Save(d) by Design

Saurabh Bhargava[*], Lynn Conell-Price[**], Richard T. Mason[*+], Shlomo Benartzi[++]

Carnegie Mellon University[*]
Consumer Financial Protection Bureau[**]
Behavioral Finance Institute for Innovation, Voya Financial[+]
UCLA Anderson School of Management[++]

This Version: September 2021

We begin by presenting novel administrative evidence from 840 401(k) plans with automatic enrollment (AE) indicating that the risk of retirement insecurity extends to a significant share of actual *enrollees*. Hypothesizing that this risk is materially affected by the initial decision to enroll at the default rate or to personalize enrollment at a higher rate, we investigate the sensitivity of initial enrollment to non-economic features of digital design that increasingly shape plan engagement. Specifically, we describe three large-scale field experiments, administered across 500 AE plans, that vary the *psychological design* (i.e., color, layout, phrasing, informational salience) of the digital interface from which employees decide to confirm, personalize, or decline enrollment. The field studies, supplemented by hypothetical choice experiments and a survey of hundreds of plan administrators, yield four findings. First, we show that modest changes to the psychological design of the interface result in sizable increases in personalized enrollment, full match take-up, and average contributions—equivalent to those predicted from a 68 to 74 percent increase in the modal match. Second, we show that marginal personalized enrollees appear to increase their initial contributions substantially and to an extent equivalent to inframarginal counterparts, implying potentially significant welfare gains due to design. Third, lab evidence indicates that design does not affect enrollment through standard economic channels of preferences/beliefs or often-cited behavioral frictions (inattention, confusion, distrust) and suggests instead that enrollment may emerge from a non-deliberative process in which design shifts affective appraisals. Finally, an industry survey shows that most plan administrators underestimate the potency of design and cannot identify optimal design elements. The findings raise new concerns about the retirement preparedness of 401(k) enrollees, highlight the potentially profound, and largely unrecognized, influence of digital design on financial decisions such as savings, and challenge basic economic assumptions underlying prevailing approaches to consumer protection and welfare analyses.

Acknowledgements: We thank Linda Babcock, Karna Basu, John Beshears, David Card, Stefano DellaVigna, Jon Guryan, Kareem Haggag, Hilary Hoynes, Prasad Krishnamurthy, Paola Guiliani, David Laibson, George Loewenstein, Ted O'Donoghue, Mark Patterson, Vikram Pathania, John Payne, Jan Potters, Silvia Saccardo, Emmanuel Saez, Jesse Shapiro, Steve Shu, Richard Thaler, Steven Thomas, and Giovanni Urga for generous insight and feedback. We also thank our collaborators at Voya Financial (Voya) including Thomas Armstrong, Lombard Gasbarro, Daniella Listro, Marilyn Morgan, and Charlie Nelson. Cassandra Taylor and Ben Schenck provided excellent project support. Mason leads the Behavioral Finance Institute for Innovation at Voya and is a Visiting Scholar at CMU, and Benartzi is a professor at UCLA and a paid consultant to Voya. Voya covered a portion of the human subject fees for the project and was provided an opportunity to review the manuscript before public release for confidentiality and factual accuracy. The authors retained full editorial control and any views, and errors, should be attributable to the authors and not to their affiliated institutions.

1 Introduction

By numerous accounts, a substantial share of working Americans save inadequately for retirement. Such inadequacy implies future risk of being unable to pay for basic retirement expenses without resort to means-tested public assistance, a return to the workforce, or bankruptcy—a scenario we hereafter refer to as *retirement insecurity*.[1] Much of the public, private, and academic discourse regarding the retirement preparedness of employees has focused on expanding access to and participation in employer-sponsored plans.[2] At first glance, this focus seems sensible given the economic attractiveness of 401(k) plans, the now predominant employer-sponsored vehicle through which approximately 90 million employees have saved $6.7 trillion. In theory, these plans should encourage enrollee savings through pre-tax contributions, portability across jobs, and, for the large majority of plans, generous matching incentives. The presumed importance of plan participation helps to explain why many cite the widespread adoption of automatic enrollment (AE), shown to sharply increase participation (Madrian and Shea 2001), as a pivotal development in the evolution of 401(k) plan administration.

We motivate the present research with new data indicating that the risk of retirement insecurity extends to a substantial share of employees currently *enrolled* in 401(k) AE plans. The evidence offers a rare snapshot of contemporaneous retirement preparedness informed by administrative records from 186,000 enrollees across 840 plans—the universe of enrollees from the small-to-medium market segment of a major US plan provider and record-keeper, Voya Financial (hereafter, Voya).[3] To characterize retirement risk, we eschew the traditional approach in the literature of estimating the share of working income the average enrollee might replace in retirement and, instead, estimate the likelihood that an enrollee will accumulate less than some minimum threshold of savings plausibly associated with retirement security. Specifically, we generated several distributions of projected savings for each enrollee by simulating the year-to-year time-path of accumulation associated with conservative and/or widely varying assumptions regarding future market returns, income growth, retirement age, and inertia with respect to contribution adjustments. For each set of assumptions, we then consulted commercial annuity-pricing calculators and research on the financial requirements of retirement to map the resulting distribution to a score indicating an enrollee's risk of retirement insecurity.

The projections suggest that anywhere from 27 to 65 percent of actual 401(k) enrollees face a non-trivial risk—arbitrarily defined as 25 percent or higher—of retirement insecurity. The width of the interval largely reflects the wide range of plausible future market returns and the financial requirements of

[1] Our definition borrows from the Elder Index, conceived by Mutchler, Li, and Xu (2016).
[2] For example, in a prominent survey of 401(k) plans, a majority of plans cited increased participation as the main objective of plan education initiatives (2019 PCSA's 63rd Annual Survey Report); policy briefs from institutes/academics commonly cite access as a central barrier to employee savings.
[3] The snapshot analysis targets firms with less than 5,000 employees. According to establishment data from the US Census, approximately 65 percent of employees are at establishments of this size.

retirement. Our preferred estimates imply that 44 percent of current 401(k) enrollees are at risk of accumulating less than $200,000 by retirement—an amount sufficient to purchase an annuity guaranteeing monthly retirement income (inclusive of Social Security) approximately equivalent to the retirement security threshold asserted by recent research (Mutchler et al. 2016). Our analyses also indicate that average retirement risk varies across plans, even after controlling for differences in plan features and demographics. We view the assessment as conservative in that it ignores the potentially substantial leakage in savings arising from loan default, unemployment, and job transition (Beshears et al. 2020). Overall, while the precise prevalence of risk depends on one's preferred assumptions, the exercise conveys that the problem of retirement insecurity, commonly understood to afflict employees outside of a 401(k) plan, also affects a substantial share of enrollees.

We hypothesize that one factor that may help explain the pattern of high (and variable) retirement risk in AE plans involves the non-economic features of plan design that shape an employee's initial enrollment. Practically, most newly-eligible employees in an AE 401(k) plan can either enroll through inaction or actively visit an online enrollment interface to confirm automatic enrollment, personalize enrollment at an adjusted contribution rate, or decline enrollment. There are at least two reasons why the initial decision to personalize enrollment may have outsized influence on an employee's long-run financial security. First, because default rates are typically set below targets recommended by financial planners and, for plans with a match, the threshold up to which contributions are matched, most employees aspiring to save at target rates or to fully take-up the match must increase their initial contribution via personalized enrollment. Second, employees have a documented propensity of insufficiently adjusting away from their initial contribution rate over time (a pattern evident in our own administrative data).[4] Low default contribution rates that fall short of the plan match and inertia in contribution adjustments effectively amplify the long-run financial significance of initial enrollment.

Notably, while federal requirements strictly govern the economic structure of 401(k) plans and the timing, and content, of plan-related disclosures, plan providers and employers retain considerable discretion in designing the (increasingly digital) enrollment experience. As a result, AE plans vary in what we term "psychological design"—the visual presentation of enrollment options (e.g., differences involving color, phrasing, layout, or the salience of previously communicated plan details), the usability and interactivity of the enrollment interface, access to decision aids such as retirement calculators, and the intensity of marketing efforts (e.g., email reminders) to encourage employees to make an active, as opposed to passive, decision to personalize enrollment. From the perspective of standard economic theory, variation in psychological design should not affect the decision to save (initially or otherwise) so

[4] Several studies have documented, or discussed, inertia in employee benefit plan decisions such as: Choi et al., 2002; Choi, Laibson, and Madrian, 2011; Beshears et al. 2016; Cronqvist, Thaler, and Yu 2018.

long as such variation does not materially change the economic incentives of enrollment or provide new decision-relevant information. This presumption, however, has been implicitly rejected by the massive private sector and human capital investment in digital consumer engagement. By an array of metrics, firms are spending extraordinary amounts to integrate digital technology across business functions, an increasing share of (financial) consumer interactions are occurring via digital channels, and there has been rapid growth in the market for laborers skilled in digital development and design.[5] While economists have studied the transformative effects of digitization on a range of economic activity (Goldfarb and Tucker 2019), they have yet to systematically investigate its potentially profound effect on consumer decisions.

Our primary aim in this paper is to investigate the influence of psychological design on the savings of 401(k)-eligible employees and to clarify the implications of these findings for savings policy, consumer protection, economic models of financial decisions, and welfare analyses. Our central evidence draws from three field experiments that randomized several aspects of the psychological design of the otherwise standardized interface from which 8,565 employees from 500 401(k) AE plans made their initial enrollment decisions. The studies were administered to the near-universe of small-to-midsize market plans for which Voya was the record-keeper from December 2016 to March 2018. To clarify the behavioral mechanisms underlying the influence of design, we supplemented the field studies with evidence from an experimental paradigm in which a distinct sample of 6,871 employees made hypothetical enrollment decisions from an online interface programmed to resemble its real-life analogue in look and functionality. Finally, to better understand the sophistication of those responsible for plan oversight and employee engagement, we surveyed several hundred plan sponsors to test the accuracy with which they could forecast the influence of design and identify successful design elements. To our knowledge, the collective scope of the field experiments, supplementary lab studies, and survey of plan executives is without precedence in the literature on 401(k) savings.

We operationalized our tests of psychological design by defining two categories of variation relative to the pre-existing commercial enrollment interface. This baseline interface displayed introductory text asking employees to select from one of three available enrollment options represented by a horizontal array of boxes displayed across the middle of the screen. Within each box, a headline described the enrollment option: "I want to enroll with other choices." [Personalize]; "Let my scheduled automatic enrollment go through." [Confirm]; "I do not wish to enroll." [Decline]. The personalized enrollment box displayed the additional explanatory subtext: "Note: This enrollment will cancel your scheduled automatic enrollment." A first category of design variation, *enhanced presentation*, involved replacing the existing headlines with more concise and instrumental language (e.g., "Do It Myself"

[5] As examples, one industry report forecasts that the private sector will spend $6.8 trillion from 2020 to 2023 to integrate digital technology across business functions while a 2020 Mckinsey & Co. Survey estimates that 65 percent of consumer interactions in the US already occur via digital platforms. In the most recent Occupational Outlook Handbook, the BLS projects that growth in the web developer and digital designer job category from 2019 to 2029 will be double the national average.

[Personalize]), adopting tone-neutral and standardized subtext within each box (e.g., "I want to personalize my enrollment by selecting a different savings rate." [Personalize]), and replacing the orange-colored select-buttons in each box with green (Personalize), yellow (Confirm), and red (Decline) buttons to generically encourage personalized enrollment. A second category, *enhanced information*, involved displaying previously communicated plan detail such as the default rate or, in one field experiment, the plan's match limit. We refer to the joint presence of enhanced presentation and information as *enhanced design*. Critically, while the design enhancements reflect modifications of the sort prescribed by practitioners and routinely deployed by firms, we intended them to be economically neutral—they did not change plan incentives, materially simplify or otherwise reduce the time-costs of enrollment, or convey previously uncommunicated plan detail or encouragements to save.

The field study, along with the supplementary evidence from the lab and surveys, yield four main findings. Our primary finding is to document the pronounced influence of enhanced design on initial enrollment. Specifically, we find that, relative to baseline, enhanced design led to a 0.09 increase in the share of personalized enrollment (0.60 baseline), a 0.62 increase in average contribution rate (5.41 baseline), a 0.98 increase in average savings rate (7.28 baseline), and a small increase in overall participation. Enhanced design also led to a substantial increase in the share of contributions exceeding the plan default and, for plans with a match, the share attaining the match. The influence of enhanced design was robust to increases in plan size and the financial stakes of the enrollment decision (as indicated by a plan's match limit or its difference from the default), suggesting the generalizability of the results. We estimate that the effect of enhanced design on average contribution (personalized enrollment) as equivalent to that predicted from a 74 (68) percent increase to the modal plan match. Accounting for the plan match and inertia, we project that, for a newly enrolled 50-year old, a design-induced shift from automatic to personalized enrollment would yield an average of 6.7 additional years of retirement security—a larger increase than that projected from the introduction of automatic enrollment.

In theory, one would expect that the marginal employee who personalizes enrollment due to non-economic shifts in design would nominally adjust their contribution in either direction of the default. Our second finding is to document that marginal enrollees, much like their inframarginal counterparts, appear to more than double their contributions relative to the default. While one cannot typically observe marginal behavior, our setting (in which an extensive-margin decision to personalize enrollment is followed by an intensive-margin decision of how much to contribute) offers unique insight into marginal saving. Specifically, assuming that design did not meaningfully change the behavior of inframarginal enrollees, we can infer the behavior of marginal personalized enrollees by comparing contribution adjustments across experimental conditions. The equivalence in marginal and inframarginal saving revealed by this comparison suggests that employees may not have well-formed preferences for how much to save (or for claiming the match) and helps to explain how savings, and consumer welfare, could

be substantially, and enduringly, increased by modest changes to design. The breadth of our plan sample offers an opportunity to assess marginal and inframarginal equivalence along a second, perhaps more important, margin—an employee's decision to actively enroll via the online interface. After presenting evidence from our industry survey suggesting that economically similar plans varied in their marketing efforts to encourage active enrollment and that such variation helps to explain plan-level differences in actual active enrollment, we tested whether active enrollment rates predicted average plan-level contribution adjustments. While only suggestive, the analysis implies that marginal active enrollees increased their contributions substantially and comparably, once again, to inframarginal counterparts.

Our third finding pertains to the mechanisms through which design affects enrollment. Of the three primary design modifications, the most successful at influencing behavior appears to be the standardization of descriptive text followed by the use of traffic-coded lights. We sought to understand the mechanisms underlying this influence by asking several thousand US employees to make hypothetical 401(k) enrollment decisions from an online paradigm programmed to resemble the experimentally varying enrollment interfaces and personalized webflow from the field. The paradigm also captured rich detail on preferences for saving and beliefs about the costs and benefits of personalized enrollment. Beyond corroborating the large and positive effect of enhanced design on personalized enrollment, the study confirmed our presumption that design did not affect behavior via standard economic channels of beliefs or preferences. And with the exception of displaying plan information, enhanced design did not appear to affect enrollment by remedying the cognitive frictions often discussed in research on nudges—e.g., limited attention to (or recall of) previously conveyed information, confusion about the choice menu, or a lack of trust. We speculate instead, and provide suggestive evidence from an additional lab study, that the decision to personalize emerges from an affective decision-process in which design influences behavior by shifting hedonic appraisals of each enrollment option—this also may help to explain why enrollment doesn't appear to be highly responsive to changes in the generosity of the plan match.

Our fourth finding addresses whether those who manage, market, and oversee plans recognize the potency of design to increase plan engagement. To test industry sophistication, we surveyed several hundred plan sponsors and asked them to forecast employee response to the tested design modifications. The survey revealed that 88 percent of respondents underestimated the positive influence of design on personalized enrollment and, when asked, in an incentive-compatible context, to rank the efficacy of specific design modifications, only 12 percent correctly identified the most influential design element from a menu of four options. Neither years of experience nor self-reported confidence in one's ranking predicted accuracy (indeed, the most confident respondents were actually the least accurate).

Collectively, we see these findings as having fundamental implications for savings policy, 401(k) plan administration, and consumer protection. First, our projections of retirement preparedness offer the most direct evidence of which we are aware that the risk of retirement insecurity extends to millions of

actual 401(k) enrollees. While discourse on savings reform has focused on improving plan access and participation, our analyses implies a significant share of at-risk employees even in the presence of automatic enrollment, a generous plan match, and minimal leakage. Second, the projected effect of enhanced design on an enrollee's long-run financial security, particularly in settings with high rates of active enrollment, positions design as a powerful, cost-effective, lever through which policymakers and plans can meaningfully improve retirement preparedness. The influence of design contrasts with the modest/ mixed response of employees to changes in plan generosity (Madrian 2013), financial education (Fernandes et al. 2014), and saving recommendations (e.g. Bhargava and Conell-Price 2021).

The potency of design, and the lack of industry sophistication with respect to such potency, also raises questions regarding consumer protection. In the absence of interventions that raise awareness of design (e.g., the promulgation of industry standards or expansion of fiduciary duty to include design) or reforms that lessen the consequence of design (e.g., higher defaults, non-contingent matches, or automatic auto-escalation), a possibility is that commercial variation in design across providers or large plans—variation far more pronounced than the modest differences we presently consider—could result in large and unintended variation in preparedness. To explore this prospect, we conducted an experiment in which we asked subjects to make a hypothetical enrollment decision after exposure to images of real-life enrollment interfaces drawn from different providers and market segments. Despite fixing economic plan features, we found substantial differences in the share of personalized enrollment and even overall participation. The exercise emphasizes the need to explicitly consider design, alongside plan incentives and information disclosures, in discussions of consumer protection in settings such as retirement savings.

These findings also offer a theoretical challenge to how economists traditionally think about consumer demand for financial products and welfare analyses of policy. In particular, the suggested equivalence between marginal and inframarginal saving, and the non-standard decision mechanisms suggested in the lab, imply that preferences may not be as stable, well-understood, and decisions not as deliberative, and reliably sensitive to economic incentives as traditionally presumed. As a result, welfare analyses of policies, such as those that encourage saving, may be biased if they do not recognize the potentially significant gains (or losses) in utility incurred by marginal individuals or if they rely on inaccurate assumptions of price elasticity. Ultimately, the findings point to the need for models of savings, and perhaps other financial decisions, that extend beyond the incorporation of common behavioral frictions to otherwise standard models of choice.

Finally, we see this research as an early step in understanding the potentially profound influence of psychological design on financial decisions increasingly made from digital settings. At first glance, our work most closely relates to the literature in economics that has catalogued the influence of small, non-economic, contextual changes on economic decisions. Much of this inquiry reflects the prescriptive program of nudging which describes how choice architects can leverage an understanding of systematic

decision biases to predictably "nudge" many into welfare-improving choices without infringing on any specific person's autonomy.[6] While the definition of a nudge has invited debate, practically, economists interested in choice architecture have documented the importance of defaults, complexity, numerical framing, reminders, and salience across a range of economic decisions.[7] Theoretically, nudges have been interpreted as affecting choice by addressing cognitive limits pertaining to self-control, attention, computation, and understanding (e.g., Datta and Mullainathan 2014). Notwithstanding the possibility that specific designs from the present research might qualify as a nudge, we interpret psychological design as encompassing a broader, and theoretically differentiated, class of non-economic modifications to the (digital) choice environment. Specifically, we see design as entailing the strategic use of visual, textual, social, interactive and gamified elements to influence behavior not only through the cognitive channels associated with nudging but through the hedonic and motivational channels suggested by our lab tests and commonly invoked in research on reward-seeking. In this sense, design relates more closely to the nascent literature in advertising and marketing on digital engagement and interdisciplinary research on the application of gameplay to domains such as health, education, and marketing.[8] The increasing prevalence of choice from digital platforms that can be easily scaled, standardized, and optimized implies that design may not only have a potent influence on consumer financial decisions but that this influence may not be as easily avoided, and its implications for welfare not as easily discerned, as the typical nudge.

2 Institutional Background and Details of Enrollment

In this section, we provide background relevant to understanding the research design and interpreting the study findings. Specifically, we first describe how the emergence of the 401(k), and recent changes to its structure, have shaped the retirement savings landscape and then describe the plan structure and enrollment experience for AE plans administered by Voya.

2.1 Overview of Employer-Sponsored Retirement Savings Plans

Employers have long-played a central role in facilitating retirement savings by US employees. In recent decades, the institutional nature of this role has shifted away from "defined benefit" plans that guaranteed regular monthly benefits during retirement to "defined contribution" plans in which retirement income is largely determined by the size and timing of employee and employer contributions. The most

[6] See Thaler and Sunstein (2003; 2008). A similar program was contemporaneously advanced by Camerer et al. (2003).
[7] Research has documented the presence of psychological frictions in decisions pertaining to health insurance (e.g., Kling et al. 2012; Bhargava et al. 2017), parental school choice (Hastings and Weinstein 2008), voting (Augenblick and Nicholson 2016), and the take-up of social benefits (e.g., Bhargava and Manoli 2015). In savings, research has shown the sensitivity of enrollment to the introduction of defaults (Madrian and Shea 2001; Chetty et al. 2014), the complexity/framing of disclosures (e.g., Goldin et al. 2020; Beshears et al. 2013), and the future timing of contributions (Thaler and Benartzi 2004).
[8] Scholars in advertising and marketing theorize that digital design affects decisions via emotion and that potent designs are those that leverage factors such as perceived personalization, speed, fun, or narratives (e.g., Baek and Yoo 2018, Seo et al. 2018). See Koivisto and Hamari (2019) for a review of research on the application of gamification to non-game domains.

prominent of these plans, the 401(k), permits eligible employees at for-profit firms to divert pre-tax earnings to an tax-deferred savings account.[9] As of 2018, of pension plans offered by US private employers 94 percent took the form of a defined contribution plan of which 87 percent took the form of a 401(k).[10] Table 1 summarizes relevant features of the employer-sponsored plan landscape for national and Voya-specific samples during a period contemporaneous with the field studies.

The emergence and subsequent popularity of 401(k) plans effectively shifted two key savings decisions—determining when and how much to contribute—from employers to employees. While high plan participation is now the norm, the initial years of the 401(k) were marked by substantial non-participation and significant disparities in participation by employee age, income and ethnicity (Madrian 2013). Two structural reforms to the 401(k) were intended, at least in part, to increase plan participation. The first was the introduction of employer matching contributions. Initially encouraged by legislation in the 1980s, the plan match entailed employers matching employee contributions, either dollar-for-dollar, or as a fraction of each dollar, up to a pre-specified limit expressed as a percentage of an employee's annual salary.[11] A 2017 industry survey found that 75 percent of firms with 401(k) plans offered some form of a match, comparable to the plan match share in the Voya sample.[12] The generosity of the plan match varies significantly by plan. For example, in the Voya sample, match limits ranged from 1 to 10 percent of annual salary with a mode of 6 percent. While economic theory predicts that the match should increase plan participation, several studies suggest that enrollment decisions may be only modestly sensitive to the match on the extensive and intensive margins (Madrian 2013).[13]

A second structural change, which had a more demonstrably positive effect on participation, was the introduction of automatic enrollment. Automatic enrollment, which plans were able to adopt after guidance from the IRS in 1998, refers to the passive enrollment of newly eligible employees, who do not actively opt out during some pre-specified period, into a plan at a default contribution rate. According to the aforementioned 2017 industry survey, 42 percent of small-to-midsize plans, representing 62 percent of assets, automatically enrolled their eligible employees. Automatic enrollment has led to sharp increases in employee 401(k) participation. In their landmark study, Madrian and Shea (2001) found that the adoption of automatic enrollment led to a 48 percentage point increase in participation among newly-eligible employees as well as a decrease in participation disparity across employee income. The effects of

[9] While the 401(k) specifically refers to plans sponsored by for-profit employers, the 457(b) refers to a functionally equivalent plan sponsored by state and local governments while the 403(b) refers to such a plan sponsored by non-profits. One can further distinguish 401(k) plans between the traditional 401(k) which allows for pre-tax contributions into a tax-deferred account (the focus of the present research) and a Roth 401(k) which allows for post-tax contributions into a yield-exempt account.
[10] Private Pension Plan Bulletin, Abstract of 2018 Form 5500 Annual Reports (EBRI 2021). Retrieved June 2021.
[11] The Tax Reform Act of 1984 encouraged matching incentives by establishing contribution limits for highly compensated employees that were tied to the contributions of less compensated employees.
[12] 2017 PLANSPONSOR Defined Contribution Survey.
[13] As one example, Engelhardt and Kumar (2007) use HRS data to estimate that a 25 cent increase in the match, per dollar contributed, would lead to a 5 percent increase in participation. Choi, Laibson, and Madrian (2011) document that over one-third of elderly employees did not take full advantage of a 401(k) plan match despite access to penalty-free withdrawal.

automatic enrollment on overall savings, however, is less clear, presumably because of typically low default contribution rates from which employees only slowly adjust (e.g., Cronqvist and Thaler 2004).[14]

2.2 Voya AE 401(k) Plans– Plan Structure and Enrollment Detail

Our field experiments were administered in collaboration with Voya, a major financial service provider and 401(k) record-keeper. At the time of the studies, Voya offered recordkeeping and other services for approximately 48,000 retirement plans covering about 4.9 million individuals spanning the tax-exempt and corporate markets. Due to commercial considerations, our research was restricted to AE plans at small-to-midsize employers, defined as having fewer than approximately 5,000 employees (nationally, roughly 65 percent of employees work in establishments of this size).[15] As summarized in Table 1, Voya plans resemble a comparable national sample with respect to the distribution of default rates (ranging from 1 to 6 percent with a mode of 3 percent) and the prevalence of matching incentives.

We turn next to describing the three enrollment decisions potentially faced by a newly-eligible employee in our sample (summarized in Figure 1). While some aspects of the enrollment experience are provider-specific, because of federal guidelines governing plan communication and the requirement that employees must be given the opportunity to personalize or decline their automatic enrollment, the generic tri-parte structure of enrollment we describe applies to AE plans across providers. Within our sample of a few hundred Voya AE plans, with the exception of potential differences in the content and timing of plan communication or marketing, there is high standardization in the enrollment experience as all plans share the same online enrollment interface and webflow.

The first enrollment decision faced by newly-eligible employees occurs during the period preceding automatic enrollment. It is during this period when federal rules require that AE plans communicate details regarding eligibility, plan features, and enrollment to potential enrollees.[16] Beyond mandated disclosures, many plans distribute more user-friendly marketing material, or administer information sessions, to further clarify plan details and encourage plan participation. Importantly, plans must explicitly inform employees, reasonably in advance of automatic enrollment, that they will be automatically enrolled at a default contribution rate and investment allocation unless they actively modify their enrollment. Voya plans typically inform employees by one or more emails that they may personalize enrollment, confirm automatic enrollment, or decline enrollment by visiting an online enrollment interface. We refer to an employee's decision to visit the interface as the active enrollment decision.

---

[14] Some plans offer enrollees the opportunity to automatically escalate contributions each year. Thaler and Benartzi (2004) document that the combination of automatic enrollment and escalation does lead to an increase in savings.

[15] See SUSB annual data tables for 2017 published by the US Census.

[16] Fiduciaries are also required to regularly update enrollees regarding plan details and inform them of any plan modifications. The specific governance of disclosures varies based on plan characteristics such as size and qualification for "Safe Harbor" status, a designation exempting certain plans with employer contributions from non-discrimination tests. Safe Harbor plans—which includes most of the plans in our field sample—are required to disclose additional detail regarding enrollment, applicable default parameters, and matching incentives reasonably in advance of the plan year or automatic enrollment.

The second enrollment decision, for employees electing to visit the interface, is the decision to personalize enrollment. After logging in and providing some basic information, employees who visit the interface arrive at a landing page from which they can make their enrollment decision (Appendix Figure A1). The landing page, highly standardized across Voya plans, briefly welcomes employees and then directs them to select one of three available enrollment options corresponding to one of three horizontally arranged boxes. The headlines of these boxes read, from left to right: "I want to enroll with other choices." [PERSONALIZE]; "Let my scheduled automatic enrollment go through." [AUTO]; "I do not wish to enroll." [DECLINE]. The personalized enrollment box also displays additional gray subtext below the headline: "Note: This enrollment will cancel your scheduled automatic enrollment." The other two boxes have no subtext. To select an option, an employee must click on the orange-colored select-button within the appropriate box. An employee's visit ends if the employee leaves the interface without making a selection (such employees will be automatically enrolled in the plan unless they revisit the site within the enrollment period), confirms automatic enrollment (the employee receives a congratulatory message), or declines enrollment (the employee is asked to confirm their decision). Employees choosing to personalize their enrollment proceed through a subsequent enrollment webflow (Appendix Figure A2).

The final enrollment decision, for employees choosing to personalize enrollment, is the election of a new contribution rate from the personalized enrollment webflow. After capturing information pertaining to employee demographics (e.g., date of birth and salary), retirement goals (e.g., retirement age, income replacement ratio), and total accumulated savings, the webflow asks employees to elect a contribution rate in the context of a retirement calculator. The retirement calculator (a common feature of enrollment portals across providers) translates the indicated contribution rate, and other inputs, into a projected, and recommended, monthly retirement income flow.[17] There are two instances in which the webflow provides explicit encouragement to increase one's contribution—when an employee elects a contribution below the match threshold in plans that offer a match that can be discerned in real-time by the portal or when an employee selects a rate below an age-specific savings target in a plan that either offers no match or offers a match that cannot be discerned by the portal.[18] After the contribution election, the webflow invites the employee to personalize other enrollment options (e.g., the investment allocation or, if available, auto-escalation) after which the webflow asks the employee to confirm their selections.

3 Risk of Retirement Security for 401(k) Enrollees

In an initial empirical analysis, we attempted to assess the risk of retirement insecurity for 401(K) enrollees and further examine heterogeneity in average risk across firms. Conceptually, our approach differs from the norm in the literature in that we sought to estimate an employee's probabilistic risk of

[17] The input field is anchored by a 6 percent contribution rate, in greyed-out text, which changed to 6 percent on 12/14/17.
[18] The visibility of the plan match to the portal depends on a number of plan-specific commercial and technical factors.

accumulating less than some lump-sum threshold associated with retirement insecurity as compared to estimating a single ratio of retirement income replacement. More specifically, our strategy involved leveraging access to administrative records of active enrollees to simulate the year-to-year time-path of accumulation for each employee from now until their projected retirement, under the assumption of stochastic market returns and uncertainty in annual contribution adjustments. The simulations capture multiple sources of real-life retirement risk—insufficient adjustment to plan contributions, poor market returns, and the risk associated with the sequencing of such returns. We repeated this exercise for a range of conservative and/or widely varying assumptions governing market return distributions, severity of contribution inertia, income growth and retirement age. Finally, we mapped each resulting distribution of projected savings to an estimated risk of retirement insecurity informed by research on the financial requirements of retirement and commercial prices for life-time annuities.

3.1 Summary of Data

Our analysis drew on administrative records for approximately 270,000 active enrollees in 1,768 Voya 401(k) plans with automatic enrollment. These data covered the universe of employees, as of February 2019, who were actively enrolled (i.e., someone with a non-zero account balance) in any AE 401(k) plan within Voya's self-defined small-to-midsize market segment. Specifically, we observed each enrollee's current account balance, contribution rate, salary, age and, if applicable, details on any plan match and auto-escalation. To construct our analytic sample, we restricted attention to 25 to 64 year-old employees for whom we had reliable salary data, and, to facilitate plan comparison, excluded any plans with fewer than 60 enrollees satisfying those criteria. The resulting sample included 185,905 enrollees across 840 AE plans. Of these plans, 51 percent automatically enrolled employees at a 3 percent contribution rate, 15 percent enrolled employees at a rate lower than 3 percent, and the remaining plans enrolled employees at a rate between 4 and 8 percent. All plans offered to auto-escalate enrollees at a default annual escalation rate of 1 percent up to a default threshold of 10 percent—42 percent of plans automatically enrolled enrollees, who did not actively opt out, into auto-escalation, while in the remaining plans, auto-escalation required an active decision to opt in.

3.2 Projecting 401(k) Asset Accumulation at Retirement

The first step in our analysis of retirement insecurity was to generate a distribution of projected accumulated assets at retirement for each enrollee by simulating the year-to-year time-path of savings assuming annual market returns drawn from an underlying market-return distribution and annual increases in contribution determined by enrollee-specific draws from an underlying likelihood of inertia (for those not discernably enrolled in auto-escalation). We repeated the exercise for a wide range of plausible and/or conservative assumptions describing the distribution of future market returns, the

severity of contribution inertia, annual wage growth, retirement age. More formally, we simulated the time-path of accumulation 1,000 times for each employee, and set of assumptions, with these equations:

$$\hat{a}_t = a_{t-1}(1 + r_t) + c_t(\hat{w}_t)$$
$$\hat{c}_t = c_{t-1} + increment_t$$
$$\hat{w}_t = (1 + g)w_{t-1}$$

where: $r_t \sim N(\mu, \sigma^2)$

$$increment_t \sim \text{Bernoulli}(1 - i)\ if\ c_t < 10\text{, otherwise } 0$$

Here, a_t denotes accumulated assets in year *t*, c_t denotes the annual percent contribution rate, w_t specifies the annual real wage (with an annual growth rate of *g*), and r_t denotes the annual investment return, drawn for all enrollees from some distribution $r_t \sim N(\mu, \sigma^2)$. To represent inertia in annual contributions, we stipulated that an enrollee contributing less than 10 percent would increase their annual contribution by 1 percentage point with probability $(1 - i)$ each year (unless they were enrolled in auto-escalation in which case they would escalate by 1 percentage point annually up to a cap of 10 percent).

We sought to project savings using conservative and/or widely varying assumptions for key parameters. Our primary projections assumed a retirement age of 65 (compared to a median age of recent retirees of 62 years), an annual increase in real wages of 1 percent, and, for anyone not in a plan with automatic enrollment into auto-escalation, a 25 percent likelihood of increasing one's contribution each year—a degree of inertia significantly less than the 9 percent we observe in the administrative data.[19] We further assumed that employees never reduced their annual contribution rate, withdrew from the plan entirely, and sustained no leakage in savings due to early withdrawals, loan defaults, cash-outs during job transitions, or gaps in employment. By one estimate, leakage from channels excluding employment gaps results in an average reduction in savings of 25 percent (Munnell and Webb 2015). Finally, we assumed a 0.10 tax rate at withdrawal and no private savings outside of the employee's 401(k).[20] The Appendix reports estimates of retirement risk under alternative assumptions of retirement age and wage growth.

To capture a wide range of plausible future market regimes, we separately projected savings for each of four potential distributions of portfolio returns. Three of these distributions drew on scenario-based projections published by the BlackRock Investment Institute (BII), the macro research arm of a large investment management firm. Specifically, we used BII's forecast for the expected annual return of

[19] We estimate contribution inertia for enrollees from the following model of an enrollee's current contribution rate, y_i: $y_i = \beta_0 + \beta_1 t_i + \gamma' X_i + \delta_p + \varepsilon_i$. Here, t_i captures tenure in years, X'_i is a vector of basic enrollee characteristics (age, sex, and log salary) and δ_p denotes plan fixed effects. The coefficient estimate of $\beta_1 = 0.09$ implies an inertial rate of 91 percent.

[20] According to household surveys, the assumption of no private savings outside of the 401(k) is approximately accurate for the large share of enrollees. Assuming that outside savings are correlated with 401(k) savings or income, those with non-trivial private savings would not be tagged as being at retirement risk through our estimates.

a mixed 60/40 global investment portfolio over a 20-year future horizon under pessimistic, baseline, and optimistic forecasts (corresponding to the 25th, 50th, and 75th percentile of their forecast distribution).[21] After adjusting for an annual inflation rate of 2 percent, the forecasts yielded an average real return of μ = 1.6, 3.5, and 5.6. For each distribution, we assumed $\sigma = 10.2$, the single volatility estimate reported from the BII report. Finally, to accommodate the unlikely possibility (judging from a multitude of forecasts) that future market returns resemble historical ones, in a fourth projection, we assumed annual returns drawn from a normal distribution with μ = 6.4 and $\sigma = 9.2$, the distribution of real returns from a 60-40 US equity-bond portfolio from 1987 to 2016.[22]

3.3 Retirement Insecurity Thresholds and Annuity Pricing

After generating distributions of projected savings for each enrollee and each set of parameter assumptions, we mapped the distribution to an employee-specific score indicating the risk of retirement insecurity. The risk score indicates the share of 1,000 simulations for which an employee's projected savings did not exceed a minimum threshold associated with retirement security. We define this threshold as the minimum level of post-tax savings, after accounting for social security benefits, required to purchase an immediately vesting lifetime annuity that would generate sufficient monthly disbursements to provide for basic retirement expenses without return to the labor market, means-tested benefit programs, or financial support from family. While we report risk scores associated with a range of plausible insecurity thresholds, the Elder Index informs our basic understanding of the minimal requirements of retirement (Mutchler et al. 2016). The Elder Index reflects an effort to estimate the minimum income that would permit a retiree to cover basic expenses for housing, food, transport, health care, and miscellaneous expenses without reliance on public assistance or family based on their household configuration and location.[23] An implication of the research is because elderly populations have much higher typical expenses than working-age adults, financial security requires levels of income that exceed federal poverty thresholds. Specifically, the 2020 Elder Index for a nationally representative single retiree in good health, residing in a rented apartment, indicates a minimum expense threshold of $25,884/year ($2,157/month), twice as large as the corresponding Federal Poverty Level ($12,760/year).

We obtained prices for a guaranteed lifetime annuity from an online pricing calculator hosted by a large US financial services firm.[24] As an example, the pricing calculator indicates that for $200,000, a 65-year-old male could purchase an annuity guaranteeing a lifetime monthly income, in real dollars, of

[21] BII forecasts the mean annual nominal return (and the interquartile range of this range), gross of fees, for a wide-range of portfolios via this online tool. Regularly updated forecasts may differ from inflation-adjusted estimates retrieved in August 2020.
[22] These parameters draw from a 2016 discussion note from Norges Bank Investment Management. The figures reflect the average return, and standard deviation, of a 60 (S&P 500) / 40 (Bloomberg US Bond Index) portfolio from 1987 to 2016 after adjusting for average inflation over that period. Accessed November 2020.
[23] Details on methods and applications of the Eder Index can be found in this 2017 CSDRA Report.
[24] Prices for an immediately vesting, fixed-income, lifetime annuity retrieved from a Charles Schwab online calculator in March 2021. Annuities priced for a 65-year old male from Minnesota, born March 23rd, 1956, vesting on March 26th, 2021.

$2,181, without bequests or guaranteed disbursements. As a conservative simplification, our calculations assume the slightly lower prices available to male annuitants apply to everyone. We further assume that all employees supplement their annuity income with Social Security benefits equivalent to $1,293, the median monthly benefit for a retired worker as of 2018.[25] The assumption of a median benefit should inflate the estimated retirement income for lower-income employees in the sample while underestimating the income of employees whose accumulation was already likely to exceed the threshold. In consideration of the minimal financial demands of retirement, the magnitude of Social Security benefits, and a desire for simple thresholds, we report retirement risk scores for three post-tax saving thresholds of $150,000, $200,000, and $250,000. These thresholds, along with Social Security Benefits, would permit the purchase of annuities upon retirement at the age of 65 guaranteeing a lifetime monthly income of $1,959, $2,181, and $2,403, respectively, an income interval reflecting 0.91 to 1.11 of the Elder Index.

3.4 Projected Retirement Risk of Employees and Plans

We now present estimates of retirement risk. We begin by describing the distribution of projected savings under our baseline assumptions regarding retirement age (65 years), contribution growth (g = 0.01, i=0.75) and investment returns, $r_t \sim \mathrm{N}(6.4, \sigma = 10.2)$. Figure 2 depicts the cumulative share of employees whose risk of retirement insecurity exceeds 10 (short-dashed line), 25 (long-dashed line), and 50 (solid line) percent for increasing thresholds up to $400,000. For example, a $200,000 threshold implies that 51 percent of employees have a risk of insecurity exceeding 10 percent; 44 percent of employees have a risk exceeding 25 percent; and 35 percent of employees have a risk exceeding 50 percent. The average employee in our sample has a 38 percent chance of failing to accumulate $200,000.

Table 2 reports the share of at-risk enrollees for each of the four market scenarios and three retirement insecurity thresholds assuming a risk threshold of 25 percent (panel A) and 10 percent (panel B). While the first panel conveys the sensitivity of retirement risk to market and insecurity threshold assumptions, across these wide-ranging parameters, 25 to 65 percent of 401(k) enrollees have a risk of insecurity in excess of 25 percent. Under our favored assumptions—an insecurity threshold of $200,000 and moderate expectations of future market returns—we estimate an at-risk share of enrollees of 44 percent. Comparing across the two panels indicates that the share of at-risk enrollees does not rise sharply with a slightly lower risk threshold, suggesting that many at-risk enrollees are at substantial risk.

Does retirement risk vary for observationally similar employees enrolled in economically similar plans? We addressed this question by examining the explained and unexplained variation in preparedness across the 840 plans in our administrative sample. As a measure of preparedness, y_i, we estimated each employee's risk of accumulating less than $200,000 under baseline assumptions. We then estimated the following regression to roughly assess how much of variation in preparedness can be explained by

[25] Figure cited from an online post by the Pension Rights Center using 2018 SSA data, retrieved in November 2020.

compositional differences in employee demographics, Z_i, and economic plan features, X_p: $y_i = \beta_0 + \alpha' Z_i + \gamma' X_p + \varepsilon_i$.[26] The analysis indicates that while employee demographics alone can explain 61 percent of the variation in preparedness, the inclusion of plan characteristics increases the model's explained variation to just 64 percent, leaving more than one-third of total variation unexplained.

Overall, the projections point to a striking degree of retirement risk for 401(k) enrollees, even in the presence of a plan match, under conservative assumptions about contribution growth, investment return, and leakage. The analysis also suggests variation in risk for observationally similar employee across similar plans and a degree of unexplained variation across plans substantially larger than that explained by economic plan features. In the rest of the paper, we explore the possibility that psychological plan design—i.e., both the design of the online enrollment interface and the intensity/efficacy of plan marketing during pre-enrollment—may help to better understand the high degree of average retirement risk and the variation in such risk across the broader 401(k) plan landscape.[27]

4 Research Design of Field Experiments

We turn next to a series of field experiments intended to test whether variation in the psychological design of 401(k) enrollment interfaces causally influence an employee's initial savings and long-run financial security. Our central evidence draws from three randomized field experiments administered across 500 401(k) AE plans from December 2016 to March 2018. While the three experiments share many details pertaining to the research design, setting, and implementation, each study was motivated by a distinct theoretical objective. The first, and primary, field experiment tested the influence of enhanced presentation (a design incorporating three non-economic modifications to the interface), enhanced information (a design proximally displaying a plan's default rate), and enhanced design (a design with both enhanced presentation and enhanced information) on enrollment. The second field study, which was administered to a limited number of plans, tested the influence of enhanced presentation in the specific context of an interface displaying information about both the plan default and its match limit. Finally, the third field study aspired to decompose the effect of enhanced presentation by separately testing employee response to each of three constituent design components.

4.1 Plan and Employee Sample

The field experiments were collectively administered to the near-universe of active 401(k) AE plans in Voya's small-to-midsize market segment as of each study's launch date.[28] Field Study #1 was

[26] Z_i is a vector of indicators that flexibly account for enrollee demographics (i.e., \$25k-salary bins, 5-year age bins, 5-year tenure bins, and gender). X_p is a vector of indicators that flexibly account for plan characteristics (i.e., the plan match, censored at 6 percent; the default rate, censored at 2 and 8 percent; and auto-escalation). Robust standard errors clustered at the plan-level.
[27] While plans in the projection analysis had standardized online enrollment interfaces, they almost certainly varied in the intensity and efficacy of their plan marketing and communication (a topic we engage further when discussing the plan survey).
[28] Voya excluded four plans for commercial reasons. Active plans are those with at least one online enrollment in the prior year.

administered to all plans excluding four plans designated for inclusion in Field Study #2, administered concurrently with the first study. These four plans were the largest plans in the sample whose match limit exceeded the default rate and for which Voya was able to secure commercial approval for the study. Field Study #3 was not contemporaneous with the other two studies and was administered to all available plans. The first two studies ran from July 21st, 2017 to March 31st, 2018 while the third study (the first study, chronologically) ran from December 6, 2016 to June 28, 2017.[29] The studies were administered to any newly eligible employee in a treated plan whose initial visit to the enrollment landing page and eventual enrollment decision fell within the study period. After excluding plans that failed to enroll an employee in at least two distinct experimental treatments and any employees inferred to have seen different landing pages across multiple visits to online portal, our final analytic sample comprised enrollment decisions from 8,565 employees across 500 plans (397 from the first study, 4 from the second study, and 308 from the third study).[30] For select analyses, we further excluded some observations for which we could not observe or infer the plan match (see below for a discussion of incomplete data and inference).

Table 3 summarizes plan and employee characteristics for the analytic sample. Overall, sample plans resembled the broader population of AE plans at small-to-midsize firms as described in Table 1 with respect to the default and plan match. Sample default rates ranged from 1 to 6 percent, with a mode of 3 percent, and 72 percent of plans offered a plan match among plans for which we could discern its presence (see discussion below). Notably, the sample was left-skewed with respect to plan size such that 80 percent of enrollment decisions were from the largest 20 percent of plans. By design, plans in the first and third field studies were similar in size, default distribution, and match generosity, while the 4 plans in the second study were larger and offered more generous matching incentives than the typical plan in the other studies. Demographically, employees were roughly similar across the studies and resembled a contemporaneous national sample of employed adults apart from featuring a higher share of men and more highly compensated employees ($63k versus $44k median earnings).[31]

4.2 Experimental Treatments

The field studies collectively test the response of 401(k) enrollment to several modifications to the design of an otherwise standardized enrollment interface. These modifications reflect the type of modest variation in digital design commonly observed across plans within a common plan provider. Critically, the design modifications were intended to involve *non-economic* adjustments to enrollment

[29] Due to implementation constraints, two plans did not enter Field Study #1 until December 9th, 2016 and one plan did not enter Field Study #2 until August 11th, 2017. For narrative clarity, we did not label the field studies chronologically.
[30] Given our analyses relied on pair-wise comparisons across experimental conditions, we excluded any plan that did not enroll at least one employee in each of two distinct conditions resulting in 77 exclusions from the first field study and 142 exclusions from the second field study. We also excluded a few hundred employees who appeared to have seen multiple versions of the landing page on separate visits due to the timing of their visits, their use of multiple browsers, or their clearing of HTTP cookies.
[31] See the 2016 Current Population Survey, published by the Bureau of Labor Statistics.

insofar as they did not alter any plan incentives or the administrative requirements of enrollment, nor did they introduce plan information not previously conveyed to employees through disclosures or marketing.

We categorize the tested design modifications into three distinct categories. The first category, *Enhanced Presentation,* includes three specific modifications to the visual appearance of the landing page: (i) simpler, punchier, and more agentic headlines for each enrollment option, (ii) standardized and emotionally-neutral descriptions of each enrollment option, and (iii) the adoption of green-yellow-red colored buttons to encourage personalized enrollment and discourage declined enrollment. The enhancements broadly reflect principles of effective design—such as simplicity, clarity, consistency, warmth, minimalism and the strategic use of color—widely advocated by practitioners and used by firms. For example, the modified headlines closely resemble those used by at least one other major financial service provider in their 401(k) enrollment interface and also resemble the language adopted by the Swedish government to encourage active portfolio choice (Cronqvist et al. 2018). As another example, the strategic use of traffic-light colors is a common online design-feature. And while it encourages saving, it does so no more strongly than encouragement routinely included in plan marketing and communication.

A second category of design modification, *Enhanced Information*, involves the display of the default rate (along with, in some cases, a plan's match limit) explicitly on the landing page. According to standard economic theory, because this information was repeatedly communicated through mandated plan disclosures and any plan marketing, heightening the salience of this information, or otherwise situating it in proximity to the decision, should not affect enrollment choice. However, beyond once again adhering to principles of successful design, as advocated by practitioners, the potential for the proximal display of decision-relevant information to influence behavior is consistent with recent research on information salience (e.g., Hossain and Morgan 2006; Dellavigna and Pollet 2009; Chetty et al. 2009). Lastly, a third category of design modification, *Enhanced Design*, refers to the joint presence of both enhanced presentation and enhanced information.

Field Study #1. We now describe the construction of each treatment condition. The first study sought to test the impact of enhanced design, enhanced presentation, and enhanced information on initial enrollment by randomizing employees to one of three experimental landing pages (Appendix Figure A3):

- BASIC. A first control condition replicated the original commercial design (Section 2), but for two small changes intended to facilitate comparisons across treatments. First, we amended the text introducing the enrollment options from, "Did you know you can save more?", to a more neutral, "Please select one of the following options:". Second, we standardized the wording of each option headline to read: "I want to enroll with different choices." [PERSONALIZE], "I want to confirm my automatic enrollment." [AUTO], and "I do not want to enroll" [DECLINE].

- ENHANCED INFORMATION. A second condition was intended to test the effect of displaying the plan default rate directly on the landing page. This condition was identical to BASIC but for a single line of gray text situated above the enrollment options: "If you confirm your automatic enrollment you will be enrolled at a savings rate of <X> percent." The differential response to this condition, relative to BASIC, indicated the influence of Enhanced Information.

- ENHANCED DESIGN. A third condition was intended to test the effect of enhanced presentation, and the effect of enhanced design, on personalized enrollment. The condition was identical to the ENHANCED INFORMATION condition but for three modifications. First, we replaced the neutral headlines describing each enrollment option with simpler headlines emphasizing the employee agency: "Do It Myself" [PERSONALIZE], "Do It for Me" [AUTO], and "I Don't Want to Save" [DECLINE]. Second, we neutralized/standardized the language describing personalized enrollment by replacing "Note: This enrollment will cancel your scheduled automatic enrollment." with "I want to personalize my enrollment by selecting a different savings rate." We also added analogous descriptions for the other two options. Finally, we replaced the orange select-buttons associated with each option with green (PERSONALIZE), yellow (AUTO), and red (DECLINE) buttons to encourage personalized enrollment. The differential response to this condition, relative to BASIC, indicated the influence of Enhanced Design, while the differential response, relative to ENHANCED INFORMATION, indicated the influence of Enhanced Presentation (in the context of default display).

Field Study #2. The second field study sought to test the influence of enhanced presentation in the specific context of a landing page that displayed both the plan default and the plan match limit by randomizing employees to one of two experimental conditions (Appendix Figure A3):

- ENHANCED INFORMATION. A first control condition resembled the analogous condition from the first field study but for additional text communicating the match limit (e.g., "Note: Your employer will match your contribution up to a rate of 6 percent").

- ENHANCED DESIGN. A second condition resembled the analogous condition from the first study but for text communicating the match limit and additional visual and textual elements to encourage match take-up. Specifically, to emphasize that employees seeking to fully take-up the match had to personalize their enrollment, a thin colored cap adorned each enrollment box with a short message reading: "Receive all of your match" [PERSONALIZE, green], "Lose some of your match" [AUTO, yellow], and "Lose all of your match" [DECLINE, red]. The differential response to this condition, relative to ENHANCED INFORMATION, revealed the influence of Enhanced Presentation (in the context of the display of both the default and match limit).

Field Study #3. Finally, the third field study aimed to separately test each of the three design modifications that comprised Enhanced Presentation from the first study: standardization, headline modification, and traffic-light colors in the context of no displayed plan information. The study randomized employees to one of the following four conditions (Appendix Figure A4):

- BASIC. A first control condition was identical to the original landing page in commercial use prior to our field studies. This condition also resembles the BASIC condition from the first field study but for the nominal differences outlined above.

- STANDARDIZED. A second condition adapted the BASIC condition by standardizing and neutralizing the sub-text used to describe each enrollment option and replacing the original introductory prompt ("Did you know you can save more?") with the more neutral prompt ("Please select one of the following options:"). The differential response to this condition, relative to BASIC, indicated the influence of standardized language.

- HEADLINES (Headlines + Standardized). A third condition adapted the STANDARDIZED design by replacing the neutral headlines describing each enrollment option with the shortened headlines described above. The differential response to this condition, relative to STANDARDIZED, indicated the influence of the modified headlines.

- ENHANCED PRESENTATION (Color + Headlines + Standardized). A final condition modified the HEADLINES condition by replacing the standard orange-colored select-buttons with green, yellow, and red colored buttons to encourage personalized enrollment. The differential response to this condition, relative to BASIC, indicated the influence of Enhanced Presentation (in the context of no display of plan information). The differential response of this condition, relative to HEADLINES, revealed the influence of the traffic-light colors.

4.3 Randomization

Employees qualified for any of the field studies proceeded through the same enrollment protocol as described in Section 2 but for random assignment to one of the experimentally varying landing pages. Treated employees who visited the online portal were randomized, based on a unique identifier associated with their browser, to a landing page using an algorithm with pre-designated assignment probabilities. The algorithm assigned employees to one of the landing pages associated with the pertinent study with equal probability but for one condition in Field Study #3 that was initially under-sampled.[32] We note that because the employee-level randomization was not stratified by plan due to operational constraints, some plans (particularly smaller ones) were not perfectly balanced across conditions.

4.4 Data and Empirical Outcomes

Our empirical analysis relies on administrative data describing plan enrollment (enrollment decision, contribution rate), employee demographics (age, gender), employment detail (e.g., hire date, tenure, salary), and plan characteristics (e.g., default rate, details of matching incentives), as well as additional detail from the personalized enrollment webflow. We focused the analysis on four outcomes of

[32] Due to sample size constraints and a desire to maximize particular pairwise comparisons, we originally opted to under-sample the BASIC condition (10 percent assignment probability versus 30 percent for the other conditions). After observing lower-than-anticipated enrollment volumes in the first quarter of 2017, we transitioned to an equal-weighted assignment across all conditions for the remainder of the study. The shift in assignment probability took effect on 4/13/2017.

central interest: (1) enrollment choice (i.e., the decision to personalize, confirm, or decline), (2) attainment of contribution rate thresholds associated with the plan's default rate or, when applicable, the plan's match limit, (3) the contribution rate, and (4) the saving rate (defined as the sum of the contribution and employer match) To ensure outcomes reflected an employee's final enrollment decision, we used administrative records generated after the close of an employee's auto-enrollment period.[33]

Despite a single record-keeper, retrieving and standardizing administrative data from several hundred small-to-midsize plans was a complicated undertaking. We encountered three challenges in which incomplete or non-standard records required us to impute data, restrict the sample for selected analyses, and in a few cases, exclude observations entirely. Our primary data challenge involved incomplete plan match records. The availability of plan match data depended on idiosyncratic factors but was positively correlated with plan size. Through explicit access to administrative data, and inference using enrollee records from the snapshot assessment, we discerned plan match detail, including the presence of a match, for 86 percent of employees (62 percent of plans) in the sample. This included explicit data for 78 percent of employees (39 percent of plans) and inference for an additional 9 percent of employees (24 percent of plans). We estimate that our inference was 87 percent accurate based on an analysis of plans for which we had both explicit data and enrollee records.[34] We excluded the residual data, along with data from 22 plans with a non-standard or ambiguous match scheme, from analyses involving the match or savings (i.e. the sum of employee and matching contributions) but retained these data for other analyses.[35] While the restricted sample featured a disproportionate share of larger plans, its employees appeared demographically similar to those in the unrestricted sample.

A second challenge involved a modest share of missing demographic data, particularly salary and gender. We addressed these missing data by imputation. Notably, we imputed approximate salary for the 11 percent of missing observations through a predictive model estimated from the data for which salary was observed.[36] In addition, we inferred gender for 22 percent of the sample from public algorithms that

[33] Recordkeeping data was not available for two plans in the sample. For these plans, we treated their online enrollment decision as final (if the employee made no active decision online, we assumed automatic enrollment at the default rate).
[34] We inferred the plan match for 115 plans using enrollee records from the 2019 snapshot assessment of preparedness (for an additional 4 plans, we obtained explicit match data as of 2019). The inference involved using a series of rules informed by plans for which we did observe match detail to map every plan's distribution of enrollee and employer contributions to a potential match formula. As a conservative measure, we only attempted the inference for plans for which there were at least 50 enrollee records. Specifically, the first step of the inference was to infer a zero plan match if we observed no employer contributions for any enrollee. For remaining plans, we calculated two plan-level statistics—the modal annual percent contribution rate at, or above, the plan's default rate, and the modal employer contribution rate for the top decile of employee contributions. We then mapped these statistics to one of the 16 match formulas that characterized nearly all plans with observable detail. For example, a modal employee contribution of 6 percent and a modal employer contribution among high contributors of 3 percent led us to infer a 50 percent plan match up to a 6 percent limit. We assume match formulas did not change from the study period to 2019.
[35] The 22 non-standard/ambiguous plans either did not report a uniform match limit for all employees (e.g., limits were tied to union status, location) or had a match limit that was not expressed as a percent contribution (e.g., a dollar-denominated limit).
[36] Our salary data included explicit administrative records for 75 percent of employees and self-reports for an additional 14 percent of employees. To impute missing records, we predicted out-of-sample values from a linear model of salary estimated using employee demographics, median plan-level salary, enrollment choice, and county-level income from the US Census.

probabilistically assigned genders to first names (which we did observe).[37] Lastly, we excluded a small number of records with missing contribution or age from any analyses requiring those fields.[38]

A third challenge was the presence of outlier contributions. While rare, we interpreted such outliers (which reached as high as 99 percent of annual salary) as short-term catch-up contributions. As a result, throughout the analysis we censored contribution rates to 20 percent of annual salary. The specific censoring threshold, selected for simplicity, effectively censored contributions at or above the 97th percentile, inclusive of non-participants. The censoring of contributions follows precedent from the existing literature and we found no evidence indicating that the findings were substantively sensitive to the particular choice of threshold or the decision to censor.

5 Evidence on Enhanced Design from the Field

We now describe the results of the three field experiments. After summarizing overall response to the field experiments, we report the marginal treatment effects associated with the three design categories of central interest across a range of enrollment outcomes. To contextualize the magnitude of the effects, we compare them to the influence of two other known predictors of employee savings, the generosity of the plan match and employee age. Next, to better understand how the effect of design varies as the financial consequences of personalized enrollment increase, and to address the generalizability of the findings to plans outside of our sample, we describe the heterogeneity of these treatment effects across a range of plan characteristics. Finally, we present evidence clarifying the specific design modifications responsible for the overall influence of enhanced design.

5.1. Overall Response and Estimates of Treatment Effects by Field Study

Table 4 summarizes average enrollment outcomes associated with the experimental conditions across the field studies. The table reveals several patterns of interest. First, it suggests substantial variation in enrollment across the tested conditions, offering initial evidence of the importance of design. Across the tested conditions, enhanced design, as well as enhanced presentation alone, generally delivered the highest shares of personalized enrollment, share of contributions exceeding the default rate and attaining full match take-up, and highest average contribution and saving. Second, much of the variation in enrollment choice across the experimental conditions involves shifts in mode of enrollment rather than differences in overall participation, except for a modest increase in participation associated with enhanced relative to basic design. We speculate that this pattern may be attributable to the low baseline rate of declined enrollment in auto-enrollment settings. Finally, differences in the share of personalized enrollment across conditions are of roughly similar magnitude to differences in the observed share of

[37] Our gender inference, using a 90 percent matching threshold, reduced the rate of missing observations from 22 to 3 percent.
[38] We observe a small share of missing data for employee contribution rates (1.7 percent missing) and age (4.3 percent missing).

employees attaining the two contribution thresholds. This correspondence is consistent with the possibility that marginal employees who personalized enrollment due to design increased their contribution rates to a non-trivial degree, a possibility to which we later return.

We more formally estimated the causal effect of each experimental treatment on enrollment through a series of regressions, estimated separately for each field study and controlling for observed plan and employee characteristics and include indicators for each treatment. As an example, to evaluate the treatments from Field Study #1, we estimated the marginal effect of the two experimental treatments, enhanced design and enhanced information, relative to the excluded baseline condition (BASIC), on enrollment outcomes, y^k, for potential enrollee *i*, from plan *j* with the following equation:

$$y^k{}_{ij} = \alpha + \gamma\, Enhanced\ Design_i + \pi\, Enhanced\ Info_i + \delta_j + X\varphi + \varepsilon_{ij}$$

The main coefficients of interest, γ, and π, capture the marginal effects of enhanced design (i.e., enhanced presentation plus enhanced information) and enhanced information alone, while $(\gamma - \pi)$ captures the marginal effect of enhanced presentation in the context of an interface displaying the plan default. The regression controls for plan-specific variation with plan fixed effects, δ_j and a vector of indicator variables, *X*, to flexibly account for the timing of enrollment (i.e., day-of-week and week-of-year). Standard errors are robust and clustered by plan to account for potential non-independence of errors within plan. We adapted the model to estimate treatment effects in the other field studies.[39]

Table 5 reports estimates for each experimental treatment across the three field studies (statistically significant estimates at $p < 0.05$ bolded). While we postpone discussion of specific treatment effects, overall the estimates corroborate our interpretation of the summary table—enrollment exhibits sensitivity to differences in design, variation in enrollment appears to be largely driven by changes in the mode of enrollment rather than overall participation, and increases in the share of personalized enrollment compare with increases in the share of contribution threshold attainment.

5.2 Enhanced Design and 401(k) Plan Enrollment

We now consider the integrated response, across the field studies, of employees to the three design categories of interest. Table 6 summarizes the marginal treatment effects, relative to the basic condition, associated with enhanced design (Panel A), enhanced presentation only (Panel B), and enhanced information only (Panel C) across the field studies. The first panel conveys our principal finding: a large and statistically significant influence of enhanced design on personalized enrollment, contribution and saving in the first field study. Enhanced design led to a 0.09 increase in the share of personalized enrollment ($p < 0.01$), a 0.07 decrease in the share of automatic enrollment ($p < 0.01$), and a

[39] Specifically, for Field Study #2 we estimated an identical regression but for the omission of the enhanced information indicator (also, we did not cluster standard errors in this estimate due to too few clusters). For Field Study #3, we estimated an analogous regression but with indicators denoting each constituent design element less one exclusion.

small, marginally significant, decrease in declined enrollment (b = -0.03, p < 0.10). The increase in the share of personalized enrollment was comparable in magnitude to the increase in the share of employees contributing in excess of the default (b = 0.08, p < 0.01) and at or above the match limit (b = 0.10, p < 0.01). Enhanced design was also associated with a 0.62 increase in the average contribution rate (p < 0.01) and a 0.98 increase in average saving (p < 0.01).[40] These estimates—along with the small and insignificant effect of enhanced design on average contribution among personalized enrollees—are consistent with the interpretation that marginal personalized enrollees adjusted their contributions non-trivially and to a degree comparable to their inframarginal counterparts.

The second panel of Table 6 isolates the effect of enhanced presentation on enrollment across varying degrees of information display associated with each field study. The first row indicates that enhanced presentation, in the context of a landing page that displayed the default (Field Study #1), led to a large and statistically significant increase in the share of personalized enrollment, offset by an equal-sized decline in automatic enrollment. As with enhanced design, enhanced presentation resulted in a substantial increase in the share of employees electing contribution rates in excess of the default or attaining the match threshold. These increases in personalized enrollment translated to a moderate but no longer statistically significant increase in overall contribution rates (b = 0.29, ns) and an increase in average saving rates (b = 0.59, p < 0.05). The remaining estimates suggest that the positive effect of enhanced presentation on personalized enrollment, and attainment of contribution thresholds, was robust to the additional display of information regarding the plan match (Field Study #2) or the absence of any informational display (Field Study #3). One might be tempted to conclude from the first two panels that an enhanced design that displayed both the default and plan match would have had led to an even more pronounced increase in enrollment than an enhanced design that displayed only the default. We caution, however, that the second field study was administered to only four large plans with match limits that exceeded the plan default. As Appendix Table A2 reports, while enhanced presentation in the context of both default and match display increased personalized enrollment in each of the four plans, the magnitude of the increase ranged from 0.05 (ns) to 0.15 (p < 0.05).

The third panel of the table isolates the effect of enhanced information on enrollment. The panel indicates that displaying the plan default modestly, but insignificantly, increased enrollment, with the exception of a marginally significant increase in the share attaining the match threshold (b = 0.04, p < 0.10). An inspection of heterogeneity across plans varying in the size of the default offers some additional clarity into the influence of default display. The analysis, reported in Appendix Table A3, reveals a more robust pattern of increased enrollment due to default display for the approximately three-quarter of

[40] We note that a strict comparison of estimates is confounded by modest compositional differences in the underlying estimation samples. Estimates of the saving rate rely on the 86 percent sub-sample for which we can observe/infer the match, while estimates of match take-up rely on a sub-sample that further excludes any plan with a zero match.

employees in a plan with a median-to-low default rate (i.e., ≤ 3 percent). There is little evidence that such display increased enrollment for plans with a higher default rate. Enhanced design, which includes the display of the plan default, exhibits a similar pattern of heterogeneous influence. While the estimates in the table are imprecise, they are at least suggestive of the possibility that, in plans with low default rates (and, consequently, plans where the default is likely below the match limit), employees are more likely to increase their contribution via personalized enrollment when the default is displayed.

Overall, Table 6 reveals the substantial positive effect of enhanced design on personalized enrollment, attainment of contribution thresholds, and overall contribution and saving rates and a modest positive effect on overall enrollment. The influence of enhanced design appears largely driven by enhanced presentation and the influence of enhanced presentation appears robust across variation in the display of plan information on the landing page. To contextualize the magnitude of the effect of enhanced design on enrollment, we can benchmark the treatment effects for design against the influence of two known positive predictors of savings–the financial generosity of the plan match and employee age. To generate these benchmarks, we estimated the marginal increase in enrollment associated with a one percentage point increase in a plan's match limit and a one-year increase in employee age conditioned on plan and employee characteristics across the several hundred plans in our sample.[41]

The benchmarking exercise indicates that the increase in enrollment attributable to enhanced design is equivalent to that predicted from substantial increases to the generosity of the plan match and advances in employee age. Specifically, the effect of enhanced design on personalized enrollment is equivalent to the increase one would predict from a 4.09 percentage point (i.e., 0.09 / 0.022) increase in a plan's match limit while its effect on average contribution is equivalent to a predicted increase of 4.46 percentage points (i.e., 0.62 / 0.139). Considering the modal plan match limit of 6 percent, these equivalencies imply that the effect of design on personalized enrollment (average contribution) is comparable to a 68 (74) percent increase in match generosity. Turning to age, the benchmarking analysis implies that the effect of design on personalized enrollment (average contribution) is equivalent to the predicted difference between two otherwise similar employees separated in age by 45 (i.e., 0.09 / 0.002) (11, i.e., 0.62 / 0.056) years. While we caution against interpreting the reference estimates as causal, particularly those involving age, the exercise nevertheless underscores the substantial economic importance of design relative to factors generally understood as predictive of employee saving.

[41] We estimated the following model for each enrollment outcome, $y^k{}_i$, for potential enrollee i, using the sample of all active enrollees randomized to a landing page without information display: $y^k{}_i = \alpha + \gamma MatchLimit_i + \beta Age_i + \theta Female_i + \Sigma\pi^s Treat^s{}_i + Z\varphi + \varepsilon_i$. The model controls for employee demographics, assignment to experimental treatment, and a vector, Z, that flexibly controls for plan-level characteristics (default rate, auto-escalation, and log median plan-level salary).

5.3 Influence of Design and Economic Consequence of Enrollment Decision

A specific question of theoretical, and practical, interest is whether design influences enrollment even in the context of plans with a (generous) match and, more generally, how the influence of design varies with the financial stakes of the enrollment decision. This inquiry is particularly important given that plans vary considerably in the generosity of their plan match. To explore this dimension of potential heterogeneity, we categorized plans using two measures of financial consequence: a plan's match limit as a percent of annual salary and the gap between a plan's default rate and match limit. While the first is a straightforward proxy for generosity, the second measure proxies the additional match an employee could potentially receive by increasing their contribution, relative to the default, via personalized enrollment.

Figure 3 displays average enrollment by experimental condition across plans varying in the financial consequence of the enrollment decision using data from Field Study #1. The panels depict the average share of personalized enrollment for the enhanced and basic design conditions across plans ordered by increasing match limit (Panel A) and difference between match limit and plan default (Panel B). The figure offers no indication that the positive influence of enhanced, relative to basic, design on personalized enrollment diminishes as the financial consequences of personalized enrollment rises. Indeed, the differential propensity of employees to personalize enrollment under enhanced design appears highest in plans where the stakes of enrollment are highest. Finally, we note one could interpret the relative slopes of the two lines in each panel as indicating that employees, under enhanced design, exhibit greater sensitivity to the financial returns associated with personalized enrollment. This illustrates the intriguing possibility that psychological design may actually complement economic plan design by amplifying employee responsiveness to a plan's underlying economic incentives.

5.4 Assessing the Generalizability of Treatments

We turn next to an analysis intended to clarify how our findings might translate to plans larger than the small-to-midsize market plans in our sample. Our strategy for extrapolating the potency of design at larger plans was to examine how the magnitude of treatment effects associated with design varied across plan size. To implement this test, we segregated plans into terciles based on size (using the number of experimental observations for each plan as a proxy for size) and then separately estimated the effect of enhanced design on personalized enrollment for each tercile from the first field study. We attempted to control for potential compositional differences across plans with the inclusion of plan, day-of-week, and week-of-year fixed-effects in the regressions. The analysis yielded the following treatment effects of design, from small to large tercile: 0.08 (se = 0.03), 0.12 (se = 0.03), and 0.09 (se = 0.04). The exercise offers no evidence that the effects of enhance design diminish as plans increase in size.

5.5 Decomposing the Effects of Enhanced Presentation

The notable influence of enhanced design, and in particular, enhanced presentation on enrollment, prompts the question of which design elements drove employee response. Field Study #3 was intended to decompose the influence of enhanced presentation into its three constituent modifications—standardization of language, modified headlines, and traffic-light colored buttons—in the context of no display of plan information.[42] Due to the structure of the study, inferences about decomposition involve pairwise comparisons across four experimental conditions: Basic, Standardization, Headlines (standardization + modified headlines), and Enhanced (standardization + modified headlines + color).

The third panel of Table 5 reports the marginal effects associated with each condition of the study and aggregate treatment effects. Overall, the estimates confirm the large and significant influence of enhanced presentation, relative to the basic design, on personalized ($b = 0.07$, $p < 0.01$) and automatic ($b = -0.08$, $p < 0.01$) enrollment. While the tests lack sufficient power for clear statistical distinctions, the results imply that the efficacy of enhanced presentation is largely driven by the standardization of language ($b = 0.04$, $p < 0.10$, personalized; $b = -0.06$, $p < 0.01$, automatic) and the adoption of traffic-light colors ($b = 0.03$, $p = 0.14$, personalized; $b = -0.02$, *ns*, automatic). The study indicates no relationship between modified headlines and enrollment with sufficient power to rule out anything in excess of a moderate positive effect (95% CI: -0.04 to 0.04).

6 Savings of Marginal and Inframarginal Personalized Enrollees

According to standard economic theory, to the extent that modest variation in design prompts some utility-maximizing employees, with well-defined preferences and beliefs, to personalize their enrollment, one would expect such employees to be largely indifferent between contributing at the plan's default rate and a rate proximal to the default. An alternative possibility is that marginal employees, compelled to personalize due to enhanced design, increase their contribution rates substantially, perhaps even as substantially as the typical inframarginal personalized enrollee. Typical personalized enrollees, as judged from the baseline condition of the first field study, increased their contributions by 4.44 percentage points, more than double the 3.38 percent average default. Considering the prevalence of plans with a match limit exceeding the plan default, such an increase in contribution implies an even more striking increase in an employee's saving rate. The alternative possibility of large increases in savings among marginal enrollees could arise if enrollment decisions were governed by non-standard mechanisms sensitive that were sensitive to design. For example, potential enrollees inattentive to decision-relevant details might substantially, and asymmetrically, adjust their contributions if design helped to remedy such inattention; or perhaps contribution decisions emerge from a heuristic process in which enrollees either

[42] As noted in Section 4, the interventions tested in the third field study feature nominal differences in wording relative to those in the first field study due in part to the chronological order in which the studies were administered (see Appendix for screenshots).

select the default rate or a rate far closer to recommended rates of savings. Depending on the model that underlies enrollment (a topic we engage in the subsequent section), the potential equivalence in saving between inframarginal and marginal enrollees introduces the possibility that small changes to design could result in increases to savings, and a shift to consumer welfare, far larger than that implied under standard economic assumptions.

Though one cannot typically observe the decisions of marginal decision-makers, the present research setting offers unique visibility into the relative savings of marginal and inframarginal enrollees along two distinct decision-making margins. Specifically, recall that to modify one's contribution, a newly plan-eligible employee in an AE 401(k) plan (in our sample, or more generally) must decide to visit the online enrollment portal (active enrollment), decide to personalize enrollment from the online interface (personalized enrollment), and finally elect a new contribution rate via the personalized enrollment webflow. Under reasonable assumptions, the present research allows us to assess the behavior of marginal enrollees along the margins of personalized enrollment—by comparing the contribution of personalized enrollees across experimental treatments—and active enrollment—by comparing enrollment outcomes of active enrollees across plans varying in their average rate of active enrollment.

6.1 Marginal and Inframarginal Equivalence – Personalized Enrollment

We first present evidence on contribution adjustments on the margin of personalized enrollment using data from the first field study. We can infer the savings behavior of marginal personalized enrollees (i.e., those induced to personalize due to enhanced design) by comparing contribution adjustments across experimental conditions if we assume that design did not affect the contribution adjustments of inframarginal personalized enrollees. The assumption that design did not affect inframarginal contribution adjustments seems reasonable given that none of the design modifications, particularly those comprising enhanced presentation, alluded to actual contributions (the assumption is also consistent with evidence on the distribution of adjustments described below).

The proposed comparison of personalized contribution adjustments reveals that personalized enrollees in the enhanced condition increased their relative contribution rate by an average of 4.57 percentage points, slightly exceeding the average 4.44 percentage point adjustment under basic design. Under the aforementioned assumption, a rough calculation implies that if enhanced design increased personalized enrollment by 0.09 relative to a baseline of 0.60 (Table 6), marginal personalized enrollees would have increased their contribution rate by an average of 5.44 percentage points, or 161 percent of the average plan default. If one were to abandon the assumption, and assume instead that design caused the average inframarginal enrollee to increase their adjustment by 10 basis points, relative to personalized enrollees in the basic condition, marginal personalized enrollees would have positively adjusted their contribution rate by an average of 4.77 percentage points, or 141 percent of the average plan default. Even

assuming a 30 basis point increase in inframarginal adjustments due to design would imply an average marginal adjustment of 3.44 percentage points, or 102 percent of the average plan default.

The apparent similarity in average personalized adjustment across experimental conditions is also conveyed by the regression estimates reported in Table 6. Specifically, Panel A of the table estimates that enhanced, relative to basic, design resulted in a nominally positive, and statistically insignificant, increase in average contribution rate among personalized enrollees of 0.24 (se: 0.25). Assuming once again that design did not directly influence the adjustments of inframarginal enrollees, the estimates imply that the 0.13 share of marginal personalized enrollees (i.e., 0.09/(0.60 + 0.09)) adjusted their contribution by an average of 6.28 percentage points (i.e., 0.24/0.13 + 4.44 average personalized adjustment in the basic condition), a 186 percent increase relative to the average plan default. Even the 95% statistical lower bound of the estimate implies an average adjustment among marginal personalized enrollees of 2.52 percentage points, a 75 percent increase relative to the average plan default.

An alternative, and perhaps more striking, test of equivalence involves comparing the full distribution of personalized adjustments across the basic and enhanced designs. Figure 4 implements this test by displaying the cumulative distribution of contribution adjustments for personalized enrollees across the basic and enhanced design conditions from the first field study. The first panel visually corroborates the similarity in contribution adjustments across experimental conditions for nearly every segment of the distribution.[43] Notably, the plot shows predominantly positive adjustments across both of the interface designs, suggesting that even marginal enrollees were disproportionately positive adjustors. A more formal non-parametric statistical test of similarity fails to reject the equality of the two distributions (Kolmogorov-Smirnov test, $p = 0.996$).

We propose one final test of marginal and inframarginal equivalence that leverages the breadth of the experimental plan sample. Specifically, suppose that the plan-level average treatment effect of design on personalized enrollment indicates the share of marginal personalized enrollees in each plan. If marginal enrollees are prone to only nominally adjust their contribution relative to the default, we should expect a negative correlation between the magnitude of plan-level treatment effects (i.e., the share of marginal enrollees) and the difference in the average within-plan personalized adjustment across basic and enhanced designs. Alternatively, if marginal personal enrollees were not prone towards smaller adjustments than inframarginal enrollees, one would not expect the size of the treatment effect to predict the differential personalized adjustment across conditions. The second panel of Figure 4 implements this test by depicting the plan-level difference in personalized adjustment under enhanced and basic design across increasingly large personalized enrollment treatment effects. To avoid noise associated with very

[43] The figure also supports the earlier assumption regarding the comparability of adjustments between inframarginal personalized enrollees under enhanced design and any personalized enrollee under basic design. To otherwise generate the depicted similarity in distributions would require that design led inframarginal personalized enrollees to increase their contribution to a degree that precisely offset the decreased contribution of marginal personalized enrollees.

small plans, the plot is restricted to plans from Field Study #1 with at least 5 employees assigned to each of the enhanced and basic conditions. Contrary to the negative correlation implied by non-equivalence, the plot shows a small positive correlation between the size of the treatment effect and the average differential adjustment across conditions.

Overall, the similarity in the distribution of contribution adjustments among personalized enrollees across experimental conditions and the cross-plan variation of such differences imply that marginal personalized enrollees increase their contributions substantially, possibly to a degree equivalent to their inframarginal counterparts. This equivalence helps to explain the previously discussed correspondence between the estimated marginal effect of design on the share of personalized enrollment and the share of employees whose contribution exceeded the default and met or exceeded the plan match. The simple explanation for these patterns is that most of the employees compelled to personalize enrollment due to design adjusted their contributions positively and to a significant degree.

6.2 Marginal and Inframarginal Equivalence – Active Enrollment

Next we consider evidence as to the potential equivalence in saving between marginal and inframarginal employees across a second consequential, and largely ignored, margin– the decision of a newly-eligible employee to actively visit the online enrollment portal during the initial enrollment period. The importance of active enrollment for initial saving is underscored by the significantly higher average contribution rate of active versus non-active enrollees (active enrollment allows the possibility of personalized enrollment which is, as discussed, associated with large increases to contribution). Notably, there is substantial variation in the share of employees who make active enrollment decisions across plans. Across plans in our experimental sample with at least 50 observations, the rate of active enrollment varied from 10 to 55 percent. A simple regression indicates that less than only 10 percent of variation in an employee's decision to actively enroll can be explained by variation in economic plan features (match limit, default rate) and employee characteristics (i.e., salary, age, gender).[44] According to standard theory, one would expect an employee who is indifferent between passive and active enrollment but is compelled towards the latter by some exogenous factor to nominally adjust their contribution relative to the plan default. Alternatively, if marginal active enrollees behaved like their inframarginal counterparts, one would expect them to substantially increase their contributions.

While we do not experimentally intervene on the margin of active enrollment, we have reason to believe that differences in psychological plan design—defined to include, beyond digital design, the success with which plans engage employees through marketing, communication, and reminders—contribute to the unexplained variation in active enrollment across plans. Indeed, conversations with plan

[44] We decomposed variance with the following regression: $y_i = \alpha + X'\gamma_i + Z'\theta_i + \varepsilon_i$. Here, y_i denotes active enrollment, X' is a vector of indicators that flexibly controls for the match limit and default rate, while Z' is a vector of indicators that flexibly controls for employee demographics (i.e., $25k salary bins, 5-year age bins, and gender). Robust errors clustered at the plan level.

sponsors, and our survey of several hundred plans (described in detail in a subsequent section), makes it clear that while most AE plans encourage active enrollment—78 percent of AE plans reported that they "somewhat" to "strongly" encourage eligible employees to make an active decision during pre-enrollment—they vary in the frequency, modality, and efficacy of such communications. For example, when asked how they sought to encourage active enrollment, 77 percent of plans cited the use of enrollment kits, 67 percent cited information sessions, 43 percent cited mailings, 24 percent cited posters and flyers, and 21 percent cited online websites. Critically, the intensity with which plans encouraged active enrollment strongly predicted the plans' actual share of active enrollment ($b = 0.16$, $p < 0.01$).[45]

The likelihood that at least some of the residual variation in active enrollment across plans is attributable to variation in aspects of design presents an opportunity to test for marginal and inframarginal equivalence on the margin of active enrollment through a cross-sectional plan analysis. Specifically, one can interpret a non-negative correlation between the rate of residual active enrollment—that is the rate of active enrollment after flexibly controlling for economic plan features (default, plan match) and employee characteristics (income, age, gender)—and enrollment outcomes among active enrollees as evidence for equivalence since such a pattern implies that across plans with increasingly larger shares of marginal active enrollees, there is no corresponding reduction in average saving.

We undertake this assessment by first categorizing plans into terciles based on their rate of residual active enrollment and then comparing enrollment outcomes across plans ranked by increasing tercile. Table 7 reports the outcome of this exercise for the 72 small-to-midsize plans with at least 50 online and offline enrollees from 2016 to 2018 for which we have data. Despite expected balance across plan characteristics (first panel), the table offers no evidence of diminished enrollment outcomes in plans with high rates of residual active enrollment (second panel). While caution is warranted in light of the assumptions required for causal inference—e.g., one could argue that unobserved differences in employee characteristics across plans might be positively correlated with enrollment—the table is consistent with the striking possibility that employees on the margin of active enrollment increase their savings as much as the average inframarginal active enrollee. This equivalence, which parallels the earlier equivalence on the margin of personalized enrollment, implies that plans could increase savings, and increase employee welfare, perhaps dramatically, through modest non-economic interventions such as email reminders that increase the rate of active enrollment or small changes to the design of the enrollment interface.

7 Clarifying Decision Mechanisms

To explore the mechanisms underlying employee response to design, and the suggested equivalence between marginal and inframarginal behavior, we tested several theory-driven hypotheses

[45] Estimate derived from a simple regression of each plan's average share of online enrollment (inferred from categorical estimates provided by plans) and self-reported intensity of encouragement measured on a 1 (not at all) to 4 (very) scale.

through an online paradigm. The paradigm involved enlisting several thousand full-time US employees to make hypothetical enrollment decisions from a digital platform programmed to closely resemble that used by employees in the field experiments. Our practical ability to replicate a functional commercial platform was facilitated by cooperation with technical and design specialists from Voya. Beyond attempting to simulate the enrollment environment, the online instrument collected background information and elicited several decision-relevant beliefs and preferences intended to help disentangle behavioral mechanisms. To conceptually organize the hypotheses of interest, we describe an economic model of enrollment informed by traditional economic assumptions but for the allowance of a series of behavioral frictions informed by interdisciplinary research. After assessing each potential mechanism using experimental data, we speculate as to an alternative model of decision-making that might more accurately describe our findings and perhaps financial decisions made from digital interfaces more broadly.

7.1 Conceptual Framework

Our framework stipulates that a utility-maximizing employee, newly eligible for a 401(k) plan, will decide to either confirm automatic enrollment or personalize enrollment at a rate greater than the default based on a deliberative comparison (allowing for some uncertainty) of the discounted flows of future utility associated with each option. For tractability, and to reflect choice patterns observed in the field, we do not consider declined enrollment and for those who personalize their enrollment, we assume a non-trivial increase in contribution rate. To set up our tests of behavioral mechanisms, we allow for the potential presence of three decision-making frictions commonly invoked in research on non-standard consumer financial decisions, often in the context of understanding the efficacy of some nudge: the possibility that choice emerges from limited capacity to attend to, or recall, decision-relevant information unless that information is made salient (e.g., Chetty et al. 2009); the possibility that choice is shaped by confusion or a lack of understanding of available options (e.g., Shue and Luttmer 2009); or the possibility that choice reflect a lack of trust in financial institutions (e.g., Agnew et al. 2012; Galiani et al. 2020).

The resulting framework specifies four distinct channels through which enhanced design might cause a respondent to personalize their enrollment. First, design might compel personalized enrollment by prompting the respondent to update their beliefs/preferences regarding the costs (e.g., the effort- and time-costs associated with the concrete act of enrolling) or benefits (e.g., the importance of increased contribution for retirement security or the minimum rate of savings required to attain such security) of increased contributions. Second, for those constrained by limited attention, memory, or recall, enhanced design may act by drawing attention to relevant plan details such as the default rate. Third, enhanced design may help to reduce confusion about the practical meaning of each enrollment option. Finally, design may increase saving by remedying some underlying distrust in financial institutions.

7.2 Research Design of Hypothetical Choice Experiment

We sought to test the proposed channels via an experimental paradigm in which online subjects were asked to make hypothetical enrollment decisions from an enrollment portal—i.e., the enrollment landing page and the subsequent personalized enrollment webflow—programmed to look and function like that encountered by employees in real-life (apart from modest stylization of the real-life webflow). Our practical ability to replicate a highly professional online portal, with functionality, including a operational retirement calculator, was facilitated by cooperation from members of Voya's design and technical teams. While we tested an expansive set of experimental variations to the design of the interface, and to the generosity of the plan match, we focus here on describing the response to the treatments most proximal to those tested in the field. Assuming that employees were responsive to design in this paradigm, we sought to gain insight into underlying mechanisms by eliciting decision-relevant beliefs and preferences informed by the model and an additional set of exploratory mechanisms.

We drew our sample from a commercial panel of online survey respondents purchased through Qualtrics.[46] As the result of initial screens on employment status and age, and after excluding respondents who failed an attention screen, the final sample included 6,871 full-time employed US adults of which 79 percent reported being eligibility for a 401(k) plan with their real-life employer. The sample is diverse across a range of demographic and financial characteristics and observably resembled the field experimental sample with the notable exception of gender (41 percent male vs. 56 percent in the field).

We administered the 10- to 15-minute online instrument during the summer of 2017 (prior to the launch of the field experiments). The instrument comprised three sections. An initial section collected demographic and financial background of the respondents. Respondents then proceeded to a second section introducing the hypothetical enrollment scenario. Specifically, respondents were asked to imagine they had recently been hired by a new manufacturing firm ("The ABC Company") and had to make a series of decisions regarding enrollment in benefit programs. On the next screen, respondents were presented with a letter (adapted from the actual language used by an anonymous large firm) communicating their eligibility for the employer's 401(k) AE plan. The letter summarized plan details including the default contribution rate, dollar-for-dollar plan match up to an experimentally varying match limit, and investment allocation (Appendix Figure A5).

Respondents were then directed to one of several experimental variants of the online interface from which they were instructed to either confirm enrollment at the default rate, personalize their enrollment at some other rate, or decline enrollment (Appendix Figures A6 and A7). The interface, which closely resembled that used in the field, featured a welcome message in gray type below which respondents were directed to choose one of the three enrollment options. Respondents electing to

[46] Qualtrics recruits its commercial panel through a variety of direct and indirect channels and did not disclose the size of subject payments. We paid Qualtrics $6 for each completed survey from a respondent satisfying the pre-specified screens.

personalize their enrollment then proceeded through a modestly simplified version of the real-life personalized enrollment webflow. The webflow collected additional financial information, elicited the desired contribution rate in the context of a functional retirement savings calculator, and asked respondents to confirm their election decision (Appendix Figure A8).[47] The webflow did not permit respondents to modify their investment allocation. The portal featured the insignia of a fictionalized provider, "Star Financial," but otherwise retained the Voya design and color aesthetic.

Following their enrollment decision, respondents proceeded to a final section of the survey that elicited several retirement- and decision-relevant beliefs. Specifically, we asked for ex-ante perceptions of the costs and benefits of personalized enrollment (i.e., the effort-costs of personalized enrollment, the importance of personalized enrollment for future financial security, and the minimal annual rate that one should contribute each year to ensure retirement security), understanding of available enrollment options, recall of plan features, and financial trust in the plan provider.[48]

Experimental Conditions. The online experiment randomized respondents to one of several variations in the design of the landing page relative to a baseline interface including designs close .[49] In the baseline condition, the hypothetical plan offered a dollar-for-dollar match up to a 6 percent annual contribution, with a default rate of 3 percent. The experimental conditions varied two dimensions of psychological design—the display of plan information (either no display, the display of the plan match and plan default, or the display of the default alone) and the visual presentation of the landing page. Specifically, the latter included three of the design elements tested in the field experiment—the standardization of descriptive text for each option, the simplification/rephrasing of option headlines, and the use of traffic-light colors. Appendix Table A4 enumerates the full set of tested conditions in the experiment and associated sample sizes. We note that the experiment tested the marginal effect of the three presentational design elements from the field (standardization, headlines, color) as well as combination of such elements (standardization + headlines, standardization + color, standardization + headlines + default display) that approximate, but do not precisely replicate, the enhanced design (standardization + headlines + color + default display) and enhanced presentation (standardization + headlines + color) conditions from the first field experiment.

7.3. Comparison of Enrollment in the Lab and the Field

Overall, respondents from the hypothetical choice paradigm exhibited a similar rank-ordering of enrollment options (0.50 personalized enrollment, 0.41 automatic enrollment, and 0.09 declined

[47] This calculator used a simplified version of the algorithm that powered the calculator in the real-life enrollment portal.
[48] A subset of experimental conditions examined the effect of design with a less generous plan match but we omit those results for brevity. The survey also included questions on exploratory mechanisms such as financial literacy and decision-making style. To minimize survey length, some modules were randomized so that not all participants answered all questions.
[49] Since the hypothetical choice experiment was conducted prior to the field studies and an internal compliance review, minor discrepancies in wording were introduced between the online paradigm and field as a result of the compliance process.

enrollment) as employees from the field (0.60 personalized enrollment, 0.22 automatic enrollment, and 0.18 declined enrollment) but were more apt to choose automatic enrollment. Both contexts generated a similar average rate of overall contribution (5.1 percent in the lab, 5.4 percent in the field), or personalized contributions (7.7 percent in the lab, 7.8 percent in the field). We caution, however, that these baseline comparisons are shaped by the specific hypothetical plan features we chose to present.

Table 8 describes the treatment effect of design on personalized enrollment and mechanism-relevant outcomes for the experimental analogues to the conditions tested in the field. The estimates emerge from a series of regressions of the same structure as those used to evaluate the field experiment. The first column reports the effect of enhanced design, presentation, and information on the share of personalized enrollment. Specifically, the first row indicates that the experimental analogue to enhanced design influenced personalized enrollment to an absolute extent (+0.08) closely resembling the field (+0.09). The remaining rows indicate the positive effect of other design elements on personalized enrollment, including the display of the plan match limit. Overall, the first column points to the substantial effect of design on enrollment, to a degree comparable to that observed in the field.

7.4 Evidence on Decision Mechanisms

The similarity in responsive of hypothetical and real-life enrollment decisions to psychological design nicely sets up tests of decision mechanisms. The remaining columns of Table 8 summarize the marginal treatment effects of design with respect to outcomes associated with the four channels specified by the model. Focusing on the experimental conditions that elicited a large enrollment response, the first set of columns imply that design did not affect enrollment through standard economic channels of preferences or beliefs. That is, design did not meaningfully shift beliefs about the relative time- or effort-costs of enrollment nor did it shift beliefs/preferences about the importance of personalized enrollment for retirement security or the contribution levels required for retirement security (while some design combinations did appear to reduce the perceived effort of personalized enrollment, the reductions were modest relative to baseline perceptions and not strongly correlated with the magnitude of the treatment effect). While our focus on average effort beliefs might obscure important shifts elsewhere in the distribution, closer inspection confirms that enhanced design did not substantively shift the share of respondents with very low (or high) perceptions of effort costs relative to baseline.

The subsequent set of columns address the possibility that design influenced enrollment by improving knowledge/recall of plan details. The columns point to significant baseline deficits in recall of the plan match and the default rate and Panels A and C suggest that the display of such information on the enrollment interface improves recall. Moreover, the display of the match (but not the default rate alone) appears to (positively) affect match recall and personalized enrollment to a similar extent, consistent with increased knowledge/awareness of the plan match leading respondents to greater personalized enrollment.

Further examination highlights that the influence of the match display on personalized enrollment was restricted to hypothetical plans in which the match limit exceeded the plan default (i.e., a plan match of 6 percent and default of 3 percent). Improved knowledge of the match does not, however, appear to be the channel through which enhanced presentation affected personalized enrollment (Panel B). While one must exercise caution in generalizing inferences about plan information from an online paradigm in which the timing and nature of plan communication differs from that in real-life, the table suggests that improvements in match knowledge/awareness may help to explain the effect of enhanced information on enrollment it does not help clarify the influence of enhanced presentation.

The final two columns address the potential mechanisms of confusion and institutional trust. Regarding the former, the table indicates that design did not influence enrollment by affecting self-reported decision clarity (which we interpret as a proxy, at least in part, for respondents' understanding of the choice menu). The final column of the table rejects the possibility that design influenced enrollment by increasing perceptions of institutional trust. Enhanced design had a consistently positive, fairly precise, but ultimately small, relative to the 5.2 baseline, effect on perceived trust. Overall, the table replicates the pronounced influence of design but fails to implicate any of the hypothesized channels – design does not influence behavior by shifting decision-relevant beliefs or preferences, improving knowledge/recall of previously communicated plan detail (with the possible exception of the match display), reducing decision confusion, or heightening institutional trust.

7.5 An Affective Framework for Plan Enrollment

Given that design does not appear to influence enrollment via the economic or behavioral channels implied by our framework of deliberative decision-making, we consider an alternative, more markedly, non-standard mechanism. One intriguing possibility is that enrollment emerges, at least in part, from a heuristic, or intuitive, decision process in which design has outsized influence. Specifically, we speculate that design may influence behavior not by providing new decision-relevant information or otherwise remedying cognitive/informational deficits in the context of a deliberative comparison of costs and benefits, but rather by influencing affective evaluations in the context of a heuristic decision process.

The potential for emotion and specifically anxiety—broadly defined as a mental-physical state that arises in response to some perceived threat—to shape judgement and choice has been asserted by an expansive literature in neuroscience, psychology, and decision science.[50] For example, scholarship on emotion commonly cites avoidance as one of the primary behavioral responses to anxiety (e.g., Hartley and Phelps 2012), while an influential perspective in decision-making describes how choice is influenced both by cognitive evaluation and anticipatory emotion (Loewenstein et al. 2001). More recently,

[50] Researchers, as an example, have found that anxiety prompts activity in brain regions associated with emotional regulation, attention, memory and executive function (e.g., Park et al. 2016, Eysenck et al. 2007, Wolf 2009).

economists and policy analysts have asserted the importance of anxiety, and in particular, financial anxiety on well-being and decision-making. For example, recent national surveys have documented the high prevalence of financial anxiety among US households, while recent research has posited the influence of anxiety across a range of domains including saving (e.g., Bhargava and Conell-Price 2021).[51]

We sought evidence on whether design may have influenced personalized enrollment via an affective channel through a simplified adaptation of the previously described online paradigm. As with that study, we asked a smaller sample of US employees (n = 145), this time recruited through Amazon Mechanical Turk, to make a personalized enrollment decision as if they were a new employee at a firm with the same plan match and default as earlier. Rather than randomizing subjects to experimental variations of a fully functional online interface, subjects were asked to make an enrollment choice after randomization to an image of either the baseline or enhanced enrollment interface from the first field experiment. Following their decision to personalize, confirm, or decline enrollment (and their contribution election if they decided to personalize) we asked subjects to anticipate the degree to which they would have felt anxious, on a scale ranging from 1 to 5, about their financial situation had they decided to personalize their enrollment.

An initial observation of note is that the experiment replicated the substantial effect of enhanced design on increased personalized enrollment seen in the field and primary lab study ($b = 0.33$, $p < 0.01$). While the effect of enhanced design was larger than that previously observed, the paradigm generated the same rank-order of enrollment choice. Accepting the validity of the paradigm, three specific findings help us evaluate the hypothesized affective channel. First, the decision to personalize enrollment was negatively correlated with the degree of anxiety respondents associated with personalization. Specifically, a one-unit increase in anxiety was linked to a 0.13 decrement in the share of personalized enrollment ($p < 0.01$). Second, assignment to the enhanced design condition led respondents to perceive personalized enrollment with less anxiety ($b = -0.36$, $p < 0.10$). Finally, using experimental assignment to enhanced design as an instrument for anxiety, we found that increasing anxiety reduced the likelihood that a respondent would personalize enrollment ($b = -0.92$, $p < 0.10$).

While we interpret these findings cautiously—e.g., to interpret the anxiety instrument as valid one must assume that enhanced design influences enrollment only through the hypothesized channel—the evidence collectively suggests the possibility that enrollment decisions are, at least in part, influenced by negative affective evaluations induced by design. Indeed, we speculate that one specific design element that may have depressed personalized enrollment is the use of the word "cancel" in the baseline condition—not because it led respondents to shift their inferences about the value of saving or because it led to confusion—but rather because it may have led to a diminished affective evaluation.

[51] For example, in their survey of 25,000 households, the 2018 National Financial Capability Study found that 53 percent of respondents strongly agreed that thinking about personal finances triggered anxiety.

8 Implications of Psychological Design for Employee Welfare and Policy

The documented influence of design, and the suggested equivalence between inframarginal and marginal savings, has potentially significant implications for the long-run well-being of employees, the optimal design of policies intended to encourage savings, and economic estimates of consumer welfare. To better understand these implications, we present three additional analyses. First, to clarify long-term effects of design on retirement security, we project the additional years of security that an employee of varying ages would be expected to achieve were they to transition from automatic to personalized enrollment. Second, to clarify the potential for optimal design to emerge from existing market forces, we explore whether those charged with overseeing 401(k) plans are cognizant of the efficacy of design and are able to identify specific features of design that could improve engagement. To generate evidence, we present the results of an original survey of several hundred plan administrators in which we assess their sophistication regarding the potency of design. Finally, we present suggestive evidence as to how the significant diversity in design that currently exists across large plan providers—differences much larger than those engaged in our field studies—might affect average employee savings.

8.1 Design and Years of Excess Retirement Security

We sought to project the long-run effect of enhanced design on retirement security by estimating the excess years of financial security that representative employees, of varying ages, might achieve were they to personalize, rather than automate, their enrollment due to design. The critical assumption of the exercise, informed by evidence on the equivalence between marginal and inframarginal saving, was to set an initial contribution rate of 7.8 percent for those who personalized enrollment and 3.4 percent for those who were automatically enrolled (the respective empirical averages of personalized and automatic enrollees across the field studies). As with the assessment of preparedness, our strategy was to first project accumulated savings for representative employees given assumptions about wage growth, contribution inertia, and market returns. For simplicity, the exercise differs from the earlier projections in that the projections are deterministic rather than probabilistic—that is, we assume constant annual market returns and represent contribution inertia with a fractional increase in the annual contribution rate. Next, we calculated the years of retirement security that one could secure with projected savings, post-tax, via the purchase of an immediate, fixed-period, annuity.[52] Our representative employee earned $50,000 annually at the onset of saving, with a 1 percent annual rate of real wage growth, and was enrolled in a plan that matched 50 percent of pre-tax contributions up to a 6 percent limit (the modal match structure of plans in our sample). As a simplified representation of inertia, we assumed that employees would increase

[52] Specifically, we report the longest fixed-period annuity that one could purchase providing $864 of monthly income assuming a 3 percent rate of return during retirement. The monthly income is the minimum income required for retirement security according to the 2020 Elder Index, after adjusting for median Social Security benefits (see above). Annuity calculations drawn from an annuity pricing calculator accessed in August 2021.

their contributions by 0.25 annually until a 10 percent contribution limit. We followed the earlier projections in assuming a 10 percent tax rate at retirement and retirement at the age of 65.

Table 9 projects the number of years of retirement security, censored at 30 years, associated with automatic and personalized enrollment for representative employees across varying savings horizons for each of four potential future market scenarios. While the table conveys the importance of the timing of initial enrollment (as one would expect given the external savings assumption) and future market returns for retirement security, it also points to the substantive influence of the mode of initial enrollment. This influence is particularly pronounced for employees who begin to save in middle age or later and for even younger employees if one abandons the assumption of high future market returns. For example, the analysis implies that an automatic enrollee with a 15-year savings window would accumulate assets lasting an estimated 7.5 retirement years, assuming moderate market returns, as compared to 14.2 years with personalized enrollment.

Overall, the table underscores the material importance of plan design for the estimated 9 percent of employees who personalized enrollment due to changes in interface design and the potentially far larger share of employees that we speculate would personalize enrollment if prompted by their plan to make an active enrollment decision. A potentially informative benchmark is to compare the projected effect of design to that of automatic enrollment. Using the same assumptions as above (including moderate market returns), we estimate that those who shift from automatic to personalized enrollment due to design would accumulate an additional 35 to 67 percent in post-tax savings, for accumulation windows spanning 10 to 25 years, a greater increase than those who shift from non-participation to default participation (assuming the modal default rate of 3 percent) under automatic enrollment.

8.2 Market for Optimal Design

The potential value of optimal 401(k) plan design raises the question as to the sophistication of those responsible for overseeing, managing, and marketing plans. In theory, for plans desirous of high employee engagement via active enrollment, optimizing plan design likely offers a less expensive alternative to increases in the plan match. On the other hand, if plans were unaware of the potency of design, or unsure of how to identify optimal design, then plans might inadvertently discourage employee savings through poor plan design. One speculates that the absence of such sophistication might warrant the introduction of regulatory oversight or lead to the emergence of third-parties specializing in design.

We formally tested for market sophistication regarding design by inviting roughly two-thousand plan administrators, fiduciaries, and human resource executives from Voya's small-to-midsize 401(k) market to complete an approximately 10-minute online survey about plan administration in December 2018. To maximize response, the survey was marketed as confidential, invitations were administered by Voya and the invited sample included plans without automatic enrollment. After capturing background

detail, the survey collected data on plan engagement, marketing, and communication. Respondents then proceeded to a module intended to assess their ability to forecast the influence of design on enrollment. Specifically, the module first asked respondents to imagine they were responsible for overseeing a 401(k) AE plan with a 3 percent default and a 50 percent match up to a 6 percent limit. Respondents were then shown an image of the plan's supposed online enrollment interface (the basic design from the first field study) and were provided hypothetical enrollment statistics (i.e., approximate averages from the field). Next, respondents were prompted to forecast how the adoption of a new interface design (the enhanced design condition from the first field study) would affect personalized enrollment both in absolute terms and relative to an increase in the generosity of the plan match. Finally, respondents were asked to rank the efficacy of specific design modifications and indicate the confidence of their ranking. We incentivized the ranking by qualifying accurate responses for a prize lottery featuring an Apple Watch.

The 637 respondents who completed the survey reflected a representative mix of plans (37 percent offered automatic enrollment; 76 percent offered a plan match) and significant experience in the benefit administration space (3 percent reported less than 1-year of experience; 63 percent reported 10+ years). An initial takeaway from the survey was that while most plans appear invested in improving engagement, plans varied significantly in the intensity and modality of their marketing efforts and in their level of engagement. For example, among AE plans, 78 percent reported encouraging employees to actively enroll during pre-enrollment to some degree.[53] The expression of this encouragement, as noted earlier, took diverse forms (e.g., information sessions, online resources, flyers/posters, letters), and the self-reported intensity of encouragement strongly predicted each plan's self-reported rate of active enrollment. The findings support the contention that the substantial variation in active enrollment in AE plans may be due, at least in part, to variation in non-economic aspects of plan administration.

Figure 5 summarizes the forecasts of design efficacy. The first panel conveys that 88 percent of administrators underestimated the potency of enhanced design. The majority of respondents actually predicted that the shift from basic to enhanced design would *decrease* personalized enrollment. When asked to estimate the increase in the plan match required to achieve an increase in contribution equivalent to that delivered by enhanced design, for respondents correctly anticipating the directional effect of enhanced design, 98 percent indicated an increase smaller than the 4.4 percent estimated from the field. The second panel depicts the accuracy with which respondents were able to rank-order specific design modifications by efficacy along with the confidence they had in their rankings. Specifically, 11.8 percent of respondents were able to identify the most influential modification, while only 4.3 percent correctly

[53] While one might reasonably ask why plans invest in employee engagement, we speculate that this interest is driven by regulatory requirements that penalize non-uniform engagement across income, reputational benefits, and the perception that greater financial security may lead to higher employee productivity.

rank-ordered all three modifications.[54] Respondent confidence did not predict increased accuracy and the most confident respondents were actually the least likely to have identified the most effective design. Further analysis indicated that self-reported years of experience in benefit administration did not predict greater ranking accuracy. Overall, the survey suggests that while most plans expend resources to encourage active enrollment, they underestimate the importance of design, particularly in relation to plan incentives, and struggle to identify successful design elements despite confidence in the ability to do so.

8.3 Naturalistic Variation in 401(k) Plan Design across Financial Providers

Given the motivating evidence on the retirement risk faced by current 401(k) enrollees, the capacity of design to decrease this risk, and the naiveté of plan administrators with respect to design, we consider whether the substantial diversity in design that characterizes the existing plan/provider landscape might contribute to variation in saving. We note that in the present research we consider modest changes to the design of otherwise standardized enrollment interfaces across non-large plans sharing the same administrator. Such standardization contrasts with the far more substantial variation that one can discern across plans from different providers or across large plans with customized interfaces from the same provider (or across large and non-large plans within the same provider, etc.). These differences extend to the visual appearance of the interface, language, as well as the degree of personalization, simplicity, usability, warmth, etc. While a thorough audit and evaluation of design across providers is beyond the scope of the current paper, we can generate preliminary insight into the consequences of design variation across the broader plan landscape from a final experiment involving hypothetical enrollment.

The study was administered in April 2019 and involved N = 302 employed US adults recruited from Amazon Mechanical Turk. Following the format of the earlier studies, we presented subjects with a stylized enrollment scenario involving a 401(k) AE plan (once again, featuring a 3 percent default and a match up to 6 percent) and then elicited an enrollment decision after randomization to an experimentally varying interface. The interfaces corresponded to images of the real-life enrollment interfaces used by a large plan administered by Voya, the standard interface from another major provider, and the basic and enhanced designs from the field studies. Notably, the large plan interface, like many existing interfaces, was largely text-based and featured a choice menu with two options—personalized and declined enrollment (a note indicated that selecting neither option would confirm automatic enrollment).

Overall, the exercise generated the same ordering of options, across conditions, as the field and earlier lab studies (0.68 personalize, 0.26 confirm, 0.06 decline) and replicated the positive effect of enhanced, relative to basic, design ($b = 0.27$, $p < 0.01$). Most intriguingly, relative to basic design, the large plan interface yielded a significantly higher share of declined enrollment ($b = 0.13$, $p < 0.01$), while

[54] We first asked respondents to identify the most effective design modification (or indicate that none of the modifications were effective). Those selecting an initial design modification were then asked to identify the second-most effective modification.

the interface from the outside provider yielded a share of personalized enrollment higher than Voya's basic design ($b = 0.19$, $p < 0.05$) but (imprecisely) lower than Voya's enhanced design ($b = -0.08$, *ns*). The exercise, while not meant to provide definitive estimates of real-life enrollment, nevertheless suggests that real-life variation in psychological plan design, beyond that considered in our field studies, could lead to significant heterogeneity in personalized enrollment and even overall plan participation.

9 Conclusion

We began our research by presenting new administrative evidence from several hundred 401(k) plans with automatic enrollment indicating that the risk of retirement insecurity, usually attributed to the lack of plan access or non-participation, extends to a significant share of actual enrollees and varies across similar plans (the precise share of risk depends on one's preferred assumptions). We posited that because of pervasively low default rates, plan match limits that usually exceed default rates, and considerable inertia in contributions over time, a newly plan-eligible employee could reduce their risk of retirement insecurity by adopting a high initial contribution rate via personalized enrollment. To test the sensitivity of this initial decision to personalize enrollment to a plan's psychological design (i.e., the non-economic features that increasingly shape how consumers engage plans on digital platforms)—we conducted a series of field experiments that randomized several thousand employees across a few hundred 401(k) AE plans to enrollment interfaces varying in their color, language, and diction and the display of plan detail.

The field experiments revealed that, despite not providing new information or changing plan incentives, small changes in design resulted in substantial increases in personalized enrollment, average contribution, the share of contributions exceeding the default rate and, for plans with a match, the share of contributions with full match take-up. We estimate that the design-induced increase in saving is equivalent to that predicted by large, and presumably expensive, increases to the plan match. As notable as the profound effect of design on saving is evidence suggesting that marginal personalized enrollees more than doubled their contributions, similar to inframarginal counterparts. Additional cross-sectional evidence hints at the equivalence of marginal and inframarginal saving along a second, perhaps even more consequential, decision-margin of active enrollment (i.e., the decision to actively visit the enrollment interface). Collectively, the effect of design on personalized enrollment, and the apparent equivalence in saving between marginal and inframarginal enrollees, implies the enormous importance of psychological design for long-run financial security and economic analyses of consumer welfare. Indeed, we projected that the adoption of enhanced design at initial enrollment could result in multiple years of additional retirement security, particularly for employees who begin to save their 40s or later.

We investigated the mechanisms underlying the response to design through an experiment in which several thousand subjects were asked to make hypothetical 401(k) enrollment decisions from experimental varying interfaces in the context of an online setting programmed to closely resemble real-

life. While the paradigm confirmed the influence of design on enrollment, detailed data on preferences and decision-relevant beliefs suggested this influence cannot be explained by standard economic channels of preferences or beliefs, nor the cognitive frictions of attention, confusion, and trust often contemplated by the literature, and cited in explanations of the efficacy of nudges. We speculate instead, and provide evidence, that enrollment decisions from (digital) menus arise from a non-deliberative process in which design shapes affective evaluations that inform choice. Integrating across evidence from the field and lab, we speculate that many employees may not have defined preferences for how much to save and that financial decisions, such as 401(k) enrollment, are influenced by non-cognitive assessments that are as sensitive to features of psychological design as they are to economic incentives.

We caution that the present research is subject to potentially important limitations. For example, our visibility into the saving landscape is restricted to the initial enrollment decisions (and a few months thereafter) of enrollees in small-to-midsize AE 401(k) plans. And while the robustness of effects across plan size and financial stakes alludes to generalizability, future work is needed to explicitly examine the influence of design in large plans, in plans without automatic enrollment, and on the long-term trajectory of contributions. With respect to the latter, we cannot rule out the possibility that an increase in loan defaults or early withdrawals in the long-run might offset short-run increases to savings.

In spite of these limits, we see our findings as having first-order implications for savings policy, welfare analyses, and consumer protection. A critical question as digital choice environments evolve is whether market dynamics will organically lead plans (or providers) to arrive at designs that encourage savings or, in the least, take-up of generous matching incentives. Our survey of several hundred plan administrators suggests that one challenge, even for plans committed to engagement, is widespread underestimation of the potency of design, an inability to identify efficacious design elements, and overconfidence in forecasting accuracy. While one can only speculate as to the adoption of evidence-based approaches to benevolent design or the continued emergence of third-party design specialists, such adoption may not take place without explicit guidance from regulators or industry advocates or a redefinition of fiduciary responsibility so that it directly encompasses design. More generally, unlike nudges—ostensibly avoidable interventions intended to facilitate choice by helping decision-makers overcome predictable cognitive errors—some digital designs could be costly for consumers to avoid, perhaps because they wield their influence not only through channels, but through the motivational and affective channels more commonly discussed in marketing, advertising, and gamification research. As individuals are increasingly asked to make financial decisions from tactically designed digital platforms, including gamified environments, considerations of psychological design may be as, or more, critical for welfare than the economic plan features and the informational disclosures presently prioritized by policymakers.

References

Agnew, Julie R., Lisa R. Szykman, Stephen P. Utkus, and Jean A. Young. 2012. "Trust, Plan Knowledge and 401(k) Savings Behavior." *Journal of Pension Economics & Finance; Cambridge* 11 (1): 1–20.

Arnsten, Amy F. T. 1998. "Catecholamine Modulation of Prefrontal Cortical Cognitive Function." *Trends in Cognitive Sciences* 2 (11): 436–47.

Augenblick, Ned, and Scott Nicholson. 2016. "Ballot Position, Choice Fatigue, and Voter Behaviour." *The Review of Economic Studies* 83 (2): 460–80.

Baek, Tae Hyun, and Chan Yun Yoo. 2018. "Branded App Usability: Conceptualization, Measurement, and Prediction of Consumer Loyalty." *Journal of Advertising* 47 (1): 70–82.

Beshears, John, James J Choi, Christopher Clayton, Christopher Harris, David Laibson, and Brigitte C Madrian. 2020. "Optimal Illiquidity," NBER Working Paper #27459, .

Beshears, John, James J. Choi, David Laibson, and Brigitte C. Madrian. 2013. "Simplification and Saving." *Journal of Economic Behavior & Organization* 95 (November): 130–45.

Bhargava, Saurabh, and Lynn Conell-Price. 2021. "Serenity Now, Save Later? Evidence on Retirement Savings Puzzles from a 401(k) Field Experiment."

Bhargava, Saurabh, George Loewenstein, and Justin Sydnor. 2017. "Choose to Lose: Health Plan Choices from a Menu with Dominated Option*." *The Quarterly Journal of Economics* 132 (3): 1319–72.

Bhargava, Saurabh, and Dayanand Manoli. 2015. "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment." *American Economic Review* 105 (11): 3489–3529.

Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin. 2003. "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism.'" *University of Pennsylvania Law Review* 151 (3): 1211.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States *." *The Quarterly Journal of Economics* 129 (4): 1553–1623.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence." *American Economic Review* 99 (4): 1145–77.

Choi, James J., David Laibson, and Brigitte C. Madrian. 2011. "$100 Bills on the Sidewalk: Suboptimal Investment in 401(k) Plans." *The Review of Economics and Statistics* 93 (3): 748–63.

Cronqvist, Henrik, and Richard H. Thaler. 2004. "Design Choices in Privatized Social-Security Systems: Learning from the Swedish Experience." *American Economic Review* 94 (2): 424–28.

Cronqvist, Henrik, Richard H. Thaler, and Frank Yu. 2018. "When Nudges Are Forever: Inertia in the Swedish Premium Pension Plan." *AEA Papers and Proceedings* 108 (May): 153–58.

Datta, Saugato, and Sendhil Mullainathan. 2014. "Behavioral Design: A New Approach to Development Policy." *Review of Income and Wealth* 60 (1): 7–35.

Dellavigna, Stefano, and Joshua M. Pollet. 2009. "Investor Inattention and Friday Earnings Announcements." *The Journal of Finance* 64 (2): 709–49.

Engelhardt, Gary V., and Anil Kumar. 2007. "Employer Matching and 401(k) Saving: Evidence from the Health and Retirement Study." *Journal of Public Economics*, Special Issue published in cooperation with the National Bureau of Economic Research and Uppsala University: Proceedings of the Trans-Atlantic Public Economics Seminar on Public Policy and Retirement Behavior June 12-14 2006, 91 (10): 1920–43.

Eysenck, Michael W., Nazanin Derakshan, Rita Santos, and Manuel G. Calvo. 2007. "Anxiety and Cognitive Performance: Attentional Control Theory." *Emotion* 7 (2): 336–53.

Fernandes, Daniel, John G. Lynch, and Richard G. Netemeyer. 2014. "Financial Literacy, Financial Education, and Downstream Financial Behaviors." *Management Science* 60 (8): 1861–83.

Galiani, Sebastian, Paul Gertler, and Camila Navajas Ahumada. 2020. "Trust and Saving in Financial Institutions." Working Paper 26809. Working Paper Series. National Bureau of Economic Research.

Goldfarb, Avi, and Catherine Tucker. 2019. “Digital Economics.” *Journal of Economic Literature* 57 (1): 3–43.
Goldin, Jacob, Tatiana Homonoff, Richard W Patterson, and William L Skimmyhorn. 2020. “How Much to Save? Decision Costs and Retirement Plan Participation,” NBER Working Paper, , July.
Hartley, Catherine A., and Elizabeth A. Phelps. 2012. “Anxiety and Decision-Making.” *Biological Psychiatry* 72 (2).
Hastings, Justine S., and Jeffrey M. Weinstein. 2008. “Information, School Choice, and Academic Achievement: Evidence from Two Experiments*.” *The Quarterly Journal of Economics* 123 (4): 1373–1414.
Hossain, Tanjim, and John Morgan. 2006. ...“...Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on EBay.” *The B.E. Journal of Economic Analysis & Policy* 6 (2).
Kling, Jeffrey R., Sendhil Mullainathan, Eldar Shafir, Lee C. Vermeulen, and Marian V. Wrobel. 2012. “Comparison Friction: Experimental Evidence from Medicare Drug Plans *.” *The Quarterly Journal of Economics* 127 (1): 199–235.
Koivisto, Jonna, and Juho Hamari. 2019. “The Rise of Motivational Information Systems: A Review of Gamification Research.” *International Journal of Information Management* 45 (April): 191–210.
Loewenstein, George F., Elke U. Weber, Christopher K. Hsee, and Ned Welch. 2001. “Risk as Feelings.” *Psychological Bulletin* 127 (2): 267–86.
Madrian, Brigitte C. 2013. “Matching Contributions and Savings Outcomes: A Behavioral Economics Perspective.” *Matching Contributions for Pensions: A Review of International Experience*, 289–310.
Madrian, Brigitte C., and Dennis F. Shea. 2001. “The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior.” *The Quarterly Journal of Economics* 116 (4): 1149–87.
Mutchler, Jan, Yang Li, and Ping Xu. 2016. “Living Below the Line: Economic Insecurity and Older Americans Insecurity in the States.” *Center for Social and Demographic Research on Aging Publications*. https://scholarworks.umb.edu/demographyofaging/13.
Park, Junchol, Jesse Wood, Corina Bondi, Alberto Del Arco, and Bita Moghaddam. 2016. “Anxiety Evokes Hypofrontality and Disrupts Rule-Relevant Encoding by Dorsomedial Prefrontal Cortex Neurons.” *Journal of Neuroscience* 36 (11): 3322–35.
Seo, Yuri, Xiaozhu Li, Yung Kyun Choi, and Sukki Yoon. 2018. “Narrative Transportation and Paratextual Features of Social Media in Viral Advertising.” *Journal of Advertising* 47 (1): 83–95.
Shue, Kelly, and Erzo F. P. Luttmer. 2009. “Who Misvotes? The Effect of Differential Cognition Costs on Election Outcomes.” *American Economic Journal: Economic Policy* 1 (1): 229–57.
Thaler, Richard H., and Cass R. Sunstein. 2003. “Libertarian Paternalism.” *American Economic Review* 93 (2): 175–79.
———. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Nudge: Improving Decisions about Health, Wealth, and Happiness. New Haven, CT, US: Yale University Press.
Thaler, Richard H., and Shlomo Benartzi. 2004. “Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving.” *Journal of Political Economy* 112 (S1): S164–87.
Wolf, Oliver T. 2009. “Stress and Memory in Humans: Twelve Years of Progress?” *Brain Research, Stress, Coping, and Disease*, 1293: 142–54.

Figure 1.
Schematic Overview of AE 401(k) Plan Enrollment

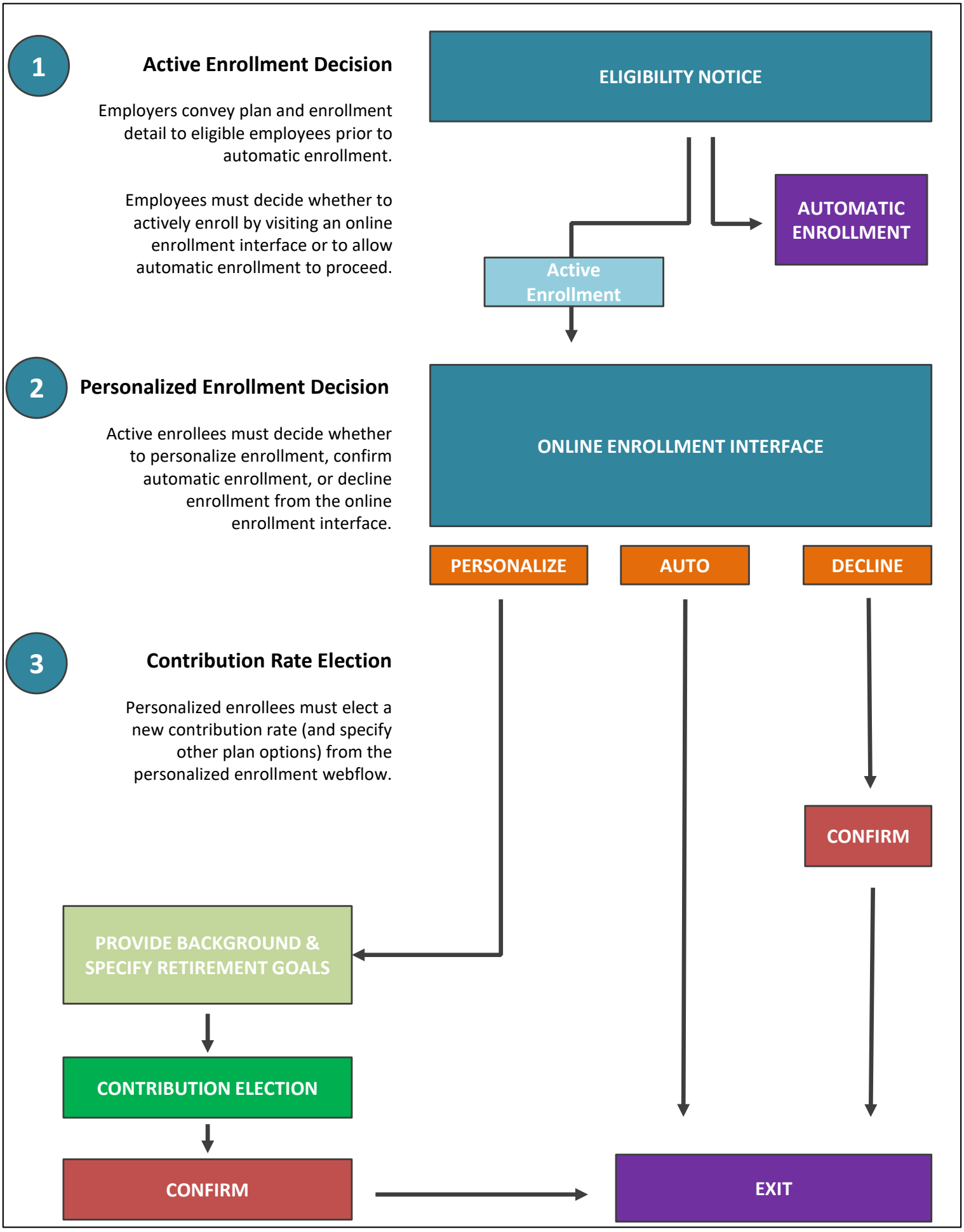
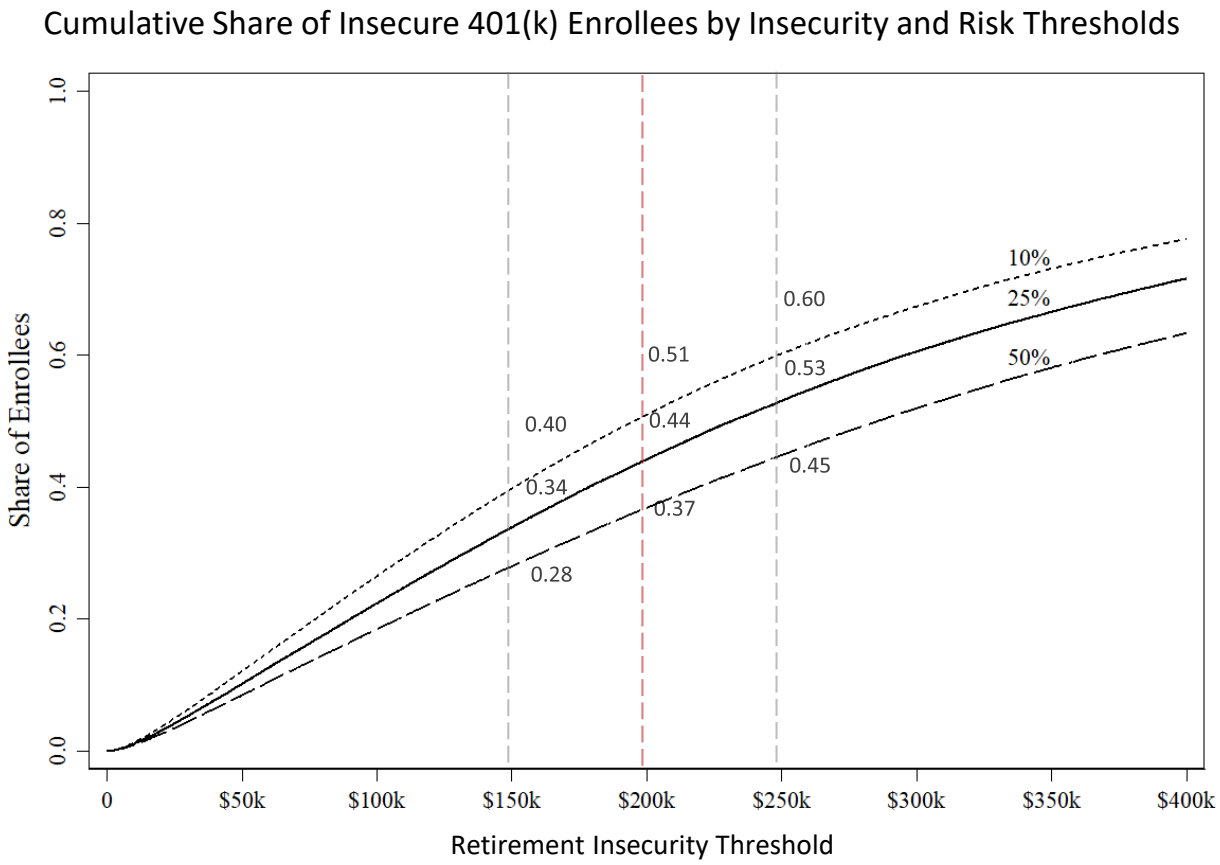


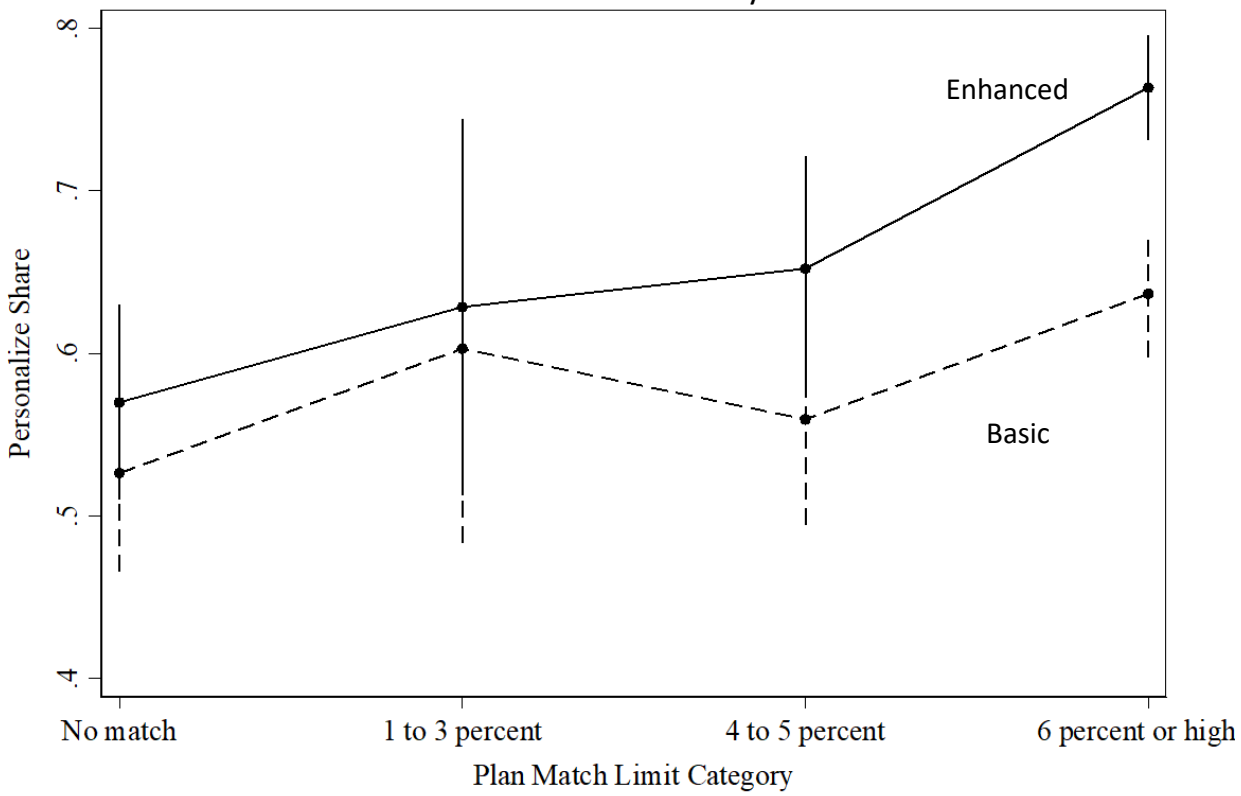
Figure 2.
Projected Risk of Retirement Insecurity for AE 401(k) Enrollees



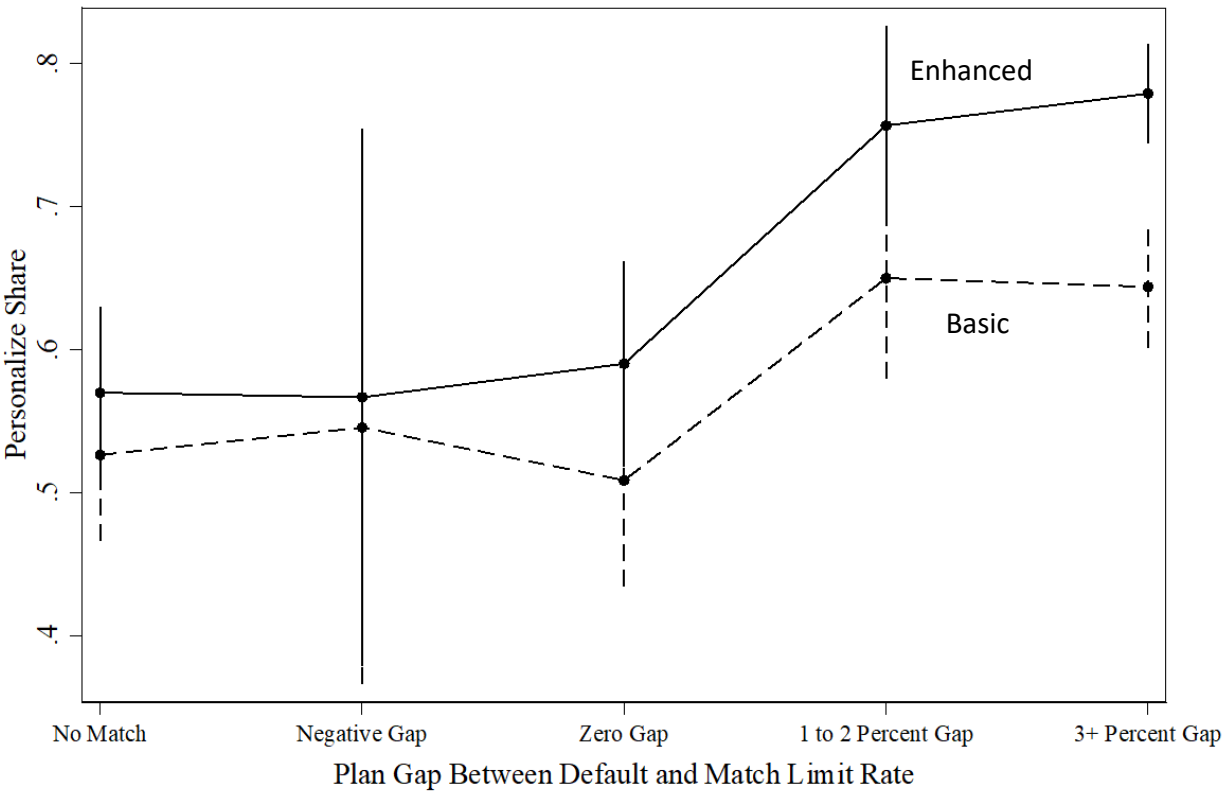
Notes: Figure depicts the cumulative share of enrollees whose projected retirement risk exceeds either 10, 25, or 50 percent for a range of retirement insecurity thresholds. Vertical lines indicate the three insecurity thresholds of central interest in the paper.

Figure 3.
Effect of Enhanced Design on Enrollment by Financial Consequence of Decision

Panel A. Personalized Enrollment by Plan Match Limit



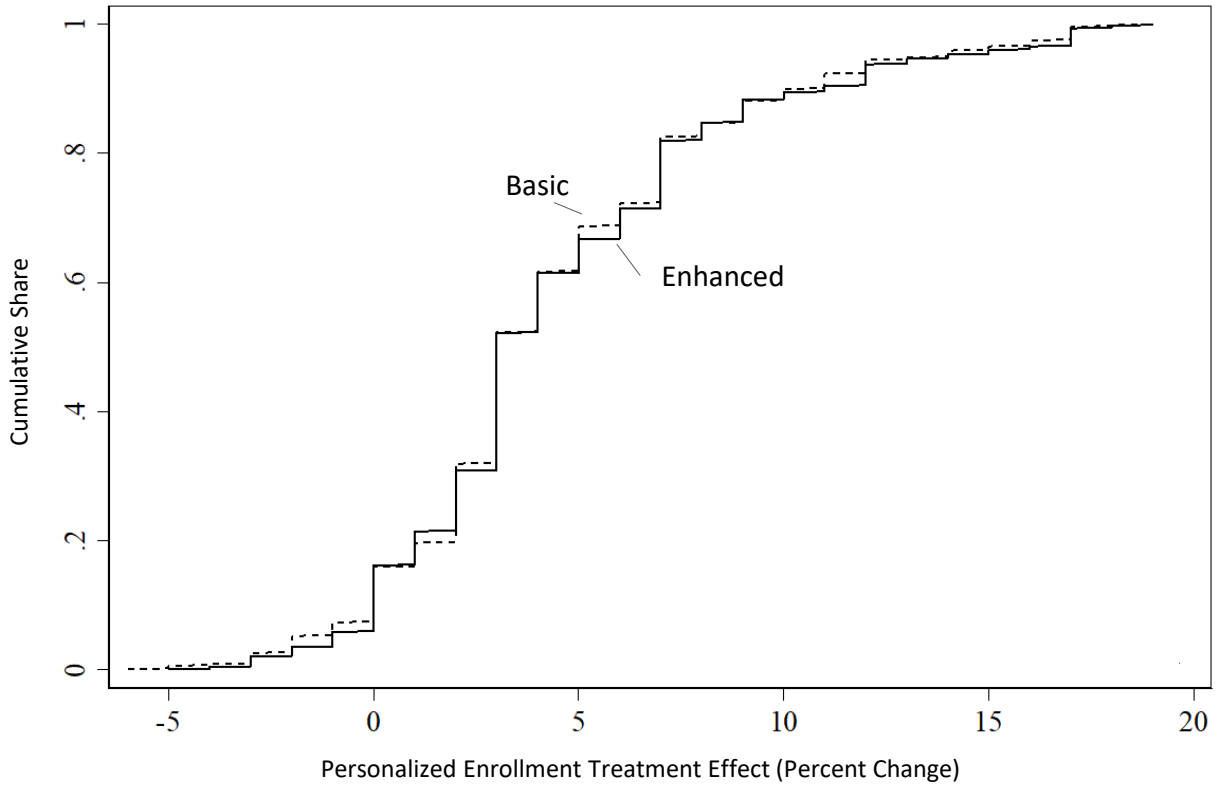
Panel B. Personalized Enrollment by Gap Between Plan Default and Match Limit



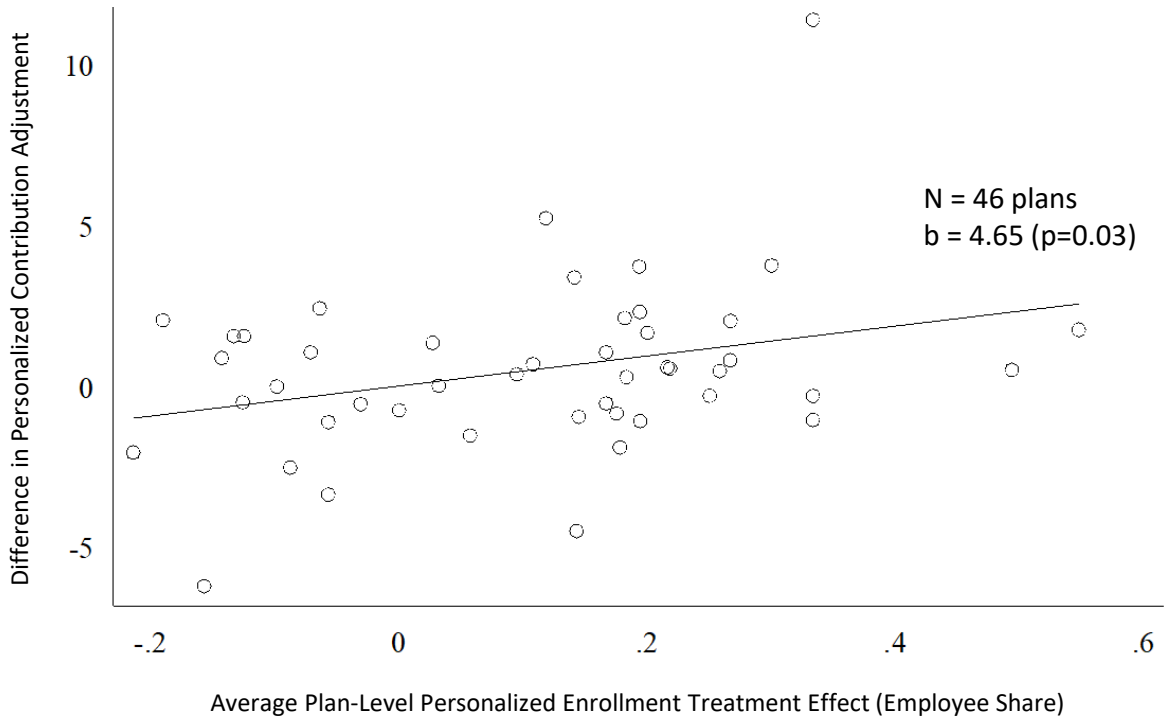
Notes: Panel A plots the share of personalized enrollment for enhanced (solid) and basic (dashed) design conditions in Field Study 1 across employees grouped by increasing plan match threshold. Error bars reflect 95% confidence intervals. Panel B depicts the same outcome for each condition with the x-axis now reflecting employees grouped by increasing size of the gap between the plan default rate and rate associated with full match take-up.

Figure 4.
Equivalence in Saving Between Marginal and Inframarginal Personalized Enrollees

Panel A. Distribution of Personalized Adjustments across Experiment Conditions



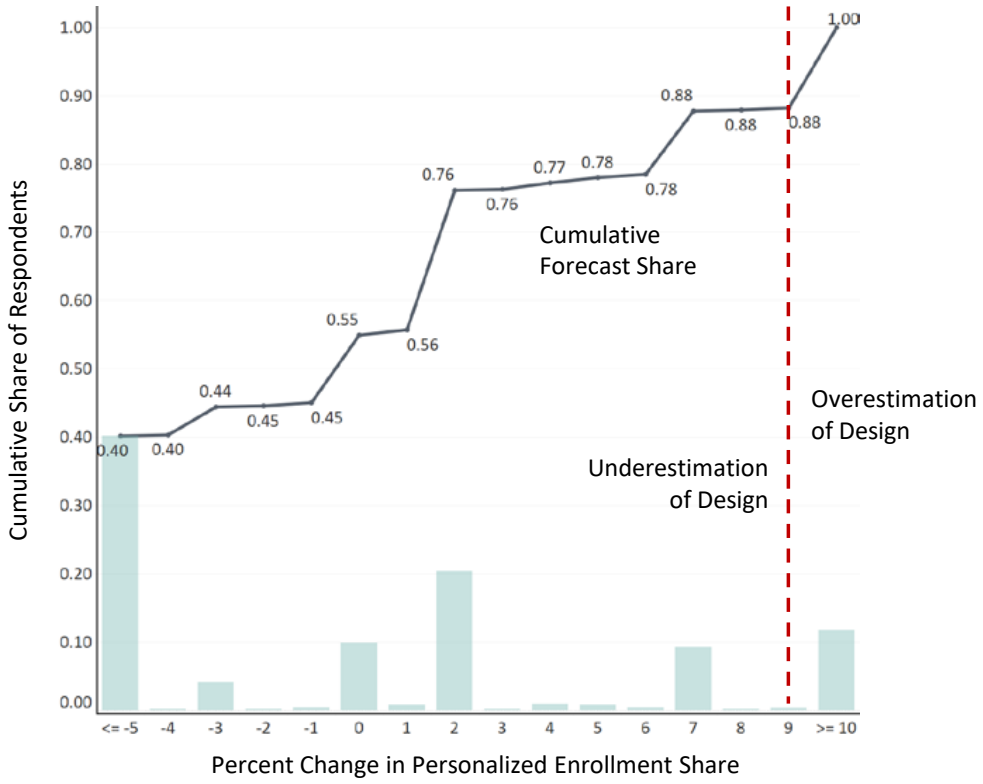
Panel B. Cross-Plan Covariation in Differential Personalized Adjustment across Experiment Conditions



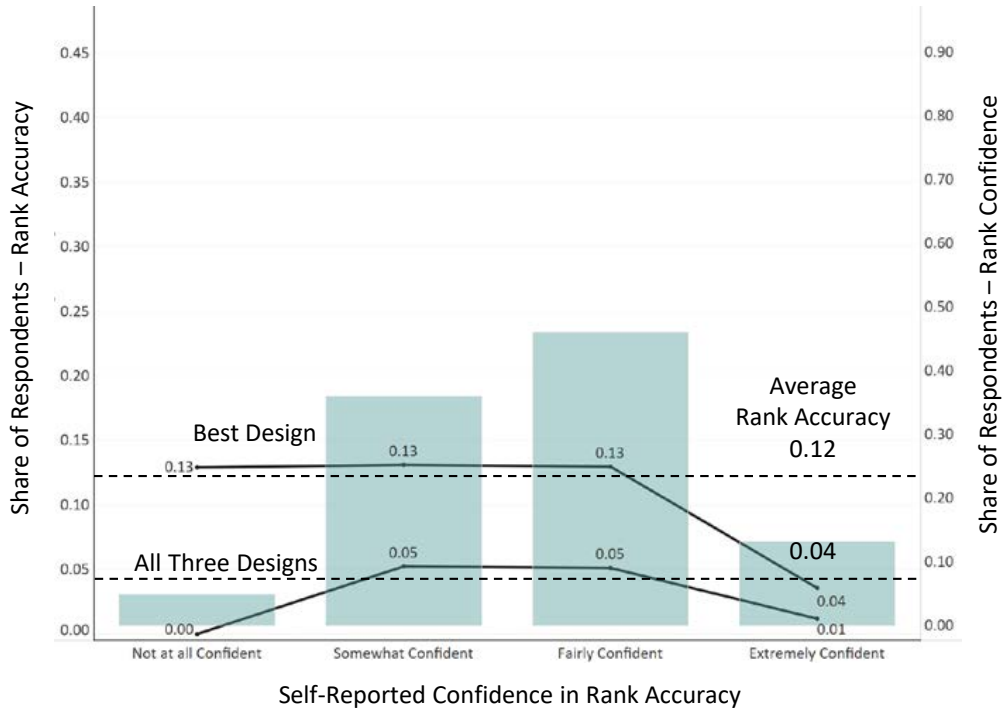
Notes: Panel A plots the cumulative distribution of personalized contribution adjustments, relative to the plan default, across experimental conditions from the first field study. Panel B depicts the covariation between the average plan-level treatment effect of Enhanced Design and the average plan-level difference in personalized adjustments across experimental conditions for the 46 plans from the first field study with at least 5 employees in each condition.

Figure 5.
Forecasts of Design Efficacy – Plan Administrator Survey

Panel A. Forecasted Change to Personalized Enrollment under Enhanced Design



Panel B. Rank Accuracy by Confidence of Plan Administrator



Notes: Panel A displays the cumulative distribution of forecasted change to the rate of personalized enrollment attributable to design, superimposed on a histogram depicting the frequency of each forecast. The vertical red-dashed x-line denotes the effect of design from the field study. Panel B displays the share of respondents accurately ranking design efficacy superimposed on a histogram of self-reported confidence. The plot separately depicts the share of respondents accurately identifying the single-best design and the share accurately rank-ordering all three design elements. The dashed x-lines indicate the average rank accuracy for each measure.

Table 1.
Overview of US Employer-Sponsored Retirement Plan Landscape

	Summary Statistics Across Different Plan Samples					
					Voya Plans	
	All Markets [Census] (2014)[1]			Small-to-Midsize [Survey] (2016)[2]	Small-to-Midsize [Census] (2016)[3]	
	Any DB	Any DC	Any 401(k)	Any 401(k)	Any 401(k)	Auto-Enroll 401(k)
Panel A. Overview						
Plan Size						
# of Plans	44,869	640,334	533,769	4,219	16791	675 [5]
Total Participants (thousands)	37,749	94,685	62,700	--	896	196
Total Plan Assets (millions)	2,985,476	5,321,958	4,399,891	469,000 [4]	49,916	9,081
Plan Share w/ Less than 1,000 Participants	0.92	0.98	0.98	--	0.997	0.956
Plan Type and Enrollment Features						
401(k) - Plan share	--	0.83	1.00	1.00	1.00	1.00
401(k) Plans w/ Auto-Enrollment - Plan share	--	--	--	0.42	1.00	1.00
401(k) Plans w/ Auto-Enrollment - Asset share	--	--	--	0.62 [6]	1.00	1.00
Automatic Escalation - Plan share	--	--	--	0.35	0.33	0.35
Panel B. Default Rates						
Less than 3 Percent	--	--	--	0.12	0.15	0.13
3 Percent	--	--	--	0.41	0.56	0.57
4 Percent	--	--	--	0.12	0.10	0.11
5 Percent	--	--	--	0.11	0.06	0.06
6 Percent	--	--	--	0.16	0.11	0.13
Greater than 6 Percent	--	--	--	0.02	< 0.01	< 0.01
Panel C. Match Generosity						
Any Match (plan share)	--	--	--	0.75	--	--
Average Match Limit	--	--	--	--	0.04	0.04 [7]
Match Limit Less than 3 Percent (plan share)	--	--	--	--	0.03	0.05 [7]
Match Limit of 3 Percent (plan share)	--	--	--	--	0.17	0.04 [7]
Match Limit Greater than 3 Percent (plan share)	--	--	--	--	0.80	0.91 [7]
Panel D. Enrollment						
Participation Rate	--	--	--	--	0.75	0.89
Average Participant Contribution Rate	--	--	--	--	0.05	0.05
Contribution ≥ Default Rate (employee share)	--	--	--	--	--	0.78
Take-up of Full Match (employee share)	--	--	--	--	0.73	0.70

Notes: This table summarizes the distribution of employer-sponsored retirement plans by plan type and size, and then summarizes plan features and average participation levels and contribution rates specifically among 401(k) plans with automatic enrollment. Data for the table was collected from a range of sources including administrative records reported by the Department of Labor, a widely-used industry survey published by PLANSPONSOR, and internal data on a relevant sample of Voya-sponsored plans provided to the authors by Voya.

[1] Source: Private Pension Plan Bulletin Abstract of 2014 Form 5500 Annual Reports Data. Extracted on 6/30/2016 Employee Benefits Security Administration U.S. DOL.

[2] Source: 2016 PLANSPONSOR Defined Contribution Survey.

[3] Source: Voya internal data as of 11/30/2016.

[4] In the absence of precise reporting, the reported estimate is a lower bound on total plan assets was determined by assuming firms have DC assets equal to the lower bound of each reported category

[5] Represents the number of auto enroll plans from the Voya study sub-set that were eligible to participate in the study. Note, not all plans may have had newly eligible employees during the study period.

[6] In the absence of precise reporting, the reported estimate is a lower bound on total plan assets was determined by assuming firms have DC assets equal to the lower bound of each reported category.

[7] Percentages expressed based on a sample of 80 plans where matching contributions were provided by the employer.

Table 2.
Estimated Risk of 401(k) Enrollee Retirement Insecurity by Market Return Assumptions

	Assumed Distribution of Future Market Returns			
Pre-Tax Accumulation Threshold [$ Monthly Income]	Low N(1.6, 10.2)	Moderate N(3.5, 10.2)	High N(5.6, 10.2)	Historic N(6.4, 9.2)
Panel A. Risk of Retirement Insecurity ≥ 25%				
$150,000 [$1,959]	0.42	**0.34**	0.27	0.25
$200,000 [$2,181]	0.54	**0.44**	0.35	0.32
$250,000 [$2,403]	0.65	**0.53**	0.42	0.38
Panel B. Risk of Retirement Insecurity ≥ 10%				
$150,000 [$1,959]	0.50	**0.40**	0.32	0.29
$200,000 [$2,181]	0.62	**0.51**	0.41	0.37
$250,000 [$2,403]	0.71	**0.60**	0.50	0.44

Notes: This table summarizes the share of 401(k) enrollees at risk of retirement insecurity for a range of asset thresholds ($150k-$250k) by row and a range of assumptions regarding market returns (Low, Moderate, High, or Very High). We fix the assumption that "at risk" means 25% or higher chance of falling below the asset threshold at retirement age 65. All estimates in this table reflect our preferred, moderate assumptions regarding market return expectations (mean 3.5, standard deviation 10.2 nominal annual return rate), contribution rate inertia of 75% (fixed annual probability of 25% (100% - inertia) of increasing contribution rate by 1% of income until reaching 10%), real wage growth of 1%, and retirement taxes of 10%. Reported monthly income includes both the annuity and estimated benefits from Social Security.

Table 3.
Summary of Plan Features and Employee Characteristics across Field Studies

	Analytic Samples Across Field Studies			
	All [Dec 2016 to March 2018]	Field Study #1 [July 2017 to March 2018]	Field Study #2 [July 2017 to March 2018]	Field Study #3 [Dec 2016 to July 2017]
Panel A. Plan-Level Summary				
General Data				
Number of Plans	500	397	4	308
Number of Employees	8,565	3,867	1,065	3,633
Employees per Plan (median)	6	5	249	5
Months per Plan (average)	4.4	3.6	8.3	3.4
Plan-Level Median Salary* (average)	67,501	68,632	63,368	64,381
Plan-Level Median Salary* (median)	58,822	60,000	62,650	56,065
Match Observed (share of employees)	0.86	0.81	1.00	0.87
Match Observed (share of plans)	0.62	0.65	1.00	0.71
Distribution of Default Rates (Plan Shares)				
Less than 3 Percent	0.15	0.15	0.00	0.11
3 Percent	0.55	0.56	0.75	0.61
4 Percent	0.11	0.10	0.25	0.12
5 Percent	0.07	0.06	0.00	0.06
6 Percent	0.11	0.12	0.00	0.10
Greater than 6 Percent	<0.01	0.00	0.00	<0.01
Distribution of Match Limits (Plan Shares | Match Observed)				
No Match	0.28	0.29	0.00	0.20
1 to 3 Percent	0.10	0.10	0.00	0.08
4 to 5 Percent	0.20	0.19	0.25	0.27
6 Percent	0.39	0.38	0.75	0.39
Greater than 6 Percent	0.04	0.03	0.00	0.06
Panel B. Employee-Level Summary				
General				
Share of Male Employees	0.56	0.56	0.52	0.57
Age (years)	40	39	39	40
Salary - Average	81,855	80,823	78,294	83,977
Salary - Mean	63,141	61,998	64,386	65,000
Participation and Contribution				
Participation Rate	0.86	0.84	0.92	0.87
Contribution Rate**	5.97	5.74	6.20	6.15
Participant Contribution Rate	6.93	6.85	6.74	7.09
Personalize Contribution Rate	7.92	7.96	7.43	8.05
Estimated Savings Rate (inclusive of Match)	8.47	7.77	9.92	8.66
Estimated Participant Savings Rate (inclusive of Match)	9.80	9.31	10.79	9.89
Estimated Personalize Savings Rate (inclusive of Match)	11.05	10.64	11.83	11.11
Match Take-Up	0.64	0.63	0.66	0.65

Notes: This table summarizes demographic, enrollment, and savings detail at the plan- and participant-level for the three field studies. This sample includes all plans for which we observed at least one employee in two different experimental conditions during the study period. Panel A summarizes unweighted plan-level characteristics. Panel B summarizes characteristics and behavior of potential enrollees who visited the online enrollment site during the study periods. Figures regarding match take-up and savings are restricted to the sample for which the match is observed and full take-up is associated with a contribution rate threshold. * Participant salary is censored at $500k and comes from three sources prioritized as follows: annualized gross compensation from administrative data, self-reported annual salary, or predicted salary based on a linear model predicting annualized gross compensation as a function of observable characteristics. ** Contribution rates are censored at 20 percent.

Table 4.
Summary of Enrollment Outcomes by Field Study and Experimental Condition

	Enrollment Choice [1,0]			Contribution Threshold [1,0]		Contribution Rate [Percent]		Saving Rate [Percent]	
	Personalize	Auto	Decline	> Default Rate	≥ Match	All	Personalized	All	Personalized
Panel A. Field Study #1									
Basic [N = 1276]	0.60	0.22	0.18	0.50	0.58	5.41	7.80	7.28	10.39
Enhanced Info (Default) [N = 1254]	0.61	0.23	0.16	0.52	0.62	5.76	8.15	7.76	10.77
Enhanced Design [N = 1337]	0.70	0.16	0.15	0.58	0.67	6.04	7.92	8.25	10.74
All Experimental Designs (Basic excluded)	0.66	0.19	0.15	0.55	0.65	5.91	8.02	8.02	10.76
# Employees	3,867	3,867	3,867	3,859	2,341	3,859	2,452	3,071	1,963
# Plans	397	397	397	397	183	397	297	263	254
Panel B. Field Study #2									
Enhanced Info (Default/Match) [N = 547]	0.72	0.19	0.08	0.68	0.63	6.09	7.58	9.74	12.05
Enhanced Design (Default/Match) [N = 518]	0.82	0.10	0.08	0.75	0.70	6.31	7.29	10.12	11.64
# Employees	1,065	1,065	1,065	1,064	1,064	1,064	822	1,064	822
# Plans	4	4	4	4	4	4	4	4	4
Panel C. Field Study #3									
Basic [N = 529]	0.65	0.23	0.12	0.55	0.60	5.96	8.03	8.20	10.97
Standardized [N = 1015]	0.69	0.17	0.14	0.59	0.65	6.14	8.13	8.72	11.25
Headlines (Standardized + Headlines) [N = 1019]	0.69	0.17	0.13	0.61	0.67	6.36	8.33	8.89	11.46
Enhanced Presentation (Standardized + Headlines + Color) [N = 1070]	0.72	0.15	0.13	0.59	0.67	6.06	7.73	8.61	10.73
All Design Variations (BASIC excluded)	0.70	0.17	0.13	0.60	0.66	6.18	8.05	8.74	11.13
All Headlines (BASIC and STANDARDIZED excluded)	0.71	0.16	0.13	0.60	0.67	6.20	8.01	8.75	11.08
# Employees	3,633	3,633	3,633	3,633	2,746	3,633	2,515	3,212	2,255
# Plans	308	308	308	308	161	308	226	221	209

Notes: This table summarizes average enrollment and contribution outcomes in the three field studies. Summary statistics are reported by experimental treatment for Field Study #1 in Panel A, Field Study #2 in Panel B, and Field Study #3 in Panel C. Figures regarding match take-up and savings are restricted to the sample for which match details are observed and, in the case of match take-up, to cases where the observed match is nonzero. "All" refers to potential enrollees who visit the online enrollment site during their automatic enrollment period, "Auto" refers to participants electing to confirm their automatic enrollment, and "Personalize" refers to participants who enroll after modifying their plan options. "> Default Rate" refers to participants who chose initial contribution rates higher than their plan's default contribution rate associated with automatic enrollment and "≥ Match" refers to participants with contribution rates corresponding to full match take-up. Data on elections is based on administrative recordkeeping data on automatic enrollment elections.

Table 5.
Marginal and Total Treatment Effects by Field Study

	Enrollment Choice [1,0]			Contribution Threshold [1,0]		Contribution Rate [Percent]		Saving Rate [Percent]	
	Personalize	Auto	Decline	> Default Rate	≥ Match	All	Personalized	All	Personalized
Panel A. Field Study #1									
Marginal Effects									
Enhanced Information (relative to Basic)	0.02	0.01	-0.03	0.03	0.04*	0.33	0.34	0.39	0.33
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.20)	(0.26)	(0.26)	(0.31)
Enhanced Presentation (relative to Enhanced Information)	**0.07*** **	**-0.07*** **	-0.00	**0.06*** **	**0.06** **	0.29	-0.10	**0.59** **	0.13
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.19)	(0.23)	(0.23)	(0.27)
Total Effect									
Enhanced Design (relative to Basic)	**0.09*** **	**-0.07*** **	-0.03*	**0.08*** **	**0.10*** **	**0.62*** **	0.24	**0.98*** **	0.46
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.22)	(0.25)	(0.26)	(0.29)
Panel B. Field Study #2									
Marginal Effect									
Enhanced Presentation	**0.10** **	**-0.09** **	-0.01	0.07*	0.07*	0.26	-0.27	0.43*	-0.31
(relative to Enhanced Information w/ Default & Match)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.11)	(0.26)	(0.18)	(0.25)
Panel C. Field Study #3									
Marginal Effects									
Standardized (relative to Basic)	0.04	**-0.06*** **	0.02	0.04	0.04	0.17	0.08	0.35	0.11
	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.35)	(0.40)	(0.42)	(0.44)
Headlines (relative to Standardized)	0.00	0.00	0.00	0.01	0.02	0.15	0.12	0.19	0.26
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.22)	(0.25)	(0.26)	(0.26)
Color (Enhanced Presentation relative to Headlines)	0.03	-0.02	-0.00	0.00	0.01	-0.20	**-0.51** **	-0.25	**-0.71** **
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.18)	(0.25)	(0.21)	(0.29)
Total Effects									
Headlines + Color (Enhanced Presentation relative to Standardized)	0.03	-0.02	-0.01	0.01	0.02	-0.05	-0.39	-0.06	-0.45
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.23)	(0.27)	(0.27)	(0.31)
Headlines + Standardized (Headlines relative to Basic)	0.04*	**-0.06*** **	0.02	0.05*	**0.06** **	0.32	0.20	0.54	0.37
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.29)	(0.36)	(0.33)	(0.40)
Enhanced Presentation (relative to BASIC)	**0.07*** **	**-0.08*** **	0.01	0.05*	**0.06** **	0.12	-0.31	0.29	-0.34
	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.30)	(0.35)	(0.35)	(0.39)

Notes: This table summarizes marginal and total treatment effects of psychological design on a series of enrollment outcomes from samples of employees in Field Study #1 (Panel A), Field Study #2 (Panel B), and Field Study #3 (Panel C). Each cell in a panel corresponds to an OLS/LPM regression estimate of the marginal effect specified by the panel and row on the outcome indicated by the column heading, after controlling for plan fixed effects and day-of-week and week-of-year variation in the timing of enrollment. Robust standard errors clustered at the plan-level are displayed in parentheses. Asterisks indicate p-values associated with tests of statistical significance: (*) $p < 0.10$, (**) $p < 0.05$, and (***) $p < 0.01$.

Table 6.
Consolidated Effect of Enhanced Design on Initial Enrollment and Saving across Field Studies

	Enrollment Choice [1,0]			Contribution Threshold [1,0]		Contribution Rate [Percent]		Saving Rate [Percent]	
Treatment Category	Personalize	Auto	Decline	> Default Rate	≥ Match	All	Personalized	All	Personalized
Panel A. Enhanced Design									
Default Information Only [Field Study #1]	**0.09*** **	**-0.07*** **	-0.03*	**0.08*** **	**0.10*** **	**0.62*** **	0.24	**0.98*** **	0.46
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.22)	(0.25)	(0.26)	(0.29)
Panel B. Enhanced Presentation									
Default Information Only [Field Study #1]	**0.07*** **	**-0.07*** **	-0.00	**0.06*** **	**0.06** **	0.29	-0.10	**0.59** **	0.13
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.19)	(0.23)	(0.23)	(0.27)
Default + Match Information [Field Study #2]	**0.10** **	**-0.09** **	-0.01	0.07*	0.07*	0.26	-0.27	0.43*	-0.31
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.11)	(0.26)	(0.18)	(0.25)
No Plan Information [Field Study #3]	**0.07*** **	**-0.08*** **	0.01	0.05*	**0.06** **	0.12	-0.31	0.29	-0.34
	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.30)	(0.35)	(0.35)	(0.39)
Panel C. Enhanced Information									
Default Information Only [Field Study #1]	0.02	0.01	-0.03	0.03	0.04*	0.33	0.34	0.39	0.33
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.20)	(0.26)	(0.26)	(0.31)
Panel D. Reference Estimates									
Increase in Plan Match Threshold of 1 Percent Point	**0.022*** **	-0.002	**-0.020*** **	**0.025*** **	**-0.063*** **	**0.139*** **	-0.044	**0.589*** **	**0.480*** **
	(0.005)	(0.003)	(0.004)	(0.005)	(0.011)	(0.040)	(0.039)	(0.065)	(0.057)
Increase in Employee Age of 1 Year	**0.002*** **	**-0.002*** **	-0.001	**0.003*** **	**0.003*** **	**0.056*** **	**0.068*** **	**0.059*** **	**0.067*** **
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.007)	(0.006)	(0.009)	(0.008)
Basic Average [Field Study #1]	0.60	0.22	0.18	0.50	0.58	5.41	7.80	7.28	10.39

Notes: This table summarizes treatment effects of psychological design on a series of enrollment outcomes estimated from separate samples from each Field Study. Panel A summarizes the treatment effects associated with enhanced design (i.e., the joint presence of enhanced presentation and enhanced information). Panel B summarizes the effect of enhanced presentation across interfaces varying in their information display. Panel C summarizes the effect of displaying the plan default on the interface. Each cell in the first three panels corresponds to an OLS or LPM regression estimate of the marginal effect specified by the panel and row on the outcome indicated by the column heading, after controlling for plan fixed effects and day-of-week and week-of-year variation in the timing of enrollment. To benchmark the treatment effects, Panel D summarizes the match and age elasticity estimated from a linear model of enrollment decisions on match limit and age during the field studies. The reference elasticity estimates include flexible controls for plan features (plan's default rate and presence of auto-escalation), gender, and day-of-week and week-of-year variation in the timing of enrollment, as well as the experimental version of the landing page an employee viewed. Please refer to text for additional details on the underlying specifications. Robust standard errors, displayed in parentheses, are clustered at the plan-level for all estimates except FS2 which are unclustered due to small number of plans. Asterisks indicate p-values associated with tests of statistical significance: (*) $p < 0.10$, (**) $p < 0.05$, and (***) $p < 0.01$.

Table 7.
Marginal and Inframarginal Equivalence in Savings - Active Enrollment

		Rate of Residual Active Enrollment		
	All Sample	Low Tercile	Med Tercile	High Tercile
Panel A. Plan Characteristics				
Number of Plans	72	24	24	24
Observations per Plan	210	227	227	174
Online Visit Rate	0.20	0.11	0.18	0.31
Plan Default Rate	3.40	3.40	3.40	3.50
Plan Match (1,0)	0.82	0.76	0.80	0.90
Panel B. Enrollment Outcomes				
Personalized Enrollment	0.61	0.60	0.59	0.64
Declined Enrollment	0.16	0.15	0.18	0.15
Contribution Rate Adjustment	2.36	2.43	2.07	2.58
Contribution Rate	5.76	5.85	5.48	5.96
Full Match Take-Up | Any Plan Match	0.62	0.54	0.59	0.73

Notes: This table summarizes plan and employee characteristics and enrollment outcomes by plan terciles ordered by increasing active enrollment rate for a sample of 72 small-to-midsize plans for which we have at least 50 online and offline enrollment observations from 2016 to 2018. Panel A summarizes plan-level features and average employee salary across plan terciles organized by increasing rate of active enrollment. Panel B summarizes enrollment outcomes of share personalizing and declining enrollment, and contribution rate across the same terciles, as well as savings rates and share fully taking up the match among those plans with a match. This panel also describes each enrollment outcome after adjusting for covariates of employee salary category, age in 10-year bins, and gender through linear regressions.

Table 8.
Clarifying Decision Mechanisms through Hypothetical Enrollment Experiment

		Potential Decision-Frictions						
		Costs and Benefits associated with Personalized Enrollment			Knowledge of Plan Detail		Confusion	Institutional Trust
	Marginal Effect Pr (Personalize = 1)	Effort and Time Costs [1-7]	PE Importance [1-7]	Min Contrib to Secure Retirement [%]	Match Recall [1,0]	Default Recall [1,0]	Decision Clarity [1-7]	Trust in Plan [1-7]
Panel A. Enhanced Design								
[Relative to Baseline]								
Standardization + Headlines + Default	**0.08****	**-0.29****	0.08	-0.59	-0.00	**0.14*****	0.00	0.07
	(0.03)	(0.12)	(0.07)	(0.56)	(0.03)	(0.03)	(0.07)	(0.09)
Standardization + Headlines + Default/Match	**0.13*****	-0.07	0.10	-0.65	0.03	**0.07****	0.10	0.17*
	(0.03)	(0.11)	(0.07)	(0.64)	(0.03)	(0.03)	(0.07)	(0.09)
Standardization + Color + Default/Match	0.07*	-0.23*	0.02	-0.40	**0.09*****	**0.11*****	-0.05	0.01
	(0.03)	(0.12)	(0.07)	(0.55)	(0.03)	(0.03)	(0.07)	(0.09)
Panel B. Enhanced Presentation								
[Relative to Baseline]								
Standardized	0.03	0.01	-0.00	0.51	-0.04	0.00	0.03	0.06
	(0.03)	(0.09)	(0.06)	(0.59)	(0.03)	(0.03)	(0.06)	(0.07)
Standardized + Headlines	**0.12*****	-0.11	**0.17*****	0.49	0.02	0.04	0.06	0.08
	(0.04)	(0.12)	(0.07)	(0.75)	(0.03)	(0.04)	(0.07)	(0.09)
Standardized + Color	0.05	-0.02	0.07	-0.62	-0.04	-0.01	0.03	0.08
	(0.04)	(0.12)	(0.07)	(0.56)	(0.03)	(0.04)	(0.07)	(0.09)
Panel C. Enhanced Information								
[Relative to no Information Display]								
Default Only	0.01	-0.06	0.05	**-1.09*****	0.01	**0.11*****	0.03	0.03
	(0.02)	(0.07)	(0.04)	(0.37)	(0.02)	(0.02)	(0.04)	(0.05)
Default & Match	**0.05****	-0.06	0.01	-0.47	**0.07*****	**0.10*****	-0.03	0.00
	(0.02)	(0.06)	(0.03)	(0.33)	(0.01)	(0.02)	(0.03)	(0.04)
Baseline Mean	0.50	4.17	6.18	8.34	0.70	0.55	3.81	5.20

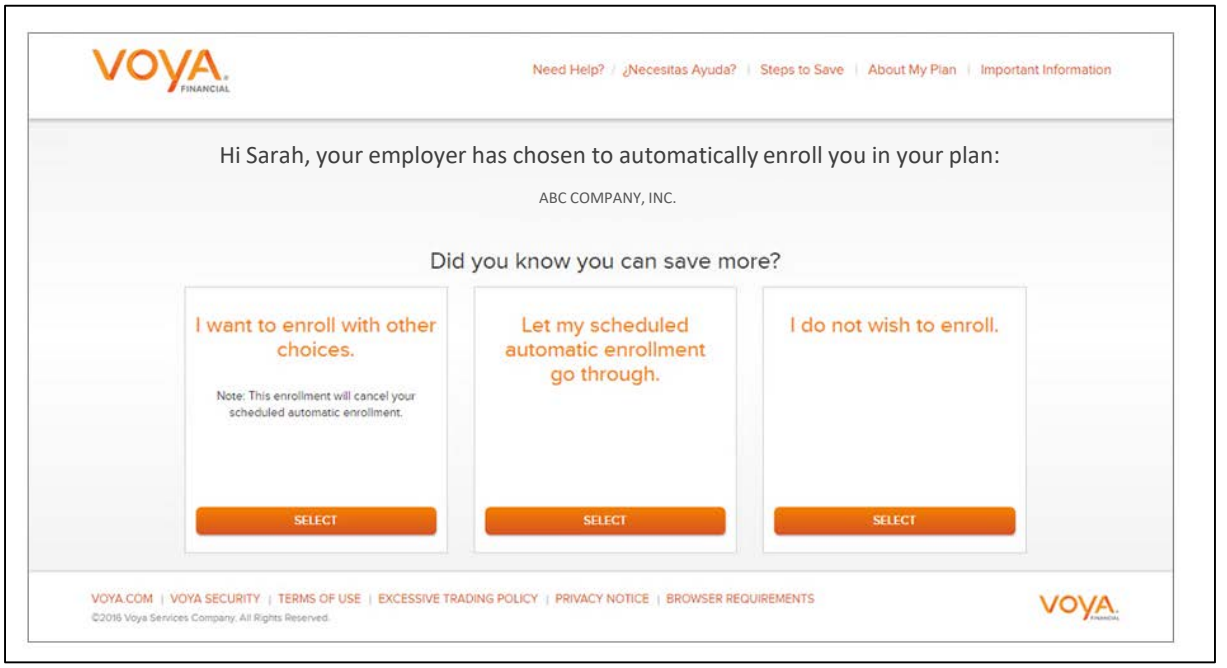
Notes: This table summarizes evidence from survey responses designed to capture potential mechanisms driving differences in experimental responses in the lab study. The point estimates are generated from OLS regressions in which the survey measure indicated in each column is regressed on indicators for different treatment variations in the online landing page described in each row. Robust standard errors are displayed in parentheses. Asterisks indicate p-values associated with tests of statistical significance: (*) $p < 0.10$, (**) $p < 0.05$, and (***) $p < 0.01$.

Table 9.
Projected Years of Retirement Security by Mode of Initial Enrollment

Initial Enrollment Parameters			Assumed Distribution of Future Market Returns			
Savings Window	Mode	Contribution [%]	Low μ =1.6	Moderate μ =3.5	High μ =5.6	Historic μ =6.4
10 years	Automatic	3.4	3.6	4.1	4.6	4.6
(start age of 55 years)	Personalized	7.8	6.4	7.5	8.1	8.8
			+79%	+84%	+77%	+90%
15 years	Automatic	3.4	6.9	7.5	9.4	10.0
(start age of 50 years)	Personalized	7.8	11.4	14.2	17.3	18.9
			+64%	+89%	+84%	+89%
20 years	Automatic	3.4	11.4	14.2	18.1	19.8
(start age of 45 years)	Personalized	7.8	18.1	24.5	≥ 30.0	≥ 30.0
			+59%	+73%	+66%	+52%
25 years	Automatic	3.4	17.3	23.5	≥ 30.0	≥ 30.0
(start age of 40 years)	Personalized	7.8	27.6	≥ 30.0	≥ 30.0	≥ 30.0
			+60%	+28%	--	--

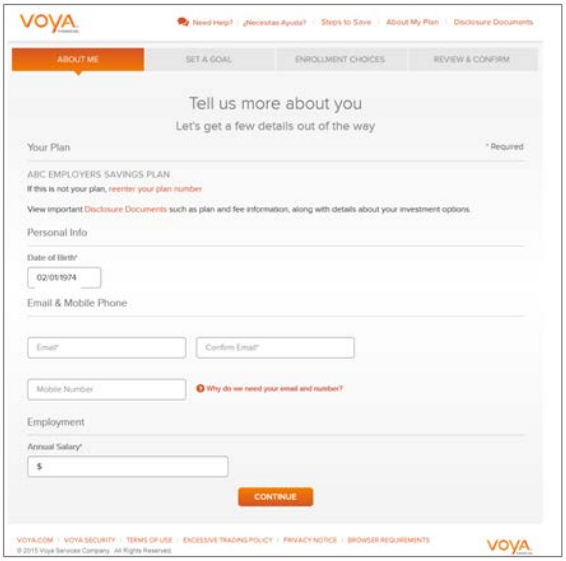
Notes: This table summarizes the estimated years of retirement security for employee archetypes defined by mode of initial enrollment, the duration of an employee's accumulation period, and the assumed average annual market return. Security estimates reflect the length of an annuity that one could purchase using accumulated post-tax savings that would provide $864 of monthly income (i.e., the minimal level of income required for retirement security after accounting for Social Security benefits) assuming a 3 percent real rate of return during retirement. Accumulation calculations assume deterministic market returns, a 0.25 annual contribution rate increase up to a cap of 10 percent, a $50,000 annual salary with a 1 percent real growth rate, a plan match of 50 percent up to a 6 percent limit, a retirement age of 65, and are rounded to the nearest $5k. Annuity calculations drawn from an online calculator accessed in August 2021: https://www.bankrate.com/calculators/investing/annuity-calculator.aspx.

Appendix Figure A1.
Online Enrollment Interface

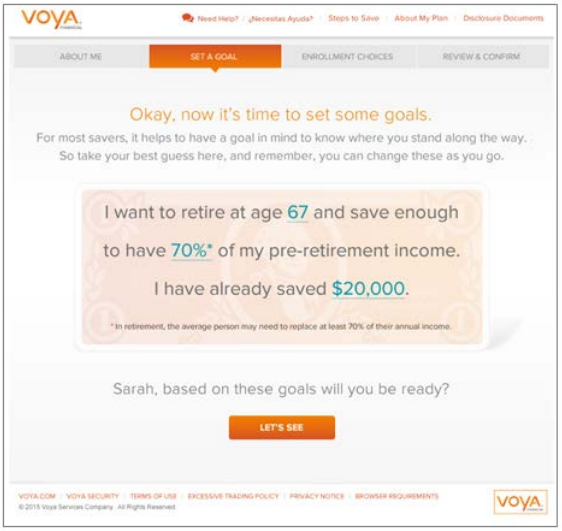


Appendix Figure A2.
Personalized Enrollment Webflow

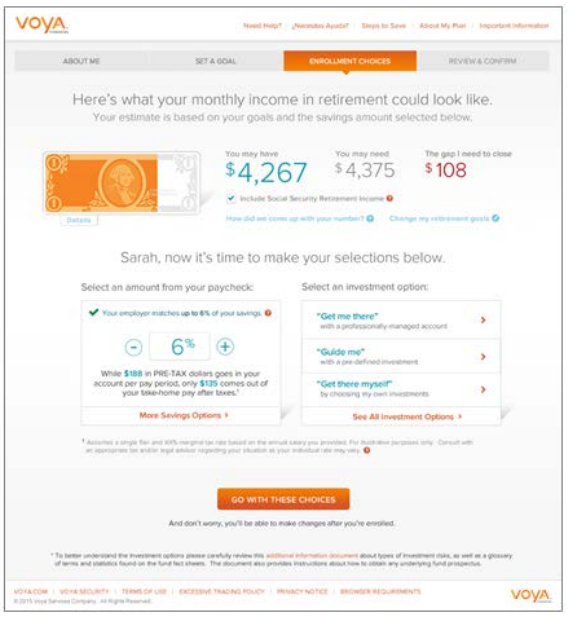
Panel A. Background Information



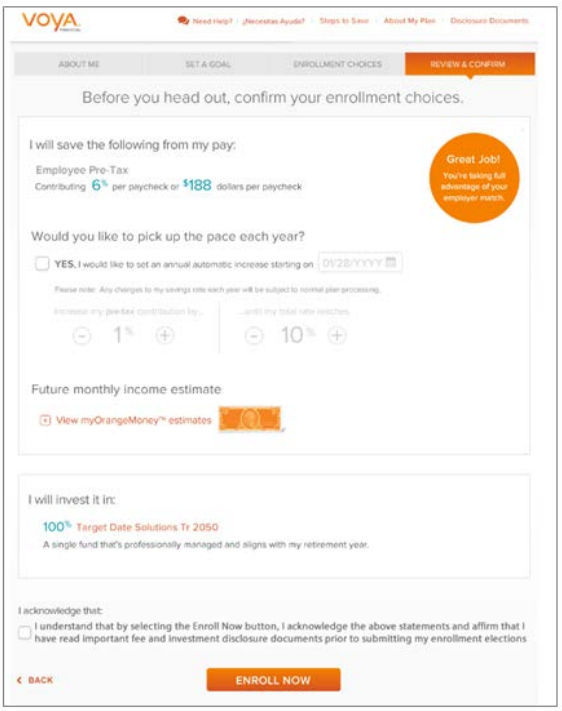
Panel B. Retirement Goals



Panel C. Retirement Calculator



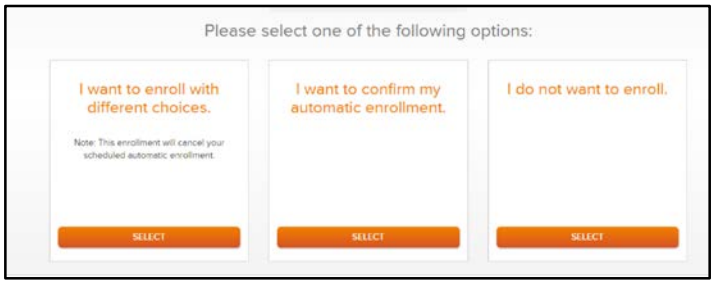
Panel D. Enrollment Confirmation



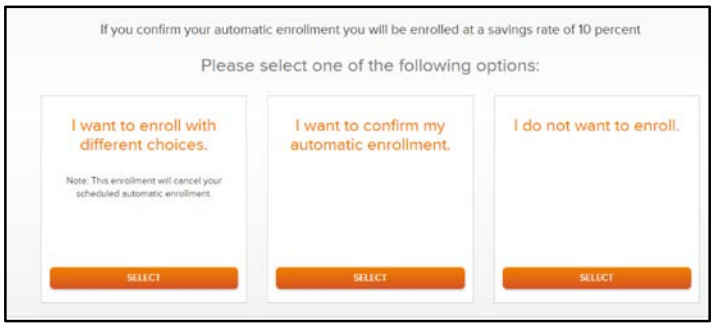
Appendix Figure A3.
Field Studies #1 and #2 – Experimental Conditions

Panel A. Snapshot of Experimental Conditions from Field Study #1

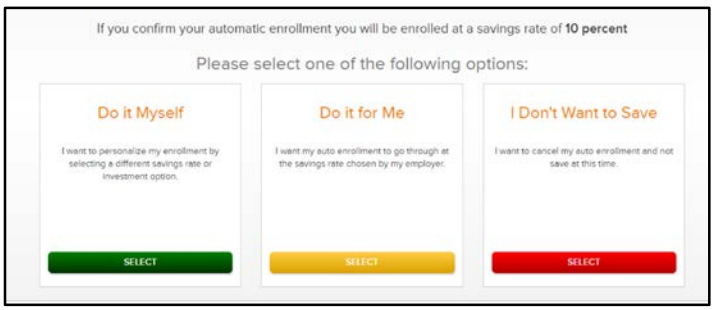
1. Basic – Original presentation of enrollment options



2. Enhanced Information – Basic condition modified to display plan default

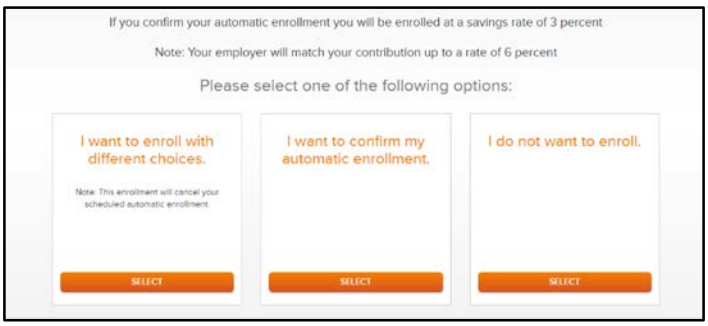


3. Enhanced Design – Enhanced Presentation condition (color + headlines + standardized), modified to display plan default

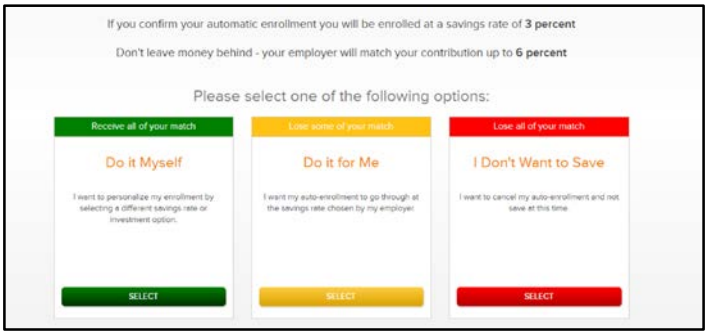


Panel B. Snapshot of Experimental Conditions from Field Study #2

1. Enhanced Information – Presentation of enrollment options originally used by studied plans, modified to display plan default and match threshold

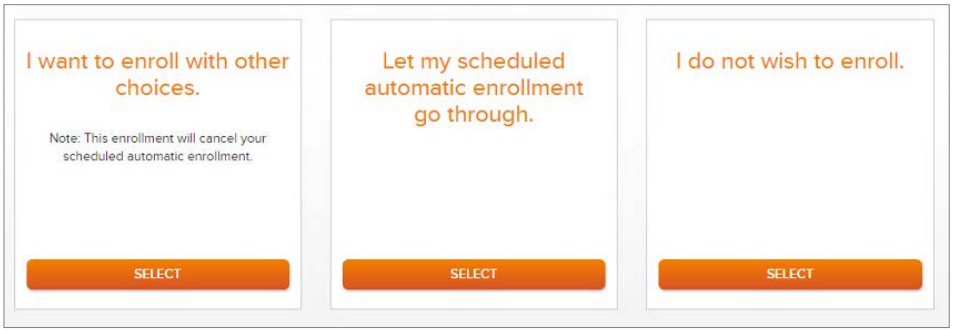


2. Enhanced Design – Enhanced Presentation condition (color + headlines + standardized), modified to display plan default, match threshold, and guidance linking PERSONALIZE to full take-up of match



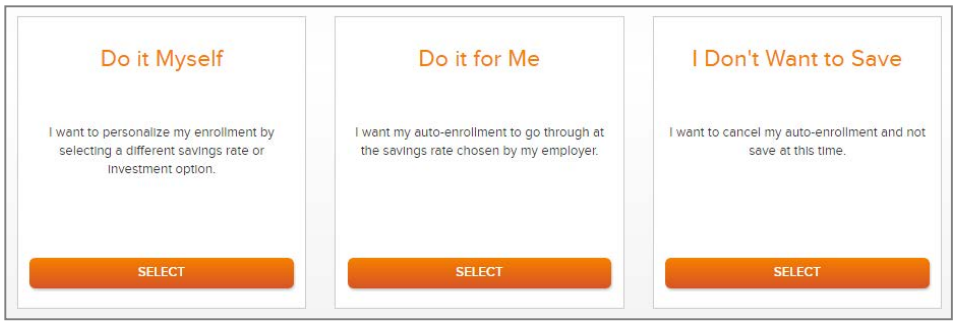
Appendix Figure A4.
Field Study #3 – Experimental Conditions

1. **Basic** – Original presentation of enrollment options [PERSONALIZE, AUTO, DECLINE]

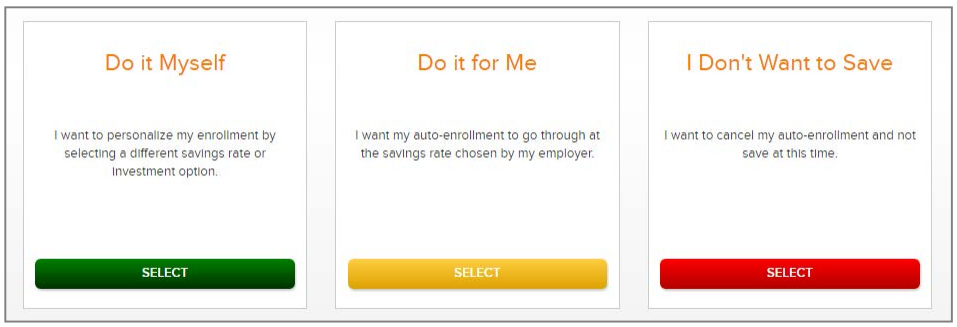


2. **Standardized** – Basic condition modified to standardize descriptions and remove warning message

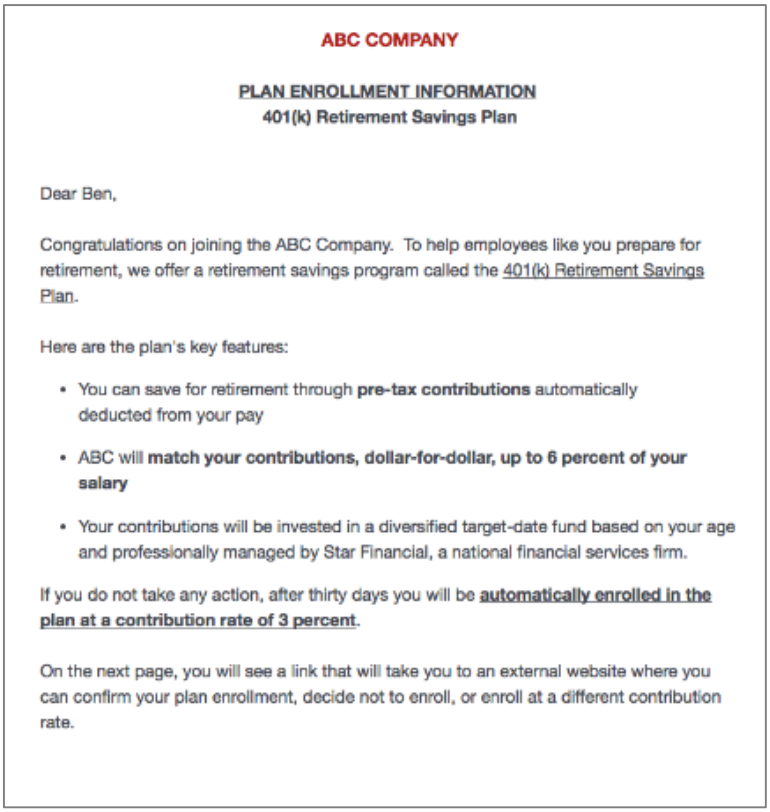
3. **Headlines + Standardized** – Standardized condition modified to include simplified headlines that emphasize decision-making autonomy



4. **Enhanced Presentation** – Headlines + Standardized condition modified to include traffic light colors that encourage active plan participation [Green (PERSONALIZE), Yellow (AUTO), Red (DECLINE)]



Appendix Figure A5.
Hypothetical Enrollment Experiment –Plan Eligibility Notice



ABC COMPANY

PLAN ENROLLMENT INFORMATION
401(k) Retirement Savings Plan

Dear Ben,

Congratulations on joining the ABC Company. To help employees like you prepare for retirement, we offer a retirement savings program called the 401(k) Retirement Savings Plan.

Here are the plan's key features:

- You can save for retirement through **pre-tax contributions** automatically deducted from your pay
- ABC will **match your contributions, dollar-for-dollar, up to 6 percent of your salary**
- Your contributions will be invested in a diversified target-date fund based on your age and professionally managed by Star Financial, a national financial services firm.

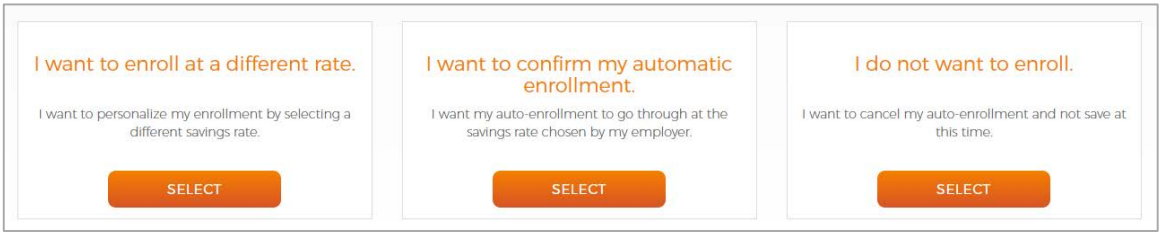
If you do not take any action, after thirty days you will be **automatically enrolled in the plan at a contribution rate of 3 percent**.

On the next page, you will see a link that will take you to an external website where you can confirm your plan enrollment, decide not to enroll, or enroll at a different contribution rate.

Content of plan eligibility notice varies to reflect whether hypothetical 401(k) plan matches annual contributions up to 3 or 6 percent (6 percent version displayed here).

Appendix Figure A6.
Hypothetical Enrollment Experiment – Variation in Presentational Design

1. Baseline – Baseline presentation of three enrollment options [PERSONALIZE, AUTO, DECLINE]



2. Language – Identical to <u>baseline</u> condition but for simplified headlines that emphasize decision autonomy

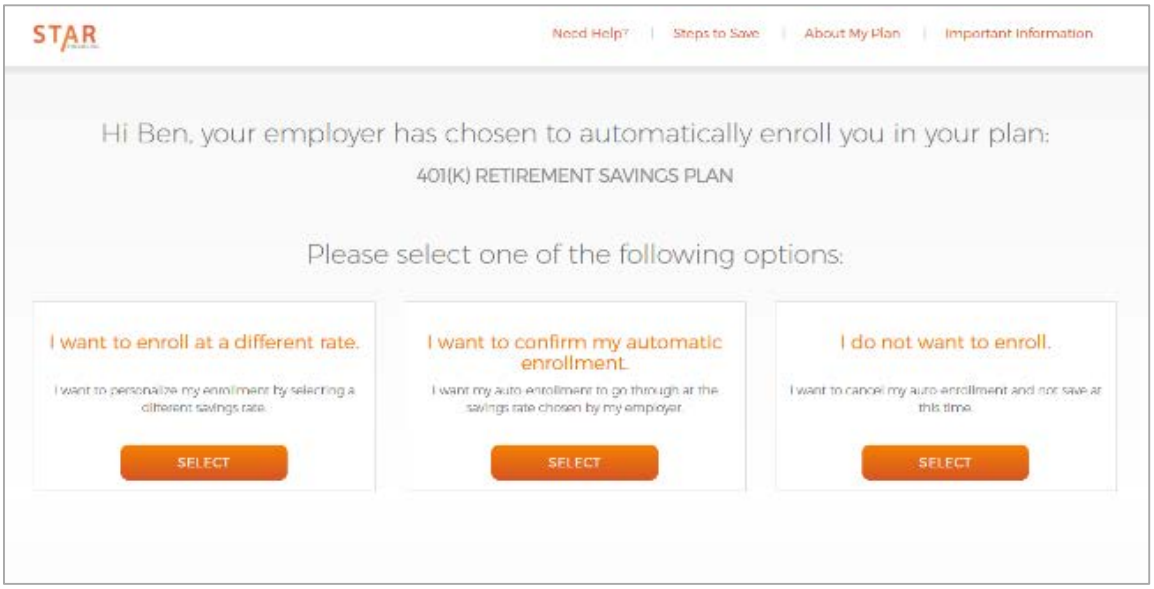
3. Order – Identical to <u>baseline</u> condition but for reordering of options [AUTO, PERSONALIZE, DECLINE]

4. Color (Green AUTO) – Identical to <u>baseline</u> condition but for the association of colors to options [Yellow (PERSONALIZE), Green (AUTO), Red (DECLINE)]

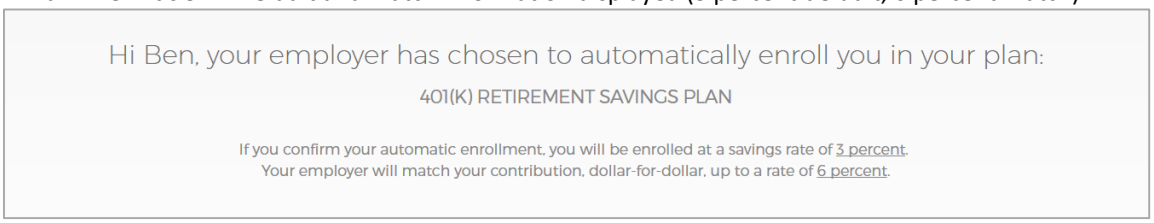
5. Color (Green PERSONALIZE) – Identical to <u>baseline</u> condition but for the association of colors to options [Green (PERSONALIZE), Yellow (AUTO), Red (DECLINE)]

Appendix Figure A7.
Hypothetical Enrollment Experiment – Variation in Plan Information

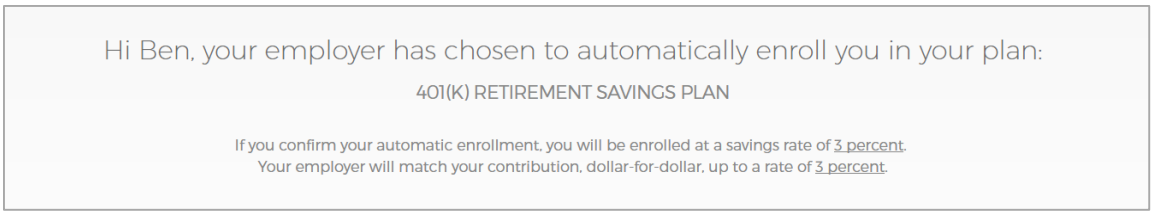
1. No Information- Default and match information not displayed (3 percent default, 6 percent match)



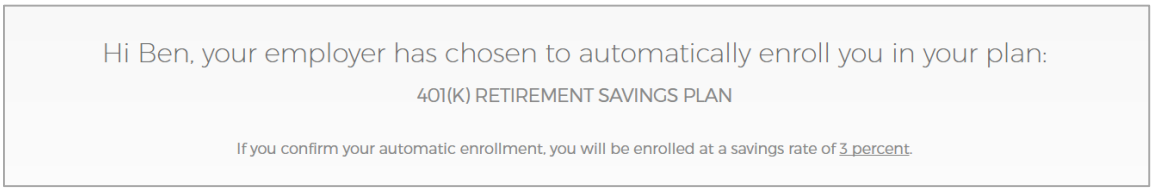
2. Full Information – Default and match information displayed (3 percent default, 6 percent match)



3. Full Information with 3 Percent Match - Default and match information displayed
(3 percent default, 3 percent match)

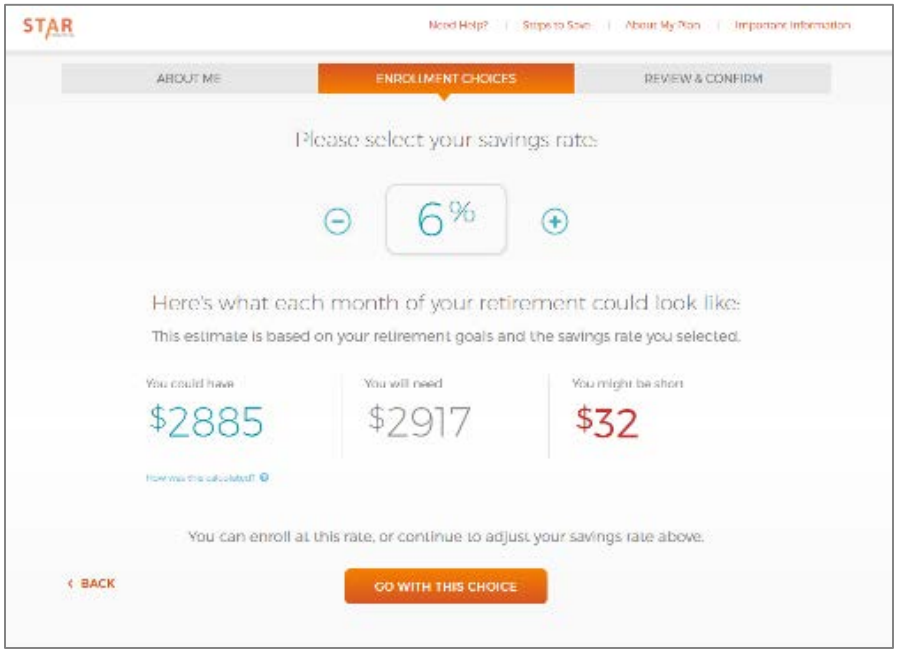


Supplementary Setting - Partial Plan Information - Default information displayed
(3 percent default, 6 percent match)

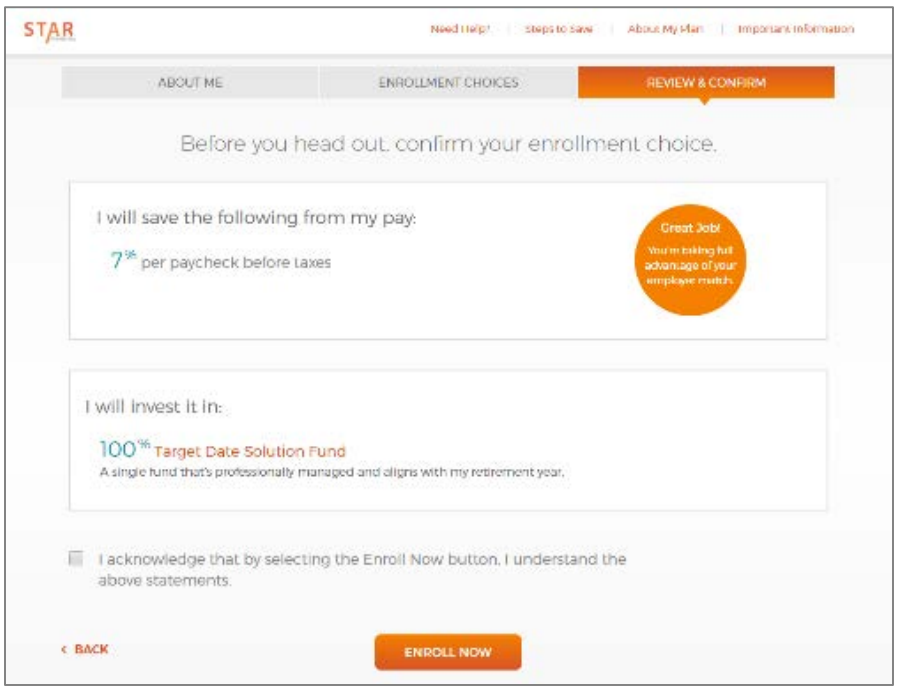


Appendix Figure A8.
Hypothetical Enrollment Experiment – Personalized Enrollment Webflow

Panel A. Selection of Contribution Rate and Retirement Calculator



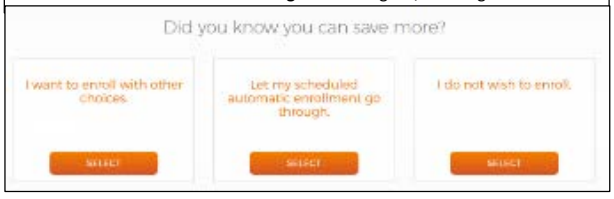
Panel B. Enrollment Confirmation



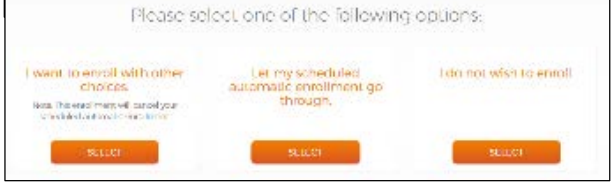
Appendix Figure A9.
Conditions in Supplemental Hypothetical Enrollment Experiment

1. Basic - Presentation of options used by studied plans prior to research

2. Basic + No PERSONALIZE Warning - Basic design w/ warning omitted



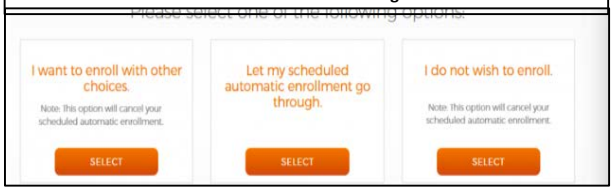
3. Basic + Revised Instructions - Basic design w/ neutral language



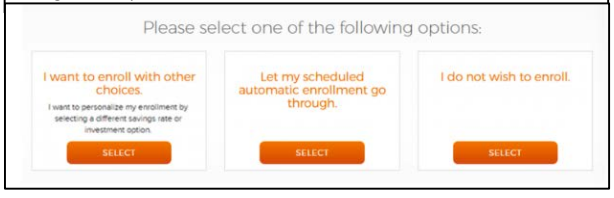
4. Basic + Revised Instructions + No PERSONALIZE Warning

5. Basic + Revised Instructions + No PERSONALIZE Warning + DECLINE

6. Basic + Revised Instructions + DECLINE Warning



7. Basic + Revised Instructions + Only PERSONALIZE Explanation - Basic design with explanation under PERSONALIZE



8. Baseline - Standard presentation of enrollment options with explanations

Appendix Table A1.
Sensitivity of Enrollee Retirement Risk to Assumed Wage Growth and Retirement Age

	Accumulation Threshold		
	Low $150,000	Moderate $200,000	High $250,00
Panel A. Retirement Age			
Age 65	**0.42**	**0.53**	**0.62**
Age 70	0.30	0.41	0.50
Panel B. Annual Real Wage Growth			
Zero Percent	0.46	0.58	0.67
1 Percent	**0.42**	**0.53**	**0.62**
2 Percent	0.38	0.48	0.57

Notes: This table summarizes the share of 401(k) enrollees at risk of retirement insecurity for a range of asset thresholds ($150k-$250k) by row and a range of assumptions regarding annual real wage growth (Low at 0%, Moderate at 1%, and High at 2%) and retirement age (65 or 70). We fix the assumption that "at risk" means 25% or higher chance of falling below the asset threshold . All estimates in this table reflect our preferred, moderate assumptions regarding market return expectations (mean 3.5, standard deviation 10.2 nominal annual return rate), contribution rate inertia of 75% (fixed annual probability of 25% (100% - inertia) of increasing contribution rate by 1% of income until reaching 10%).

Appendix Table A2.
Transcription of Field Experimental Conditions

					Enrollment Options					
			Plan Information		Headlines and Color			Description		
Experimental Conditions	N	Instructions	Default	Match	Personalize	Auto	Decline	Personalize	Auto	Decline
Panel A. Field Study #1										
Basic	1276	Please select one of the following options:	--	--	I want to enroll with different choices.	I want to confirm my automatic enrollment.	I do not want to enroll.	Note: This enrollment will cancel your scheduled automatic enrollment.	--	--
Enhanced Info (Default Info)	1254	Please select one of the following options:	If you confirm your automatic enrollment you will be enrolled at a savings rate of [%] percent.	--	I want to enroll with different choices.	I want to confirm my automatic enrollment.	I do not want to enroll.	Note: This enrollment will cancel your scheduled automatic enrollment.	--	--
Enhanced Design (Default Info)	1337	Please select one of the following options:	If you confirm your automatic enrollment you will be enrolled at a savings rate of [%] percent.	--	Do it Myself (Green)	Do it for Me (Yellow)	I Don't Want to Save (Red)	I want to personalize my enrollment by selecting a different savings rate or investment option.	I want my auto enrollment to go through at the savings rate chosen by my employer.	I want to cancel my auto enrollment and not save at this time.
Panel B. Field Study #2										
Enhanced Info (Default/Match Info)	547	Please select one of the following options:	If you confirm your automatic enrollment you will be enrolled at a savings rate of [%] percent.	Note: Your employer will match your contribution up to a rate of [%] percent.	I want to enroll with different choices.	I want to confirm my automatic enrollment.	I do not want to enroll.	Note: This enrollment will cancel your scheduled automatic enrollment.	--	--
Enhanced Design (Default/Match Info)	518	Please select one of the following options:	If you confirm your automatic enrollment you will be enrolled at a savings rate of **[%] percent**.	Note: Your employer will match your contribution up to a rate of **[%] percent.**	Do it Myself (Green)	Do it for Me (Yellow)	I Don't Want to Save (Red)	I want to personalize my enrollment by selecting a different savings rate or investment option.	I want my auto-enrollment to go through at the savings rate chosen by my employer.	I want to cancel my auto enrollment and not save at this time.
Panel C. Field Study #3										
Basic	529	Did you know you can save more?	--	--	I want to enroll with other choices.	Let my scheduled automatic enrollment go through.	I do not wish to enroll	Note: This enrollment will cancel your scheduled automatic enrollment.	--	--
Standardized	1015	Please select one of the following options:	--	--	I want to enroll with different choices.	I want to confirm my automatic enrollment.	I do not want to enroll.	I want to personalize my enrollment by selecting a different savings rate or investment option.	I want my auto-enrollment to go through at the savings rate chosen by my employer.	I want to cancel my auto-enrollment and not save at this time.
Headlines	1019	Please select one of the following options:	--	--	Do it Myself	Do it for Me	I Don't Want to Save	I want to personalize my enrollment by selecting a different savings rate or investment option.	I want my auto-enrollment to go through at the savings rate chosen by my employer.	I want to cancel my auto-enrollment and not save at this time.
Enhanced Presentation	1070	Please select one of the following options:	--	--	Do it Myself (Green)	Do it for Me (Yellow)	I Don't Want to Save (Red)	I want to personalize my enrollment by selecting a different savings rate or investment option.	I want my auto-enrollment to go through at the savings rate chosen by my employer.	I want to cancel my auto-enrollment and not save at this time.

Notes: This table summarizes the content of the landing page and sample size in each experimental condition of the three field studies. Experimental treatments vary in the text of the instructions positioned above the three enrollment options [Instructions], whether the plan default and match limit rates are displayed [Plan Information], the headlines and color associated with the three enrollment options [Headlines and Color], and the presence and content of subtext describing these options [Subtext].

Appendix Table A3.
Effect of Enhanced Presentation by Individual Plan - Field Study #2

		Plan Details		Enrollment Choice			Contribution Threshold		Annual Contribution Rate		
Anonmyized Plan	N	Match Limit [Percent]	Default Rate [Percent]	Change [1,0]	Auto [1,0]	Decline [1,0]	> Default Rate [1,0]	≥ Match [1,0]	All [Percent]	Participants [Percent]	Change [Percent]
Plan 1	117	5	3	0.05	-0.05	0.01	0.05	0.09	0.88	0.97	0.73
				(0.06)	(0.05)	(0.03)	(0.06)	(0.06)	(0.77)	(0.74)	(0.75)
Plan 2	451	6	3	**0.13*** **	**-0.14*** **	0.01	**0.10** **	**0.09** **	-0.04	0.04	**-0.94** **
				(0.04)	(0.04)	(0.03)	(0.05)	(0.05)	(0.40)	(0.39)	(0.44)
Plan 3	351	6	3	0.06	-0.07**	0.01	0.04	0.03	0.31	0.39	0.13
				(0.04)	(0.03)	(0.01)	(0.04)	(0.05)	(0.36)	(0.35)	(0.37)
Plan 4	146	6	4	0.15**	-0.05	-0.10*	0.13	0.10	0.39	-0.36	-0.68
				(0.07)	(0.05)	(0.06)	(0.08)	(0.08)	(0.66)	(0.63)	(0.70)
All Plans	1065			**0.10*** **	**-0.09*** **	-0.01	**0.07*** **	**0.07** **	0.23	0.20	-0.30
				(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.24)	(0.23)	(0.25)
Basic + Info Average				0.72	0.19	0.08	0.68	0.63	6.09	6.64	7.58

Notes: This table summarizes treatment effects of enhanced presentation on a series of enrollment outcomes estimated from a sample of all employees in Field Study #2 and separately for each of the four plans in this study. Each cell in the main panel corresponds to an OLS/LPM regression estimate of the marginal effect of enhanced presentation for the sample specified by the on the outcome indicated by the column heading. Please refer to text for additional details on the underlying specifications. Average outcomes for employees in Field Study #2 who viewed the Basic condition with default and match information displayed are presented in the bottom panel. Robust standard errors are displayed in parentheses. Asterisks indicate p-values associated with tests of statistical significance: (*) $p < 0.10$, (**) $p < 0.05$, and (***) $p < 0.01$.

Appendix Table A4.
Effect of Enhanced Design on Enrollment by Plan Default Rate - Field Study #1

	Personalized Enrollment [1,0]			Contribution Rate > Default Rate [1,0]			Contribution Rate ≥ Match Limit [1,0]			Contribution Rate [Percent]		
	Default Contribution Rate			Default Contribution Rate			Default Contribution Rate			Default Contribution Rate		
Treatment Category	Any	≤ 3 Percent	> 3 Percent	Any	≤ 3 Percent	> 3 Percent	Any	≤ 3 Percent	> 3 Percent	Any	≤ 3 Percent	> 3 Percent
Enhanced Design	**0.09*** **	**0.10*** **	**0.08** **	**0.08*** **	**0.09*** **	0.07	**0.10*** **	**0.11*** **	0.08	**0.62*** **	**0.76*** **	0.21
	(0.02)	(0.02)	(0.04)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	(0.05)	(0.22)	(0.25)	(0.46)
Enhanced Presentation (w/ Default Info)	**0.09*** **	**0.10*** **	**0.08** **	**0.08*** **	**0.09*** **	0.07	**0.10*** **	**0.11*** **	0.08	0.29	0.31	0.22
	(0.02)	(0.02)	(0.04)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	(0.05)	(0.19)	(0.22)	(0.41)
Enhanced Information	0.02	0.03	-0.02	0.03	0.03	0.01	0.04*	0.05*	0.02	0.33	**0.46** **	-0.01
	(0.02)	(0.02)	(0.04)	(0.02)	(0.03)	(0.04)	(0.02)	(0.03)	(0.05)	(0.20)	(0.22)	(0.48)
Average for Basic Condition	0.60	0.61	0.55	0.50	0.56	0.33	0.58	0.55	0.68	5.41	5.31	5.71
Number of Employees	3,867	2,886	981	3,859	2,881	978	2341	1776	565	3,859	2,881	978
Number of Plans	397	287	110	397	287	110	183	131	52	397	287	110

Notes: This table summarizes treatment effects of psychological design on a series of enrollment outcomes estimated from a sample of employees in Field Study #1, for all plans and separated by plans with lower default contribution rates (≤ 3 percent) or higher default contribution rates (> 3 percent). Each cell in the main panel corresponds to an OLS/LPM regression estimate of the marginal effect specified by the panel and row on the outcome indicated by the column heading, after controlling for plan fixed effects and day-of-week and week-of-year variation in the timing of enrollment. Please refer to text for additional details on the underlying specifications. Average outcomes for employees in each regression who viewed the Basic condition are displayed in the bottom panel. Robust standard errors, displayed in parentheses, are clustered at the plan-level. Asterisks indicate p-values associated with tests of statistical significance: (*) $p < 0.10$, (**) $p < 0.05$, and (***) $p < 0.01$.