0.405	0.429	0.075	0.443	0.518	0.519	0.068	0.527	0.480	0.494	0.070	0.503

*p<0.1; **p<0.05; ***p<0.01

Notes: This table reports the results of linear regressions of the contribution rate observed in the administrative data on the model-predicted values. Column (1) includes only a constant and the predicted rate. Column (2) adds the two survey responses used to construct the prediction. Column (3) includes only demographic controls: log income, age, age squared, tenure, and tenure squared. Column (4) includes the predicted rate, survey responses, and demographic controls.

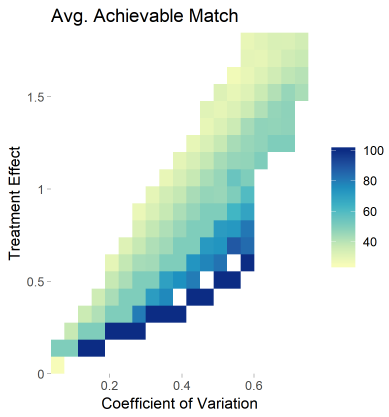
Figure C.14: Possibilities



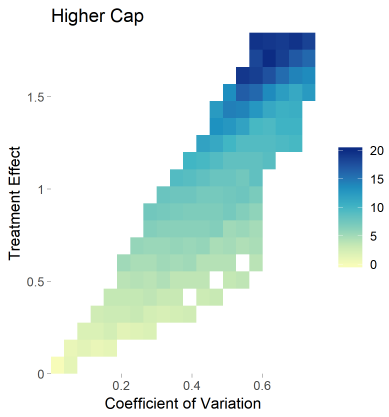
(a) Full sample (b) Full Sample (c) Passes Test Retest

Notes: These figures illustrate the equity and efficiency characteristics of the set of plans which cost 3% of payroll, are either single tier or two tier, and have a match rate that is non-increasing in employee contribution. In both panels, efficiency is measured using the increase in employee saving. In panel (a) the the equity dimension is represented using the coefficient of employer contributions.

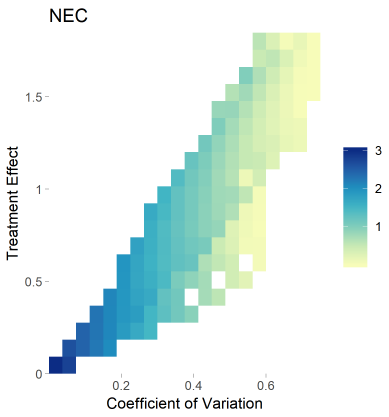
Figure C.15: Contour plots



(a) Average Achievable Match

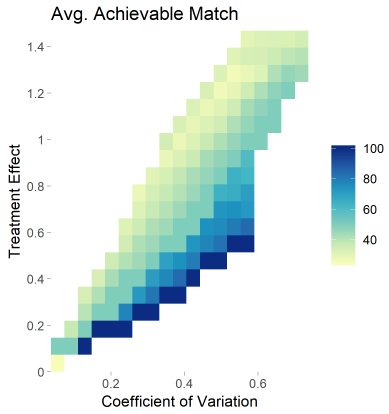


(b) Top Cap

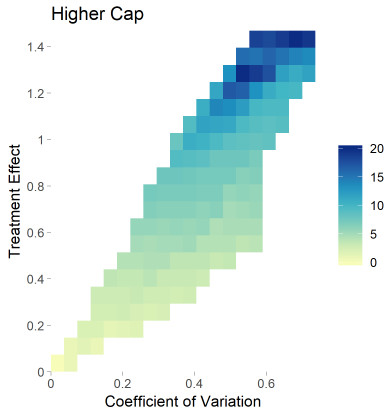


(c) NEC

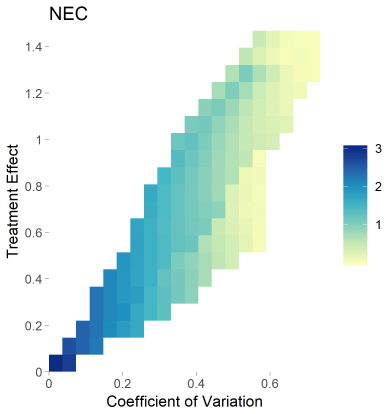
Main Specification



(d) Average Achievable Match

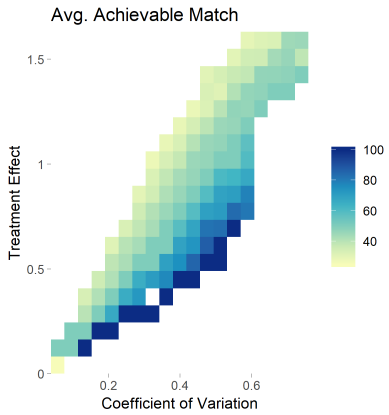


(e) Top Cap

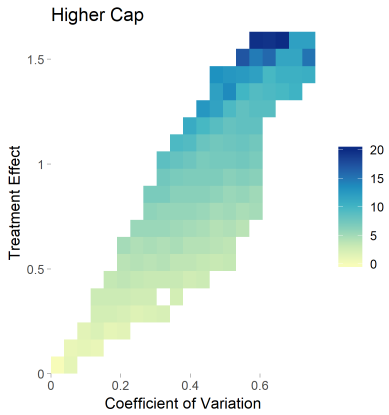


(f) NEC

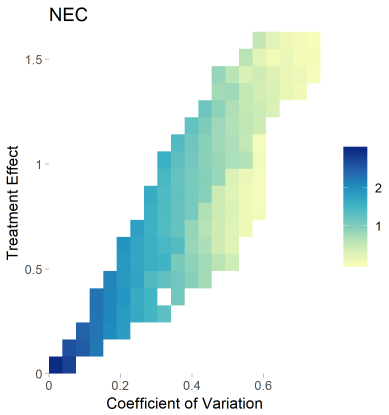
Rationalizable Sample



(g) Average Achievable Match



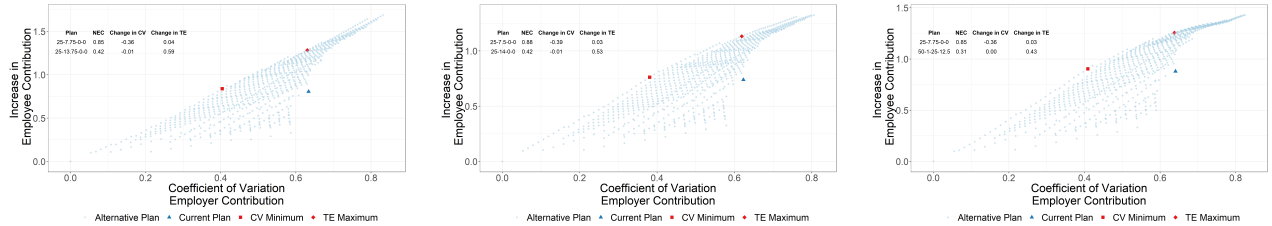
(h) Top Cap



(i) NEC

Passes Test Retest

Figure C.16: Moving to the frontier for plans which have a 50% up to 6%



(a) Full sample (b) Rationalizable Sample (c) Passes Test Retest

Notes: This figure shows the set of equity and efficiency characteristics for the set of plans which are as costly as the 50% up to 6% of pay. We show four alternative plans. We show the double tier and single tier plans which maximizes employee saving, subject to no increase in inequality. These are the red and blue diamonds. We also show the double tier and single tier plans which minimize inequality, subject to no decrease in employee saving. These are the red and blue squares. The characteristics of plans which achieve these frontier points, as well as the changes in equity and efficiency they generate are given in the legend.

Table C.8: The effect of moving to the frontier: five most common plans: rationalizable sample

Panel A: No Auto-Enrollment

	Minimize Inequality				Maximize Savings			
Current Plan	Plan	NEC	Baseline CV	Change CV	Plan	NEC	Baseline Savings	Change Savings
100-3-50-5	25-6.25-0-0	1.86	0.59	−0.36	50-1.75-25-20	0.75	6.84	0.71
50-6-0-0	25-7.5-0-0	0.88	0.62	−0.24	25-14-0-0	0.42	6.96	0.39
100-1-50-6	25-7.75-0-0	1.24	0.61	−0.28	25-18.25-0-0	0.68	6.98	0.49
100-5-0-0	25-6.75-0-0	2.52	0.60	−0.40	50-6.75-25-19.5	0.64	6.90	0.79
50-4-0-0	25-4.75-0-0	0.62	0.60	−0.24	25-8.25-0-0	0.19	6.69	0.35

Panel B: Auto-Enrollment

	Minimize Inequality				Maximize Savings			
Current Plan	Plan	NEC	Baseline CV	Change CV	Plan	NEC	Baseline Savings	Change Savings
100-3-50-5	50-1-25-6	2.14	0.26	−0.13	50-1.25-25-11.25	1.45	8.96	0.38
50-6-0-0	50-1-25-7	0.94	0.32	−0.10	50-1-25-9.75	0.58	9.07	0.19
100-1-50-6	50-1-25-7.25	1.39	0.29	−0.10	50-1-25-10.75	0.99	9.10	0.20
100-5-0-0	50-1.25-25-6.5	2.92	0.27	−0.16	50-3.75-25-14	1.61	9.03	0.48
50-4-0-0	50-1-25-4.5	0.59	0.27	−0.08	50-1-25-6.25	0.27	8.81	0.19

Notes: This table shows, for the five most common plans, the counterfactual plan which is equally costly and either minimizes inequality (subject to no fall in employee saving), or maximizes saving (subject to no increase in inequality). Panel A assumes no auto-enrollment; Panel B assumes auto-enrollment.

Table C.9: The effect of moving to the frontier: five most common plans, sample passing test retest

Panel A: No Auto-Enrollment

	Minimize Inequality				Maximize Savings			
Current Plan	Plan	NEC	Baseline CV	Change CV	Plan	NEC	Baseline Savings	Change Savings
100-3-50-5	25-6.75-0-0	1.80	0.60	−0.35	50-2-25-20	0.74	6.56	0.69
50-6-0-0	25-7.75-0-0	0.85	0.64	−0.23	50-1-25-12.5	0.31	6.68	0.38
100-1-50-6	25-8-0-0	1.21	0.62	−0.27	50-2.25-25-14	0.39	6.71	0.44
100-5-0-0	25-7-0-0	2.47	0.61	−0.40	50-6.75-25-20	0.66	6.62	0.83
50-4-0-0	25-5-0-0	0.59	0.61	−0.23	25-8-0-0	0.23	6.36	0.37

Panel B: Auto-Enrollment

	Minimize Inequality				Maximize Savings			
Current Plan	Plan	NEC	Baseline CV	Change CV	Plan	NEC	Baseline Savings	Change Savings
100-3-50-5	50-1.25-25-6.5	2.01	0.28	−0.12	50-3.5-25-10	1.03	8.58	0.46
50-6-0-0	50-1.5-25-7.25	0.78	0.34	−0.09	50-2-25-9.5	0.37	8.70	0.24
100-1-50-6	50-2-25-7.5	1.11	0.31	−0.09	50-2.5-25-10	0.69	8.75	0.26
100-5-0-0	50-1.25-25-7	2.83	0.30	−0.16	50-6-25-11.75	1.28	8.66	0.60
50-4-0-0	50-1-25-4.75	0.55	0.30	−0.07	50-1.5-25-6	0.21	8.38	0.18

Notes: This table shows, for the five most common plans, the counterfactual plan which is equally costly and either minimizes inequality (subject to no fall in employee saving), or maximizes saving (subject to no increase in inequality). Panel A assumes no auto-enrollment; Panel B assumes auto-enrollment.