



Using Behavioral Insights to Increase Retirement Savings

Trial Design and Findings

Samia Amin

Greg Chojnacki

Aravind Moorthy

Irma Perez-Johnson

*Matt Darling**

*Jaclyn Lefkowitz**

*ideas42

April 2017

Submitted to:

U.S. Department of Labor
Chief Evaluation Office
200 Constitution Ave., NW
Washington, DC 20210
(202) 693-5911
Project Officer: Megan Lizik
Contract Number: DOLQ129633249/DOL-OPS-16-U-00126

Submitted by:

Mathematica Policy Research
1100 1st Street, NE
12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763
Project Director: Samia Amin
Reference Number: 50291

Recommended Citation:

Amin, Samia, Greg Chojnacki, Aravind Moorthy, Irma Perez-Johnson, Matt Darling, and Jaclyn Lefkowitz. "Using Behavioral Insights to Increase Retirement Savings: Trial Design and Findings." Washington, D.C.: Mathematica Policy Research, April 2017.

Acknowledgments

We gratefully acknowledge the many colleagues and partners whose generous contributions of time and expertise made this study possible. We are indebted to our partners in the Human Resources Division of the U.S. Department of Labor (DOL-HR) and the Employee Benefits Security Administration (EBSA), without whom this behavioral trial would not be possible. In particular, we want to thank Deborah Dudley and Veronda Nicholas at DOL-HR and Debra Golding and Samuel Tseng at EBSA for their instrumental role in program design, implementation, and data collection for this trial. We thank Megan Lizik, Jessica Lohmann, Celeste Richie, Molly Irwin, and Demetra Nightingale at DOL's Chief Evaluation Office for their guidance and support of this behavioral trial. We appreciate the Financial Literacy and Education Commission's input on this effort. We thank Eldar Shafir at ideas42 for his advice on intervention design.

At Mathematica Policy Research, we recognize Sophie MacIntyre for the data analysis she conducted for this project and Mikia Manley for her creativity and skill in helping develop the template and graphics for this report. We thank Ken Fortson for his input on the design of the trial, the evaluation, and the analyses. We are grateful to Michael Ponza for his careful review of the draft version of this report and for his helpful suggestions. We also acknowledge Amy Taylor and William Garrett for their careful editing of this report and Jennifer Baskwell for her valuable production assistance.

All errors or omissions are the responsibility of the authors.

The DOL Behavioral Interventions Project

The Department of Labor Behavioral Interventions (DOL-BI) project was launched to explore the potential of using behavioral science to improve the performance and outcomes of Labor programs. It is sponsored by the DOL Chief Evaluation Office and executed by Mathematica Policy Research and ideas42. The project team has designed, implemented, and rigorously tested three sets of behavioral trials in selected Labor programs. This report describes the findings of one of these efforts, conducted in partnership with the Employee Benefits Security Administration and the Department of Labor's Human Resources division. The project team also developed interventions and executed trials in partnership with the Occupational Safety and Health Administration to boost workplace safety and with the Employment and Training Administration to help unemployed workers with their job search efforts.

Access reports, briefs, presentations, and infographics on these trials, as well as additional tools for applying behavioral insights by visiting <https://www.dol.gov/asp/evaluation/BIStudy/>.

This report was prepared for the Chief Evaluation Office of the U.S. Department of Labor (DOL) by Mathematica Policy Research and ideas42 under contract number DOLQ129633249/DOL-OPS-16-U-00126. The views expressed are those of the authors and should not be attributed to DOL, nor does mention of trade names, commercial products, or organizations imply endorsement of same by the U.S. government.

Contents

I	INTRODUCTION.....	1
	Trial synopsis	2
	Understanding the context.....	2
	Research questions and trial design.....	4
	Report roadmap	5
II	INTERVENTION DESIGN	6
	Understanding the problem and its context.....	6
	Diagnosing why people do not participate in the TSP	8
	Designing the intervention	9
	Supporting implementation	13
III	EVALUATION DESIGN	14
	An experimental design	14
	A rigorous analytic approach	16
IV	FINDINGS.....	19
V	DISCUSSION AND LESSONS LEARNED	26
	Discussion of results.....	26
	Lessons learned	27
	REFERENCES	29
	APPENDIX A BEHAVIORAL INTERVENTION DESIGN.....	31
	APPENDIX B EVALUATION DESIGN AND ANALYTIC METHODS.....	35
	APPENDIX C BASELINE EQUIVALENCE.....	44
	APPENDIX D DETAILED TABLES ON STUDY FINDINGS	48

Tables

II.1	Characteristics of the analysis sample at baseline	7
III.1	Sample description and sample sizes at random assignment for each phase	14
III.2	Data collection timelines for each phase	16
III.3	Final analysis sample sizes by phase.....	16
III.4	Research questions and relevant analyses	17
B.1	Characteristics of the sample at baseline	37
B.2	Attrition in Phase 1	39
B.3	Attrition in Phase 2	39
B.4	Sample sizes in each phase.....	40
B.5	Description of analyses	41
C.1	Baseline equivalence between T1 and Control in Phase 1	45
C.2	Baseline equivalence between T1 and Control in Phase 2	46
C.3	Baseline equivalence between T2 and Control in Phase 2	46
C.4	Baseline equivalence between T2 and T1 in Phase 2	47
D.1	Impacts of T1 and T2 (Analyses 1, 2, and 3).....	49
D.2	Impact of T2 vs. T1 in Phase 2 after the reminder (Analysis 4).....	50
D.3	Impact of T1 in Phase 2 before the reminder (Analysis 5)	50
D.4	Impact of T2 in Phase 2 before the reminder (Analysis 6)	50
D.5	Impact of the reminder in Phase 2 (Analysis 7)	51
D.6	Impact of T1 by phase (Analysis 8)	51
D.7	Summary of subgroup analysis findings	53
D.8	Email receipt in Phase 1.....	54

Figures

I.1	Evaluation design.....	5
II.1	Using behavioral science to improve programs: Six steps.....	6
II.2	Behavioral diagnosis and intervention design.....	9
II.3	Original email text for Phases 1 and 2.....	11
II.4	Original email attachment for Phases 1 and 2: EPP instructions	12
IV.1	Impact of original email and social norms variant on percentage of employees who started contributing at least 5% to TSP	19
IV.2	Impact of original email and social norms variant on overall contribution rates	20
IV.3	Impact of original email and social norms variant on percentage of employees who started contributing to TSP (any level)	21
IV.4	Comparison of impacts for the original email versus the social norms variant in Phase 2 after reminder	22
IV.5	Effect of Phase 2 reminders	23
IV.6	Impacts by phase, limiting differences in sample composition	24

Exhibits

A.1	Behavioral map of TSP benefit change process.....	32
A.2	Social norms variant email for Phase 2	33
A.3	Reminder text for original email.....	34

I. Introduction

Preparing for retirement is important. For many workers in the United States, a comfortable retirement may depend on the savings decisions they make now. Failing to save today can have very real consequences as people age, reducing the comforts they get to enjoy during retirement and their ability to cope with health and financial shocks.

Despite these high stakes, we often struggle to plan for our retirement. We have limited attention and the urgent demands of the here and now can deplete our ability to focus on the future. The present and near future are palpable for us while the distant future seems vague and abstract.¹ Struggling to visualize our needs and wants during retirement, we may find it difficult to devote time now to making retirement choices.

Even when we decide to think about retirement, actually putting money aside can be hard to do. Behavioral scientists explain that this is because we are biased toward the present.² We find immediate rewards more attractive than the same or better rewards in the future.^{3, 4, 5} Although we know there are clear benefits to starting to save early in our career, those rewards are largely confined to the future. Instead, many of us focus on the immediate cost of having less money now.

We also focus disproportionately on the time and effort involved in making savings decisions. Our natural tendency to avoid complex tasks that feel unpleasant or burdensome becomes more acute. If we associate decisions about retirement savings with the stress of handling lots of information—understanding current finances, calculating future needs, and comparing different savings plans—we may be too daunted to act, even though we want to.^{6, 7} And so our tendencies to get distracted, procrastinate, and be deterred by small hassles may sabotage our good intentions.

Fortunately, behavioral science doesn't just provide insights into why these types of problems can occur. It also offers ways in which they might be addressed. The way choices are framed can help people understand better and take action. Small changes informed by behavioral science may have large effects.

In 2014, the U.S. Department of Labor's (DOL's) Chief Evaluation Office contracted with Mathematica Policy Research and ideas42 to explore the potential of using insights from behavioral science to improve the performance and outcomes of DOL programs. In this study, the DOL Behavioral Interventions (DOL-BI) team partnered with the Human Resources Division of DOL (DOL-HR) and the Employee Benefits Security Administration (EBSA) to explore whether low-cost, behaviorally informed emails would result in more DOL employees saving, or saving more, for retirement. We conducted two phases of tests, in the fall of 2015 and spring of 2016, which examined different behavioral messages and reminder strategies. Our tests yielded strong positive results: The most effective intervention we tested increased the share of employees saving at least 5 percent of their income by 7.5 percentage points.

About Behavioral Science

Behavioral science studies how people make decisions and act in a complex world. It draws on decades of research in the social sciences to provide a more realistic model of how we make decisions and act in real life. Other approaches commonly assume that we consider all available information, weigh the pros and cons of each option, optimize our choices, and then reliably act on them. In practice, however, people often decide and act with imperfect information or fail to act altogether, even when they may want to. Behavioral interventions test whether aligning policies, programs, and products to these human tendencies can result in improved outcomes.

Trial synopsis

This report presents our findings on the effects of emails designed to encourage DOL employees to increase their contributions to the Thrift Savings Plan (TSP)—a program similar to private sector 401(k) plans—and take full advantage of the available employer match. These emails were sent to employees contributing less than 5 percent of their salary to the TSP, the amount needed to receive the employer’s full matching contribution. The messages built on retirement planning resources that are already available to federal employees and drew on behavioral research to address cognitive barriers that may prevent people from saving for retirement.

Thrift Savings Plan

A retirement savings plan for federal government agency employees that offers a 5 percent employer match.

Understanding the context

Federal law requires DOL and other agencies to match the first 5 percent of salary that an employee contributes to the TSP. DOL has sought to find effective strategies to encourage employees to take full advantage of this offer. These strategies are intended to supplement broader ongoing efforts to promote retirement savings among all federal employees (Box I.1). The most notable of these broader initiatives was new legislation enacted in 2009 that automatically enrolled new federal employees in the TSP program when they first join an agency, at 3 percent of their salary. DOL has supplemented these efforts by sending out letters and emails to its employees reminding them of the wide range of benefits available to them, including the TSP.

Box I.1. Strategies used to encourage greater TSP participation among federal employees

Initiatives targeted at all federal civilian employees:

- **Auto-enrollment.** Under the Thrift Savings Plan Enhancement Act of 2009, starting in August 2010, newly hired federal civilian employees were automatically enrolled in the TSP, with their contributions set by default at 3 percent of salary unless they opted out or chose a different contribution level. This change did not apply to current employees.
- **Educational materials.** In 2014, the TSP team released a series of educational materials organized around the theme “Take Five for Your Future” to highlight the ease of changing TSP contribution levels (in five minutes), the range of investment options (five core and five life-cycle funds), and the benefit employees gained by contributing at least 5 percent of their annual salary to TSP (the full 5 percent employer match). The U.S. Office of Personnel Management also shares basic information on the TSP and provides links to TSP resources.
- **EBSA.** EBSA sends out an annual letter to federal employees with TSP plan information and deferral limits.
- **In-person and webinar information sessions.** The TSP is discussed at new-employee orientations, financial planning seminars, and retirement planning seminars held monthly at DOL. The TSP team also holds frequent on-site workshops (TSP from A to Z; TSP Pre-Separation Workshop: To Retirement and Beyond) and annual informational webinars during open-enrollment season.

Initiatives targeted specifically at DOL employees:

- **Personalized letter on benefit elections.** Every year, DOL employees receive a letter reminding them of their benefit elections.
 - **DOL-HR outreach.** DOL-HR occasionally sends out emails on a wide range of employee benefits that include information on the TSP.
-

Despite these efforts, more than one quarter of eligible DOL employees were contributing less than 5 percent of their salary to the TSP as of the summer of 2015.⁸ This is not entirely surprising. Although automatic enrollment has the advantage of increasing the proportion of employees who start saving, it can still lead to undersaving, because default amounts are typically set fairly low.⁹ The same inertia that keeps

28%

of DOL employees contribute less than 5 percent to the TSP, not taking advantage of the full employer match

people from opting out of auto-enrollment can also prevent them from increasing their savings rate. DOL employees are not unusual in their failure to maximize employer contributions. Several other agencies, including the Department of Defense (DOD), face these challenges (see Box I.2).

EBSA and DOL-HR were eager to explore whether applying insights from behavioral science using low-cost interventions (such as emails) could help address this problem. Behavioral interventions have been effective in similar contexts, and our discussions with DOL-HR and EBSA partners suggested they could be here as well. Although there were many good reasons why people might not participate in the TSP (such as limited financial resources, urgent financial needs, or availability of other appealing investment options), a host of behavioral

factors might also play an important role: Employees may underestimate the importance of savings and the consequences of insufficient retirement funds; they may have lowered or stopped their contributions in times of financial hardship and forgotten to revisit these decisions once finances improved; or they may have been deterred by the belief that changing contribution levels would be cumbersome.

Box I.2. Increasing military service members' TSP participation rates using behavioral insights

Just 44 percent of military service members participate in the TSP, compared to 87 percent of civilian federal employees. DOD and the White House Social and Behavioral Sciences Team (SBST) partnered to design and test the following series of interventions to increase TSP participation:

- One-time email to nonenrollees.** SBST and DOD tested eight email variants that combined different features informed by behavioral insights. These included clearly outlining the steps to enroll, framing the option to enroll as a “Yes” or “No” choice, and stressing either the short-term or long-term benefits of saving. In April 2015, more than 800,000 service members received either one of the eight email variants or no email (control group). On average, email recipients had an enrollment rate 0.77 percentage points higher than the control group. The most effective email, which outlined the long-term benefits of saving for retirement and clearly specified the steps to enroll, increased TSP enrollment by 2.1 percentage points.
- Choice prompts for nonenrollees.** SBST and DOD required service members to make a “Yes” or “No” choice for TSP enrollment when completing routine paperwork during orientation at a new military base. This intervention led to a 4.3 percentage point increase in the rate of TSP enrollment.
- TSP reenrollment for those already enrolled.** After a system change, 140,000 service members had to re-enroll to remain in the Roth TSP plan to which they had already been contributing. SBST and DOD tested a redesigned email drawing on behavioral insights against DOD’s standard email. The new message urged service members to make a fresh start in the new year with their finances and avoid losing out on the chance to save. It also specified the steps to re-enroll, included a personalized greeting, and re-emphasized the key message in a postscript. The redesigned email increased enrollments by 5.2 percentage points compared to the standard email.

Starting in 2018, new military service members will be enrolled automatically in the TSP under the 2016 National Defense Authorization Act.

Sources: Social and Behavioral Sciences Team (SBST), 2015 Annual Report; SBST, 2016 Annual Report; SBST, “Service member TSP enrollment.” Available at <https://sbst.gov/>. Accessed October 9, 2016.

In partnership with DOL and EBSA, we decided to design and rigorously test behaviorally informed emails that would do the following:

- Focus exclusively on the TSP benefit
- Explicitly target those DOL employees who currently were not participating in the TSP or were contributing less than 5 percent
- Leverage existing informational materials developed by the TSP and build on lessons from DOD's efforts to boost service members' participation in the TSP (see Box I.2)

Since this was the first time DOL would send out emails focused exclusively on the TSP, we wanted to rigorously test the effectiveness of this approach by rolling it out as an experiment and analyzing the data on employee contributions to see whether it worked.

Research questions and trial design

In conducting this trial, we had four main research questions:

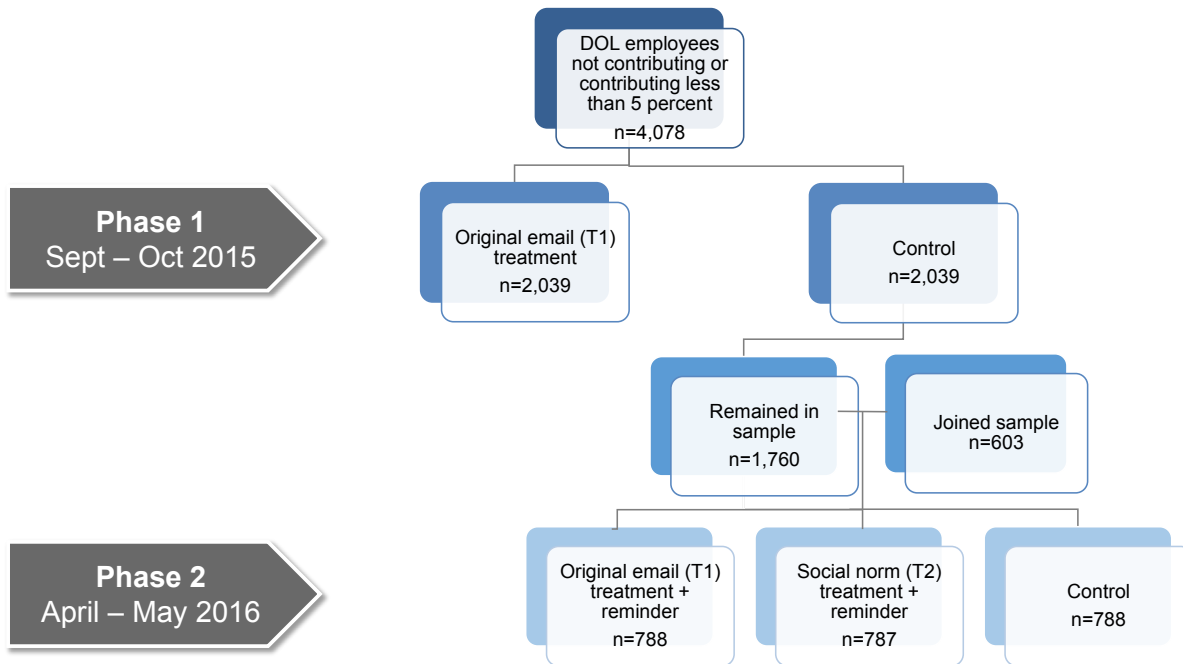
- **Can emails informed by behavioral science encourage more DOL employees to increase their savings?** Do they prompt DOL employees to increase their contributions enough to qualify for the full employer match? Do they prompt employees not participating in the TSP to start contributing?
- **What email communication strategies are most effective at improving savings?** Are emails that invoke social norms (for example, by citing what their peers are doing) more effective than emails that do not? Does sending reminder emails change the impact? Does timing matter?
- **For whom do these emails work?** Do impacts vary for different employee subgroups—for example, older versus younger employees, those earning more versus those earning less?
- **What lessons might help inform similar efforts in other contexts?** What infrastructure is necessary for sending these targeted communications? What considerations are important when seeking to design these strategies?

We conducted a behavioral diagnosis and developed multiple iterations of prototypes to inform our email design. The prototypes consisted of behaviorally informed email messages that (1) made retirement salient and the benefits of participating in the TSP easy to understand and visualize, (2) urged recipients not to lose out on matched savings, (3) made action steps clear, and (4) reduced hassle factors that might deter employees from acting. We also developed a variant that invoked social norms to encourage saving and two reminder messages based on each variant of the initial email.

We tested the effectiveness of these email messages using a randomized controlled trial conducted in two phases (see Figure I.1). In Phase 1, conducted September–October 2015, we sent a new email to half of the more than 4,000 DOL employees who either were not participating in the TSP or were participating but contributing less than 5 percent of their salary to the TSP. Email recipients were randomly selected and nonrecipients served as the business-as-usual (control) group that did not receive any messages. In Phase 2, conducted April–June 2016, we tested two variants of the original email message (with and without references to social norms) with employees who had not received an intervention email in Phase 1 and were still contributing less than 5 percent of salary to the TSP, as well as recent hires and others who began contributing less than 5 percent after the end of Phase 1. We then sent reminder messages six weeks after

the initial emails. For both phases we measured the effects on TSP savings behavior using DOL records of retirement contributions.

Figure I.1. Evaluation design



Note: Across the trial's two phases, 4,681 employees were randomly assigned to a control group or one of the treatment groups.

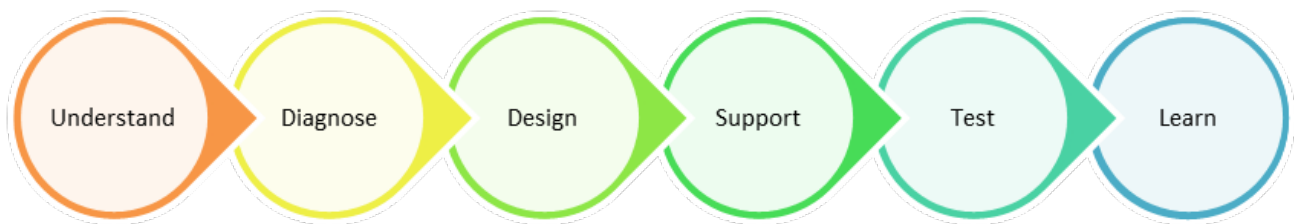
Report roadmap

This report describes how we designed and implemented these interventions and the accompanying evaluation, what we found, and why our findings matter. Chapter II and Appendix A provide more details on our process for developing these behavioral interventions, for those interested in designing similar interventions and tests. Chapter III and Appendices B and C describe our evaluation design so readers can assess the validity of our findings. We also discuss our experimental design, target population, data collection, and analysis approaches. Chapter IV and Appendix D discuss the findings. Chapter V concludes by summarizing what we learned, how our results can be used, and next steps in exploring how behavioral interventions can promote savings and increase retirement security.

II. Intervention Design

In developing interventions and a trial to test possible ways of increasing TSP participation among DOL employees, we followed six steps that form the core of our approach (Figure II.1). We began by deepening our understanding of the specific problem we were attempting to solve and the context in which it occurs. Then we diagnosed potential behavioral barriers that may be contributing to the problem, designed an intervention that addressed those barriers, and provided support for its implementation. Finally, we tested the effectiveness of the intervention using a rigorous trial design and learned from our experimental findings. In this chapter, we discuss the first four steps in this process; the remainder are discussed in the following chapters.

Figure II.1. Using behavioral science to improve programs: Six steps



Understanding the problem and its context

Designing a well-targeted and effective intervention requires a thorough understanding of the problem being addressed. We began by working with our partners at EBSA and DOL-HR to determine their goals. For example, did they want to increase the number of people who were contributing 5 percent of their salary and receiving all possible matched funds, or did they want to increase the average savings rate for the full population? Although these goals are related, having a final outcome in mind is key for developing a well-targeted intervention. An intervention intended to increase the number of employees who save can look very different from one meant to increase the average savings rate. We determined that EBSA's and DOL-HR's main priorities were (1) getting noncontributors to start participating in the TSP program and (2) increasing people's contribution rate to at least 5 percent.

Next, we investigated and tried to learn about the people whose behavior we wished to change, the contexts in which these people make decisions, the steps they have to take to make the desired choice, and feasible ways of influencing those choices given our implementing partners' capabilities and constraints. Although we explored these issues in some depth, below we highlight only those details that are helpful for understanding our behavioral diagnosis.

Identifying the target population. DOL-HR and EBSA identified two priority groups of DOL employees for our efforts: (1) employees not participating in the TSP (noncontributors) and (2) employees contributing less than the 5 percent of salary needed to gain the full employer match (undercontributors). On the basis of demographic and income information we received from DOL, we were able to develop a basic understanding of who the individuals in these two groups were (Table II.1). The median age was 45, and

the groups were nearly equally split between men and women. A little less than a third worked in the national office, and the remainder were based in regional offices. Over two thirds were union members. When comparing noncontributors to undercontributors, a few differences emerged. Unsurprisingly, noncontributors had a larger share of employees hired before auto-enrollment began in August 2010. They also had higher median salaries and a slightly higher median age, were a bit more likely to work at the national office, and were a little less likely to be union members. Compared to Phase 1 study participants, those from Phase 2 had higher salaries and a higher proportion of people hired before auto-enrollment began (see Appendix B, Section 1.B).

Table II.1. Characteristics of the analysis sample at baseline

Characteristic	Phase 1			Phase 2		
	Total	Non-contributors	Under-contributors	Total	Non-contributors	Under-contributors
Sample size (N)	4,011	1,721	2,290	2,302	938	1,364
Contribution rate (%)	1.5	0.0	2.7	1.6	0.0	2.7
Salary (\$)	63,536	67,636	61,486	78,209	81,902	77,490
Age	45.4	46.8	44.1	45.1	45.7	44.6
National office (%)	30.8	33.3	28.9	32.1	34.4	30.6
Hired before auto-enrollment began (August 2010) (%)	36.6	47.2	28.6	54.3	66.4	46.0
Female (%)	51.5	52.4	50.9	51.2	51.0	51.3
Union membership (%)	78.8	77.4	79.9	78.3	77.3	79.0
Race/ethnicity						
Asian or Pacific Islander (%)	3.9	2.8	4.7	4.0	3.3	4.5
Black (%)	35.7	39.7	32.7	36.4	40.2	33.8
White (%)	49.5	47.5	50.9	48.3	45.9	49.9
Hispanic (%)	9.7	9.0	10.3	4.3	4.7	4.0
Other race/ethnicity or multiple (%)	1.2	1.0	1.4	7.0	5.9	7.8

Notes: The contribution rate row displays the average value for each group. The Salary and Age rows each display the median values for each group. All other rows display the proportion of the group that has the corresponding characteristic.

What are the contexts in which people make decisions? The first instance in which employees are faced with decisions regarding retirement savings is when they join DOL. Orientations for new staff include sessions on the TSP, and at this stage people can choose to make their savings election by filling out paper forms. Employees who joined DOL after August 2010 automatically have 3 percent of their salary directed to the TSP unless they choose otherwise. Those who joined DOL before that date were not affected. Soon after joining, employees receive access to the online payroll system (known as the Employee Personal Page), which they can use to update their choices at any time, as well as conduct other tasks such as accessing their pay stubs. Elections made in one pay cycle go into effect in the next pay cycle. If people lose their password for accessing the online payroll system, they can request a reset, but this password is sent to them only in hard copy via postal mail. After orientation, all TSP changes have to be made online. The TSP and DOL-HR conduct a number of activities to encourage people to periodically revisit their savings decisions (see Box I.1). However, after an initial prompt to make a savings election when joining DOL,

employees do not face another decision point where taking action is urgent. This is important to note because, when it comes to retirement savings, people tend not to perceive a difference between saving today and saving tomorrow.^{10, 11} This lack may encourage them to delay enrollment, especially if they face hassles, even those that seem small.

What are key operational features and constraints relevant for the design of the intervention?

Given competing demands on staff resources and system capabilities, only low-cost interventions that required minimal staff effort to implement were feasible. For instance, we could not make any changes to the Employee Personal Page website, to TSP policies or procedures, or to the actual savings options available to employees. Moreover, in designing the intervention, we were sensitive to the fact that a recent security breach of government employees' personal information might have made employees wary of email communications related to their personal finances. Hence, it would be important for communications to come from a trusted source. We were also aware that only employing agencies may send emails to employees about increasing contributions.

Diagnosing why people do not participate in the TSP

We worked closely with EBSA and DOL-HR staff to identify potential factors that may lead employees to contribute less than 5 percent of their salaries to the TSP. Tapping many different methods (Box II.1), we developed a "behavioral map" of the process by which DOL employees could alter their savings rate (see Appendix A, Exhibit A.1). We then used this map to create hypotheses about the behavioral barriers that potentially deter DOL employees from beginning to save or increasing their savings rate to 5 percent. Refining our focus through ongoing discussions with EBSA and DOL-HR staff, we honed in on four central themes:

Methods used to investigate bottlenecks:

- Analyzed administrative data
- Obtained feedback and input from DOL-HR staff and EBSA
- Examined materials shared with DOL employees (including the TSP's "Take Five" materials)
- Conducted "walkthrough" of the DOL payroll website where employees change their TSP elections
- Solicited DOL employee perspectives



Inattention and misperception. Employees may not actively think about retirement planning or realize its benefits because they are focused on the present and the immediate future and not paying enough attention to their long-term needs and desires. They may be unaware of or forget that they are foregoing matching funds.



Present bias. Even if they focus on retirement, employees might not begin saving because they are (1) unduly sensitive to the immediate cost of giving up money now, (2) concerned about the time and effort involved in changing their savings elections, and (3) undervaluing the benefits of savings that accrue in the future.



Information overload and fear of a permanent decision. Employees may fail to make a choice because they are overwhelmed by the volume of information or the complexity of the decisions to be made. They may incorrectly believe that their choice will be difficult to revisit or undo.

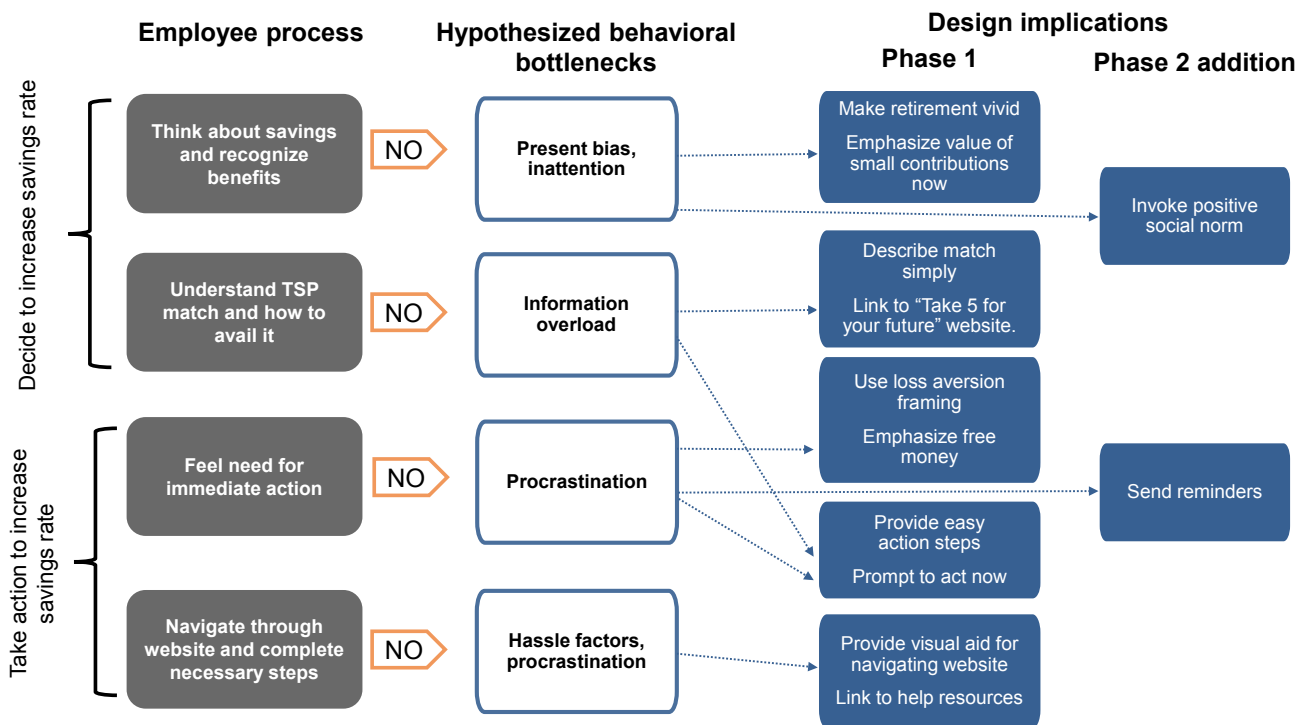


Hassle factors and procrastination. Employees may view the process of changing the contribution levels to be a hassle and may procrastinate.

Designing the intervention

As discussed, we determined that an email intervention was the most practical intervention strategy. And existing studies, including others also targeting TSP participation, suggested that email interventions could work.^{12, 13} We reviewed the literature to see what strategies worked in other contexts and may be suitable in ours. We made sure to integrate those strategies that proved effective in boosting TSP participation in the DOD Social and Behavioral Sciences Team (DOD-SBST) trials (described in Box I.2). These included emphasizing the long-term benefits of retirement savings and clearly listing needed action steps. The email designs went through several revisions as we received feedback from EBSA, DOL-HR, and the Financial Literacy and Education Commission on the specific language that might be most effective for DOL employees. Additionally, we user tested our final message with three DOL employees randomly selected from the target population. This allowed us to confirm that our messages were clear, easy to understand, and likely to be read.

Figure II.2. Behavioral diagnosis and intervention design



The email developed for Phase 1 (Figure II.3) did the following:



Made retirement vivid and emphasized the concrete benefits of retirement savings. Studies suggest that helping individuals visualize themselves in the future makes them more likely to save.¹⁴ We used a graphic that contrasts a retirement with many consumption options against one that has fewer options (Figure II.3). The email also included a link to the TSP's Take Five educational resources, which explain the benefits of beginning to save early.



Included both positive and negative framing by highlighting the “free money” that employees “miss out on” if they contribute less than 5 percent to the TSP. This message draws on behavioral research suggesting people are more strongly motivated to avoid losses (loss aversion) than to achieve equal-sized gains^{15, 16} as well as studies indicating that people tend to consume as much of a free good as is socially acceptable.¹⁷



Encouraged employees to “act now” by describing how to change their contribution level in five simple steps. An email attachment uses images of the DOL payroll website to illustrate these steps (Figure III.4). Studies show that we tend to focus on the present moment¹⁸ and delay completing tasks that seem unpleasant.^{19, 20} Showing the ease of changing one's contribution can counter the tendency toward procrastination.²¹



Reduced actual and perceived hassle by giving people easy pathways to access the TSP, ask for help, and recover their passwords. Although steps such as logging in with a password may seem minor, research shows that even small tasks can be perceived as large hassles that discourage individuals from completing an action.^{22, 23, 24}

In Phase 2, we tested two additional strategies:



We included a variant of the original email that added descriptive social norms by informing employees that the majority of their coworkers were currently saving 5 percent or more for retirement (Appendix A, Exhibit A.2). This message drew on research showing that social pressure can have a strong influence on retirement savings behavior²⁵ and that people tend to be more comfortable choosing options that everybody else is selecting.²⁶ Since at least one study had shown that effects can be negative if the savings rate highlighted in such messages is lower than what people had believed it to be (triggering a “boomerang” effect), we chose to evaluate whether framing the norm to have a more general tone would be more successful.²⁷ The language we used was as follows: “Most DOL employees contribute 5 percent or more to their TSP retirement accounts. You can join them by easily changing your contribution right now to make sure you're getting the money you've earned.” (see Appendix A, Exhibit A.2).



We sent reminder emails to employees who failed to take action after the initial email. Studies suggest that reminders can increase the salience of savings-related decisions and increase short-term and possibly long-term savings rates.²⁸ Our reminders also emphasized the reversibility of employees' savings decisions (Appendix A, Exhibit A.3). Research shows that people like to reserve the right to have many options²⁹ and are anxious when making seemingly big decisions.

Figure II.3. Original email text for Phases 1 and 2

[Subject line: You are missing out on free money every pay period – here’s how to get it]

How will you spend your retirement?



Emphasizes concrete benefits of retirement savings using consumption framing

Combines negative and positive framing

A simple choice now can give you more choices to enjoy in retirement. But if you aren't contributing at least 5% to the Thrift Savings Plan (TSP), you are missing out on free money every pay period! Each time you contribute money to your TSP retirement account, DOL will match your contributions up to 5% of your salary. Getting the full match today may open up more choices for you to enjoy in retirement. The sooner you start, the more you can make your money work for you.

Illustrates ease of changing contribution

You can easily change your contribution right now to make sure you're getting the money you've earned.

How to quickly change your contribution now

1. Log into your [Employee Personal Page](#) and click "TSP" on the left hand side. Forgot your user ID or password? Click [here](#).
2. Click "self-service" in the upper right corner.
3. Click "change," then enter the percent you want to contribute (for example, 5%) and the pay period you would like to start your new contribution amount. *****Don't worry, this decision isn't final. You can change your contribution at ANY time.*****
4. Click "continue," then, "yes," and then "submit" to finish.
5. Sit back and watch your money grow!

Reduces real and perceived hassles by providing links and guidance

Prompts to act now

Remember, setting aside a little money today can [make a big difference in the future](#). Start [now](#) by trying at least 5% this pay period.

Figure II.4. Original email attachment for Phases 1 and 2: EPP instructions

Just a few clicks can give you more choices to enjoy in retirement.

Complete these **5 simple steps** now to get your full match from DOL and avoid missing out on free money. Don't worry, this decision isn't final. You can change your contribution at ANY time.

1. Log in to your [Employee Personal Page](#), click "TSP," then "self-service" to get started.
2. Click "change" to adjust your contribution.
3. Enter the amount or percent you want to contribute, when you want it to start, and click continue.
4. You're almost done. Check over your contribution then click "yes."
5. Click "submit" and watch your money grow!

Supporting implementation

In addition to developing prototypes of the emails that were sent, we provided additional planning and implementation support to help maximize the intervention's effectiveness. This included working with EBSA and DOL-HR to determine how and when to send these email communications.

Logistics. We determined that having DOL-HR send the emails would make them more credible and noticeable—they owned the mailing lists and had in-house capacity to send out mass mailings. It would also maximize the odds of employees opening the emails, because the intervention was conducted soon after the government employee information database was hacked and employees were likely to be wary of unsolicited emails about their finances. We worked with DOL-HR to optimize the format of the emails for their email system. We also helped identify the randomly selected list of DOL employees to whom they should be sent.

Timing. We drew on marketing literature and input from DOL and EBSA partners to determine the optimal times in the week and day to send the emails.³⁰ We avoided Monday and Friday, when employees were likely to be focused on beginning or wrapping up their work week. We sent emails first thing in the morning so people could attend to them when they are focused on dealing with incoming email. We also examined the DOL-HR and TSP calendar of events to make sure that our intervention would not coincide with open season for benefit changes, when participants were likely to be receiving information from other sources about their benefits and when our implementing partners were likely to be very busy.

As we discuss in later chapters, despite this planning, delays in getting departmental approvals led to the intervention coinciding with multiple potentially influential events in Phase 1 that were not anticipated or salient during trial design. These included the following:

- **Pope's visit to Washington DC.** The email was scheduled to be released on the day of the Pope's visit. As that date approached, DOL-HR alerted us that extended security arrangements impeding access to DOL offices were likely to result in many DOL employees telecommuting or taking leave on that day.
- **A potential government shutdown due to an impasse in budget negotiations in Congress.** The email was sent around the time when people were worried about a potential shutdown of government for an indefinite amount of time. DOL-HR and EBSA reported that employees appeared to be worried about a shutdown. The previous government shutdown in 2013, due to an impasse in budget negotiations, had resulted in government employees temporarily losing pay. In subsequent months, some government departments had to place staff on furlough. The specter of a government shutdown was likely to be accompanied by fears of personal financial loss that might make employees less receptive to saving for the future.
- **Religious holidays.** Jewish and Muslim holidays occurred just a few days before the Phase 1 email was sent, increasing the likelihood of some employees being out of the office or just catching up on work after an absence. In either event, they might be less likely to attend to nonurgent emails.

The possible effect of these events on the trial's Phase 1 results was one factor that motivated fielding a second test of the original intervention six months later. In the next chapter, we discuss how we designed and implemented the evaluation to assess the effectiveness of both phases of email interventions.

III. Evaluation Design

To determine whether these emails worked, we designed and implemented a two-phase experimental trial. Phase 1 was designed to assess the effects of sending a single behaviorally informed email urging DOL employees to fully participate in the TSP. Phase 2 was designed to examine the effectiveness of an alternate initial email that invoked social norms and of adding a timely reminder. We also used this opportunity to learn about the effect of trial timing on the intervention’s efficacy. This chapter describes our evaluation design and analysis approach; Appendix B provides additional technical details.

An experimental design

Conducting a randomized controlled trial allowed us to learn whether our behaviorally informed emails caused DOL employees to participate in the TSP and/or contribute more. Random assignment ensures that members of the *treatment* and *control* groups have similar observable and unobservable characteristics, on average. Since the only difference between them is whether they received the intervention (in this case, our emails), we can safely assume that any observed differences in their outcomes were caused by the intervention and not other factors that we may have failed to take into account. For EBSA, these other factors could include changes in the employees’ actual or expected earnings, differences in the characteristics of employees who choose to participate versus those who do not, or the effects of additional efforts by the TSP or others to increase savings.

Study sample. Our study sample consisted of Federal Employees Retirement Services (FERS)–covered DOL employees who were not contributing to TSP (noncontributors) or were contributing less than 5 percent of their salary (undercontributors). This included 4,078 individuals in Phase 1 and 2,363 employees in Phase 2. Study participants included federal DOL employees nationwide—in both the national and regional offices. (See Appendix B for more details on the study sample.)

Table III.1. Sample description and sample sizes at random assignment for each phase

	Study participants	Sample size at random assignment
Phase 1 Fall 2015	All Federal Employees Retirement Services (FERS)–covered employees who were either not contributing or contributing less than 5 percent	4,078 (T1: 2,039; C: 2,039)
Phase 2 Spring 2016	All FERS-covered employees who had not received an intervention email in Phase 1 and were still contributing less than 5 percent of salary to the TSP, as well as recent hires and others who began contributing less than 5 percent after the end of Phase 1	2,363 (T1: 788; T2: 787; C:788)

Random assignment. We conducted random assignment at the individual level. In the fall of 2015, DOL shared an anonymized list of employees who were not participating in TSP or who were participating but contributing less than 5 percent of their salary, as well as baseline information on employee contributions to TSP, salaries, locations, hire dates, ages, genders, ethnicities, and their union membership (referred to as “bargaining unit status”). We first vetted these lists to make sure we retained only those employees who were noncontributors and undercontributors. Then we conducted a stratified random assignment to ensure balance on key variables (See box below and read more about stratification rationale and procedures in Appendix B, Section I.C.).

Stratified random assignment ensured balance on:

- Annual salary
- Hired date
- Union membership
- Whether employees were contributing to the TSP at all at the trial’s start

Study groups. In Phase 1, we randomly assigned DOL employees to one of two groups: (1) a treatment group that received a single email encouraging participation in the TSP program and (2) a control group that did not receive this email (See Figure III.1). In Phase 2, we randomly assigned individuals in the control group from the first phase—as well as any new hires and other DOL employees saving less than 5 percent—to three groups:

1. **T1 treatment group**, which received the original email plus a reminder email six weeks later;
2. **T2 treatment group**, which received a variant of the original email with an additional social norm message, plus a reminder email; and
3. **Control group**, which did not receive any of our emails and experienced business as usual.

In addition, the treatment groups in Phase 2 had longer to respond to the intervention, relative to Phase 1 participants, before we gathered final TSP contribution data. In Phase 1, endline data were gathered 37 days after intervention emails were sent; in Phase 2, endline data were gathered 66 days after the intervention emails were sent (Table III.2 below). The longer timeline for Phase 2 was because the intervention included two rounds of emails instead of one.

Outcomes of interest. We examined the effects of our intervention emails on the following outcomes:

- **Contributes to TSP:** whether an individual contributes any amount to TSP (for noncontributors only)
- **Contributes at least 5 percent to TSP:** whether an individual contributes at least 5 percent of their salary to TSP (for the full sample)
- **Contribution rate:** the individual’s contribution rate as a percentage of their salary (for the full sample)

DOL-HR provided administrative data on these outcomes at four points during each trial phase: once before sending the emails, then at three points in time after the original email was sent. For Phase 2, two of these rounds of data collection happened after the reminder email was sent. The timelines for data collection in each phase are shown below.

Table III.2. Data collection timelines for each phase

Activity	Phase 1 Dates	Phase 2 Dates
Collected baseline data	9/5/2015	3/19/2016
First email sent	9/24/2015	4/5/2016–4/6/2016
Collected first round of outcome data (preliminary)	10/3/2015	4/16/2016
Collected second round of outcome data (midline)	10/17/2015	5/14/2016
Reminder email sent	Not applicable	5/17/2016
Collected third round of outcome data (endline)	10/31/2015	6/11/2016

For Phase 1 only, we were also able to measure exposure to our email, by examining read receipts, out of office responses, and deletions without reading. A summary of these measures is provided in Appendix D and indicates that roughly a quarter of employees who were sent the email in Phase 1 opened it and returned a read receipt. DOL-HR also reported back that they did not have a valid email address on file for roughly 5 percent of the employees targeted by the trial.

A rigorous analytic approach

Below we describe the analyses we conducted. Even when random assignment procedures are followed carefully, events may compromise the equivalence of the study sample and potentially bias findings. We confirmed that attrition in the study sample was low and that the study groups were equivalent at baseline.

Checking for attrition. We began by assessing whether attrition might have compromised our random assignment design. Attrition happens when members of the initial research sample that had been randomly assigned are not part of the final analysis sample (In this example, this could occur because employees leave DOL, or data on their TSP contributions are not available in follow-up data). As discussed in Appendix B., Section I.C, we determined that overall attrition rates were low—approximately 1.6 percent in Phase 1 and 2.6 percent for Phase 2—and differential attrition was also minimal. Our final analysis sample sizes are shown in Table III.3.

Table III.3. Final analysis sample sizes by phase

Phase	Total	T1	T2	C
Phase 1	4,011	2006	NA	2005
Phase 2	2,302	764	765	773

Confirming baseline equivalence. Even though attrition rates were low in our random assignment study, we checked whether the groups were similar in observable baseline characteristics (Appendix C provides baseline equivalence results). In Phase 1, none of the baseline characteristics we examined differed significantly. In Phase 2, the proportion of employees in the national office was larger in the control group than in either of the two treatment groups—a difference that could have occurred by chance. There were no other significant differences among the Phase 2 study groups.

Estimating impacts. To estimate impacts and answer the research questions we identified in the first chapter, we conducted eight analyses on the full samples and a number of subgroup analyses. Table III.4 below shows the comparisons we conducted to answer each research question (For details on our estimation models, see Appendix B.II).

Table III.4. Research questions and relevant analyses

Research Questions	Relevant analyses
<p>Can emails informed by behavioral science encourage more DOL employees to increase their savings? Do they prompt DOL employees to increase their contributions enough to receive the full employer match? Do they prompt employees not participating in TSP to start contributing?</p>	<p>Analysis 1: T1 in Phase 1 (no reminder) vs. control Analysis 2: T1 in Phase 2 after reminder email vs. control Analysis 3: T2 in Phase 2 after reminder email vs. control</p>
<p>What email communication strategies are most effective at improving savings?</p> <ul style="list-style-type: none"> -Are emails that invoke peer norms more effective than emails that don't? -Does sending reminder emails change the impact? -Does the timing of emails matter? 	<p>Analysis 4: T2 vs. T1 in Phase 2 Analysis 5: T2 in Phase 2 before reminder vs. control Analysis 6: T1 in Phase 2 before reminder vs. control Analysis 7: Pooled Phase 2 treatment after the reminder vs. before the reminder Analysis 8: T1 in Phase 2 vs. T1 in Phase 1</p>
<p>For whom do these emails work? Do impacts vary for different employee subgroups—for example, older versus younger employees, men versus women?</p>	<p>Effects among subgroups (described below)</p>
<p>What lessons might help inform similar efforts in other contexts? What infrastructure was necessary to send these targeted communications? What considerations were important when seeking to evaluate these strategies?</p>	<p>Review of implementation findings from study team</p>

Unlike the remaining analyses, analyses 7 and 8 in Table III.4 are non-experimental and do not leverage random assignment (See Appendix Table B.5 for additional detail on the analyses). They are simply exploratory analyses that are designed to provide additional information about how the intervention worked. Analysis 7 compares impacts of the interventions before and after the reminders, and we cannot separate out the effects of the reminders from the lagged effects of the intervention itself. Similarly, Analysis 8 compares the effects of the original email in Phase 1 with that in Phase 2, but since the analysis sample is defined in part based on employees’ decisions following random assignment, we cannot interpret any differences in impacts as the effect of timing alone. (In particular, control group employees who left DOL or increased their TSP contributions above 5 percent before Phase 2 baseline were not included in the Phase 2 sample.)

In addition to examining the effects of our emails for the full study sample, we examined their effects among subgroups with the following characteristics:

- Salary above and below the median for DOL employees
- Age above and below the median for DOL employees
- Located in the national office or elsewhere
- Hired before or after August 2, 2010 (when auto-enrollment was introduced)
- Member of a union or not
- Noncontributor vs. undercontributor to TSP

IV. Findings

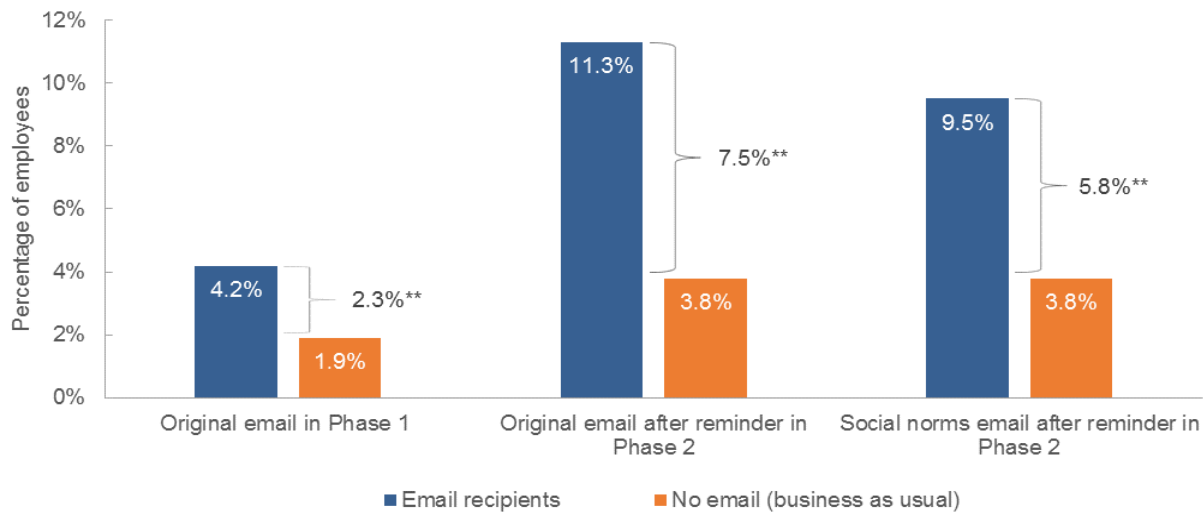
Results for both phases of our trial provide strong evidence that these interventions encouraged more individuals to save and thus enhance their retirement security. Below we summarize what impacts these interventions had and for whom (Appendix D provides detailed impact estimates).

LOW-COST, BEHAVIORALLY INFORMED EMAILS INCREASED THE NUMBER OF DOL EMPLOYEES RECEIVING THE FULL EMPLOYER MATCH BY UP TO 7.5 PERCENTAGE POINTS.

Our findings indicate that emails informed by behavioral science encouraged more DOL employees to avail themselves of the full employer match and increased overall TSP contribution rates.

Both emails increased the number of DOL employees receiving the full employer match. Both variants of our intervention email (the original email [T1] and the social norms variant [T2]) had statistically significant impacts on the proportion of employees contributing at least 5 percent of their income to TSP in both trial phases (see Figure IV.1). The original email improved this outcome by 2.3 percentage points when sent on its own in Phase 1 and by 7.5 percentage points when combined with a reminder and given more time to affect contribution levels in Phase 2. Similarly, the social norms variant, which also included a reminder email, improved this outcome by 5.8 percentage points in Phase 2.

Figure IV.1. Impact of original email and social norms variant on percentage of employees who started contributing at least 5% to TSP

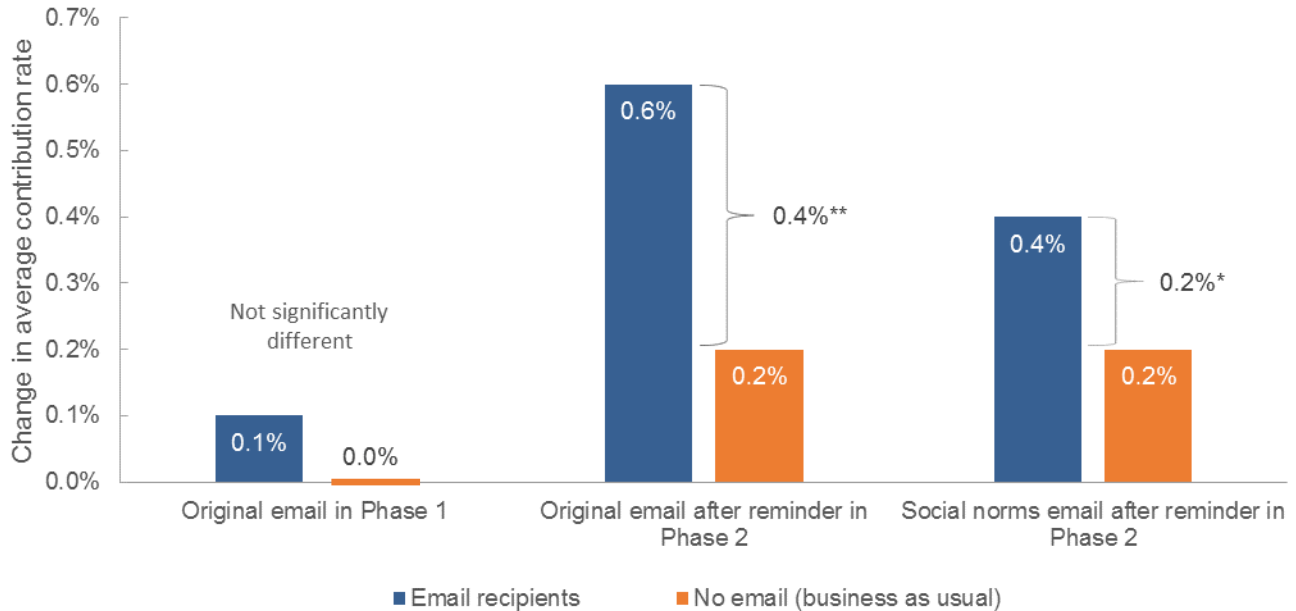


Source: DOL-HR administrative data. See Appendix D for detailed tables.

Note: Significance levels: * $p < 0.05$, ** $p < 0.01$. Impact may not equal the difference shown between treatment and control due to rounding. Note that the effects of the original email in Phase 1 and the original email in Phase 2 should not be directly compared because the sample composition and contextual factors differs across the two phases.

These interventions also increased the overall contribution rates in Phase 2. In Phase 2, the original email and the social norms variant increased contribution rates by 0.4 percentage points and 0.2 percentage points, respectively (see Figure IV.2). The original email had positive results for this outcome in Phase 1, but these were not statistically significant.

Figure IV.2. Impact of original email and social norms variant on overall contribution rates

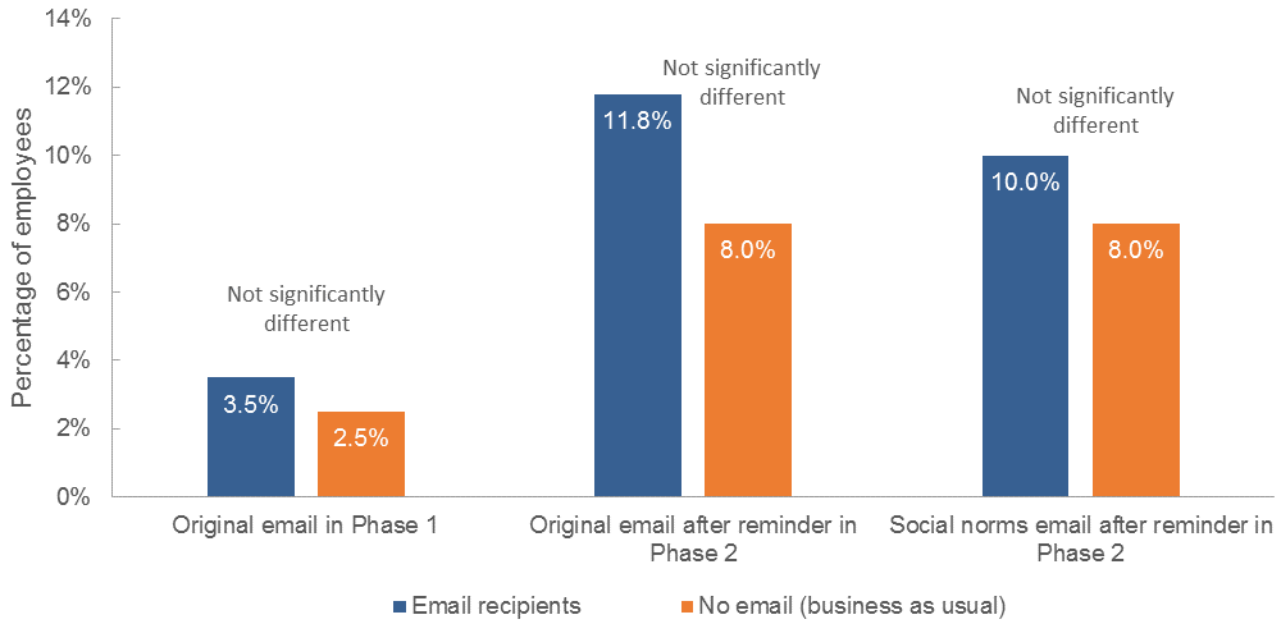


Source: DOL-HR administrative data. See Appendix D for detailed tables.

Note: Significance levels: * $p < 0.05$, ** $p < 0.01$. Impact may not equal the difference shown between treatment and control due to rounding. Note that the effects of the original email in Phase 1 and the original email in Phase 2 should not be directly compared because the sample composition and contextual factors differ across the two phases.

Intervention emails did not result in employees contributing to the TSP if they were not already doing so. Although treatment group members for the original email in Phases 1 and 2, and for the social norms variant in Phase 2, were between 1.0 and 3.9 percentage points more likely to start contributing to the TSP than control group members, none of these results were statistically significant (see Figure IV.3).

Figure IV.3. Impact of original email and social norms variant on percentage of employees who started contributing to TSP (any level)



Source: DOL-HR administrative data. See Appendix D for detailed tables.

Note: Significance levels: * $p < 0.05$, ** $p < 0.01$. Impact may not equal the difference shown between treatment and control due to rounding.

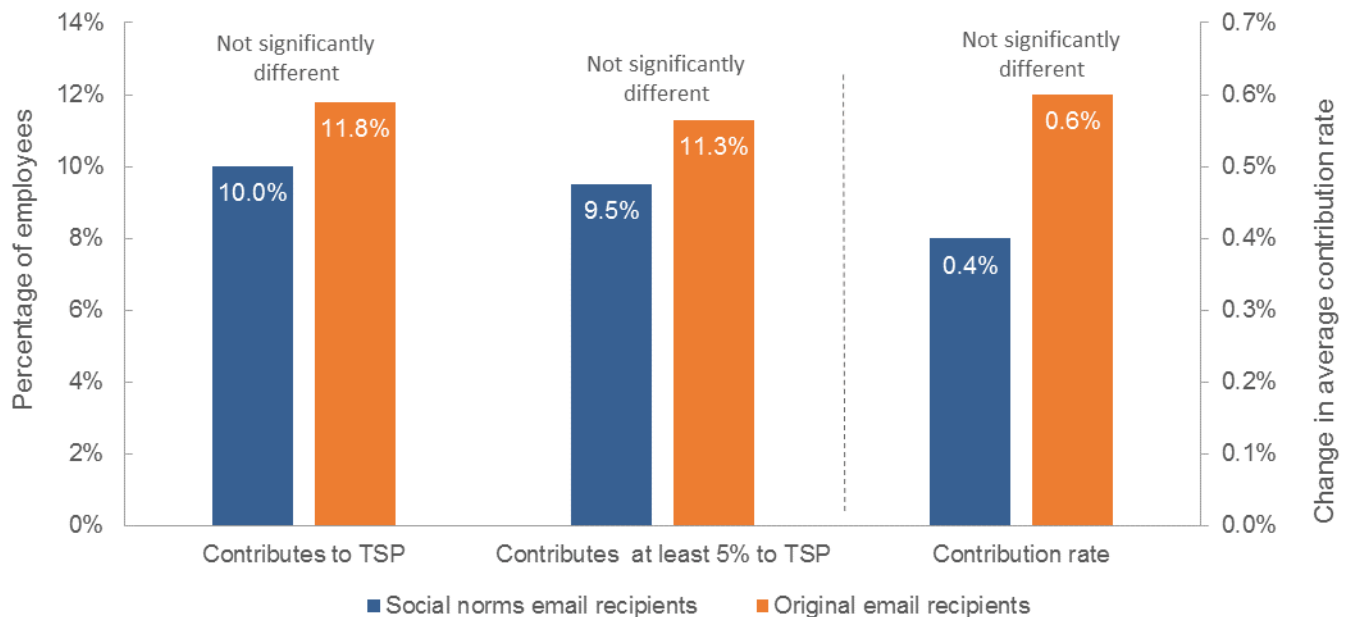
These findings are worth comparing to those of the DOD-SBST experiment designed to increase military service member participation in the TSP (see Box I.2). As in our trial, the DOD-SBST experiment succeeded in changing employees’ TSP retirement savings choices. It increased enrollment in the Roth TSP among nonenrolled service members. Although our intervention did not cause people to start contributing to TSP if they were not already doing so, it did cause those already contributing to increase their savings to at least 5 percent. Notably, the magnitude of results we found was much larger. Although the most effective of the eight emails tested in the DOD-SBST trial increased the number of military service members participating in the TSP by 2.1 percentage points, the most effective version of the interventions we tested with DOL employees increased the proportion of employees receiving the full employer match by 7.5 percentage points.

One reason our trial was relatively less effective among noncontributors may be the baseline participation rates. DOL employee participation rates for the TSP at baseline were almost twice those of DOD’s, leaving a smaller potential margin for improvement and raising the possibility that those not participating were those who were the most difficult to influence and/or who may face greater financial or other barriers to saving.

INVOKING SOCIAL NORMS DID NOT APPEAR TO INCREASE THE MESSAGE'S EFFECTIVENESS.

In Phase 2 of the trial, we tested the original intervention email against a variant that included language invoking social norms (see Appendix A, Exhibit A.2). In Phase 2, the original email (without this social norms language) increased impacts by a greater margin than the social norms alternative, but these differences were not statistically significant and could have been due to chance alone (see Figure IV.4). This finding was observed consistently for subgroups that we examined (see subgroup findings in Appendix D, Table D.7).

Figure IV.4. Comparison of impacts for the original email versus the social norms variant in Phase 2 after reminder



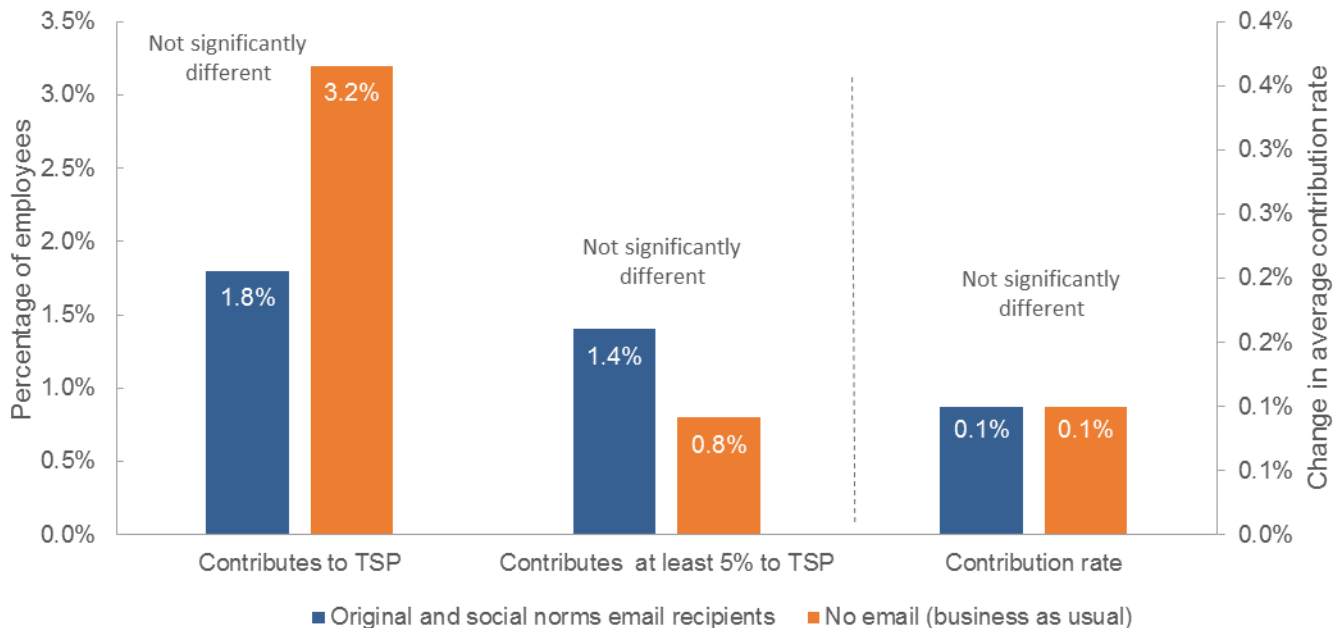
Source: DOL-HR administrative data. See Appendix D for detailed tables.

Note: Significance levels: *p<0.05, **p<0.01. Impact may not equal the difference shown between treatment and control due to rounding.

REMINDERS DID NOT APPEAR TO INCREASE THE EFFECTS OF THE INITIAL EMAILS.

In Phase 2, we paired each variant of the email message with a reminder sent six weeks after the original email. We gathered outcome data before and after the reminder for each intervention and conducted an exploratory analysis to assess whether the reminders seemed to increase the impact of the original emails. To isolate the effects of the reminders, we pooled the samples for both Phase 2 interventions (that is, the samples for the original email and the social norms variant) and compared the outcomes for this pooled sample to the outcomes for the Phase 2 control group between our midline follow-up (before reminders were sent) and endline follow-up (after reminders were sent). As Figure IV.5 shows, we found no statistically significant differences (See Appendix D, Tables D.3–D.4 for additional analyses with similar results).

Figure IV.5. Effect of Phase 2 reminders



Source: DOL-HR administrative data. See Appendix D for detailed tables.

Note: Significance levels: *p<0.05, **p<0.01. Impact may not equal the difference shown between treatment and control due to rounding.

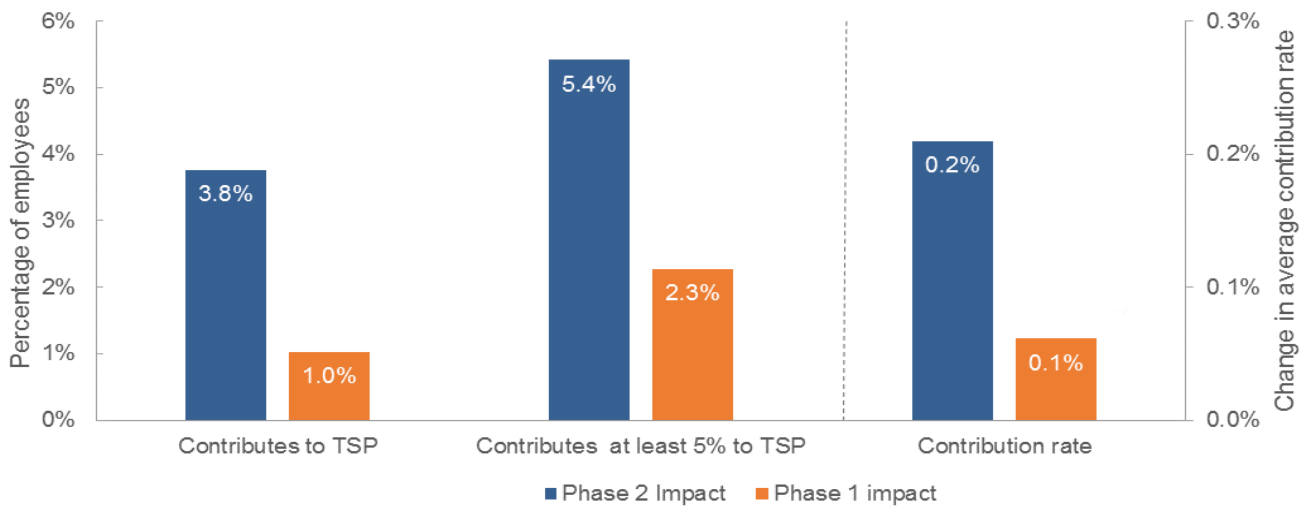
TIMING LIKELY INFLUENCED THE EMAILS' IMPACT. THE PHASE 2 EMAILS HAD LARGER IMPACTS ON ALL SAVINGS OUTCOMES EXAMINED.

As indicated in Figures IV.1–3, we generally measured larger impacts in Phase 2 than in Phase 1, even when considering the same email message. These differences appear to be driven by changes in the circumstances of each trial as well as changes in the composition of the sample between phases. (Phase 2 trial participants were more likely to be long-tenure employees.)

To understand the effect of timing—that is, when in the year the email was sent—we compared the impact of the original email at the end of Phase 1 to its impact in Phase 2 before the reminder email. To reduce the influence of changes in sample composition, we excluded from the Phase 1 sample those employees who left DOL before the end of Phase 2, and we excluded from Phase 2 those employees who were not present in the Phase 1 sample. (These are primarily new DOL-hires or people who became eligible for the trial's second phase by reducing their TSP contributions below 5 percent.)

Phase 2 impacts were larger than those for Phase 1 for all outcomes studied. Our exploratory analyses suggest that in Phase 2, the original email increased the proportion contributing any amount by 2.8 percentage points more than in Phase 1. Similarly, in Phase 2, the original email increased the proportion contributing at least 5 percent by 3.1 percentage points more than in Phase 1 and increased the contribution rate by 0.1 percentage points more than in Phase 1.

Figure IV.6. Impacts by phase, limiting differences in sample composition



Source: DOL-HR administrative data. See Appendix D for detailed tables.

Note: Because this analysis is exploratory in nature, tests of statistical significance are not reported.

Although these findings are not causal, they do suggest that the timing of our trials may have influenced the magnitude of the impacts found. In Phase 1, during the week that the email was sent, the Pope visited Washington DC, and many federal employees may have worked from home or taken leave. Given that 31 percent of the study sample consisted of federal employees in the Washington DC area and approximately 5 percent of the treatment sample sent an “out of office” status message in response to receiving the email, the Pope’s visit may have affected the treatment group’s responsiveness. As discussed later in this chapter, regional office staff had higher response rates relative to national office staff in Phase 1 but not in Phase 2.

In addition, news stories that speculated about the possibility of a government shutdown during congressional budget discussions were circulating around the time of the Phase 1 trial. This may have caused members of the study sample to be more preoccupied with their immediate financial security and less able to focus on saving for the future. It may have also led to a preference for keeping their assets in more liquid accounts, rather than retirement savings accounts.



THE EFFECTS OF OUR EMAILS WERE STRONGEST FOR THOSE ALREADY CONTRIBUTING AT LEAST SOME AMOUNT TO THE TSP AND FOR DOL EMPLOYEES WHO WERE UNION MEMBERS.

We looked at different subgroups to find out whether this intervention works better for some populations than for others (see Chapter III for subgroups examined and Appendix D, Table D.7 for summary results). In aggregate, the patterns we observed in the overall findings persisted among subgroups: The reminder email did not appear to have a significant marginal effect in any subgroup, and Phase 2 impacts were higher than Phase 1 impacts across all subgroups.

Our emails’ effects were concentrated among those who were already contributing some amount at baseline and among DOL employees in bargaining units. In Phase 1, younger employees and employees outside the national office appeared to be more responsive to the intervention, but this finding did not persist in Phase 2. This may have been because one of the contextual events described for Phase 1—the Pope’s visit—affected employees only at the national office. Interestingly, there were no notable patterns associated with employees’ salary levels or whether they were hired after the automatic enrollment policy was enacted.

V. Discussion and Lessons Learned

This chapter discusses the results we found, how they fit in with findings from similar trials, and what they suggest for next steps. We also examine the lessons we learned in the course of developing and implementing this trial.

Discussion of results

Our results suggest that sending targeted emails that draw on insights from behavioral science can be a compelling and low-cost strategy for promoting retirement security. These emails encouraged those DOL employees who were already participating in the federal TSP program to increase their contributions and maximize their employer match.

Low-cost behaviorally informed emails generated substantial impacts. When timed right and the most effective variant of the intervention was used, these impacts were as large as 7.5 percentage points, doubling the rate at which employees accessed the full employer match. This is a sizable impact, especially considering that the large proportion of DOL employees who already participated in the TSP and received the full employer match left little room for improvement. These impacts also exceed those observed in the DOD-SBST trials targeting TSP enrollment among military service members.

Impacts on contribution rates could translate into meaningful savings over the long term. Knowing the average increase in overall contribution rates (0.4 percent) allows us to roughly estimate the potential value of savings resulting from our intervention. If the 764 DOL employees who received the original intervention email in Phase 2 maintained this level of savings until retirement, they could gain a meaningful amount of savings in 20 years—approximately \$11,500 per person on average. This translates to an additional \$8.8 million in savings for these email recipients. (This estimate applies the 0.4 percent overall contribution rate increase to the median salary of \$78,162 that we observed among email recipients for 20 years, assuming a 6 percent annual rate of return—see Appendix D for further details on these estimates.)

Our interventions have broad relevance and scalability. Our results are especially promising because of the low-touch, low-cost nature of the intervention. Many employers already have the capability of sending mass emails and are likely accustomed to doing so. There is also a growing menu of affordable technology options for improving how these mass emails are managed (including personalizing them and tracking responses). These points suggest that implementing such strategies and examining their results may be doable for many employers and that these findings may be relevant for a broad spectrum of people interested in improving retirement security.

Further investigation on what made these emails effective may be useful. In reflecting on the magnitude of the impacts we observed, we are cognizant that these emails may have been especially effective because they are the first of their kind received by DOL employees. Prior to this intervention, DOL employees had not received any emails focused exclusively on TSP retirement benefits. Receiving a TSP-specific email for the first time may have helped make retirement and the benefits of the program more salient in the minds of those already contributing to it and may have prompted them to increase their contributions. Research shows that salience rather than self-control determine procrastination and spending behavior, especially when it comes to saving.³¹

However, it is important to recall that in addition to making retirement and the TSP program more salient, our original intervention included many other important behavioral strategies. These included careful targeting to a clearly defined audience, including visuals that helped people picture their choices in retirement, a simple and clear description of the full match, language urging recipients not to lose out on matched savings, making action steps clear, and providing visual aids and useful links to reduce hassle factors.

Because these strategies were bundled together, we cannot disentangle which specific elements worked or were most important—all we know is that together they produced significant and substantial results for people already contributing to the TSP. For interested agencies or employers that have the opportunity to test emails with large numbers of employees, comparing different elements of the strategies we adopted may be worthwhile. In addition, understanding the reasons why some employees are not saving for retirement and continuing to look for effective strategies to engage them will be important. Finally, examining how the effectiveness of such emails changes as they are used more frequently may also be valuable.

Lessons learned

The DOL-BI team drew on our experiences with this and other trials we conducted for DOL to identify lessons learned. For those who might be interested in pursuing such tests, we have further reflected on implementation lessons that we learned in the course of developing this trial. (For a broader discussion of the implementation lessons learned from this trial and two other trials conducted as part of this project, please see the associated implementation findings report.³²)

Our email interventions were relatively quick to develop and implement. Once we pinned down the goals of the intervention and the target population, a series of targeted meetings and document reviews enabled us to develop an intervention rapidly. Reviewing all available documentation and data and identifying outstanding questions pertinent to trial design allowed us to engage efficiently with partners.

Email interventions can often be readily tailored to implementer constraints. Being flexible and responsive to the constraints faced by our implementing partners was critical. We tailored both the email itself and its implementation procedures to these constraints. DOL-HR and EBSA faced many competing priorities at the time this trial was underway, including operationalizing new policies to tighten oversight of benefits advisors. We carefully navigated the tradeoffs between increasing email effectiveness and minimizing burden on staff. So we prioritized those features that were essential and did without those that would be nice to have but strain staff resources. For example, personalizing emails by using a first name is known to improve response rates but would have been much more labor-intensive than sending out a mass email without a salutation, so we opted for the latter.

When the ideal fix was not an option, we reached for alternate solutions. Ideally, we would have liked to get employee perspectives on benefit choices and benefit elections directly. DOL-HR staff were rightly reluctant to take any actions that might exacerbate employee concerns about their privacy. So, instead, we adopted a user perspective in thinking about the benefit change process and conducted a walkthrough of the online payroll system. We tapped Chief Evaluation Office, DOL-HR, and EBSA staff perspectives to learn about what they had heard from other DOL employees and their own personal

experiences as beneficiaries of the system. Finally, we provided a script to DOL-HR staff to enable them to conduct user testing on our behalf.

Low-tech solutions can sometimes be an effective stand-in. In our walkthrough of the online payroll system, we found that navigation could be a challenge (links and navigation buttons were hard to find and not always intuitively labeled or placed). Making changes to the TSP system was not feasible and could not be randomly assigned. We did not have the resources to develop a video tutorial either. So as the best alternative, we developed a one-page visual guide using screenshots. User testing suggested that these were well received.

Capitalizing on opportunities to learn when changes are introduced can pay large dividends. Given our results from Phase 1, DOL-HR was committed to sending our email to the control group. We recognized that as an opportunity to explore lingering questions about the optimal design of the emails. We proposed testing of reminders and social norms and highlighted the potential to use the second phase to learn about the effects of timing. Our efforts were unexpectedly rewarded by the same intervention yielding results that were almost four times larger.

Different groups may need different approaches. Knowing that noncontributors are less responsive to these emails on savings compared with undercontributors, DOL-HR and EBSA can now focus on exploring further whether noncontributors' choices can be influenced by alternate strategies. The first step could be to further investigate whether the reasons why noncontributors do not save are structural (such as financial constraints) or behavioral (such as over optimism about the future). If the latter, it may be worth testing alternative approaches (such as in-person follow-up from staff, more frequent and personalized communications, or committing to "save more tomorrow"). Whichever strategy is selected, it would be invaluable to test it rigorously. With more evidence on whether behaviorally informed emails to save more are effective, and how different people respond to them, we will be able to help more American workers prepare for a comfortable retirement.

References

- ¹ Trope, Y., and N. Liberman. "Construal-Level Theory of Psychological Distance." *Psychological Review*, vol. 117, no. 2, 2010, pp. 440–463.
- ² O'Donoghue, T., and M. Rabin. "Doing It Now or Later." *American Economic Review*, vol. 89, no. 1, 1999, pp. 103–124.
- ³ Kahneman, D. T., and A. Tversky. "Intuitive Prediction: Biases and Corrective Procedures." *Management Science*, vol. 12, 1979, pp. 313–327.
- ⁴ Laibson, D. "Golden Eggs and Hyperbolic Discounting." *The Quarterly Journal of Economics*, vol. 112, no. 2, 1997, pp. 443–477.
- ⁵ Green, L., N. Fristoe, and J. Myerson. "Temporal Discounting and Preference Reversals in Choice Between Delayed Outcomes." *Psychonomic Bulletin & Review*, vol. 1, no. 3, 1994, pp. 383–389.
- ⁶ Madrian, B., and D. Shea. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics*, vol. 116, no. 4, 2001, pp. 1149–1187.
- ⁷ Dymond, S., and B. Roche. "A Contemporary Behavior Analysis of Anxiety and Avoidance." *The Behavior Analyst*, vol. 32, no. 1, 2009, pp. 7–27.
- ⁸ Data provided to the study team by U.S. Department of Labor in July 2015.
- ⁹ Thaler, R., and S. Benartzi. "The Behavioral Economics of Retirement Savings Behavior." 2007.
- ¹⁰ Frederick, S., G. Loewenstein, and T. O'Donoghue. "Time Discounting and Time Preference: A Critical Review." *Journal of Economic Literature*, 2002, pp. 351–401.
- ¹¹ Laibson, D. (1997).
- ¹² Social and Behavioral Sciences Team. "Annual Report 2015." Washington, DC: Executive Office of the President, National Science and Technology Council, 2015.
- ¹³ Goda, G. S., C. F. Manchester, and A. Sojourner. "What Will My Account Really Be Worth? An Experiment on Exponential Growth Bias and Retirement Saving." NBER Working Paper No. 17927. Cambridge, MA: National Bureau of Economic Research, 2012, 1–43.
- ¹⁴ Hershfield, H. E., D. G. Goldstein, W. F. Sharpe, J. Fox, L. Yeykelis, L. L. Carstenses, and J. N. Bailenson. "Increasing Savings Behavior Through Age-Processed Renderings of the Future Self." *Journal of Marketing Research*, vol. 48, 2011, pp. S23–S37.
- ¹⁵ Kahneman, D., J. L. Knetsch, and R. H. Thaler. "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias." *The Journal of Economic Perspectives*, vol. 5, no. 1, 1991, pp. 193–206.
- ¹⁶ Ellsberg, D. "Risk, Ambiguity, and the Savage Axioms." *The Quarterly Journal of Economics*, 1961, pp. 643–669.
- ¹⁷ Boshi, S., M. Lavie, and A. Weiss. "Quasi-Free Goods and Social Norms: The Effects of Quantity Restrictions and Scrutiny." 2014. Available at <http://ssrn.com/abstract=2574131> or <http://dx.doi.org/10.2139/ssrn.2574131>
- ¹⁸ Frederick et al. (2002) and Laibson D. (1997).

- ¹⁹ Tversky, A., and E. Shafir. "Choice Under Conflict: The Dynamics of Deferred Decision." *Psychological Science*, vol. 3, no. 6, 1992, pp. 358–361.
- ²⁰ Blunt, A. K., and T. A. Pychyl. "Task Aversiveness and Procrastination: A Multi-Dimensional Approach to Task Aversiveness Across Stages of Personal Projects." *Personality and Individual Differences*, vol. 28, no. 1, 2000, pp. 153–167.
- ²¹ Lusardi, A., P. A. Keller, and A. M. Keller. "New Ways to Make People Save: A Social Marketing Approach." NBER Working Paper No. 14715. Cambridge, MA: National Bureau of Economic Research, 2009.
- ²² Bertrand, M., S. Mullainathan, and E. Shafir. "A Behavioral-Economics View of Poverty." *The American Economic Review*, vol. 94, no. 2, 2004, pp. 419–423.
- ²³ Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu. "The Role of Simplification and Information in College Decisions: Results from the H&R Block FAFSA Experiment" (No. w15361). Cambridge, Massachusetts: National Bureau of Economic Research, 2009.
- ²⁴ Madrian, B. C. "Matching Contributions and Savings Outcomes: A Behavioral Economics Perspective." Working paper no. 18220. Cambridge, Massachusetts: National Bureau of Economic Research, 2012.
- ²⁵ Duflo, E., and E. Saez. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *The Quarterly Journal of Economics*, vol. 118, no. 3, 2003, pp. 815–842.
- ²⁶ Banerjee, A. V. "A Simple Model of Herd Behavior." *Quarterly Journal of Economics*, vol. 107, 1992, pp. 797–817.
- ²⁷ Beshears, J., J. J. Choi, D. Laibso, B. C. Madrian, and K. L. Milkman. "The Effect of Providing Peer Information on Retirement Savings Decisions." *The Journal of Finance*, vol. 70, no. 3, 2015, pp. 1161–1201.
- ²⁸ Karlan, D., M. McConnell, S. Mullainathan, and J. Zinman. "Getting to the Top of Mind: How Reminders Increase Saving." NBER Working Paper No. 16205. Cambridge, MA: National Bureau of Economic Research, 2010.
- ²⁹ Gilbert, D. T., and J. E. Ebert. "Decisions and Revisions: The Affective Forecasting of Changeable Outcomes." *Journal of Personality and Social Psychology*, vol. 82, no. 4, 2002, p. 503.
- ³⁰ Foreman, J. "Insights from MailChimp's Send Time Optimization System." 2014. Available at <http://blog.mailchimp.com/insights-from-mailchimps-send-time-optimization-system/>. Accessed August 1, 2015.
- ³¹ Karlan et al. (2010).
- ³² Lefkowitz, J., M. Darling, G. Chojnacki, I. Perez-Johnson, S. Amin, M. Manley. "Using Behavioral Science to Strengthen Labor Programs: Lessons Learned." Washington, D.C.: Mathematica Policy Research, May 2017.

APPENDIX A

BEHAVIORAL INTERVENTION DESIGN

Exhibit A.1. Behavioral map of TSP benefit change process

TSP contribution change process

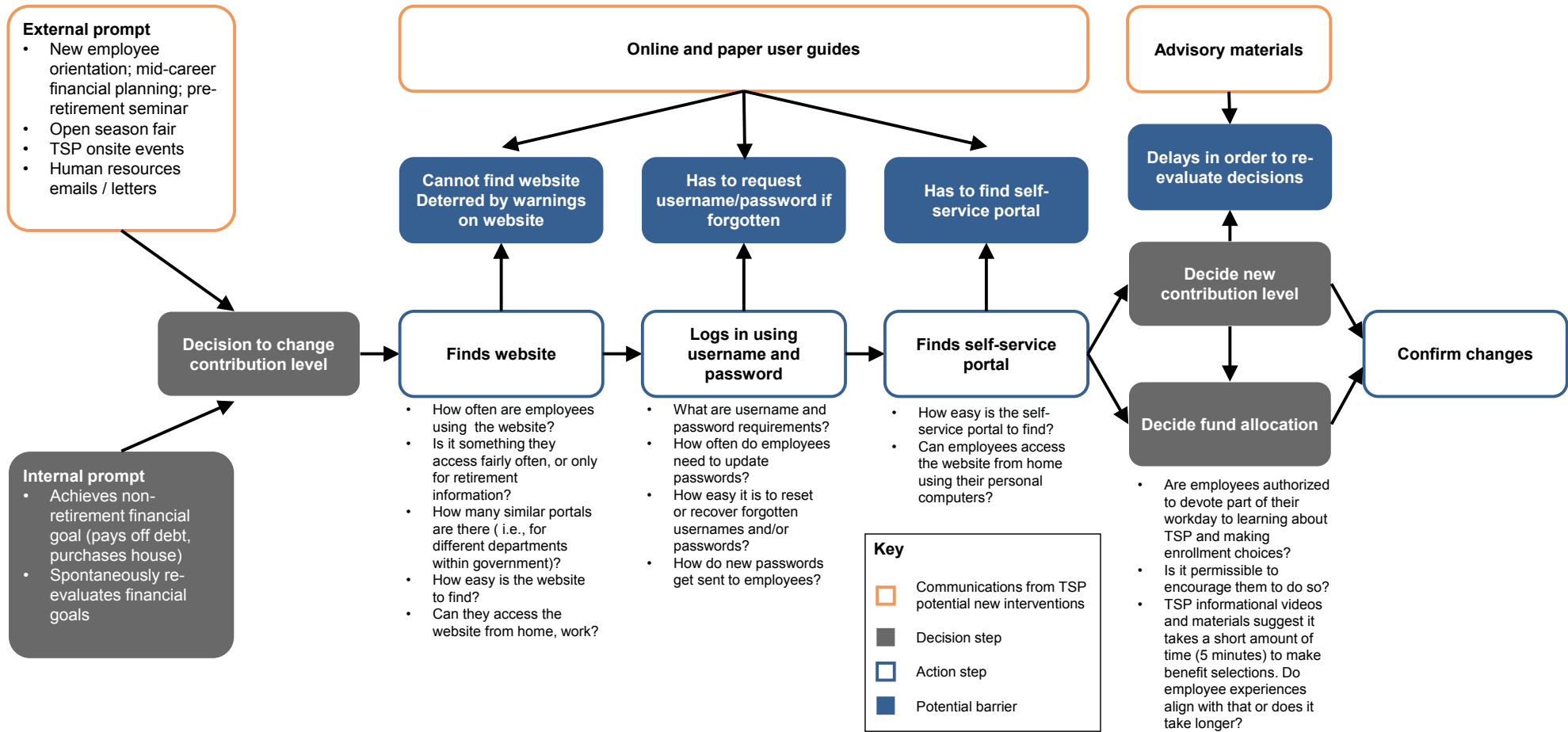


Exhibit A.2. Social norms variant email for Phase 2

Email Subject: You are missing out on free money every pay period - here's how to get it

How will you spend your retirement?



A simple choice now can give you more choices to enjoy in retirement. But if you aren't contributing at least 5% to the Thrift Savings Plan (TSP), you are missing out on free money every pay period! Each time you contribute money to your TSP retirement account, DOL will match your contributions up to 5% of your salary. Getting the full match today may open up more choices for you to enjoy in retirement. The sooner you start, the more you can make your money work for you.

Most DOL employees contribute 5% or more to their TSP retirement accounts. You can join them by easily changing your contribution right now to make sure you're getting the money you've earned.

Uses the behavior of others as an incentive

How to quickly change your contribution now

1. Log into your [Employee Personal Page \(https://www.nfc.usda.gov/epps/eplogin.aspx\)](https://www.nfc.usda.gov/epps/eplogin.aspx) and click "TSP" on the left hand side. Forgot your user ID or password? Click [here \(http://www.labornet.dol.gov/me/neow/documents/EPP-Quick-Reference-Guide.pdf\)](http://www.labornet.dol.gov/me/neow/documents/EPP-Quick-Reference-Guide.pdf).
2. Click "self-service" in the upper right corner.
3. Click "change," then enter the percent you want to contribute (for example, 5%) and the pay period you would like to start your new contribution amount. ****Don't worry, this decision isn't final. You can change your contribution at ANY time.****
4. Click "continue," then "yes," and then "submit" to finish.
5. Sit back and watch your money grow!

Remember, setting aside a little money today can [make a big difference in the future \(https://www.tsp.gov/takeFIVE/\)](https://www.tsp.gov/takeFIVE/). Start [NOW \(https://www.nfc.usda.gov/epps/eplogin.aspx\)](https://www.nfc.usda.gov/epps/eplogin.aspx) by trying at least 5% this pay period.

Need more help navigating the Employee Personal Page? Use the attached guide that shows you how to change your contribution level. You can also contact your servicing human resources office if you have any further questions.

*Office of Worklife, Leave and Benefits Policy & Programs
Human Resources Center*

Exhibit A.3. Reminder text for original email

Kept action salient by sending reminders

Email Subject: FW: You are missing out on free money every pay period – here's how to get it

On April 5th, we sent you an email about the benefits of increasing your TSP contributions. If you aren't contributing at least 5%, you are missing out on free money every pay period.

Don't forget that you can change your contribution at any time.

How will you spend your retirement?



[Followed by whichever version of the email (original or social norms variant) the individual had received in the first place].

APPENDIX B

EVALUATION DESIGN AND ANALYTIC METHODS

This appendix provides technical details on our evaluation design, data, and analytic methods, to supplement the information provided in Chapter 2.

I. Experiment Design

A. Two phase approach

The trial is a randomized controlled trial divided into two phases. In Phase 1, we tested the effects of an email intervention designed to encourage employees to increase their retirement savings by conducting a randomized controlled trial among all DOL employees contributing less than 5 percent of their salaries to the TSP. In Phase 2, we examined the effects of a modified email that included messaging that invoked social norms and sent a reminder email approximately 6 weeks after the initial email was sent. To compare the effects of the second email to the original email under the same circumstances, we also sent the original email to a separate randomly-selected group of employees in the second phase, and sent reminders to recipients of that email.

B. Data and sample selection

In the trial, we used administrative data obtained from DOL human resources records on all Federal Employees Retirement System–covered DOL employees who were contributing less than 5 percent of their salaries to TSP. This group of roughly 4,000 employees represents about 28 percent of all DOL employees covered by the Federal Employee Retirement System. In Phase 1, all employees meeting the contribution criterion were included in the trial. In Phase 2, all employees contributing less than 5 percent at the beginning of the phase were included in the trial, with the exception of those who had been assigned to the treatment group in Phase 1. This included members of the Phase 1 control group, along with new hires contributing less than 5 percent and any other employees whose contributions had dropped below 5 percent since the beginning of Phase 1.

In each phase of the study, we collected data at baseline—before the email was sent—and at three subsequent points in time after the emails were sent. In Phase 2, two of these subsequent sets of data were drawn before a reminder email was sent out, and the last set was drawn after the reminder email had been sent.

The data we received contained information about employees' contributions to TSP, salaries, locations, hire dates, ages, genders, ethnicities, and union membership. In each phase, to prepare the data for random assignment, we computed the percentage of each employee's salary that was contributed to TSP and removed individuals who were contributing 5 percent or more or were marked as ineligible to contribute.

Table B.1 describes characteristics of the analysis sample at baseline for each phase of the study. Characteristics are shown for the entire sample and are also shown separately for noncontributors—individuals who did not contribute any portion of their salaries to TSP—and undercontributors—individuals who made contributions that were greater than zero but less than 5 percent of their salaries.

Table B.1. Characteristics of the sample at baseline

Characteristic	Phase 1			Phase 2		
	Total	Non-contributors	Under-contributors	Total	Non-contributors	Under-contributors
Sample size (N)	4,011	1,721	2,290	2,302	938	1,364
Contribution rate (%)	1.5	0.0	2.7	1.6	0.0	2.7
Salary (\$)	63,536	67,636	61,486	78,209	81,902	77,490
Age	45.4	46.8	44.1	45.1	45.7	44.6
National office (%)	30.8	33.3	28.9	32.1	34.4	30.6
Hired before August 2010 (%)	36.6	47.2	28.6	54.3	66.4	46.0
Female (%)	51.5	52.4	50.9	51.2	51.0	51.3
Union membership (%)	78.8	77.4	79.9	78.3	77.3	79.0
Race/ethnicity						
Asian or Pacific Islander (%)	3.9	2.8	4.7	4.0	3.3	4.5
Black (%)	35.7	39.7	32.7	36.4	40.2	33.8
White (%)	49.5	47.5	50.9	48.3	45.9	49.9
Hispanic (%)	9.7	9.0	10.3	4.3	4.7	4.0
Other race/ethnicity or multiple (%)	1.2	1.0	1.4	7.0	5.9	7.8

Notes: The Contribution rate row displays the average value for each group. The Salary and Age rows each display the median values for each group. All other rows display the proportion of the group that has the corresponding characteristic.

As shown in Table B.1 above, there are two key differences in characteristics between the samples included in each trial phase. Participants in Phase 2 had substantially higher salaries, on average, than Phase 1 participants; they were also much more likely to have been hired before August 2010. Differences in salary reflect growth in salaries between September 2015 and March 2016, as well as the higher average salaries of employees who were not in the Phase 1 control group but joined the Phase 2 sample because they reduced their TSP contributions to below 5 percent at some point after the beginning of Phase 1.

C. Stratified random assignment

After receiving baseline data, we assigned members of the study sample to different treatment conditions at the individual level using stratified random assignment. This process is designed to ensure that individuals are assigned to groups at random but that groups are balanced in terms of key characteristics known as stratification variables.

When choosing the number of stratification variables to include, our goal was to include characteristics that were likely to affect post-trial contribution levels to TSP while ensuring a sufficient number of observations in each stratum—that is, a sufficient number of observations having a unique combination of values among the stratification variables. We decided to use four binary stratification variables, which yielded 16 strata. In Phase 1, the average stratum size was approximately 254, and the smallest stratum had 24 members. In Phase 2, the average stratum size was approximately 148, and the smallest stratum had 26 members (In our randomization procedure, we ensured that the number of individuals in each stratum did not differ between treatment arms by more than one).

We chose to use the following stratification variables because they appeared the most likely to affect post-trial savings behavior:

- **Salary above and below the median.** We anticipated that salary would be one of the most important factors influencing retirement savings decisions. For example, those with larger salaries may be more willing to devote a larger share of their incomes to retirement savings.
- **Hired before or after August 1, 2010.** Employees who were hired during or after August 2010 were automatically enrolled in TSP, and employees hired before August 2010 had to choose explicitly to enroll in TSP. Therefore, we hypothesized that undercontributing or noncontributing employees hired before this date may differ in their propensity to change their retirement savings decisions relative to those hired after this date. For example, one might expect that noncontributors hired during or after August 2010 may be more difficult to influence, since they had explicitly chosen not to contribute.
- **Member of a union (bargaining unit) or not.** Union membership may affect perceptions of job security and may be correlated with other factors that could affect retirement savings decisions, such as the types of jobs that employees hold.
- **Noncontributor or undercontributor to TSP.** Those who do not contribute any amount to TSP may differ in important ways from those who contributed a small portion of their salaries. For example, those contributing a small portion may be more familiar with the online system for making changes to their contributions and may have spent more time considering the benefits and costs of contributing to retirement savings.

In Phase 1, we assigned members of the study sample to two treatment arms—a treatment group and a control group. Members of the treatment group were intended to receive the T1 email intervention, and members of the control group were not. In Phase 2, we assigned members to three treatment arms—a T1 treatment group, a T2 treatment group, and a control group. (see sections II and III of the report for more details on each treatment arm.)

D. Outcomes

Our study focuses on three outcomes:

1. **Contributes to TSP:** whether an individual contributes any amount to TSP
2. **Contributes at least 5 percent to TSP:** whether an individual contributes at least 5 percent of their salary to TSP
3. **Contribution rate:** the individual's contribution rate as a percentage of their salary

While the second and third outcomes are assessed for the entire study sample, the first outcome, "contributes to TSP," is assessed only among those who did not contribute any amount to TSP at baseline.

In addition to these outcomes, in Phase 1 we include an analysis of whether individuals received the intervention email and indicated that they had read the email. We were not able to conduct this analysis in Phase 2, because we did not receive information about whether emails were received or read for this phase.

E. Attrition

Because the data used for this trial came from DOL administrative records that included all current employees, there was no sample attrition due to employees' refusals to provide data. Instead, the sources

of sample loss were (1) no longer being an active employee of DOL at endline and (2) having missing or invalid data at baseline or endline.

Tables B.2 and B.3 report the number of employees in the initial and final samples for each phase and identify the sources of sample loss. In Phase 1, about 1.6 percent of the initial sample was lost in the final sample; in Phase 2, about 2.6 percent was lost. The majority of losses were due to sample members no longer being actively employed by DOL. In Phase 1, the differential attrition rate—that is, the difference between the attrition rates of the treatment arms—was less than 0.1 percent. In Phase 2, differential attrition rates for each of the three study groups were 1.1 percent or less.

Table B.2. Attrition in Phase 1

Sample description	T1 (% retained)	Control (% retained)	Total (% retained)
Initial sample	2,039 (100.0)	2,039 (100.0)	4,078 (100.0)
Active employee at end of phase	2,009 (98.5)	2,005 (98.3)	4,014 (98.4)
Has nonmissing baseline data	2,009 (98.5)	2,005 (98.3)	4,014 (98.4)
Has nonmissing outcome data	2,007 (98.4)	2,005 (98.3)	4,012 (98.4)
Data are valid	2,006 (98.4)	2,005 (98.3)	4,011 (98.4)
Final sample	2,006 (98.4)	2,005 (98.3)	4,011 (98.4)

Notes: Numbers presented are sample sizes at each stage of the trial. Numbers in parentheses are the percentages of the initial sample remaining at each stage.

Table B.3. Attrition in Phase 2

Sample description	T1 (% retained)	T2 (% retained)	Control (% retained)	Total (% retained)
Initial sample	788 (100.0%)	787 (100.0%)	788 (100.0%)	2,363 (100.0%)
Active employee at end of phase	770 (97.7%)	772 (98.1%)	774 (98.2%)	2,316 (98.0%)
Has nonmissing baseline data	770 (97.7%)	772 (98.1%)	774 (98.2%)	2,316 (98.0%)
Has nonmissing outcome data	764 (97.0%)	765 (97.2%)	773 (98.1%)	2,302 (97.4%)
Data are valid	764 (97.0%)	765 (97.2%)	773 (98.1%)	2,302 (97.4%)
Final sample	764 (97.0%)	765 (97.2%)	773 (98.1%)	2,302 (97.4%)

Notes: Numbers presented are sample sizes at each stage of the trial. Numbers in parentheses are the percentages of the initial sample remaining at each stage.

F. Contamination

As mentioned above, for Phase 1, DOL-HR provided us with data on whether individuals received the intervention email and indicated that they had read the email. We were able to use these data to determine the extent of contamination—that is, the extent to which (1) individuals assigned to a treatment arm did not receive the intended treatment or (2) individuals assigned to the control arm received one of the treatment interventions. Contamination reduces estimates of the effectiveness of an intervention relative to the intervention's actual effects. For Phase 2, we did not receive information about whether emails were received or read and were therefore unable to conduct similar tests.

In Phase 1, there is evidence of sample contamination. Eighty-seven treatment group members did not have a valid email address at which to receive the email (see Table D.8), and we could not confirm whether an additional 10 members were sent the email. In addition, 7 members of the control group were sent the

email. In total, 104 sample members—2.6 percent of the final sample—were contaminated. In Phase 2, we did not receive email receipt data and cannot calculate the portion of the sample that was contaminated.

Our analysis includes all members of the final sample indicated in Tables B.2 and B.3 above, including those who were contaminated. We chose to keep contaminated members in Phase 1 for two reasons. First, removing contaminated sample members in Phase 1 could bias our findings if the characteristics of those members were different from the characteristics of others. Second, because we were unable to identify contaminated sample members in Phase 2, keeping these members in Phase 1 led to a consistent treatment of both phases of the trial. Because of this decision, our findings should be considered “intent to treat” estimates—that is, they measure impacts among those intended to be treated rather than among those who were actually treated. Impacts among those actually treated could potentially be larger than the impacts we report.

G. Study samples

Table B.4 reports the final sample sizes in each phase by treatment arm.

Table B.4. Sample sizes in each phase

	Full study sample	Noncontributors only
Phase 1		
T1	2,006	864
Control	2,005	857
Total	4,011	1,721
Phase 2		
T1	764	313
T2	765	311
Control	773	314
Total	2,302	938

Notes: Full study sample column values are used to calculate impacts for the “Contributes at least 5 percent to TSP” and “Contribution rate” outcomes. “Noncontributors only” values are used to calculate impacts for the “Contributes to TSP” outcome.

II. Analytic Methods

To examine impacts of the T1 and T2 email interventions, we conducted the analyses listed in Table B.5.

Table B.5. Description of analyses

Analysis #	Effect studied	Groups compared	Data Used
1	T1 in Phase 1	T1 (Phase 1) vs. control	Endline (Outcomes 1&2) Endline, controlling for baseline (Outcome 3)
2	T1 in Phase 2 after the reminder email	T1 (Phase 2) vs. control	Endline (Outcomes 1 & 2) Endline, controlling for baseline (Outcome 3)
3	T2 in Phase 2 after the reminder email	T2 vs. control	Endline (Outcomes 1 & 2) Endline, controlling for baseline (Outcome 3)
4	T2 relative to T1	T2 vs. T1 (Phase 2)	Endline (Outcomes 1 & 2) Endline, controlling for baseline (Outcome 3)
5	T1 in Phase 2 before the reminder email	T1 (Phase 2) vs. control	Midline (Outcomes 1 & 2) Midline, controlling for baseline (Outcome 3)
6	T2 in Phase 2 before the reminder email	T2 vs. control	Midline (Outcomes 1 & 2) Midline, controlling for baseline (Outcome 3)
7	Phase 2 reminder email (exploratory)	Pooled Phase 2 treatment (T1 and T2) vs. control	Endline, controlling for midline (Outcomes 1, 2 & 3)
8	T1 in Phase 2 relative to Phase 1 (exploratory)	T1 vs. C (Phase 1), T1 vs. C (Phase 2) This analysis restricts the analysis samples to include only employees who were present at Phase 1 baseline and at Phase 2 endline, in order to assess differences in impact by phase without the influence of changes in sample composition between phases.	Endline (Phase 1 Outcomes 1 & 2) Midline (Phase 2 Outcomes 1 & 2) Endline, controlling for baseline (Phase 1 Outcome 3) Midline, controlling for baseline (Phase 2 Outcome 3)

Notes: In Phase 1, endline data were collected 5 weeks after the intervention email was sent. In Phase 2, endline data were collected 9 weeks after the intervention email was sent and 3 weeks after the reminder email was sent. In Phase 2, midline refers to the data collected before the reminder, 6 weeks after the intervention email was sent.

The study’s principal experimental impact estimates, calculated in analyses 1 through 3, use a simple test of proportions when measuring impacts on the binary TSP outcomes (measures of whether an employee contributes to the TSP at all and whether she contributes at least 5 percent of her salary to the TSP). The test statistic in these estimated impacts on binary outcomes is calculated as follows:

$$(1) \quad z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_p \hat{q}_p \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

where

$$(1a) \quad \hat{p}_p = \frac{x_1 + x_2}{n_1 + n_2}$$

x_1 and x_2 are the total number of successes in the relevant treatment and control group, and

$$(1b) \quad \hat{q}_p = 1 - \hat{p}_p$$

Analyses 4 through 6 in the table above also estimate impacts on the binary outcomes mentioned above using the model presented in equations 1, 1a, and 1b. (In analysis 4, the proportions in each of two treatment groups are compared, rather than proportion in a treatment group versus the one in a control group.)

We estimate impacts on employees' contribution rates in Analyses 1 through 6 using a linear regression model that includes employees' baseline contribution rates to improve the precision of estimates.¹

These analyses are based on the following model:

$$(2) \quad \text{contribution rate}_{i,t_1} = \alpha + \beta_1 \text{treatment}_i + \beta_2 \text{contribution rate}_{i,t_0} + \varepsilon_i$$

where *contribution rate*_{*i,t*₁} refers to the value of individual *i*'s TSP contribution rate in the later time period being studied, *t*₁; while *contribution rate*_{*i,t*₀} refers to the value of the same outcome in the earlier period, *t*₀. *treatment*_{*i*} refers to the treatment status of the individual, and ε_i is the error term associated with individual *i*. β_1 is the effect of treatment status and is the coefficient of interest in each analysis.

The model measures the effects of treatment by comparing TSP contribution rates between a group receiving a treatment and a group that did not receive a treatment but is otherwise similar. The model adjusts estimates to account for correlation between pretreatment outcomes and treatment status and provides standard errors that account for autocorrelation in individuals over time. Heteroskedasticity-robust standard errors are estimated.

Analyses 1, 2, 3, 5, and 6 all study the impacts of an intervention email by comparing the TSP contribution outcomes of employees in one of the treatment groups to those in the control group. They differ in whether they study effects in Phase 1 or Phase 2, and—for Phase 2—in whether they study impacts before or after the reminder email was sent.

Analysis 4 studies the relative impacts of the two interventions in Phase 2. Analysis 7 explores how the reminder email may have affected impacts by comparing the pooled T1 and T2 treatment groups to the control group before and after the reminder email was sent. Because Analysis 7 compares outcomes at endline to those at midline (after the initial email but before the reminder), it is possible to include midline outcomes in the impact estimation model, so the model shown in equation 2 is used to estimate impacts on each of the three contribution outcomes.

Analysis 8 provides exploratory evidence on the role that the timing of the two trial phases played in the intervention's impact estimated in each phase. We estimate the impact of T1 in each phase as described for analyses 1 and 5, but we restrict the sample to minimize the extent of change in sample composition between phases. Specifically, the analysis sample for the Phase 1 impact in Analysis 8 compares treatment and control outcomes only among employees still employed by DOL as of Phase 2 endline. Similarly, the

¹ Baseline rates of the "contributes to the TSP" and "contributes at least 5 percent" outcomes are not used when estimating impacts on these outcomes, because all employees in the relevant analysis samples for these outcomes have zero values at baseline, by construction. That is, the analysis sample for the former includes only employees who were not contributing to the TSP at baseline, and all trial participants had contribution rates less than 5 percent at the start of the trial.

analysis sample for the Phase 2 impact in Analysis 8 compares outcomes between the Phase 2 T1 and control group employees only among those who were present at the beginning of Phase 1. By restricting the analysis sample to include only employees present in the sample at the start of Phase 1 and still employed at the end of Phase 2, we removed from the Phase 2 sample those employees who were hired by DOL between September 2015 and March 2016 as well as those who joined the sample because they reduced their contribution levels below 5 percent over that period. Finally, we removed Phase 1 sample members who left DOL at some point between September 2015 and June 2016. As a result, we may expect to reduce the influence of changes in sample composition on the measured impact in each phase, thereby gaining more information about the influence of different circumstances on the impacts measured in each phase.

All analyses except Analysis 7 and Analysis 8 leverage the benefits of the randomized controlled trial design—that is, they estimate impacts with a high level of rigor because they compare outcomes between randomly assigned groups whose outcomes are expected to differ only because of the effect of the intervention they received. However, Analysis 7 estimates the effect of the reminder by comparing the endline outcomes of the treatment and control groups to their outcomes at midline, after the treatment group has already received an initial email that was not sent to the control group. For this reason, Analysis 7 will not identify the effects of the reminder email separately from any lagged effects of the intervention email, and findings from this analysis should be treated as suggestive. We discuss this further in the context of the findings below. Similarly, Analysis 8 compares the effects of T1 in Phase 1 and Phase 2. Since random assignment was not used to assign employees to the T1 groups in the two phases and the baselines and endlines used for each group occur in different time periods, we cannot be sure that the two groups are indeed equivalent. Therefore we treat these analyses as purely exploratory.

Aside from analyzing effects for the full study sample, we examined effects among subgroups with the following characteristics:

- Salary above and below the median
- Age above and below the median
- Located in the national office or elsewhere
- Hired before or after August 1, 2010
- Member of a union or not
- Noncontributor or undercontributor to TSP²

² Subgroup analyses could not be conducted among undercontributors for the “contributes to TSP” outcome, because this outcome was assessed only among those who did not contribute to TSP at baseline.

APPENDIX C
BASELINE EQUIVALENCE

In this appendix we provide detailed tables and additional information to supplement the discussion of baseline equivalence in Chapter 3.

The validity of our impact estimates depends on whether the treatment and control groups in the analysis sample had similar characteristics at baseline. To help verify that groups were similar, we assessed whether the two groups showed statistically significant differences in observable baseline characteristics.

Table C.1 compares the baseline measures of the treatment group to the control group for Phase 1. For Phase 2, Table C.2 compares T1 to control, Table C.3 compares T2 to control, and Table C.4 compares T2 to T1.

In Phase 1, none of the baseline characteristics we examined differed significantly. In Phase 2, the proportion of employees in the national office is larger in the control group than in either of the two treatment groups.

Because we do not find a pattern of differential impacts by office location across outcomes in Phase 2 (see Table D.7), no other systematic baseline differences were observed, and because this was a randomized controlled trial with low overall and differential attrition, we concluded that imbalance on office location was not a substantial threat to the validity of Phase 2 impact estimates.

Table C.1. Baseline equivalence between T1 and Control in Phase 1

Characteristic	T1 Treatment (%)	Control (%)	Difference (%)	p-value
Contribution rate	1.5	1.5	0.0	0.446
Noncontributor	43.1	42.7	0.3	0.834
Salary less than median	46.1	45.9	0.2	0.886
Age less than sample median	49.9	49.5	0.4	0.813
National office	30.3	31.2	-0.9	0.531
Hired before August 2010	36.5	36.6	-0.1	0.964
Female	52.2	50.8	1.5	0.352
Belonged to union	78.8	78.9	-0.1	0.945
Race/Ethnicity				
Asian or Pacific Islander	4.0	3.7	0.3	0.626
Black	35.2	36.2	-1.0	0.502
White	49.8	49.2	0.6	0.716
Hispanic	9.8	9.7	0.0	0.962
Other race/ethnicity or multiple	1.2	1.1	0.1	0.773

Notes: Sample sizes are 2,006 (T1 treatment) and 2,005 (control). The Difference column shows the arithmetic difference between values for the two groups. Because values in the Difference column are rounded to the nearest tenth after being subtracted, they may not always be the same as the differences between the rounded values for each group. Significance levels: *p<0.05, **p<0.01.

Table C.2. Baseline equivalence between T1 and Control in Phase 2

Characteristic	T1 Treatment (%)	Control (%)	Difference (%)	p-value
Contribution rate	1.6	1.6	0.1	0.341
Noncontributor	41.0	40.6	0.3	0.890
Salary less than median	47.4	48.0	-0.6	0.810
Age less than sample median	52.1	48.5	3.6	0.160
National office	28.9	36.2	-7.3**	0.002
Hired before August 2010	54.7	54.2	0.5	0.842
Female	50.7	53.8	-3.2	0.215
Belonged to union	78.3	78.1	0.1	0.949
Race/Ethnicity				
Asian or Pacific Islander	4.1	4.0	0.0	0.962
Black	34.3	38.3	-4.0	0.103
White	48.7	48.8	-0.1	0.975
Hispanic	5.1	3.2	1.9	0.066
Other race/ethnicity or multiple	7.9	5.7	2.2	0.092

Notes: Sample sizes are 764 (T1 treatment) and 773 (control). The Difference column shows the arithmetic difference between values for the two groups. Because values in the Difference column are rounded to the nearest tenth after being subtracted, they may not always be the same as the differences between the rounded values for each group. Significance levels: *p<0.05, **p<0.01.

Table C.3. Baseline equivalence between T2 and Control in Phase 2

Characteristic	T2 Treatment (%)	Control (%)	Difference (%)	p-value
Contribution rate	1.6	1.6	0.1	0.312
Noncontributor	40.7	40.6	0.0	0.990
Salary less than median	48.1	48.0	0.1	0.966
Age less than sample median	49.3	48.5	0.8	0.763
National office	31.2	36.2	-5.0*	0.039
Hired before August 2010	54.1	54.2	-0.1	0.973
Female	49.0	53.8	-4.8	0.060
Belonged to union	78.6	78.1	0.4	0.840
Race/Ethnicity				
Asian or Pacific Islander	4.1	4.0	0.0	0.967
Black	36.6	38.3	-1.7	0.493
White	47.3	48.8	-1.5	0.569
Hispanic	4.6	3.2	1.3	0.174
Other race/ethnicity or multiple	7.5	5.7	1.8	0.164

Note: Sample sizes are 765 (T2 treatment) and 773 (control). The Difference column shows the arithmetic difference between values for the two groups. Because values in the Difference column are rounded to the nearest tenth after being subtracted, they may not always be the same as the differences between the rounded values for each group. Significance levels: *p<0.05, **p<0.01.

Table C.4. Baseline equivalence between T2 and T1 in Phase 2

Characteristic	T2 Treatment (%)	T1 Treatment (%)	Difference (%)	p-value
Contribution rate	1.6	1.6	0.0	0.957
Noncontributor	40.7	41.0	-0.3	0.900
Salary less than median	48.1	47.4	0.7	0.777
Age less than sample median	49.3	52.1	-2.8	0.271
National office	31.2	28.9	2.3	0.324
Hired before August 2010	54.1	54.7	-0.6	0.816
Female	49.0	50.7	-1.6	0.523
Belonged to union	78.6	78.3	0.3	0.890
Race/Ethnicity				
Asian or Pacific Islander	4.1	4.1	0.0	0.996
Black	36.6	34.3	2.3	0.345
White	47.3	48.7	-1.4	0.592
Hispanic	4.6	5.1	-0.5	0.629
Other race/ethnicity or multiple	7.5	7.9	-0.4	0.767

Note: Sample sizes are 765 (T2 treatment) and 764 (T1 treatment). The Difference column shows the arithmetic difference between values for the two groups. Because values in the Difference column are rounded to the nearest tenth after being subtracted, they may not always be the same as the differences between the rounded values for each group. Significance levels: * $p < 0.05$, ** $p < 0.01$.

In addition to the findings in these tables, which report results for the full study samples in each phase, we also conducted baseline equivalence tests among the subgroups defined in the previous section, to determine whether subpopulations of interest showed systematic differences.

In Phase 1, no subgroup analyses showed significant differences between groups. In Phase 2, several subgroup analyses corroborated the finding in the full sample that the proportion of employees in the national office was higher in the control group than in either treatment group. While a few subgroups in Phase 2 also showed differences in gender, ethnicity, baseline contribution rate, or age, none of these findings appeared in the vast majority of subgroups, and there were no notable patterns for these findings across subgroups.³

³ In Phase 2, 36 baseline equivalence tests were run for each of the 13 variables being compared; one for each of the 12 subgroups and for each of the 3 comparisons between sample arms (T1 vs. control, T2 vs. control, and T1 vs. T2). Aside from the national office variable, which was significantly different between groups in 7 of the 36 tests, no other variable was significantly different in more than 3 of the 36 tests.

APPENDIX D

DETAILED TABLES ON STUDY FINDINGS

In this section we present detailed tables of impact estimates and summaries of key findings organized by our research questions.

I. RQ1: Can emails informed by behavioral science encourage more DOL employees to increase their savings? (Analyses 1, 2, and 3)

Table D.1. Impacts of T1 and T2 (Analyses 1, 2, and 3)

Outcome	Treatment		Control		Impact (%)	p-value
	Proportion (%)	N	Proportion (%)	N		
Impact of T1 in Phase 1 (Analysis 1)						
Contributes to TSP	3.5	864	2.5	857	1.0	0.211
Contributes at least 5% to TSP	4.2	2,006	1.9	2,005	2.3**	0.000
Change in contribution rate	0.1	2,006	0.0	2,005	0.1	0.072
Impact of T1 in Phase 2 after the reminder (Analysis 2)						
Contributes to TSP	11.8	313	8.0	314	3.9	0.106
Contributes at least 5% to TSP	11.3	764	3.8	773	7.5**	0.000
Change in contribution rate	0.6	764	0.2	773	0.4**	0.001
Impact of T2 in Phase 2 after the reminder (Analysis 3)						
Contributes to TSP	10.0	311	8.0	314	2.0	0.380
Contributes at least 5% to TSP	9.5	765	3.8	773	5.8**	0.000
Change in contribution rate	0.4	765	0.2	773	0.2*	0.040

Note: The sample for “Contributes to TSP” excludes baseline contributors. The remaining outcomes use the full study sample. A sensitivity analysis pooled T1 and T2 in Phase 2 and compared the share of employees contributing to the TSP in the pooled treatment group to the share contributing to the TSP in the control group. This pooled impact estimate was also not statistically significant. Significance levels: * $p < 0.05$, ** $p < 0.01$.

Key Findings

- T1 significantly increased the proportion of employees contributing at least 5 percent of their income to TSP in both Phase 1 and Phase 2 (2.3 percentage point increase in Phase 1 and 7.5 percentage point increase in Phase 2).
- In Phase 2, T1 also significantly increased the overall contribution rate by 0.4 percentage points, though it had no statistically significant impact in Phase 1.
- T2 caused a 5.8 percentage point increase in the proportion of employees contributing at least 5 percent and a 0.2 percentage point increase in the contribution rate.
- Neither T1 nor T2 changed the proportion of employees who contributed any amount to TSP in either phase.

our calculations of per-person savings reported in Chapter V apply the 0.4 percentage-point increase in the overall contribution rate observed in Phase 2 of the trial to the median salary of \$78,162 that we observed in the study treatment group that received the original email in Phase 2. We assumed 20 years of saving, since the median age of this sample is 45 years, and a six percent annual rate of return, yielding an estimate of \$11,501 per person. (The direct calculation is $\sum_{i=1}^{20} (0.004 \times \$78,162) \times (1.06)^{(i-1)} = \$11,501$) To calculate total savings we multiplied this estimate of per-person savings with 764 (the number of employees contributing less their salary who received the original email in Phase 2 of the trial in 2016).

II. RQ2: What email communication strategies are most effective at improving savings?

1. Are emails that invoke peer norms more effective than emails that don't?

Table D.2. Impact of T2 vs. T1 in Phase 2 after the reminder (Analysis 4)

Outcome	T2 Treatment (%)	T1 Treatment (%)	Impact (%)	p-value
Contributes to TSP	10.0	11.8	-1.9	0.458
Contributes at least 5% to TSP	9.5	11.3	-1.7	0.272
Change in contribution rate	0.4	0.6	-0.2	0.173

Note: The sample for "Contributes to TSP" excludes baseline contributors. The remaining outcomes use the full study sample. Significance levels: *p<0.05, **p<0.01.

Key Findings

- T1 impacts are greater in magnitude compared to those of T2, but these differences are not statistically significant.
- There is no evidence that an email that invokes peer norms is more effective than an email that does not.

2. Does sending reminder emails change the impact? (Analyses 5, 6, and 7)

Table D.3. Impact of T1 in Phase 2 before the reminder (Analysis 5)

Outcome	T1 Treatment (%)	Control (%)	Impact (%)	p-value
Contributes to TSP	9.2	4.8	4.5*	0.029
Contributes at least 5% to TSP	9.3	3.0	6.3**	0.000
Change in contribution rate	0.5	0.1	0.3**	0.002

Notes: The sample for "Contributes to TSP" excludes baseline contributors and has 314 treatment and 314 control group members. The remaining outcomes use the full study sample with 766 treatment and 773 control group members. Significance levels: *p<0.05, **p<0.01.

Table D.4. Impact of T2 in Phase 2 before the reminder (Analysis 6)

Outcome	T2 Treatment (%)	Control (%)	Impact (%)	p-value
Contributes to TSP	9.3	4.8	4.5*	0.027
Contributes at least 5% to TSP	8.7	3.0	5.7**	0.000
Change in contribution rate	0.4	0.1	0.2**	0.008

Notes: The sample for "Contributes to TSP" excludes baseline contributors and has 312 treatment and 314 control group members. The remaining outcomes use the full study sample with 768 treatment and 773 control group members. Significance levels: *p<0.05, **p<0.01.

Table D.5. Impact of the reminder in Phase 2 (Analysis 7)

Outcome	T1 and T2 Treatment (%)	Control (%)	Impact (%)	p-value
Contributes to TSP	1.8	3.2	-1.2	0.336
Contributes at least 5% to TSP	1.4	0.8	1.1	0.057
Contribution rate	0.1	0.1	0.0	0.611

Notes: The sample for “Contributes to TSP” excludes baseline contributors and has 624 treatment and 314 control group members. The remaining outcomes use the full study sample with 1,529 treatment and 773 control group members. Treatment and control group values are differences between the midline and endline periods. Significance levels: *p<0.05, **p<0.01.

Key Findings

- The reminder email did not have a statistically significant marginal impact on the effects of the interventions (Analysis 7).
- In comparing Analyses 5 and 6 to Analyses 2 and 3, both T1 and T2 appear to have had an impact on the proportion of the sample who contributed any amount before the reminder that was no longer evident after the reminder. These results are not due to reduced contributions in either treatment group after the reminder but to increased contributions in the control group.

As stated previously, Analysis 7 cannot separate the impact of the reminder email from any lagged effect of treatment group members’ receiving the initial intervention email. For example, if contributions made after the reminder was sent were partially a delayed reaction to the initial intervention email, and not solely due to receiving a reminder, the results shown in Table D.5 would not reflect the effect of the reminder alone. However, assuming that the true effect of the reminder is not negative (that is, it did not reduce the level of contribution) and that the lagged effect of the intervention email is also not negative (that is, recipients did not reduce their contributions in response to the intervention email as time passed), the results could only overstate the effect of the reminder—that is, the reminder effect is equal to or smaller than what the analysis suggests.

3. Does the timing of emails matter? (Analysis 8)

Table D.6. Impact of T1 by Phase (Analysis 8)

Outcome	T1 Treatment (%)	Control (%)	Impact (%)	p-value
Phase 1				
Contributes to TSP	3.5	2.5	1.0	0.211
Contributes at least 5% to TSP	4.1	1.9	2.3**	<0.001
Change in contribution rate	0.1	0.0	0.1	0.082
Phase 2				
Contributes to TSP	7.2	3.4	3.8	0.065
Contributes at least 5% to TSP	7.3	1.9	5.4**	<0.001
Change in contribution rate	0.3	0.1	0.2**	0.005

Notes: The sample for “Contributes to TSP” excludes baseline contributors and has 486 Phase 2 and 1,705 Phase 1 group members. Analyses of the remaining outcomes use a sample of 1,159 employees in Phase 2 and 3,963 employees in Phase 1. Significance levels: *p<0.05, **p<0.01.

Key Findings

- Impacts of T1 were larger in Phase 2 than in Phase 1 for all outcomes studied.
- In Phase 2, the estimated impact on proportion of employees contributing any amount was 2.8 percentage points higher, the proportion contributing at least 5 percent was 3.1 percentage points larger, and the contribution rate was 0.1 percentage points larger.

In Analysis 8, we compared the impact of T1 and the end of Phase 1 to its impact in Phase 2 before the reminder email, restricting the analysis samples to minimize the influence of changes in the sample composition between phases. In Phase 1, treatment group members had been exposed to the intervention for 37 days at the time the endline data were recorded; in Phase 2, they had been exposed to the intervention for 39 days.

These findings suggest that the timing of the interventions may have played an important role in determining the magnitude of their impacts. In Phase 1, during the week that the email was sent, the Pope visited Washington DC, and many federal employees worked from home. Since 31 percent of the study sample consisted of federal employees in the Washington DC area, and approximately 5 percent of the treatment sample sent an “out of office” status message in response to receiving the email, the Pope’s visit may have affected the treatment group’s responsiveness. In addition, news stories that speculated about the possibility of a government shutdown were circulating at the time of the Phase 1 intervention. This may have caused members of the study sample to be less willing to contribute a larger share of their salaries to retirement savings accounts, in favor of keeping their assets in more liquid accounts.

Additionally, there may be broader seasonal or macroeconomic trends that prompted trial participants to be more likely to increase their TSP contributions in Phase 2. Finally, the differences in patterns between these findings and the findings of analysis 1 and 5 (the comparable full-sample T1 impacts in Phases 1 and 2) indicate that comparisons of the main impact estimates between phases should be interpreted with caution, given that they appear to be influenced by changes in the types of employees who are included in the sample for each phase.

III. RQ3: For whom do these emails work?

Table D.7 presents a summary of findings for key subgroups of interest and highlights results that were statistically significant.

Key Findings

- Findings within subgroups were largely consistent among the subgroups we studied: The effects of each intervention found for the overall study sample were generally observed across subgroups, the reminder email did not appear to have a significant marginal effect in any subgroup, and Phase 2 impacts were higher than Phase 1 impacts across all subgroups.
- Effects were concentrated among those who were already contributing some amount of their salary at baseline, and among union members.
- Although effects were concentrated among younger employees and those outside the national office in Phase 1, these findings weren’t replicated in Phase 2.
- There were no notable patterns associated with employees’ salary levels or whether they were hired after the automatic enrollment policy was enacted.

Table D.7. Summary of subgroup analysis findings

Subgroup	T1 in Phase 1 (Analysis 1)			T1 in Phase 2 after reminder (Analysis 2)			T2 in Phase 2 after reminder (Analysis 3)			T2 vs. T1 in Phase 2 after reminder (Analysis 4)			Impact of reminder in Phase 2 (Analysis 7)			Impact of phase for T1 (Analysis 8)		
	Contributes any amount	Contributes at least 5%	Overall contri- bution rate	Contributes any amount	Contributes at least 5%	Overall contri- bution rate	Contributes any amount	Contributes at least 5%	Overall contri- bution rate	Contributes any amount	Contributes at least 5%	Overall contri- bution rate	Contributes any amount	Contributes at least 5%	Overall contri- bution rate	Contributes any amount	Contributes at least 5%	Overall contri- bution rate
Baseline contribution																		
Noncontributor	NA			NA	X		NA			NA			NA			NA	X	X
Contributes < 5%	NA	X	X	NA	X	X	NA	X	X	NA			NA			NA	X	X
Salary																		
Below median		X			X	X		X	X								X	X
Above median		X			X			X									X	X
Age																		
Below median	X	X	X		X	X		X			X					X	X	X
Above median					X	X		X	X							X	X	X
Location																		
National office				X	X			X								X	X	X
Other offices	X	X	X		X	X		X								X	X	X
Hire date																		
Before August 2010		X			X	X		X								X	X	X
During or after August 2010		X			X	X		X								X	X	X
Belongs to a union?																		
Yes		X	X		X	X		X								X	X	X
No					X			X										

Notes: X = significant impact, p<0.05; NA = subgroup not applicable to analysis. The sample for “Contributes to TSP” excludes baseline contributors.

IV. Email receipt analysis

In Phase 1, we received data describing whether treatment group members indicated that they had read the emails they were sent. In addition, among treatment group members who did not indicate that they had read the emails, the data indicated whether those group members did not have an email address or deleted the email without reading it. Table D.8 summarizes the data we received, among all study sample members initially assigned to the treatment group. In Phase 2, no similar data were received.

Table D.8. Email receipt in Phase 1

Outcome	Proportion (%)	N
Sent read receipt	26.8	547
Did not send read receipt	73.2	1,492
Did not have an email address	4.3	87
Deleted email without reading	3.3	67

Note: Values in the Proportion column are proportions of the final treatment sample with the corresponding outcome.

Key Findings

- Only 26.8 percent of the treatment group sent an indication that they had read the email.
- The majority of those who didn't send a read receipt did not have an identifiable reason for not reading the email, such as not having an email address or deleting the email without reading it. Because sending a read receipt is typically a voluntary activity, it is not clear whether these individuals did or did not read the email.

www.mathematica-mpr.com

**Improving public well-being by conducting high quality,
objective research and data collection**

**PRINCETON, NJ ■ ANN ARBOR, MI ■ CAMBRIDGE, MA ■ CHICAGO, IL ■ OAKLAND, CA ■
TUCSON, AZ ■ WASHINGTON, DC ■ WOODLAWN, MD**

MATHEMATICA
Policy Research

Mathematica® is a registered trademark
of Mathematica Policy Research, Inc.

ideas **42**