The Effects of Income Tax Timing on Retirement Investment Decisions

Shane R. Stinson, Ph.D.*
Associate Professor
Culverhouse School of Accountancy
The University of Alabama
srstinson@cba.ua.edu

Marcus M. Doxey, Ph.D.
Assistant Professor
Culverhouse School of Accountancy
The University of Alabama

Timothy J. Rupert, Ph.D.
Professor
D’Amore-McKim School of Business
Northeastern University

*corresponding author

Acknowledgements: We gratefully acknowledge Chelsea Rae Austin, Andy Cuccia, Diana Falsetta, Scott Jackson, Anne Magro, Mary Marshall, Michael Roberts, Donna Bobek Schmitt, and workshop participants at the Behavioral Tax Symposium, the University of Alabama, Northeastern University and the University of South Carolina for their thoughtful comments and suggestions. Any errors or omissions are our own.
The Effects of Income Tax Timing on Retirement Investment Decisions

Abstract: In a controlled online experiment, we find evidence that the timing of taxes on retirement income affects taxpayers’ investment decisions such that tax-deferred plan participants under-adjust for future tax burdens and overestimate their future wealth compared to Roth participants. As a result, when presented with a specific, after-tax monetary goal, Roth account holders invest more in higher-risk, higher-return assets than tax-deferred account holders. We investigate four aspects of this investment context that could eliminate these differences: 1) changing the retirement goal to a non-specific “do-your-best” goal, 2) reframing specific goals in nominally higher pre-tax dollars, 3) prompting participants to explicitly estimate future taxes due on retirement savings, and 4) providing performance feedback. While we find each intervention reduces differences between Roth and tax-deferred plan participant behaviors, only goal reframing and explicitly prompting future tax-estimations encourage tax-deferred plan participants to invest in risky assets to the same degree as Roth participants.
I. INTRODUCTION

In this study, we use an online experiment to investigate how the timing of income taxes on retirement savings affects retirement plan participants’ investment risk profiles. In an era of rising life expectancies and an aging population, many taxpayers face significant challenges in retirement planning. Their task is obvious, yet elusive; they must save enough money to sustain them over a longer life expectancy than prior generations while facing increasing living costs and heightened constraints on government aid. Adding to the challenge, employers have largely transitioned from defined benefit plans to defined contribution plans that, while less expensive to manage and fund, shift the burden of planning and investing to employees (Anderson 2013; U.S. Department of Labor 2018).¹

To encourage individual retirement savings, the federal government has authorized several tax-preferred defined contribution plans.² These plans either defer taxes on contributions (i.e., the taxpayer receives a current deduction or exclusion for contributions, effectively investing pre-tax dollars that are fully taxed upon withdrawal), as with a traditional IRA or 401(k) plan, or tax contributions up-front (i.e., the taxpayer invests after-tax dollars and pays no taxes on withdrawals), as with a Roth IRA or Roth 401(k) plan. Tax-deferred and Roth retirement plans each offer their own risks and rewards, requiring individuals to carefully evaluate the potential tax consequences of their investment choices far before retirement. Regardless of the plan type participants choose, they must effectively invest their savings within the plan to achieve their retirement goals. This generally means that, to the extent participants’ future capital requirements exceed their direct contributions and current wealth, they must seek investments with adequate risk and upside potential to make up the difference.

¹ U.S. Department of Labor statistics indicate that defined contribution plans comprise approximately 94 percent of all employer-sponsored retirement plans and hold 66 percent of employer-sponsored plan assets, amounting to roughly $5.7 trillion as of 2016 (the most recent reporting period available).
² The Employee Retirement Income Security Act of 1974 first established traditional Individual Retirement Arrangements (IRAs). Section 401(k) defined contribution plans were then established under the Revenue Act of 1978 to allow employees to save for retirement through pre-tax salary reductions. More recently, Roth IRA and Roth 401(k) plans were established by the Taxpayer Relief Act of 1997 and the Economic Growth and Tax Relief Reconciliation Act of 2001, respectively, although Roth 401(k)s were not officially implemented until 2006.
While research has examined the investment choices plan participants make when allocating investments (e.g., Benartzi and Thaler 1999, 2001, 2007), many questions remain about the quality of their asset selections. Research shows that individuals often invest too conservatively for their age, forgoing the potential earnings premium attributable to long-term stock investment in favor of bonds and other assets offering less downside risk (e.g., Haliassos and Bertaut 1995; Madrian and Shea 2001). Similarly, researchers have shown that plan participants struggle with a number of basic retirement savings concepts and decisions. Despite evidence that plan participants often do not invest as well as they could, a compounding factor unexamined by previous studies on retirement investing is the effect of tax timing on retirement savings decisions. To add to this literature, we investigate the effects of tax timing (i.e., tax-deferred or Roth plans) on the allocation of funds between investments of disparate risk.

Given equivalent investment returns and constant tax rates (which we maintain in our experimental design), the expected after-tax values of tax-deferred and Roth retirement accounts are equivalent. However, a participant’s nominal account value—generally the only balance provided on participants’ account statements—is always higher in a tax-deferred plan than in a Roth plan with equivalent inputs. Further, a tax-deferred plan’s nominal balance is always higher than its realizable after-tax value, which is typically not estimated for plan participants due to the inherent uncertainty of future tax rates and personal tax idiosyncrasies. In contrast, the nominal balance of a Roth retirement account is always equivalent to its after-tax value because contributions consist of after-tax dollars and investment returns are not taxed.4

Based on these differences, we use the anchoring and adjustment heuristic (Tversky and Kahneman 1974) to hypothesize that participants in tax-deferred retirement accounts will anchor on their nominal balances and predict their future, after-tax account balances to be higher than actual. Our results comparing

---

3 In particular, participants tend to view default options in automatic enrollment plans as implicit advice (Madrian and Shea 2001), over-extrapolate prior returns when allocating investments (Choi, Laibson, Madrian, and Metrick 2009), and sometimes even fail to take advantage of nearly costless arbitrage opportunities in their retirement plans (Choi, Laibson, and Madrian 2011).

4 Throughout this article, we assume participants make only qualified distributions and do not incur additional taxes or penalties beyond their normal, marginal income tax rate.
participants’ balance predictions to their ending balances support this assertion. We also predict and find that participants in tax-deferred retirement accounts with specific (i.e., numerical) retirement account balance goals are more likely to invest in lower-risk, lower-return investments because such goals appear easier to achieve when anchoring on a nominally higher account balance. In contrast, because nominal and after-tax balances are equivalent in Roth accounts, we find Roth participants provide more modest estimates of their current and future after-tax account balances and are more likely to choose a higher-risk investment allocation to achieve their specific after-tax retirement goals.\(^5\)

In addition to documenting differences in investment selections between tax-deferred and Roth plan participants, we investigate four aspects of the retirement investing context that might reduce or eliminate these differences. First, we focus on retirement savings goals. The retirement planning industry often encourages people to set goals for retirement income based on their spending habits (i.e., after-tax consumption), but not all retirement savers develop such specific goals. Therefore, we investigate whether holding unspecified, “do-your-best” goals may actually be beneficial for participants in tax-deferred accounts if it reduces the overconfidence prompted by anchoring on pre-tax balances and comparing that position to a specific, after-tax goal. As generally predicted by goal theory, we find that Roth participants invest more conservatively with an unspecified goal than under baseline conditions with a specific monetary goal. However, contrary to traditional goal theory predictions, tax-deferred participants increase risky investment with an unspecified goal to a level equal to that of their Roth counterparts (but still significantly less than Roth participants with a specific goal).

Second, because we hypothesize that the differences between Roth and tax-deferred participants are driven by overconfidence stemming from anchoring and under-adjusting behavior in the tax-deferred

\(^5\) Our intent is not to prescribe that retirement plan participants fully invest in riskier assets without adequately considering the consequences of such a strategy. However, conventional wisdom among financial planners suggests investing a larger portion of one’s assets in securities with greater volatility and upside potential relatively early in life while gradually decreasing portfolio risk as one draws nearer to retirement. In our study, we restrict the age of our participants to obtain a sample with relatively homogenous exposure to retirement savings options and proximity to statutory age limits for qualified retirement distributions. We further construct our investment options so that the higher-risk asset has a higher annual expected return, consistent with common tradeoffs between stocks and bonds.
condition, we also investigate whether nominally increasing the goal to a pre-tax number makes it appear
harder to achieve and reduces tax-deferred participants’ overconfidence. We find that reframing the goal
for tax-deferred participants significantly increases their risk assumption to a level that, while still
nominally below the Roth level of risk assumption, is statistically indistinguishable.

Third, we investigate whether explicitly prompting tax-deferred participants to estimate their final
tax burden alleviates their under-adjustment for taxes and prompts them to invest in risk at the same levels
as Roth participants. We find that prompting participants in a tax-deferred plan to estimate their eventual
tax burden causes them to increase their investment risk profile and eliminates investment differences
compared to Roth participants. Further, the tax-estimation intervention increases the perceived difficulty of
achieving the investment goal, giving further support to the hypothesis that participants in tax-deferred
plans may suffer from overconfidence as a result of anchoring on higher nominal account balances and
under-adjusting for taxes without an explicit tax-estimation task.

Finally, periodic performance feedback that allows participants to evaluate their investment
allocations is another prominent feature of the retirement planning process, and it is conceivable that
feedback could help alleviate the differences in investment choices caused by tax timing. Fiduciary websites
(e.g., TIAA-CREF) often have tools that allow participants to forecast their retirement income or account
balance to check whether they are “on track” to reach a specific retirement goal (e.g., a projected balance
or estimated monthly income). Periodic feedback should inform subsequent investment decisions as plan
participants strive to achieve specific retirement savings goals. While we find that feedback alone is
insufficient to eliminate the differences in risk assumption between Roth and tax-deferred participants,
feedback does complement the effects of reframing goals in pre-tax dollars and explicitly prompting future
tax estimations for tax-deferred plan participants.

This study makes three distinct contributions to the accounting, taxation, investment, and goal
setting literatures. First, although research has empirically shown that retirement investors generally under-
invest in equity risk (e.g., Haliassos and Bertaut 1995; Madrian and Shea 2001), no research has examined
whether differences in tax timing for alternative retirement plans impact risk allocation. We find that Roth
(tax-deferred) plans encourage higher (lower) risk allocation when investors adopt specific, quantitative goals. However, participants in tax-deferred plans will adopt higher risk allocations when the drivers of these differences (under-adjustment and overconfidence) are reduced by increasing the apparent difficulty of the goals or by prompting participants to explicitly adjust for taxes. Our results also suggest that performance feedback, while insufficient to eliminate the risk assumption differences in isolation, is helpful when combined with goal re-framing or tax-estimation.

Second, large-sample studies show an unexplained drop in consumption near retirement indicating that retirees tend to experience a negative shock shortly after retirement that was heretofore unexplained (e.g., Banks, Blundell, and Tanner 1998). Our results suggest that tax-deferred plans (currently the most common retirement tax structure) may contribute to this phenomenon, as we find that participants in tax-deferred plans are significantly more likely to overestimate their after-tax retirement savings. This evidence of systematic tax effects on risk assumption could inform future tax policy as legislators strive to encourage effective retirement savings strategies.

Third, we identify a setting in which an unspecified goal does not reduce performance (i.e., goal-seeking behavior) for a simple task. Specifically, tax-deferred plan participants, who are otherwise hindered by biased assessments of progress toward specific goals, do not reduce their goal-seeking behavior with an unspecified goal. Such a finding is uncommon in the goal setting literature, which has consistently found specific goals to be more effective in motivating goal-seeking efforts and improving task performance on simple tasks (c.f., Latham and Locke 1991; Locke 1968; Locke, Shaw, Saari, and Latham 1981). Therefore, this study identifies an important boundary condition on the axiom that specific goals are more effective than non-specific goals for simple tasks.

---

6 Though experimental tasks vary widely in complexity, simple tasks are loosely defined as those that “require less information search, are easier to focus on, and are less likely to induce information overload” (Mone and Shalley 1995, p. 245). We discuss relative task complexity and its role in the goal setting literature later in this manuscript.
II. THEORY AND HYPOTHESES

Future Economic Equivalence of Tax-Deferred and Roth Retirement Plans

The expected realizable value of a tax-deferred or Roth retirement account can be expressed using the future value of an annuity with appropriately timed tax adjustments:

\[
\text{RothBal}_{\text{NOM,n}} = \text{RothBal}_{\text{AT,n}} = \left[ C \cdot (1 - \text{T}_{\text{PRE-RET}}) \right] \cdot \left[ \frac{((1+r)^n - 1)}{r} \right], \quad (1)
\]

\[
\text{TDBal}_{\text{NOM,n}} = C \cdot \left[ \frac{((1+r)^n - 1)}{r} \right], \quad (2a)
\]

\[
\text{TDBal}_{\text{AT,n}} = \text{TDBal}_{\text{NOM,n}} \cdot (1 - \text{T}_{\text{RET}}), \quad (2b)
\]

where:

- \( C \) = pre-tax annual contribution,
- \( r \) = pre-tax rate of return on invested capital,
- \( n \) = number of periods,
- \( \text{T}_{\text{PRE-RET}} \) = ordinary income tax rate before retirement,
- \( \text{T}_{\text{RET}} \) = ordinary income tax rate after retirement,
- \( \text{RothBal}_{\text{NOM,n}} \) = Roth retirement account balance after \( n \) periods,
- \( \text{TDBal}_{\text{NOM,n}} \) = pre-tax tax-deferred retirement account balance after \( n \) periods, and
- \( \text{TDBal}_{\text{AT,n}} \) = after-tax tax-deferred retirement account balance after \( n \) periods.

As demonstrated above, any difference in the expected realizable value of a tax-deferred retirement account \( \left( \text{TDBal}_{\text{AT,n}} \right) \) compared with a Roth account \( \left( \text{RothBal}_{\text{AT,n}} \right) \) with equivalent inputs and investment outcomes is driven entirely by expected differences in pre- and post-retirement tax rates \( \left( \text{T}_{\text{PRE-RET}} \text{ and } \text{T}_{\text{RET}} \right) \).

As a result, both tax educators and financial advisers typically encourage individuals to pursue Roth accounts when their incomes and tax rates are lowest (often early in life) and gradually increase their holdings in tax-deferred accounts as their economic status improves (presumably later in one’s career), which presupposes an eventual drop in tax rates for many retirees. However, while individual expectations for earnings capacity likely vary, the nation’s political climate and fluctuating deficit further compound the uncertainty around future tax rates. Current debates over the national debt and government spending suggest that individual tax burdens could increase in the future. Thus, the inherent uncertainty of future tax rates produces an idiosyncratic and political confound that limits our ability to examine the effects of tax timing differences on retirement planning strategies in the real world.
We assert that the tax timing differences that define tax-deferred and Roth retirement accounts warrant separate consideration in a controlled experiment to better understand their effects on retirement planning. Further, under constant average tax rates \( (T_{\text{PRE-RET}} = T_{\text{RET}}) \)—a reasonable assumption given the inherent uncertainty in future tax rates—tax timing differences between the two account types are irrelevant to their realizable after-tax values.\(^7\) Despite the future after-tax equivalence of tax-deferred and Roth retirement accounts under our simplifying assumptions, we note that the nominal balance of a tax-deferred account is strictly larger than that of a Roth retirement account at any given time because Roth contributions are taxed before retirement (i.e., \( TDBal_{\text{NOM,n}} > RothBal_{\text{NOM,n}} \)). We expect the higher nominal balance of the tax-deferred account to affect plan participants’ judgments and subsequent investment decisions during retirement planning, even when tax-timing differences are irrelevant under constant tax rates. As discussed in the following sections, these judgments and decisions are critical to achieving retirement goals, such as meeting savings targets and maintaining planned consumption levels in retirement.

**Level of After-Tax Balance Predictions**

As previously demonstrated, the nominal balance of a Roth retirement account is the same as its after-tax balance by statutory definition, while the nominal balance of a tax-deferred retirement account is always higher than its after-tax balance assuming a positive, non-zero tax rate. The added step of discounting a tax-deferred account balance for tax effects arguably makes calculating its expected after-tax balance more difficult than determining the after-tax value of a Roth account. Making matters worse, extensive research shows that many people do not put sufficient effort into retirement planning and investment management (e.g., Benartzi and Thaler 1999; see Benartzi and Thaler 2007 for a review of the literature) even before considering the added effort needed to properly account for tax effects. The increased cognitive effort needed to account for tax effects with a tax-deferred retirement account should make

\(^7\) We note that this equality of future after-tax values for tax-deferred and Roth accounts from the standpoint of individuals likely does not hold from the perspective of taxing authorities. Rather, given equivalent inputs, positive returns, and a constant tax rate, the undiscounted nominal value of taxes assessed on contributions and investment earnings for a tax-deferred account is strictly greater than the tax assessed on annual contributions to a Roth account.
participants more likely to resort to heuristics (Tversky and Kahneman 1974; Kahneman 2011) when managing their accounts.

The tax calculation for a tax-deferred retirement account, which begins with a nominal balance and expected future growth rate that must be adjusted downward for taxation, naturally lends itself to an anchoring and adjustment process. Research into the anchoring and adjustment heuristic shows a tendency toward systematic under-adjustment (Tversky and Kahneman 1974; Hardman 2009; Epley and Gilovich 2001, 2004, 2006). In this scenario, we expect participants in tax-deferred retirement accounts to systematically overestimate their after-tax account value because of the relatively high anchor provided by their nominal balance. We formalize this prediction in the following hypothesis:

**H1:** Participants in tax-deferred retirement accounts will provide systematically higher estimates of future after-tax account balances than participants in Roth retirement accounts.

**Risk Assumption to Achieve Specific Retirement Savings Goals**

While H1 predicts that tax-deferred plan participants will be relatively more likely to under-adjust for taxes and thereby overestimate the future after-tax value of their retirement assets, whether that difference in perception results in a systematic difference in retirement planning behavior or outcomes further depends on the nature of participants’ savings goals. Carver and Scheier (1998) describe the goal setting and striving process as a negative feedback loop in which people hold a desired end state in mind (the goal) and assess the current state of the world in relation to that goal. If there is a discrepancy between the current state and the goal, people take steps to reduce that discrepancy. In the context of retirement savings, one’s primary objective is to accumulate enough funds to secure a reasonably comfortable future, and this accumulation of wealth occurs through a combination of plan contributions and subsequent investment in appreciating assets. Thus, to the extent a retirement plan participant’s contributions are limited by statute (e.g., mandated contribution limits to IRA and 401(k) plans) and/or budgetary constraints,

---

8 The use of the anchoring and adjustment heuristic has been shown to increase when a relevant anchor is involved (Chapman and Johnson 1994; Strack and Mussweiler 1997) and when the anchor is self-generated (Epley and Gilovich 2001, 2006, 2004). The nominal balance in our setting is relevant to the after-tax value and is the logical starting point for participants, and thus likely to be a self-generated anchor.
the primary mode by which to reduce any discrepancy between his or her current state and future goal is to seek investments with sufficient risk and upside potential.⁹

However, consistent with goal theory (e.g., Latham and Locke 1991; Locke and Latham 2002), retirement savings goals can vary considerably in their specificity, which, in turn, may affect the manner in which people strive to meet their goals. For example, absent any personal reflection or professional advice, one might simply conclude that “more is better” and have no clear idea of what level of savings would be “enough” to maintain a desired standard of living in retirement. At the opposite extreme, one might seek professional advice or devise his or her own monetary goal that provides a clearer target by which to measure savings performance.

Goal theory (Locke 1968, 1996; Locke and Latham 2002) holds that specific, difficult goals consistently produce greater effort and higher subsequent performance than unspecified goals (i.e., “do-your-best” goals), assuming the specific goal is not so difficult that participants abandon it.¹⁰ According to Locke’s review (1968), this tenet was originally based on the results of numerous laboratory studies (e.g., Locke and Bryan 1966a, 1966b; Bryan and Locke 1967). A subsequent review of field studies investigating goal theory finds that nearly all such studies further support Locke’s assertions (Latham and Yukl 1975), as does a 1990 meta-study by Locke and Latham. More recently, Mone and Shalley (1995, p. 243) report that the “results of over 500 studies indicate that specific difficult goals lead to higher task performance than vague, easy, or do-your-best goals.”

Laboratory studies generally observe the differences caused by goal-specification toward the end of an activity where effort falls off for participants in unspecified goal conditions but not for participants in specific goal conditions, who continue to strive to meet their goal (Locke 1968; Mace 1935). This pattern implies that the participants with a specific goal in these studies knew, with some degree of accuracy, when

⁹ The practical implication is that failing to assume sufficient risk, while simultaneously being optimistically biased in estimates of future account balances (H1), could lead to retirement shortfalls for participants in a tax-deferred retirement account, consistent with the consumption behavior of recent retirees noted in Banks et al. (1998).

¹⁰ All else equal, risk assumption may be considered “effortful” to the extent the retirement plan participant is naturally risk-averse.
they were short of the goal and continued to work toward it, while participants with unspecified goals reduced their efforts earlier because they could interpret the goal as being met sooner or in different ways (Locke 1968; Locke and Latham 2002).

Compared to Roth plan participants, we posit that tax-deferred account holders are likely to perceive a smaller discrepancy between their current state and a specific retirement goal and therefore will be less likely to take actions to reduce the discrepancy (i.e., invest more in a higher-risk, higher-return asset). Compounding this problem in real-world settings is the fact that retirement account fiduciaries can only provide account statements with known nominal balances and, because of the inherent uncertainty of investment and individual idiosyncrasies in tax situations, generally cannot provide detailed guidance for expected after-tax balances. Thus, participants receive repeated exposure to nominal account balances, and in tax-deferred accounts, these balances are higher and likely to appear closer to participants’ specific after-tax retirement savings goals, making those goals appear easier to achieve.

Research further shows that people adapt behavior to meet goals when they fall short (Campion and Lord 1982; Podsakoff and Farh 1989), but strive less as they approach or meet their goals (Cianci, Schaubroeck, and McGill 2010; Carver and Scheier 1998). Therefore, participants in tax-deferred accounts who perceive themselves as closer to a specific goal (by anchoring on a higher nominal balance and under-adjusting for taxes) may be less likely to take on additional investment risk to earn higher returns than participants in Roth accounts who perceive themselves as farther from their specific goals. As a result, we expect that participants in tax-deferred accounts will invest more in lower-risk, lower-return assets compared with participants in Roth accounts when both have specific retirement goals. Stated formally:

11 Additionally, taxes reduce the payoff differential between risky investments with higher expected returns and less risky investments with lower expected returns (Falsetta, Rupert, and Wright 2013). For example, consider the two mutual funds featured in our experimental design with the payoff distributions shown in Exhibit 1. Mutual Fund A exhibits lower risk and has an expected pre-tax annual return of 8.87 percent. Mutual Fund B exhibits higher risk, but has an expected pre-tax annual return of 12.30 percent. Falsetta et al. (2013) show that as tax rates increase, investors are less likely to choose a riskier investment (Mutual Fund B) because the net benefit shrinks. In this example, there is a reduced net benefit of investing in B only in the tax-deferred account condition where investment income is taxed, while in the Roth condition, the investor captures the entire expected benefit of investing in B over A and reaps the whole reward for taking on additional risk. It is only the difference in nominal starting points (i.e., pre-tax vs. after-tax contributions) between tax-deferred and Roth plans that secure equivalent future values despite this difference in after-tax returns.
**H2:** Given specific retirement goals, participants in tax-deferred retirement accounts will invest less in higher-risk, higher-return assets than participants in Roth retirement accounts.

**Reducing Differences in Investment Behavior Caused by Tax-Timing**

As we hypothesize in H1 and H2, tax-deferred retirement accounts create a setting in which participants may systematically over-estimate their progress toward a specific goal because they anchor on their pre-tax balance and under-adjust for taxes due upon withdrawal. In this section, we investigate whether current features or new initiatives in the retirement investing context can reduce or eliminate differences in investing behavior. We examine four such aspects of the investing context that might reduce this discrepancy.

*Adopting Unspecified Retirement Savings Goals for Tax-Deferred Plan Participants*

As noted above, goal theory generally holds that specific, difficult goals improve participant motivation and performance compared to unspecified, “do-your-best” goals. One noted exception to this rule is rooted in task complexity, which generally refers to the demands a task places on one’s cognitive resources (e.g., Mone and Shalley 1995, p. 245). We are aware of only two studies suggesting that performance on more complex tasks can be improved with general goals relative to specific goals. First, Kanfer and Ackerman (1989) find that general goals produced the best performance on a simulated, but realistic air traffic control task. In addition, Mone and Shalley (1995) find that participants with general goals outperformed participants with specific goals on a complex human resources hiring task where participants were charged with making hiring and salary decisions for 23 positions in eight different departments based on 46 applicant files, each containing data on 10 different dimensions.

On the continuum of task complexity, we hold that choosing investments from a limited set of mutual funds (as is the case both in many employer sponsored plans and in our experimental design) is likely simple compared to those instances where general goals have previously been shown to increase performance (i.e., air traffic control and multi-faceted hiring decisions). However, we acknowledge that

12 Anecdotal evidence suggests the typical information summary on mutual funds provided to plan participants consists of investment type or mix, selected periodic return history (e.g., 1, 5, and 10 year returns), fee rates, and a
tax deferral—the very source of under-adjustment and decreased risk assumption predicted in H1 and H2—likely requires greater computational effort from plan participants and should thereby increase the overall complexity of retirement savings compared to Roth plans. As a result, if either plan type was more amenable to an unspecified goal than the other, prior studies involving more complex tasks (e.g., Kanfer and Ackerman 1989; Mone and Shalley 1995) suggest it would be the tax-deferred plan.

Moreover, our predictions in H1 and H2 suggest that, compared to Roth plan participants, tax-deferred plan participants are hindered by poorly calibrated self-assessments of their progress toward a specific goal. Because an unspecified retirement goal lacks a clear measure of achievement, we assert that participants in a tax-deferred plan will be less likely to overestimate their investment position relative to this more ambiguous goal, and will therefore be less likely to reduce or insufficiently increase their risk profile over time than when they have specific goals. Thus, contrary to prior goal setting studies involving simple tasks (but consistent with two studies featuring highly complex tasks), we predict that an ambiguous, but important goal can alleviate the potentially harmful effects of poor goal calibration among tax-deferred retirement plan participants. Absent a specific goal, tax-deferred participants will still likely anchor on a higher nominal balance, and so will still be more likely to over-estimate their final account balance as predicted in H1. However, since they should no longer perceive themselves to be closer to a specific goal, we predict they will be more likely to invest in higher-risk, higher-return assets.

Setting Specific Goals on a Pre-Tax Basis for Tax-Deferred Plan Participants

So far, our discussion has centered on potential heuristic behaviors that may put retirement investors in tax-deferred plans at a disadvantage relative to Roth participants provided both groups have specific savings goals. However, removing a specific retirement goal as prescribed in the previous section may not be practical or feasible in situations where tax-deferred plan participants independently choose to set goals. Given that (1) specific goals are a common fixture of retirement planning and (2) tax-deferred third-party summary rating, such as a Morning Star rating. For example, TIAA-CREF allows investors to screen investments by product type (mutual fund, variable annuities, and fixed annuities), asset class, and overall Morningstar rating.
retirement savings accounts comprise the vast majority of current retirement assets (e.g., Holden and Schrass 2017a), it is important to explore other mechanisms that might mitigate this pattern in the real world.\(^{13}\) One potential avenue is to gross up a specific goal from an after-tax to a pre-tax basis for tax-deferred plan participants. Doing so increases the absolute dollar amount of the goal and eliminates the need for tax-deferred participants to (under-)estimate their future after-tax wealth. This, in turn, should both make the goal appear harder to achieve and avoid the root cause of poor goal calibration among tax-deferred plan participants, thereby increasing risk assumption.

\textit{Nudging Tax-Deferred Plan Participants to Estimate Tax Liability}

Another alternative approach to eliminating the differences in risk assumption between Roth and tax-deferred participants is to ask tax-deferred participants to explicitly estimate the taxes due on their investment balance, thereby forcing them to “do the math” to determine their consumable wealth. This “nudge” toward calculating taxes has several benefits over transforming the explicit goal into a pre-tax goal. First, it has a theoretical basis, as it should help prevent participants in the tax-deferred plan from under-adjusting for taxes. Thus, it directly addresses the under-adjustment piece of the anchoring and adjusting heuristic we posit is responsible for the original hypothesized discrepancy between Roth and tax-deferred plans. Second, pre-tax goals may prove difficult to provide to plan participants because of the numerous factors determining private effective tax rates. Instead, if participants are asked to estimate their own eventual tax-burden (or to do so with the help of a tax-professional), they may be able to more accurately account for their specific position than an impersonal advisor. Finally, this approach allows for different

\(^{13}\) As of the second quarter of 2017, research from the Investment Company Institute (ICI) indicates that 28 percent ($7.4 trillion) of all U.S. retirement assets are held in 401(k)s and other employer-sponsored defined contribution plans, while another 32 percent ($8.4 trillion) is held in IRAs (Holden and Schrass 2017b). The ICI further reports that more U.S. households hold tax-deferred than Roth IRAs (28 vs. 20 percent), and the average household wealth held in tax-deferred IRAs is more than double that of Roth IRAs (Holden and Schrass 2017a). While the ICI does not separately report Roth and tax-deferred balances for employer-sponsored plans, we note that Roth treatment has only been available for employer-sponsored plans since 2006 (see footnote 2) and is generally available only in conjunction with a tax-deferred plan because current law only allows employer matching contributions to be made in tax-deferred accounts (i.e., an employer can match an employee’s Roth contributions, but the match must be placed in a tax-deferred account). Thus, despite recent survey evidence that more than half of employers currently offer a Roth 401(k) option (Collinson 2017), tax-deferred accounts remain the most commonly offered employer-sponsored plan.
individual styles of retirement planning, as it can be used both with plan participants who choose to set explicit goals and those who choose not to set explicit goals.

*Investment Performance Feedback Relative to a Specific Retirement Savings Goal*

While we hypothesize that, all else equal, participants in tax-deferred retirement accounts will anchor on higher nominal balances and under-adjust for taxes, it is possible that performance feedback could prevent or eliminate these differences over time. Performance feedback is a mechanism commonly used in real-world settings to educate retirement plan participants on their progress toward a specific goal. For example, many online retirement portfolio management systems offer some form of visual feedback showing participants whether they are on or off track to meet their expected retirement needs, thereby forcing participants to estimate those needs (i.e., establish a specific goal) in some fashion beforehand.\(^\text{14}\) Additionally, investors may periodically consult with a financial adviser who could provide similar, if not more detailed, feedback.

Following recent psychology research (Van-Dijk and Kluger 2011), we apply Higgins’ regulatory focus theory (1997, 1998) to classify the investment task and the predicted effects of feedback on performance within this task. According to Higgins’ theory, goals that involve security needs, obligations and responsibilities, and framing in loss vs. non-loss terms serve to activate a prevention regulatory focus, which sensitizes people to negative outcomes associated with missing a goal.\(^\text{15}\) These characteristics naturally lend themselves to our setting, as retirement planning services and websites often refer to “nest eggs,” “securing your retirement,” and “comfortable retirement,” phrasing that plainly evokes security needs. Research on myopic loss aversion further suggests that investors aim at avoiding losses (i.e., a loss

\(^\text{14}\) For example, the FutureAdvisor (www.futureadvisor.com) platform compares a user’s current investment portfolio to a “recommended path” based on commonly recommended investing practices and Monte Carlo simulations of future market conditions. In addition, popular retirement investment companies (e.g., Vanguard) offer online tools intended to help users determine if their current savings and investment habits will likely cover their estimated retirement needs. These needs are typically expressed in terms of a total balance or monthly income necessary to sustain one’s current standard of living, with various assumptions and adjustments made for expected changes in spending habits (e.g., travel plans, health needs, home downsizing, dependents leaving the home, etc.).

\(^\text{15}\) On the other hand, goals that involve nurturance needs, aspirations, creativity, and gain vs. non-gain framing activate a promotion focus (Higgins 1997; Van-Dijk and Kluger 2004). Promotion regulatory focus sensitizes people to positive outcomes associated with meeting or making progress toward those goals (Higgins 1997; Van-Dijk and Kluger 2004).
vs. non-loss frame) rather than maximizing gains (Benartzi and Thaler 1999). Planning for retirement also requires careful consideration, ongoing monitoring, and attention to detail. Van-Dijk and Kluger (2011) demonstrate that tasks requiring such qualities are perceived as prevention tasks.

Van-Dijk and Kluger (2004) show that negative feedback “is congruent with prevention focus” (p. 117), with congruence leading to increased motivation and subsequent performance. Conversely, the authors show that positive feedback decreases motivation and performance compared with negative feedback for prevention tasks. Extending these findings to a retirement planning context, we similarly expect that negative feedback will cause initially low-performing participants in retirement accounts to increase their risk taking in an attempt to increase investment returns and reduce observed discrepancy with their goal (Carver and Scheier 1998).16 We further expect positive feedback to have a dampening effect on equity risk preferences, as high-performing participants who are initially closer to achieving their specific goals may believe they can reduce risk and still avoid missing their target. Thus, after receiving feedback in a multi-period setting, initially high- and low-performing retirement plan participants should both adjust their respective risk portfolios toward a point of convergence.

III. METHODS

Experimental Design

As noted above, we are primarily interested in how differences in tax structure and timing affect individuals’ ability to accurately plan for retirement (H1) and their risk profiles for retirement investment allocations (H2), as well as what features of this investing context (e.g., goal specificity, goal framing, feedback) may reduce these differences. To answer these research questions, we designed a survey-based experiment in Qualtrics to test the effects of tax timing on the aggregate investment risk of participants’ savings allocations. The basic experimental design is a 2 X 2 [(Tax Manipulation: Roth vs. Deferred) X Goal Manipulation: Specific Goal vs. Unspecified Goal)]. We then incorporated variables nested within

---

16 Obviously, feedback might also encourage a plan participant to change his or her level of contributions to retirement savings. However, we hold contribution levels fixed in our study, as they may be further limited by statute and/or budgetary constraints in real life and must still be assigned an investment risk profile once they are introduced to a plan.
these core conditions to test the various features of the retirement investment context that might reduce differences in investing behavior. Before describing each of our experimental manipulations in detail, we present a summary of our research design in Exhibit 1.

(Insert Exhibit 1 Here)

**Tax Manipulations**

We first assigned participants to one of two tax conditions: *Deferred* or *ROTH*. *Deferred* participants received pre-tax endowments each period to invest. Any earnings achieved through investment remained untaxed until the end of the final round, when the full value of the account (i.e., contributions as well as investment earnings) was taxed at a stated rate of 19 percent, which we disclosed prior to the task.\(^{17}\) In contrast, *ROTH* participants were taxed on their endowed funds each period, at the same stated rate of 19 percent, before allocating their remaining contributions to investments. Having already invested after-tax dollars, our *ROTH* participants were not taxed on any subsequent returns accumulated over the course of the experiment.

**Goal Manipulations**

After assigning participants to our *ROTH* and *Deferred* conditions, we randomly assigned retirement goals at one of two levels of specificity. Participants randomly assigned to our baseline specific goal (*SG*) conditions were given a final account balance target of $104,000 after taxes.\(^{18}\) Nested within *Deferred*, we also include a specific goal condition that grosses up the $104,000 target by our 19% tax

\(^{17}\) We chose a 19 percent tax rate to avoid trivializing the tax computation. While some current marginal tax rates are round numbers, not all are, and they typically become less round when translated to a personal effective tax rate. Additionally, by choosing 19 percent, the closest round figure to use in an estimation of taxes would be 20 percent. If participants in the *Deferred* condition accurately estimated their taxes after rounding up to 20 percent, they would tend to underestimate their after-tax account balance (albeit slightly) rather than overestimating as we both predict and observe.

\(^{18}\) As shown in Exhibit 2, the returns attributable to each investment option follow a fixed schedule common to all conditions. We presented participants in each condition with a retirement goal equal to approximately 96 percent of the maximum after-tax balance attainable from 100 percent investment in our higher-risk, higher-return asset over all trading rounds. Based on the expected return distributions provided for our investment options and fixed return schedules shown in Exhibit 1, participants in each condition had to invest approximately 76 percent of their retirement wealth in the riskier asset in order to meet this goal.
rate to produce a pre-tax target of $128,000 (PTG).\textsuperscript{19} To encourage buy-in for our assigned goal amounts, participants received their goals after answering a brief set of questions about their health, family situation, and desired post-retirement activities. While participants in the unspecified goal (UG) condition completed the same questionnaire, they did not receive a target account balance.

\textit{Nudge Manipulation}

For those assigned to a \textit{Deferred} retirement plan structure, we further assigned at random a prompt requiring participants to explicitly estimate their final tax liability after each investment round.

\textit{Feedback Manipulations}

At the midpoint of the experiment, we nested an additional feedback condition (present or absent) within conditions featuring a specific monetary retirement goal. Specifically, we provided participants in \textit{FEEDBACK} with graphical information showing that they were either “on target” or “behind target” to reach their investment goal.\textsuperscript{20} We present excerpts of our feedback manipulations in Exhibit 2.

(Insert Exhibit 2 Here)

\textit{Experimental Task}

In all conditions, the experiment consisted of two investing rounds, each representing five years of investment. In each investing round, participants received an endowment in the form of five “annual” retirement contributions and allocated these funds between the two investment options depicted in Exhibit 3: relatively low-risk Mutual Fund A and relatively high-risk Mutual Fund B. Though Mutual Fund B carries higher risk, it also offers a higher expected annual return than Mutual Fund A, which mimics the traditional tradeoffs between equity and bond investment.

(Insert Exhibit 3 Here)

\textsuperscript{19}This manipulation is nested within \textit{Deferred} because pre- and post-tax balances are identical under a Roth taxation pattern.

\textsuperscript{20}In an extensive pilot study, we observed that participants selected a median risk profile consisting of 30 percent (70 percent) investment in our low-risk (high-risk) investment option. In our final design, we assigned negative performance feedback to all \textit{FEEDBACK} participants who allocated more than 30 percent of their funds to the low-risk investment in the first round to achieve approximate balance between positive and negative feedback in our analyses.
At the end of each experimental “year,” participants received information on the realized returns of their investment choices, and their earnings and principal carried forward to allow for the periodic compounding of returns from prior rounds. At the midpoint, between years five and six (and following any performance information provided in FEEDBACK conditions), all participants had the opportunity to re-allocate their ongoing contributions, similar to the way that taxpayers may periodically review their retirement investments and adjust their investment allocations.

**Primary Dependent Variables**

Our primary dependent variables are participants’ selected investment risk profiles (i.e., the percentage of dollars allocated to high- and low-risk investments in each condition) and estimated future after-tax account balances. We base our tabulated analyses on the second round of investment for each condition. The hypothesized theoretical mechanism underlying the differences between Roth and deferred accounts depends on the deferred participants anchoring on a higher nominal balance, which is not possible until balances have been established in round one. Additionally, performance feedback can only affect the second round of investment, and therefore primarily analyzing round two decisions allows a common point of measurement across all conditions both with and without performance feedback interventions. However, for reference we include round-by-round investment patterns in the figures accompanying our tabulated analyses. In untabulated analyses, we also find that our results hold when using the average level of risky investment across both rounds as the dependent variable.

Finally, we examine participants’ average perceived difficulty of achieving their goals for retirement before and after the first round of investment. The perceived difficulty measures speak to the causal explanation described in our development of H2, namely that by anchoring on a higher nominal balance, DEFERRED participants perceive a specific goal as closer and easier to achieve than do ROTH participants.

---

21 Investment returns for each asset followed a fixed schedule (Exhibit 2) that we developed by randomly pre-selecting returns from the investment performance distributions communicated to participants. The schedule of returns is common to all conditions.
**Covariate—Risk Preference**

The instrument also captures participants’ general investment risk preferences with a series of seven questions similar to those used by many investing and retirement plan websites to measure risk appetite. These seven questions, shown in Exhibit 4, include varied response labels (e.g., “strongly agree” to “strongly disagree” and “conservative” to “aggressive”) on Likert-type scales, and two of them are reverse-coded. We used an exploratory factor analysis to produce a factor score for each participant’s risk preference. All seven items loaded on a single factor, which we use as a covariate in statistical tests of our hypotheses, where appropriate. This factor explained 48.4 percent of the total variance in our collected responses, and Cronbach’s alpha for the six items retained in the factor score is 0.794.

(Insert Exhibit 4 Here)

**Participants**

We recruited participants using Amazon’s Mechanical Turk platform for human intelligence tasks (hereafter, MTurk). We incentivized participants with a fixed payment of $1.51 for completing the instrument and variable compensation in direct proportion to their personal performance in the investment task. Our variable compensation featured a bonus for exceeding the assigned target in our $G$ conditions and a similar bonus structure for accumulated balances in excess of endowed funds for $U_G$ conditions. On average (median), our participants completed the study in 18.5 (16.5) minutes and earned $2.22 in total compensation, yielding a mean (median) effective hourly wage of $7.18 ($8.08).

Our samples include currently employed U.S. citizens between the ages of 28 and 43, who are thus likely eligible to contribute to some form of tax-favored retirement plan or account. We screened participants who failed early attention or comprehension checks before the investing exercise; this step minimized payments and subject time commitments for unusable data and avoided post-experiment

---

22 The Institutional Review Board of the authors’ respective universities approved the procedures of the experiment.

23 Our interest in retirement savings strategies could accommodate a sample ranging from the age of majority (18 in most states) to the age at which qualifying retirement distributions can be made (generally 59 ½). However, we have chosen to limit our sample to approximately the middle of this range in order to minimize variation in our participants’ familiarity with retirement planning options and prospects for retirement within the next 10 years.
rejections, which are costly for MTurk workers (see Buchheit, Doxey, Pollard, and Stinson 2018). Out of 643 complete responses, six observations showed evidence of coming from virtual private network server farms, which research suggests is often indicative of foreign responses, and were eliminated (see Dennis, Goodson, and Pearson 2018). Of the remaining 637 observations, 66 participants failed a post-experimental manipulation or comprehension check and are excluded from analyses. Additionally, we excluded 23 participants providing final account balance predictions below $15,000 or above $350,000. Our eliminations resulted in a final sample of 548 complete and valid responses. Table 1, Panel A displays the breakdown of participants per cell.

(Insert Table 1 Here)

Descriptive statistics in Table 1, Panel B suggest that our sample is appropriate for the retirement planning task. Participants have a median age of 34.6 years, income of $60,000-$79,999, and a bachelor’s degree. All participants report having saved for retirement in some fashion. Eighty-nine percent report having used a tax-preferred defined contribution savings plan, 31 percent have personal experience with a Roth 401(k) or Roth IRA, and 79 percent have used a tax-deferred 401(k) or IRA. This is consistent with recent survey evidence that tax-deferred accounts remain the dominant mode of retirement savings. Despite the increased availability of Roth retirement plans in recent years, households continue to hold greater shares of their wealth in tax-deferred IRAs, and tax-deferred defined contribution plans are more commonly offered by employers (e.g., Collinson 2017; Holden and Schrass 2017a, 2017b).

Account balances at these levels are beyond the mathematically possible outcomes in the best- and worst-case scenarios for our experimental task. Assuming the worst possible return in each period, the minimum final account balance would be $20,245, and all participants were above this level at the midpoint of the experimental task when participants made their final account balance predictions. Assuming the best possible outcome in all periods, the maximum final account balance would be $314,148 – at the midpoint of the task, no participants were able to achieve a balance higher than $261,596 assuming perfect outcomes over the remainder of the task. We surmise these extreme predictions represent typos or a fundamental misunderstanding of the task.
IV. RESULTS

Hypothesis One—Account Balance Estimates

H1 states that, all else equal, DEFERRED participants will predict higher final account balances than ROTH participants. We test H1 by comparing predicted final account balances as a percentage of actual final account balances, using ANCOVA to control for participants’ risk preferences. The ANCOVA includes the baseline ROTH SG, DEFERRED SG conditions and the ROTH UG and DEFERRED UG conditions in order to examine whether the prediction holds across the two most commonly employed styles of retirement planning. As shown in Table 2, Panels A and B, DEFERRED 10-year ending account balance estimates comprise a significantly higher percentage of their actual final account balances than ROTH estimates after the second investment round (both p ≤ 0.001), regardless of goal specificity (interaction p = 0.275).

(Insert Table 2 Here)

As shown in Panel C of Table 2, we also compute a dichotomous indicator variable for predictions that are above the participant’s final account balance (i.e., over-predictions). The chi-square test presented in Panel C indicates that DEFERRED SG participants are more likely to predict a final account balance larger than actual (65 percent) compared to ROTH SG participants (32 percent, p = 0.004). We observe similar results for conditions with an unspecified goal (8 vs. 45 percent, p < 0.001).

Taken together, the results in Table 2 provide consistent, significant evidence in support of H1. DEFERRED participants predict systematically higher final, tax-adjusted account balances and appear more likely to over-predict their final account balances than their ROTH counterparts, regardless of goal-specificity. Consumption patterns showing surprise retirement income shortfalls have been identified in archival data and only partially explained (Banks et al. 1998). Our results suggest that such unexpected shortfalls may be more common with tax-deferred retirement savings, where plan participants are more susceptible to anchoring on pre-tax nominal balances and under-adjusting for future tax effects, resulting in a higher likelihood of over-estimating their retirement wealth.
Hypothesis Two – Risk Allocation of Investments

H2 predicts that DEFERRED SG participants will invest less in a high-risk, high-return asset than ROTH SG participants. To test H2, we examine participants’ round two investments in Mutual Fund B (the riskier of the two available assets) for our two baseline conditions using an ANCOVA with tax condition as the factor and risk preference as the covariate. As shown in Table 3, Panel B, the two tax conditions significantly differ (p=0.003), with ROTH SG participants investing more in Mutual Fund B (68.9 percent) than DEFERRED participants (49.2 percent).

(Insert Table 3 Here)

To complement our main test of H2 we examine the predicted theoretical mechanism underpinning the difference in investment allocation using a process measure – participants’ perceived difficulty of reaching either their specific goal or an acceptable retirement balance. Consistent with our tests of H1, we test perceived difficulty for our ROTH SG, DEFERRED SG, ROTH UG, and DEFERRED UG conditions using an ANCOVA with tax condition and goal specificity as the factors and risk preference as the covariate. Table 4, Panel B presents the ANCOVA results and Figure 1 depicts the means plot. An unspecified goal significantly reduces perceived difficulty (p = 0.034, one-tailed equivalent) and there is a significant disordinal interaction between tax condition and goal specificity (p = 0.003). Planned contrasts in Panel C reveal that ROTH SG participants perceive the specific retirement goal to be significantly more difficult than DEFERRED SG participants after the initial investing round (5.9 vs. 5.1 on a nine-point scale, p = 0.013). Importantly, the difference reverses with an unspecified goal (4.7 vs. 5.4, p = 0.024). Taken together, the results in Tables 3 and 4 suggest specific goals are perceived as a greater challenge by Roth plan participants, who provide less optimistic estimates of their savings progress and subsequently assume greater equity risk over time than their counterparts in an otherwise equivalent tax-deferred plan.

(Insert Figure 1 Here)

(Insert Table 4 Here)
Additional Interventions

Consistent with H2, our baseline ROTH SG participants proved more effective in estimating future after-tax balances and assuming equity risk to increase future wealth than their DEFERRED SG counterparts. Our remaining analyses consider additional approaches to eliminating the differences in risk assumption introduced by differences in tax timing. Figure 2 plots the simple average and round-by-round investment in Mutual Fund B, originally reported in Table 1, for our baseline conditions as well as each of the remaining experimental cells listed in Exhibit 1. Though we test each additional intervention in more detail in the following sections, Figure 2 suggests performance feedback and prompts requiring participants to estimate future tax burdens helped produce tax-deferred investment patterns more consistent with the ROTH SG group. However, at first glance, none of our additional interventions appears to have put tax-deferred participants in a position to assume statistically greater equity risk than those in the ROTH SG group. Moreover, neither of our interventions for ROTH participants increased risk assumption compared to the baseline ROTH SG group.

(Insert Figure 2 Here)

Examining Different Levels of Goal Specificity for Different Tax Conditions

In examining whether some aspects of the retirement investing decision context may reduce or eliminate the difference in investing that we find across our baseline tax timing conditions, we predict that DEFERRED participants with an unspecified goal will invest more in a high-risk, high-return asset than those with a specific goal. We also predict that ROTH participants with a specific goal will invest more in a high-risk, high-return asset than those with an unspecified goal, which is in accordance with traditional goal theory. To test these predictions, we first examine second-round risky investment for our ROTH SG, DEFERRED SG, ROTH UG, and DEFERRED UG conditions using an ANCOVA with tax condition and goal specificity as the factors and risk preference as the covariate. We present our results in Table 5 and a plot of the sample means in Figure 3.

(Insert Figure 3 Here)

(Insert Table 5 Here)
Consistent with our previous tests of H2 and our predictions for the added intervention of an unspecified goal, the results presented in Figure 3 and Panel B of Table 5 provide significant support (one-tailed equivalent p=0.033) for a disordinal interaction between goal condition and tax condition. We further explore this interaction with a series of planned contrasts in Panel C. As expected, we find that DEFERRED UG participants invested in risk at a higher rate than their DEFERRED SG counterparts, though the difference is not statistically significant (57.6% vs. 49.2%, p = 0.749). Nevertheless, this contrasts sharply with the reduction in performance in the ROTH UG group. Consistent with traditional goal theory, we find higher investment in risk in the ROTH SG group than in the ROTH UG group (68.9% vs. 55.8%, p = 0.022).

Also consistent with our analyses in Table 3, comparing ROTH SG and DEFERRED SG again shows that the two tax conditions significantly differ (p = 0.007), with ROTH SG participants investing more in Mutual Fund B (68.9 percent) than DEFERRED SG participants (49.2 percent). Further, comparing DEFERRED and ROTH participants within the UG condition shows that the two tax conditions no longer significantly differ (p = 0.443), and the means pattern reverses with the DEFERRED UG participants investing nominally more in the risky asset than the ROTH UG participants (57.6 percent vs. 55.8 percent) in round two.

Taken together, our results thus far suggest that an unspecified goal increased the proportion of funds tax-deferred participants invest in the riskier asset compared to those given a specific goal, which is contrary to the predictions of goal theory and our own findings for Roth plan participants. Our results identify a new boundary condition to the axiom that specific goals are “better” than unspecified goals for non-complex tasks (Latham and Locke 1991; Chen and Latham 2014). Specifically, this tenet of goal-theory may not be true when task participants (1) work toward a goal that is important and subject to non-satiation, and (2) are systematically biased in their ability to compare their progress to a specific goal.

However, despite the nominal increase in the DEFERRED UG participants’ risk assumption and elimination of the between-group differences in ROTH UG and DEFERRED UG participants’ investment decisions, we note that this is partially due to the reduction in performance for ROTH UG participants. Comparing the best performance conditions from both groups in Table 5, Panel C indicates that removing
the specific goal for DEFERRED UG participants still does not raise their performance to the level of ROTH SG participants (57.6% vs. 68.9%, p = 0.044). Thus, while providing tax-deferred participants with an unspecified goal encouraged greater risk assumption than our baseline condition featuring a specific goal, it did not completely close the gap between tax-deferred and Roth investment patterns predicted in H2.

**Goal Framing**

A second aspect of the decision context that we consider is whether reframing a specific goal as a nominally higher, pre-tax goal will alleviate the differences in risk assumption between ROTH and DEFERRED participants. Figure 4 presents the round-by-round investment in the DEFERRED PTG condition along with the baseline ROTH SG and DEFERRED SG conditions. To examine the potential impact of goal framing, we first compare DEFERRED SG participants with DEFERRED PTG participants in Table 6, Panel B. Consistent with our prediction, results indicate that reframing the specific goal as a nominally higher, pre-tax goal increases DEFERRED participant’s risk assumption (62.2 percent vs. 49.2 percent, p = 0.067, one-tailed).

(Insert Figure 4 Here)

(Insert Table 6 Here)

To examine goal framing further, we compare ROTH SG participants with DEFERRED PTG participants in Table 6, Panel C. Despite the increase in risk assumption, results indicate that reframing the specific goal on a pre-tax basis does not fully alleviate the differences in risk assumption between ROTH SG and DEFERRED PTG participants. DEFERRED PTG participants remain more likely to over-predict their final balance ($\chi^2 = 3.871, p = 0.049$, untabulated), continue to believe that their goal is less difficult (4.76 vs. 5.92, F = 389.246, p < 0.001, untabulated), and accordingly invest in risk at marginally lower levels (62.2 percent vs. 68.9 percent, p = 0.084, one-tailed) than the ROTH SG participants.

**Nudge to Estimate Future Tax Burden**

A potential “nudge” that could be incorporated into the retirement investment setting is to prompt DEFERRED participants to explicitly adjust for taxes in the hopes that this will alleviate the differences in risk assumption compared to ROTH participants. Figure 5 presents round-by-round investment in the
DEFERRED NUDGE condition along with the baseline ROTH SG and DEFERRED SG conditions. To examine this potential policy recommendation, we first compare DEFERRED NUDGE participants with DEFERRED SG participants in Table 7, Panel B using the previously applied ANCOVA approach. Consistent with our predictions, results indicate that requiring participants to estimate future tax burdens at the time of investment significantly increases DEFERRED participants’ risk assumption (63.0 percent vs. 49.2 percent, p = 0.044).

(Insert Figure 5 Here)

(Insert Table 7 Here)

To examine our predictions for the effects of the nudge further, we compare ROTH SG participants with DEFERRED NUDGE participants in Table 7, Panel C. Despite the increase in risk assumption, results indicate that the added prompt does not fully alleviate the differences in risk assumption between ROTH SG and DEFERRED NUDGE participants. Rather, DEFERRED NUDGE participants continue to invest in risk at marginally lower levels than the ROTH SG group (62.2 percent vs. 68.9 percent, p = 0.081, one-tailed).

Performance Feedback

The final aspect of the decision context that we consider is whether feedback will eliminate differences in risk preference between ROTH and DEFERRED participants. Figure 6 presents the round-by-round investment in the ROTH FEEDBACK and DEFERRED FEEDBACK conditions along with the baseline ROTH SG and DEFERRED SG conditions. To test our predictions, we employ an ANCOVA on participants’ second-round investment in Mutual Fund B with tax condition and feedback as factors and risk preferences as a covariate. Table 8 presents the ANCOVA results.

(Insert Figure 6 Here)

(Insert Table 8 Here)

As shown in Panel B of Table 8, the main effect of tax condition remains significant (p = 0.001), but the main effect of feedback is not (p = 0.333). In addition, the interaction of tax condition and feedback is significant (p = 0.049). Planned contrasts in Panel C confirm that feedback does not significantly affect ROTH participants (p = 0.485), but does significantly increase risk assumption among DEFERRED
participants (p = 0.039). However, comparing DEFERRED FEEDBACK participants to ROTH SG participants shows that the two still significantly differ (p = 0.053).

Next, we examine the main and interactive effects of each intervention for tax-deferred conditions compared to the baseline DEFERRED SG group. As shown in the ANCOVA reported in Table 9, our nudge, pre-tax goal framing, and performance feedback interventions all yielded significant increases in risk assumption relative to the baseline group (all p-values ≤ 0.026). Moreover, our insignificant interactions when combining feedback with a tax estimation prompt (p = 0.236) or a pre-tax goal frame (p = 0.764) indicate that the effects of each intervention are likely additive.

Finally, we examine the effects of multiple interventions on tax-deferred plan participants’ investment selections relative to the ROTH SG baseline group with a series of ANCOVAs in Table 10 and a means plot in Figure 7. Unlike our previous analyses in which any one intervention failed to bring tax-deferred participant’s second-round investment choices in line with their ROTH SG counterparts, we find in Panel B of Table 10 that the use of multiple interventions in the DEFERRED PTG FEEDBACK condition produced statistically indistinguishable investment in our risky asset (68.9% vs. 68.3%, p = 0.967). Similarly, we find in Panel C that participants in the DEFERRED NUDGE FEEDBACK condition also invested in risk at the same rate as the ROTH SG group (68.9% vs. 65.1%, p = 0.329). Thus, it is only with significant intervention compared to our baseline DEFERRED SG condition that we can encourage tax-deferred plan participants to invest in risk at approximately the same levels as Roth participants presented with specific savings goals.

(Insert Figure 7 Here)

(Insert Table 10 Here)

V. CONCLUSION

Current trends of rising populations, life expectancies, and healthcare costs and declining certainty over the availability and magnitude of government- and employer-provided sources of retirement income have made financial independence an increasingly important goal for individuals who wish to retire.
securely and comfortably. Several government incentives designed to encourage retirement savings are currently administered through the income tax system, including the preferences afforded to qualifying tax-deferred and Roth retirement accounts. However, to the extent retirement plan participants’ expected future capital requirements exceed their contributions (due to financial constraints, government-mandated limits, or both) and current wealth, they must seek investments with adequate risk and upside potential to make up the difference. Given that prior research suggests many individuals invest too conservatively for their age (e.g., Haliassos and Bertaut 1995), leaving little margin for error in achieving reasonable retirement goals at reasonable savings rates, understanding how retirement plan tax structures impact individuals’ investment judgments is a topic with both significant practical and policy implications.

Accordingly, we conducted an experiment designed with the primary purpose of investigating individual retirement planning judgments and allocations of funds between investments of disparate risk under the tax timing applied to tax-deferred and Roth retirement accounts. In addition, we also examined other features of the decision context by manipulating retirement goal specificity, introducing a nudge to estimate tax liability, reframing savings goals on a pre-tax basis, and further examining the influence of periodic performance feedback in the presence of specific, monetary retirement savings goals. By comparing asset allocations and tax-adjusted cumulative performance across our experimental conditions, we provide evidence that the nature and timing of taxes on retirement income affect the choices made by plan participants, and that significant interventions may be necessary to produce consistent choices across plans under otherwise equivalent economic conditions. Further, we show that goal specificity can moderate these effects in a surprising way; unspecified retirement goals elicit greater goal-seeking behavior (i.e., risk assumption) from participants in a tax-deferred plan relative to participants in a Roth plan.

This study contributes to the accounting, investment, taxation, and goal setting literatures. Consistent with the use of an anchoring and adjustment heuristic (Tversky and Kahneman 1974; Epley and Gilovich 2001, 2004, 2006), we find that repeated exposure to a higher nominal balance leaves tax-deferred participants believing a specific retirement goal is easier to achieve than participants in a Roth plan. Compared to Roth participants with a specific goal, tax-deferred participants with a specific goal
erroneously predict higher ending account balances, invest less in higher-risk, higher-return assets, and are more surprised by their final balance. In contrast, tax-deferred participants with an unspecified goal invest nominally more in higher-risk assets, but still significantly less than Roth participants with a specific goal. This finding provides evidence of an important boundary condition on a core tenet of goal setting theory – that unspecified goals harm goal-seeking efforts and subsequent performance compared with specific goals.

Our study demonstrates a setting in which replacing a specific goal with an unspecified goal did not harm goal-seeking behavior in a group that was less accurate in estimating its progress toward a specific goal. In addition, while providing tax-deferred participants with pre-tax monetary goals, explicit prompts to estimate future tax burdens, or performance feedback alone did not alleviate differences compared to Roth participants with a specific retirement goal, we were able to encourage similar levels of risky investment when presenting tax-deferred participants with multiple interventions in concert.

Our findings imply that the tax-adjusted nature of Roth retirement plans aids in retirement planning, choosing investments, and responding appropriately to performance feedback. This could have serious consequences for retirement preparedness in any economy where retirement accounts are dominated by tax-deferred plans (such as the United States and the United Kingdom). For example, Banks et al. (1998) find a significant drop in consumption just prior to and at the start of retirement for households in the United Kingdom. Advanced economic models, including additional adjustments for planned vs. unplanned retirements, could not fully explain the magnitude of the drop. Our findings, including the systematic over-prediction and resulting negative surprise experienced by participants in a tax-deferred plan, may help explain this drop in consumption. If retirees with tax-deferred plans suddenly find the after-tax values of their retirement accounts are not as high as they previously believed, a drop in consumption would be rational.

While we find evidence that Roth plans may be more effective for investors with specific goals, and a concurrent study by Cuccia, Doxey, and Stinson (2018) shows that investors have a strong ceteris paribus preference for Roth tax structures, only 28 percent of our sample have real experience investing in a Roth retirement account. This is consistent with other studies of retirement savings patterns (e.g.,
Collinson 2017; Holden and Schrass 2017a, 2017b). Such an apparent disconnect between documented appetites for Roth retirement savings and actual utilization of Roth plans may be due to artificial barriers to entry imposed by tax law (e.g., contribution limits, matching restrictions, etc.), a relative lack of access to employer-provided Roth plans, or both. Overall, we believe our findings could inform future tax policy as legislators strive to encourage effective retirement savings strategies and foster greater financial independence among retirees.
REFERENCES


## EXHIBIT 1

### Experimental Design

<table>
<thead>
<tr>
<th>Condition</th>
<th>Plan Type</th>
<th>Goal Type</th>
<th>Changes from Baseline Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROTH SG</td>
<td>Roth</td>
<td>Specific</td>
<td>N/A</td>
</tr>
<tr>
<td>DEFERRED SG</td>
<td>Tax-Deferred</td>
<td>Specific</td>
<td>N/A</td>
</tr>
<tr>
<td>ROTH UG</td>
<td>Roth</td>
<td>Unspecified</td>
<td>• Participants not given an explicit monetary goal</td>
</tr>
<tr>
<td>DEFERRED UG</td>
<td>Tax-Deferred</td>
<td>Unspecified</td>
<td>• Participants not given an explicit monetary goal</td>
</tr>
<tr>
<td>DEFERRED PTG</td>
<td>Tax-Deferred</td>
<td>Specific</td>
<td>• Monetary goal expressed in pre-tax dollars</td>
</tr>
<tr>
<td>DEFERRED NUDGE</td>
<td>Tax-Deferred</td>
<td>Specific</td>
<td>• Participants prompted to estimate final tax liability after each round</td>
</tr>
<tr>
<td>ROTH FEEDBACK</td>
<td>Roth</td>
<td>Specific</td>
<td>• Performance feedback between rounds</td>
</tr>
<tr>
<td>DEFERRED FEEDBACK</td>
<td>Tax-Deferred</td>
<td>Specific</td>
<td>• Performance feedback between rounds</td>
</tr>
<tr>
<td>DEFERRED PTG FEEDBACK</td>
<td>Tax-Deferred</td>
<td>Specific</td>
<td>• Goal expressed in pre-tax dollars • Performance feedback between rounds</td>
</tr>
<tr>
<td>DEFERRED NUDGE FEEDBACK</td>
<td>Tax-Deferred</td>
<td>Specific</td>
<td>• Participants prompted to estimate final tax liability after each round • Performance feedback between rounds</td>
</tr>
</tbody>
</table>
EXHIBIT 2

Performance Feedback Used in the Experiment

Panel A: Negative feedback when participant allocates more than 30% to Mutual Fund A in the first round

Based on your previous investment allocations and prior market trends, your investment company indicates that you are currently behind target to meet your final retirement goal of $104,000 (after taxes) [$128,000 for the control condition]. Specifically, if you do not adjust your allocations between Mutual Fund A and B and both funds continue to perform at their historical averages, your projected retirement account balance at the end of Year 10 will be less than your goal.

Panel B: Positive feedback when participant allocates 30% or less to Mutual Fund A in the first round

Based on your previous investment allocations and prior market trends, your investment company indicates that you are currently on target to meet your final retirement goal of $104,000 (after taxes) [$128,000 for the control condition]. Specifically, if you do not adjust your allocations between Mutual Fund A and B and both funds continue to perform at their historical averages, your projected retirement account balance at the end of Year 10 will meet or exceed your goal.
**EXHIBIT 3**

Parameters of the Investment Decision Used in the Experiment

**Investment Decision:**

<table>
<thead>
<tr>
<th>Mutual Fund A</th>
<th>Mutual Fund B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1% chance of a 2% loss</td>
<td>5% chance of a 20% loss</td>
</tr>
<tr>
<td>5% chance of a 0% gain</td>
<td>12% chance of a 5% loss</td>
</tr>
<tr>
<td>55% chance of a 7% gain</td>
<td>45% chance of a 10% gain</td>
</tr>
<tr>
<td>30% chance of a 12% gain</td>
<td>20% chance of a 20% gain</td>
</tr>
<tr>
<td>9% chance of a 16% gain</td>
<td>18% chance of a 30% gain</td>
</tr>
</tbody>
</table>

Prior 10-Year Average Return = 8.87%  
Prior 10-Year Average Return = 12.30%

**Actual Outcomes for Each Period:**

<table>
<thead>
<tr>
<th>Period</th>
<th>Outcome for Fund A – Lower Risk</th>
<th>Outcome for Fund B – Higher Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7% gain</td>
<td>10% gain</td>
</tr>
<tr>
<td>2</td>
<td>7% gain</td>
<td>5% loss</td>
</tr>
<tr>
<td>3</td>
<td>12% gain</td>
<td>30% gain</td>
</tr>
<tr>
<td>4</td>
<td>12% gain</td>
<td>30% gain</td>
</tr>
<tr>
<td>5</td>
<td>12% gain</td>
<td>20% gain</td>
</tr>
<tr>
<td>6</td>
<td>0% gain</td>
<td>20% loss</td>
</tr>
<tr>
<td>7</td>
<td>7% gain</td>
<td>10% gain</td>
</tr>
<tr>
<td>8</td>
<td>12% gain</td>
<td>30% gain</td>
</tr>
<tr>
<td>9</td>
<td>7% gain</td>
<td>10% gain</td>
</tr>
<tr>
<td>10</td>
<td>7% gain</td>
<td>10% gain</td>
</tr>
</tbody>
</table>

Average Return = 8.3%  
Average Return = 12.5%
EXHIBIT 4

Questions Used to Measure Risk Preferences

1. Generally, I prefer an investment with little or no fluctuation in value, and I’m willing to accept a lower return on this investment. *1=Strongly Disagree to 7=Strongly Agree>*

2. During market declines, I tend to sell portions of my riskier assets and invest the money in safer assets. *1=Strongly Disagree to 7=Strongly Agree>*

3. The chart below shows the greatest one-year loss and the highest one-year gain on three hypothetical investments of $10,000. Given the potential gain or loss in any one year, I would choose the investment with the:*  

   4. The value of many types of investments can go up or down significantly, especially in the short-term. I am willing to ride out the following level of decline in my investment account: *1=Down 0-5%; 2=Down 5-10%; 3=Down 10-15%; 4=Down 15-20%, 5=Down more than 20%>*

5. I believe that at least a portion of my long-term investments should be held in stocks. *1=Strongly Disagree to 7=Strongly Agree>*

6. I am willing to accept more risk to potentially achieve higher returns. *1=Strongly Disagree to 7=Strongly Agree>*

7. What is your investment style for retirement savings? *1=Conservative; 2=Moderately Conservative; 3=Moderate; 4=Moderately Aggressive; 5=Aggressive>*


**Adaptation based on similarly worded questions across multiple retirement advisory services (e.g. Vanguard, Charles Schwab, TIAA-CREF).
TABLE 1

Observed Cell Means and Demographic Statistics

Panel A: Observed Cell Means

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>R1 %B (S.D.)</th>
<th>R2 %B (S.D.)</th>
<th>Avg. %B (S.D.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roth SG</td>
<td>38</td>
<td>60.5 (26.6)</td>
<td>68.9 (25.6)</td>
<td>64.7 (24.5)</td>
</tr>
<tr>
<td>Deferred SG</td>
<td>37</td>
<td>54.6 (26.1)</td>
<td>49.2 (26.9)</td>
<td>51.9 (23.7)</td>
</tr>
<tr>
<td>Roth UG</td>
<td>38</td>
<td>53.9 (31.8)</td>
<td>55.8 (30.3)</td>
<td>54.9 (30.3)</td>
</tr>
<tr>
<td>Deferred UG</td>
<td>42</td>
<td>60.2 (27.8)</td>
<td>57.6 (27.4)</td>
<td>58.9 (25.3)</td>
</tr>
<tr>
<td>Deferred PTG</td>
<td>37</td>
<td>61.4 (30.8)</td>
<td>62.2 (28.5)</td>
<td>61.8 (26.9)</td>
</tr>
<tr>
<td>Deferred Nudge</td>
<td>46</td>
<td>65.7 (23.8)</td>
<td>63.0 (23.2)</td>
<td>64.4 (21.4)</td>
</tr>
<tr>
<td>Roth Feedback</td>
<td>79</td>
<td>58.7 (29.3)</td>
<td>66.3 (27.9)</td>
<td>62.5 (24.7)</td>
</tr>
<tr>
<td>Deferred Feedback</td>
<td>79</td>
<td>51.9 (26.5)</td>
<td>59.11 (23.9)</td>
<td>55.5 (22.7)</td>
</tr>
<tr>
<td>Deferred PTG Feedback</td>
<td>81</td>
<td>56.0 (25.2)</td>
<td>68.3 (22.9)</td>
<td>62.2 (20.3)</td>
</tr>
<tr>
<td>Deferred Nudge Feedback</td>
<td>71</td>
<td>67.3 (24.7)</td>
<td>65.1 (25.1)</td>
<td>66.2 (20.5)</td>
</tr>
</tbody>
</table>

Panel B: Demographic Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Median / Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (minutes)</td>
<td>18.5</td>
</tr>
<tr>
<td>Age (years)</td>
<td>34.6</td>
</tr>
<tr>
<td>Income</td>
<td>$60,000 – 79,999</td>
</tr>
<tr>
<td>Education</td>
<td>Bachelor’s</td>
</tr>
<tr>
<td>Tax-preferred Experience</td>
<td>89%</td>
</tr>
<tr>
<td>Roth Experience</td>
<td>31%</td>
</tr>
<tr>
<td>Tax-deferred Experience</td>
<td>79%</td>
</tr>
</tbody>
</table>
### TABLE 2
Tests of H1

Panel A: Observed Mean Predicted % of Final Balance for Selected Cells

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROTH SG</td>
<td>38</td>
<td>83.4%</td>
<td>25.4%</td>
</tr>
<tr>
<td>ROTH UG</td>
<td>38</td>
<td>64.4%</td>
<td>27.5%</td>
</tr>
<tr>
<td>DEFERRED SG</td>
<td>37</td>
<td>101.8%</td>
<td>28.4%</td>
</tr>
<tr>
<td>DEFERRED UG</td>
<td>42</td>
<td>95.3%</td>
<td>26.3%</td>
</tr>
</tbody>
</table>

Panel B: ANCOVA on Predicted % of Final Balance

<table>
<thead>
<tr>
<th>Variation Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Preference</td>
<td>0.445</td>
<td>1</td>
<td>0.445</td>
<td>6.357</td>
<td>0.013</td>
</tr>
<tr>
<td>Tax Condition</td>
<td>2.253</td>
<td>1</td>
<td>2.253</td>
<td>32.177</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Goal Condition</td>
<td>0.772</td>
<td>1</td>
<td>0.772</td>
<td>11.031</td>
<td>0.001</td>
</tr>
<tr>
<td>Tax × Goal</td>
<td>0.084</td>
<td>1</td>
<td>0.084</td>
<td>1.202</td>
<td>0.275</td>
</tr>
<tr>
<td>Error</td>
<td>10.501</td>
<td>150</td>
<td>0.068</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: Likelihood of Over-Predicting the Final Account Balance

<table>
<thead>
<tr>
<th>Goal Condition</th>
<th>Tax Condition</th>
<th>Over-Predicted</th>
<th>Under-Predicted</th>
<th>Percent Over-Predicted</th>
<th>$\chi^2$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>ROTH</td>
<td>12</td>
<td>26</td>
<td>32%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DEFERRED</td>
<td>24</td>
<td>13</td>
<td>65%</td>
<td>8.321</td>
<td>0.004</td>
</tr>
<tr>
<td>UG</td>
<td>ROTH</td>
<td>3</td>
<td>35</td>
<td>8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DEFERRED</td>
<td>19</td>
<td>23</td>
<td>45%</td>
<td>13.954</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

*Risk Preference* = un-rotated principal component factor score of seven items measuring investment risk preferences. Higher values indicate greater risk appetite.

*Goal Condition* = SG vs. UG experimental conditions.

*Tax Condition* = ROTH vs. DEFERRED experimental conditions.

*Predicted Final Balance* = participants’ 10-year, after-tax balance predictions (i.e., balance at the end of round two).

*Predicted % of Final Balance* = 10-year balance prediction / final balance.

*One-tailed equivalent.*
# TABLE 3

## Tests of H2

### Panel A: Observed Mean R2 %B for Selected Cells

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROTH SG</td>
<td>38</td>
<td>68.9%</td>
<td>25.6%</td>
</tr>
<tr>
<td>DEFERRED SG</td>
<td>37</td>
<td>49.2%</td>
<td>26.9%</td>
</tr>
</tbody>
</table>

### Panel B: ANCOVA on R2 %B

<table>
<thead>
<tr>
<th>Variation Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Preference</td>
<td>0.550</td>
<td>1</td>
<td>0.550</td>
<td>8.861</td>
<td>0.004</td>
</tr>
<tr>
<td>ROTH SG vs. DEFERRED SG</td>
<td>0.606</td>
<td>1</td>
<td>0.606</td>
<td>9.758</td>
<td>0.003</td>
</tr>
<tr>
<td>Error</td>
<td>4.473</td>
<td>72</td>
<td>0.062</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Risk Preference* = un-rotated principal component factor score of seven items measuring investment risk preferences. Higher values indicate greater risk appetite.  
*R2 %B* = participants’ round two percentage investment in Mutual Fund B. Participants chose from 0-100% in 10% increments prior to both rounds.
TABLE 4

Perceived Difficulty

Panel A: Observed Mean Perceived Difficulty for Selected Cells

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROTH SG</td>
<td>38</td>
<td>5.9</td>
<td>1.8</td>
</tr>
<tr>
<td>ROTH UG</td>
<td>38</td>
<td>4.7</td>
<td>1.5</td>
</tr>
<tr>
<td>DEFERRED SG</td>
<td>37</td>
<td>5.1</td>
<td>1.5</td>
</tr>
<tr>
<td>DEFERRED UG</td>
<td>42</td>
<td>5.4</td>
<td>1.6</td>
</tr>
</tbody>
</table>

Panel B: ANCOVA on Perceived Difficulty

<table>
<thead>
<tr>
<th>Variation Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Preference</td>
<td>0.001</td>
<td>1</td>
<td>0.001</td>
<td>&lt;0.001</td>
<td>0.985</td>
</tr>
<tr>
<td>Goal Condition</td>
<td>8.818</td>
<td>1</td>
<td>8.818</td>
<td>3.374</td>
<td>0.034^a</td>
</tr>
<tr>
<td>Tax Condition</td>
<td>0.133</td>
<td>1</td>
<td>0.133</td>
<td>0.051</td>
<td>0.822</td>
</tr>
<tr>
<td>Goal × Tax</td>
<td>23.123</td>
<td>1</td>
<td>23.123</td>
<td>8.848</td>
<td>0.003</td>
</tr>
<tr>
<td>Error</td>
<td>391.976</td>
<td>150</td>
<td>2.613</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: Selected Planned Contrasts b

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Sum of Squares</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROTH SG vs. DEFERRED SG</td>
<td>13.196</td>
<td>5.050</td>
<td>0.013</td>
</tr>
<tr>
<td>ROTH UG vs. DEFERRED UG</td>
<td>10.400</td>
<td>3.980</td>
<td>0.024</td>
</tr>
</tbody>
</table>

Risk Preference = un-rotated principal component factor score of seven items measuring investment risk preferences. Higher values indicate greater risk appetite.
Goal Condition = SG vs. UG experimental conditions.
Tax Condition = ROTH vs. DEFERRED experimental conditions.
Perceived Difficulty = participant response to the question, “How easy or difficult do you believe it will be to reach your 10-year goal of $104,000 from this point?” on a 9-point Likert scale anchored by “impossible” and “guaranteed.” Larger values represent greater perceived difficulty.

^a One-tailed equivalent.

b To perform each the selected planned contrasts, we assigned a weight of 1 to the first cell listed, −1 to the “vs.” cell, and 0 to the two cells not listed in the comparison title.
TABLE 5

Effects of Goal Specificity on R2 %B Relative to Baseline Conditions

Panel A: Observed Mean R2 %B for Selected Cells

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROTH SG</td>
<td>38</td>
<td>68.9%</td>
<td>25.6%</td>
</tr>
<tr>
<td>ROTH UG</td>
<td>38</td>
<td>55.8%</td>
<td>30.3%</td>
</tr>
<tr>
<td>DEFERRED SG</td>
<td>37</td>
<td>49.2%</td>
<td>26.9%</td>
</tr>
<tr>
<td>DEFERRED UG</td>
<td>42</td>
<td>57.6%</td>
<td>27.4%</td>
</tr>
</tbody>
</table>

Panel B: ANCOVA on R2 %B

<table>
<thead>
<tr>
<th>Variation Source</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Preference</td>
<td>2.153</td>
<td>1</td>
<td>2.153</td>
<td>34.585</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Goal Condition</td>
<td>0.123</td>
<td>1</td>
<td>0.123</td>
<td>1.977</td>
<td>0.162</td>
</tr>
<tr>
<td>Tax Condition</td>
<td>0.391</td>
<td>1</td>
<td>0.391</td>
<td>6.283</td>
<td>0.013</td>
</tr>
<tr>
<td>Goal x Tax</td>
<td>0.214</td>
<td>1</td>
<td>0.214</td>
<td>3.432</td>
<td>0.033^a</td>
</tr>
<tr>
<td>Error</td>
<td>9.339</td>
<td>150</td>
<td>0.062</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: Selected Planned Contrasts b

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Sum of Squares</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROTH SG vs. ROTH UG</td>
<td>0.331</td>
<td>5.317</td>
<td>0.022</td>
</tr>
<tr>
<td>DEFERRED SG vs. DEFERRED UG</td>
<td>0.006</td>
<td>0.102</td>
<td>0.749</td>
</tr>
<tr>
<td>ROTH SG vs. DEFERRED SG</td>
<td>0.462</td>
<td>7.422</td>
<td>0.007</td>
</tr>
<tr>
<td>ROTH UG vs. DEFERRED UG</td>
<td>0.037</td>
<td>0.592</td>
<td>0.443</td>
</tr>
<tr>
<td>ROTH SG vs. DEFERRED UG</td>
<td>0.611</td>
<td>9.813</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Risk Preference = un-rotated principal component factor score of seven items measuring investment risk preferences. Higher values indicate greater risk appetite.

Goal Condition = SG vs. UG experimental conditions.

Tax Condition = ROTH vs. DEFERRED experimental conditions.

R2 %B = participants’ round two percentage investment in Mutual Fund B. Participants chose from 0-100% in 10% increments prior to both rounds.

^a One-tailed equivalent.

b To perform each the selected planned contrasts, we assigned a weight of 1 to the first cell listed, −1 to the “vs.” cell, and 0 to the two cells not listed in the comparison title.
TABLE 6
Effects of Goal Framing on R2 %B Relative to Baseline Conditions

Panel A: Observed Mean R2 %B for Selected Cells

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROTH SG</td>
<td>38</td>
<td>68.9%</td>
<td>25.6%</td>
</tr>
<tr>
<td>DEFERRED SG</td>
<td>37</td>
<td>49.2%</td>
<td>26.9%</td>
</tr>
<tr>
<td>DEFERRED PTG</td>
<td>37</td>
<td>62.2%</td>
<td>28.5%</td>
</tr>
</tbody>
</table>

Panel B: ANCOVA on R2 %B, DEFERRED SG vs. DEFERRED PTG

<table>
<thead>
<tr>
<th>Variation Source</th>
<th>Type III</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Preference</td>
<td>0.637</td>
<td>1</td>
<td>0.637</td>
<td>9.241</td>
<td>0.003</td>
</tr>
<tr>
<td>DEFERRED SG vs. DEFERRED PTG</td>
<td>0.158</td>
<td>1</td>
<td>0.158</td>
<td>2.292</td>
<td>0.067a</td>
</tr>
<tr>
<td>Error</td>
<td>4.893</td>
<td>71</td>
<td>0.069</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: ANCOVA on R2 %B, ROTH SG vs. DEFERRED PTG

<table>
<thead>
<tr>
<th>Variation Source</th>
<th>Type III</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Preference</td>
<td>0.546</td>
<td>1</td>
<td>0.546</td>
<td>8.195</td>
<td>0.005</td>
</tr>
<tr>
<td>ROTH SG vs. DEFERRED PTG</td>
<td>0.129</td>
<td>1</td>
<td>0.129</td>
<td>1.943</td>
<td>0.084a</td>
</tr>
<tr>
<td>Error</td>
<td>4.793</td>
<td>72</td>
<td>0.067</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Risk Preference = un-rotated principal component factor score of seven items measuring investment risk preferences. Higher values indicate greater risk appetite.

R2 %B = participants’ round two percentage investment in Mutual Fund B. Participants chose from 0-100% in 10% increments prior to both rounds.

a One-tailed equivalent.
TABLE 7
Effects of Tax Estimation Prompt on R2 %B Relative to Baseline Conditions

Panel A: Observed Mean R2 %B for Selected Cells

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROTH SG</td>
<td>38</td>
<td>68.9%</td>
<td>25.6%</td>
</tr>
<tr>
<td>DEFERRED SG</td>
<td>37</td>
<td>49.2%</td>
<td>26.9%</td>
</tr>
<tr>
<td>DEFERRED NUDGE</td>
<td>37</td>
<td>63.0%</td>
<td>23.2%</td>
</tr>
</tbody>
</table>

Panel B: ANCOVA on R2 %B, DEFERRED SG vs. DEFERRED NUDGE

<table>
<thead>
<tr>
<th>Variation Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Preference</td>
<td>0.539</td>
<td>1</td>
<td>0.539</td>
<td>9.602</td>
<td>0.003</td>
</tr>
<tr>
<td>DEFERRED SG vs. DEFERRED NUDGE</td>
<td>0.237</td>
<td>1</td>
<td>0.237</td>
<td>4.222</td>
<td>0.043</td>
</tr>
<tr>
<td>Error</td>
<td>4.486</td>
<td>80</td>
<td>0.056</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: ANCOVA on R2 %B, ROTH SG vs. DEFERRED NUDGE

<table>
<thead>
<tr>
<th>Variation Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Preference</td>
<td>0.469</td>
<td>1</td>
<td>0.469</td>
<td>8.704</td>
<td>0.004</td>
</tr>
<tr>
<td>ROTH SG vs. DEFERRED NUDGE</td>
<td>0.107</td>
<td>1</td>
<td>0.107</td>
<td>1.993</td>
<td>0.081^a</td>
</tr>
<tr>
<td>Error</td>
<td>4.364</td>
<td>81</td>
<td>0.054</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Risk Preference = un-rotated principal component factor score of seven items measuring investment risk preferences. Higher values indicate greater risk appetite.

R2 %B = participants’ round two percentage investment in Mutual Fund B. Participants chose from 0-100% in 10% increments prior to both rounds.

^a One-tailed equivalent.
### TABLE 8

Effects of Performance Feedback on R2 %B Relative to Baseline Conditions

#### Panel A: Observed Mean R2 %B for Selected Cells

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roth SG</td>
<td>38</td>
<td>68.9%</td>
<td>25.6%</td>
</tr>
<tr>
<td>Deferred SG</td>
<td>37</td>
<td>49.2%</td>
<td>26.9%</td>
</tr>
<tr>
<td>Roth Feedback</td>
<td>79</td>
<td>66.3%</td>
<td>27.9%</td>
</tr>
<tr>
<td>Deferred Feedback</td>
<td>79</td>
<td>59.1%</td>
<td>23.9%</td>
</tr>
</tbody>
</table>

#### Panel B: ANCOVA on R2 %B

<table>
<thead>
<tr>
<th>Variation Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Preference</td>
<td>1.696</td>
<td>1</td>
<td>1.696</td>
<td>27.904</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Tax Condition</td>
<td>0.641</td>
<td>1</td>
<td>0.641</td>
<td>10.553</td>
<td>0.001</td>
</tr>
<tr>
<td>Feedback</td>
<td>0.057</td>
<td>1</td>
<td>0.057</td>
<td>0.942</td>
<td>0.333</td>
</tr>
<tr>
<td>Tax × Feedback</td>
<td>0.238</td>
<td>1</td>
<td>0.238</td>
<td>3.909</td>
<td>0.049</td>
</tr>
<tr>
<td>Error</td>
<td>13.855</td>
<td>228</td>
<td>0.061</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Panel C: Selected Planned Contrasts

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Sum of Squares</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roth SG vs. Roth Feedback</td>
<td>0.030</td>
<td>0.488</td>
<td>0.485</td>
</tr>
<tr>
<td>Deferred SG vs. Deferred Feedback</td>
<td>0.262</td>
<td>4.314</td>
<td>0.039</td>
</tr>
<tr>
<td>Roth SG vs. Deferred Feedback</td>
<td>0.160</td>
<td>2.635</td>
<td>0.053</td>
</tr>
</tbody>
</table>

*Risk Preference* = un-rotated principal component factor score of seven items measuring investment risk preferences. Higher values indicate greater risk appetite.

*Tax Condition* = Roth vs. Deferred.

*Feedback* = Feedback present vs. absent.

*R2 %B* = participants’ round two percentage investment in Mutual Fund B. Participants chose from 0-100% in 10% increments prior to both rounds.

*a To perform each the selected planned contrasts, we assigned a weight of 1 to the first cell listed, −1 to the “vs.” cell, and 0 to the two cells not listed in the comparison title.
TABLE 9

Effects of All Interventions on Tax-Deferred Plan Participants

Panel A: Observed Mean R² %B for Selected Cells

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEFERRED SG</td>
<td>37</td>
<td>49.2%</td>
<td>26.9%</td>
</tr>
<tr>
<td>DEFERRED FEEDBACK</td>
<td>79</td>
<td>59.1%</td>
<td>23.9%</td>
</tr>
<tr>
<td>DEFERRED PTG</td>
<td>37</td>
<td>62.2%</td>
<td>28.5%</td>
</tr>
<tr>
<td>DEFERRED PTG FEEDBACK</td>
<td>81</td>
<td>68.3%</td>
<td>22.9%</td>
</tr>
<tr>
<td>DEFERRED NUDGE</td>
<td>46</td>
<td>63.0%</td>
<td>23.2%</td>
</tr>
<tr>
<td>DEFERRED NUDGE FEEDBACK</td>
<td>71</td>
<td>65.1%</td>
<td>25.1%</td>
</tr>
</tbody>
</table>

Panel B: ANCOVA on R² %B

<table>
<thead>
<tr>
<th>Variation Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Preference</td>
<td>1.516</td>
<td>1</td>
<td>1.516</td>
<td>26.790</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Prompt</td>
<td>0.282</td>
<td>1</td>
<td>0.282</td>
<td>4.976</td>
<td>0.026</td>
</tr>
<tr>
<td>Goal Frame</td>
<td>0.441</td>
<td>1</td>
<td>0.441</td>
<td>7.793</td>
<td>0.006</td>
</tr>
<tr>
<td>Feedback</td>
<td>0.284</td>
<td>1</td>
<td>0.284</td>
<td>5.020</td>
<td>0.026</td>
</tr>
<tr>
<td>Feedback × Prompt</td>
<td>0.080</td>
<td>1</td>
<td>0.080</td>
<td>1.410</td>
<td>0.236</td>
</tr>
<tr>
<td>Feedback × Goal Frame</td>
<td>0.005</td>
<td>1</td>
<td>0.005</td>
<td>0.090</td>
<td>0.764</td>
</tr>
<tr>
<td>Error</td>
<td>19.469</td>
<td>344</td>
<td>0.057</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Risk Preference = un-rotated principal component factor score of seven items measuring investment risk preferences. Higher values indicate greater risk appetite.
Prompt = Explicit estimation of taxes due on retirement balance (NUDGE) vs. no estimation.
Goal Frame = SG vs. PTG experimental conditions.
Feedback = Feedback present vs. absent.
R² %B = participants’ round two percentage investment in Mutual Fund B. Participants chose from 0-100% in 10% increments prior to both rounds.
### TABLE 10

Multiple Intervention Effects on R2 %B Relative to *ROTH SG*

#### Panel A: Observed Mean R2 %B for Selected Cells

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ROTH SG</em></td>
<td>38</td>
<td>68.9%</td>
<td>25.6%</td>
</tr>
<tr>
<td><em>DEFERRED PTG FEEDBACK</em></td>
<td>81</td>
<td>68.3%</td>
<td>22.9%</td>
</tr>
<tr>
<td><em>DEFERRED NUDGE FEEDBACK</em></td>
<td>71</td>
<td>65.1%</td>
<td>25.1%</td>
</tr>
</tbody>
</table>

#### Panel B: ANCOVA on R2 %B, *ROTH SG* vs. *DEFERRED PTG FEEDBACK*

<table>
<thead>
<tr>
<th>Variation Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Preference</td>
<td>0.198</td>
<td>1</td>
<td>0.198</td>
<td>3.599</td>
<td>0.060</td>
</tr>
<tr>
<td><em>ROTH SG</em> vs. <em>DEFERRED PTG FEEDBACK</em></td>
<td>0.001</td>
<td>1</td>
<td>0.001</td>
<td>0.002</td>
<td>0.967</td>
</tr>
<tr>
<td>Error</td>
<td>6.393</td>
<td>116</td>
<td>0.055</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Panel C: ANCOVA on R2 %B, *ROTH SG* vs. *DEFERRED NUDGE FEEDBACK*

<table>
<thead>
<tr>
<th>Variation Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Preference</td>
<td>0.454</td>
<td>1</td>
<td>0.454</td>
<td>7.550</td>
<td>0.007</td>
</tr>
<tr>
<td><em>ROTH SG</em> vs. <em>DEFERRED NUDGE FEEDBACK</em></td>
<td>0.058</td>
<td>1</td>
<td>0.058</td>
<td>0.963</td>
<td>0.329</td>
</tr>
<tr>
<td>Error</td>
<td>6.379</td>
<td>106</td>
<td>0.060</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Risk Preference* = un-rotated principal component factor score of seven items measuring investment risk preferences. Higher values indicate greater risk appetite.

*R2 %B* = participants’ round two percentage investment in Mutual Fund B. Participants chose from 0-100% in 10% increments prior to both rounds.
FIGURE 1

Means for Perceived Difficulty in Baseline Conditions

Perceived Difficulty

5.9
5.8
5.6
5.4
5.2
5.0
4.8
4.6
4.4
4.2
4.0

SG
UG

5.1
5.4

4.7

ROTH
DEFERRED
FIGURE 2

Investment in Risky Asset by Experimental Condition

Panel A: Overall Average %B

Panel B: Round-by-Round %B
FIGURE 3
Unspecified Goal Effects on Round-by-Round %B

R1 R2

Roth SG
Roth UG
Deferred SG
Deferred UG
FIGURE 4
Goal Framing Intervention’s Effect on Round-by-Round %B

70
65
60
55
50
45

61.4
60.5
62.2
68.9
54.6
49.2

R1 R2

%B

ROTH SG
DEFERRED SG
DEFERRED PTG
FIGURE 5
Estimation Intervention’s Effect on Round-by-Round %B

- Roth SG
- Deferred SG
- Deferred Nudge
FIGURE 6

Feedback Intervention’s Effect on Round-by-Round %B

- **ROTH SG**
- **ROTH FEEDBACK**
- **DEFERRED SG**
- **DEFERRED FEEDBACK**

![Graph showing the effect of feedback intervention on round-by-round %B]
FIGURE 7

Effect of Multiple Interventions on Round-by-Round %B

67.3
68.9
68.3
65.1
60.5
56.0

Roth SG

Deferred PTG
Feedback

Deferred Nudge
Feedback

R1 R2