Do Tax Deferred Accounts Improve Lifecycle Savings? Experimental Evidence

John Duffy[*] Yue Li[†]

July 20, 2023

Abstract

In an individual decision-making experiment, we investigate the impact of Tax Deferred Accounts (TDAs). We design six treatments to study various channels through which TDAs may affect decisions. Across both student and Mturk samples, we consistently find that TDAs significantly increase retirement wealth compared to environments with only one non-tax advantaged, liquid saving account. This increase is primarily explained by the requirement of making retirement saving decisions precede consumption decisions. Educating participants by providing a tax calculator has minimal effects. Our results highlight the effectiveness of TDAs in enhancing retirement preparedness and the significance of the order of consumption/saving decisions.

Keywords: Tax Deferred Accounts, Lifecycle Model, Consumption, Savings, Retirement Planning, Behavioral and Experimental Economics.

JEL Codes: C91, C91 D15, D91, E21, H55.

[*]Department of Economics, University of California, Irvine. Email: duffy@uci.edu.

[†]Department of Economics, University of Albany, SUNY, Email: yli49@albany.edu.

1 Introduction

To encourage private savings for retirement, the U.S. government allows individuals to defer earned income from immediate taxation if it is contributed to an Individual Retirement Account (IRA), 401(k), and similar types of tax deferred accounts (TDAs). While TDAs were originally developed to be a complement to defined benefit retirement (pension) plans, the decline in those plans in the U.S. has raised the profile of tax deferred, defined contribution plans, making these TDAs the major means by which households supplement their social security benefits in retirement. According to the Investment Company Institute, as of the first quarter 2023, U.S. households held $22.3 trillion in TDAs, comprising 63% of U.S. total retirement assets. TDAs for retirement savings are also found outside of the U.S. among nearly all OECD countries (Yoo and De Serres, 2004).

A considerable amount of research has been directed toward getting workers to sign up for TDAs including nudges, subsidies, automatic enrollments and default investment options. In this paper we take a step back and ask the bigger question of whether TDAs are in fact, beneficial to workers as vehicles for saving for retirement when the complexity of signing up and making regular contributions is eliminated. Specifically, we ask whether the presence or absence of TDAs matters for households' consumption, retirement asset positions and lifetime utility. To answer such questions, we purposefully minimize the frictions associated with enrolling and contributing to TDAs. For instance, we educate subjects about the tax advantages of TDAs and in some cases provide them with a tax calculator. This simplified environment enables us to focus on the broader question of whether and how TDAs affect lifecycle consumption and saving decisions relative to their absence.[1]

That question is difficult to answer using field data since TDAs have been around for a while now – the first 401(k) dates to 1981 – making counterfactual analysis of lifecycle saving behavior with and without TDAs difficult to study. The existing counterfactual studies are simulation studies that evaluate the effects of TDAs using models with fully rational agents, for example Nishiyama (2011), Ho (2017) and Horneff et al. (2021, 2023). However, the empirical and experimental literature on intertemporal saving decisions finds mixed support for the rational agent model (see e.g. Arifovic and Duffy (2018)) and so we do not presently know the impact of TDAs when households may be less than fully rational.

To fill this gap in the literature, we conduct individual decision-making experiments to investigate the impact of TDAs on consumption saving decisions. We design six different treatments to study the various channels through which TDAs may affect individual decisions. The main features of the six treatments are summarized in Table 1. Our baseline treatment (T1) includes two savings vehicles, a TDA and a regular saving account (RA) which does not have any tax advantage and is fully liquid. Treatment T2 eliminates the TDA, but is otherwise identical to T1. T3 is similar to T2 but in T3 we reduce taxes on interest income so that government revenues are theoretically the same in T1 for a more honest comparison of the impact of TDAs on savings behavior and for understanding the effect of providing a simple tax incentive alone on savings. Treatment T4 is based on T1 but removes the tax calculator for TDA incentives from the decision screen, allowing us to determine the impact of providing additional education. Treatments T5 and T6 are based on the no TDA treatment T2, but include a second savings vehicle, a commitment account (CA), which disallows early withdrawals before retirement but has no tax advantages relative to the RA. In T5, subjects are required to make a CA saving choice first followed by a consumption choice and the residual going to the RA. In T6, the consumption choice is made first followed by the CA saving decision. The difference between T6 and T2 shines a light on the role of the commitment-device aspect of TDAs, and the difference between T5 and T6 reveals the effect of requiring that retirement saving decisions precede consumption decisions. Additionally, we collect data on subjects' cognitive abilities, risk preferences and financial literacy to account for possible departures from the rational choice benchmark associated with individual characteristics.

Our experiment is designed to mimic important features of real-world, lifecycle saving decisions with or without TDAs. Subjects supply labor inelastically. They face a life-cycle earning profile that is hump-shaped in the first 15 periods, representing "working phase" of their lives, and they receive a constant pension amount in retirement. Subjects always have access to a perfectly liquid RA. If given access to a second savings vehicle, a TDA or CA, subjects can only make contributions to that account in the working phase and take withdrawals from that account in retirement.[2] We impose a progressive income tax schedule, set TDA contribution limits that mimic U.S. rules, and incorporate longevity risks.

Importantly, our experiment is incentivized; subjects' monetary payments depend upon and vary with their consumption/saving decisions over the lifecycle. Within this setup, we are able to solve the model under each regime so as to provide a benchmark for optimal (fully rational) behavior which we use to compare with our experimental data. As mentioned before, according to the rational choice theory, the provision of a tax calculator or a non-tax-advantaged CA, or the change of decision orders should not affect decisions.

As a preview, for our student subjects (referred to as the "lab sample"), we find that TDAs substantially increase "retirement wealth", defined by their net worth at the end of the working phase. On average, subjects in T1 save about 17% of their endowment during the working phase, while subjects in T2 and T3 save only about 8%-10%. At the end of the working phase, the average net worth in T1 is 429.3k greater than that in T2. As we show, the primary contributing mechanism for this net worth gap is the requirement that saving decisions are made first (before consumption), accounting for 49% of the gap, followed by the commitment-device aspect of TDAs, which accounts for 36% of the gap. The effect of educating subjects by providing a tax calculator on TDA contributions is minimal and insignificant. However, we observe that, on average, subjects assigned to T4 exhibit a faster decumulation of their assets in the retirement phase compared with subjects in T1, suggesting the absence of the tax calculator plays a role in subjects' TDA withdrawal decisions. We also acknowledge that other educational devices might be useful, for example see Duflo and Saez (2003); Hu et al. (2013); Clark et al. (2014); Goda et al. (2014) and Choi et al. (2017). Lowering taxes on interest income (T3 versus T2) yields little change in saving rates. Regarding welfare implications, we do not find large differences across our six treatments, with the sole exception of T5 (where the CA saving choice is made first) where subjects have significantly lower expected lifetime utility than do subjects in the other five treatments. The absence of treatment effects on expected lifetime utility mask the distributional differences that subjects in T1 have lower lifetime utility compared to those in T2 and T3 if the experiment ends early, and they experience higher lifetime utility if the experiment extends to a later period.

Finally, we note that laboratory studies with student subjects may not be so representative of the general population, as students may not have much experience with saving for retirement and their demographic backgrounds may differ from a broader cross-section of the population. Thus, we also report on a replication study where we recruit subjects from Amazon's Mechanical Turk (MTurk) online workforce, who are, on average, older and more racially diverse than our student sample. We have these MTurkers participate in the *same* treatments of our original study. Consistently across both samples, we observe a substantial increase in retirement wealth with TDAs, an insignificant educational effect and an important order effect, whereby subjects tend to save more if they are required to make a saving decision first. The replication of our main findings using the MTurk sample suggests that our findings with undergraduate students may generalize to other subject populations. Still, we acknowledge the limits of extrapolating from experimental evidence to real-world behavior.

2 Related Literature

This paper contributes to two strands of literature. The first uses field experiment methods to study factors impacting TDA uptake. This literature generally finds that changing the TDA default from opting in to opting out is a powerful tool for increasing participation in TDAs and contributions to TDAs (for example, see Madrian and Shea (2001); Choi et al. (2004); Beshears et al. (2010); Chetty et al. (2014); Goda et al. (2020)). The evidence for other means of increasing TDA uptake are more mixed. For instance, Duflo et al. (2006)

find that taxpayers' IRA contributions react more to a match than to dollar equivalent tax credits, likely due to differences in information and framing. Choi et al. (2011) find that 36% of employees over 59 1/2 do not utilize the free lunch of an employer match, and an educational survey does not produce significant improvement. Using administrative data from eleven companies that added a Roth contribution option, Beshears et al. (2017) find that the introduction of Roth plans (where contributions are made from after-tax income so that withdrawals are tax free) result in no change in contribution rates, suggesting that employees are either neglecting the different tax properties of Roth plans versus traditional plans, or have the tendency to allocate a constant amount to retirement savings (partition dependence). Beshears et al. (2020) conduct an experiment using the American Life Panel, and find that higher early withdrawal penalties attract more commitment account deposits, indicating some participants are partially- or fully-sophisticated present-biased.

This paper complements this literature by comparing a controlled laboratory environment with a TDA to environments without a TDA. The goal is to understand the broader implications of introducing TDAs in the presence of likely behavioral irrationality. Conducting laboratory experiments allows us to introduce changes to the decision-making environment one-by-one, and decompose the overall effects of TDAs into various channels.

A second strand of literature studies life-cycle consumption and saving decisions in a controlled laboratory environment. Within this strand, our work is closely related to several recent papers that study the effects of tax incentives on retirement savings. For instance, Bohr et al. (2023) compare traditional retirement accounts to Roth retirement accounts in a setting where both treatments share the same predicted behaviors and find no significant differences between the two treatments. The next three papers do not induce utility like ours, and thus lack benchmark theory predictions for comparison. Blaufus and Milde (2021) find that after-tax pensions supported by a traditional account are 25% lower than those by an economically equivalent Roth account, indicating substantial tax misperceptions. Behavioral differences across two treatment arms are preserved with the recurrent provision of information nudges about tax savings on contributions and taxation on withdrawals individually, but disappear with the joint provision of both types of information. Cuccia et al. (2022) offer subjects a choice between traditional and Roth accounts that are economically equivalent. They find that non-economic attitudes and preferences have a significant impact on plan choices, while additional information nudges are required for changes in tax rates to have a significant impact. Blaufus et al. (2023) examine the effect of changing tax treatments on TDA contribution and withdrawals, and find that savings rates are sensitive to the contribution tax refund but do not react to the tax rate on withdrawals. Our laboratory environment differs from the aforementioned experiments in three key aspects. First, we provide subjects with a second savings vehicle that yields the same return but differs in liquidity and tax treatment. Second, we introduce random termination in the retirement phase to mimic longevity risks. Third, we consider changing the order of consumption and saving decisions.

3 Theoretical Framework

Our theoretical model is similar in some respects to Duffy and Li (2019), but we add survival risk and a second savings vehicle. Agents live for a certain $Jr - 1$ periods as workers and receive endowment income e_j in periods $j = 1, ..., Jr - 1$. Starting from the first period of retirement (Jr), agents face survival risk, modeled using a constant probability, s, that the retiree survives from one period to the next. For each period in retirement, living agents receive endowment income e_{Jr}, which can be viewed as a social security benefit. Over their lifetimes, agents make consumption and saving decisions in each period. They always have access to a RA and depending on the treatment, they may also have access to a TDA or CA. The CA is identical to the RA, except that it is illiquid during the working periods like the TDA. Below we first present the decision problem for the baseline TDA treatment, and later we will clarify how we modify the problem to fit the other treatments.

Both TDA and RA share certain characteristics, such as an initial balance of zero, a non-borrowing constraint, and a constant pre-tax interest rate of r. However, they differ in several key aspects. First, RA is a liquid account that generates interest income, which is subject to immediate taxation. Second, TDA is illiquid, allowing contributions solely during the working phase and permitting withdrawals exclusively in retirement. Third, TDA contributions are made with pre-tax income, and so it is exempt from income taxes in the contribution period. The contribution amount for each working period j is capped at the lesser the current endowment income e_j and a statuary limit $\overline{q}_j$. Lastly, any withdrawal from TDA is treated as taxable income.

Let a_j denote the period $j-1$ ending RA balance, and b_j denote the period $j-1$ ending TDA balance. These two variables combined with the period number, j form a state vector for each agent. Each period, agents choose their consumption (c_j), and their contribution to the TDA (q_j), which takes on a negative value if a withdrawal is made. The government collects taxes according to a known progressive tax function $T(e_j, ra_j, q_j)$, and this tax function may vary by treatment.

For the TDA treatment, we can write the recursive problem for workers as follows:

$$V_j(a_j, b_j) = \begin{cases} \max_{\{c_j, q_j\}} u(c_j) + \beta V_{j+1}(a_{j+1}, b_{j+1}) & \text{if } j < Jr - 1 \\ \max_{\{c_j, q_j\}} u(c_j) + \beta V_r(a_{j+1}, b_{j+1}) & \text{if } j = Jr - 1 \end{cases}$$

subject to

$$c_j + q_j + a_{j+1} = e_j + (1+r)a_j - T(e_j, ra_j, q_j) \tag{1}$$

$$b_{j+1} = b_j(1+r) + q_j \tag{2}$$

$$0 \leq q_j \leq \min\{e_j, \overline{q}_j\} \tag{3}$$

$$a_{j+1} \geq 0, \quad b_{j+1} \geq 0, \quad a_1 = b_1 = 0 \tag{4}$$

where $V_j(\cdot)$ denotes the value function of workers for period j, and $V_r(\cdot)$ denotes the value function for retirees. $u(\cdot)$ is the utility function, and β is the discount factor. Equation (1) defines the budget constraint. Equation (2) calculates TDA ending balance. Condition (3) describes the TDA contribution restriction.

The recursive problem for retirees is:

$$V_r(a_j, b_j) = \max_{\{c_j, q_j\}} u(c_j) + \beta s V_r(a_{j+1}, b_{j+1})$$

subject to equations (1), (2), (4), and

$$0 \geq q_j \geq -b_j(1+r)$$

The agent problem for the non-TDA treatments is similar to one presented above, but with three modifications. First, in treatments with a CA, we use q_t to denote the contribution to and withdrawal from the CA, and b_t to denote the CA balance. In treatments without a CA, both q_t and b_t are set to 0. Second, there are no tax incentives for saving for retirement: $T(e_j, ra_j, q_j)$ is replaced with $T(e_j, ra_j + rb_j, 0)$. Third, there is no limit on CA contribution. It is worth noting that the addition of CA does not alter the optimal path of consumption and net worth for a rational agent.

To facilitate quantitative analysis on savings, we introduce three statistics. First, we employ the contribution to endowment ratio (q_j/e_j) to gauge the tendency to deposit into an illiquid retirement account. Second, we utilize the total saving to endowment ratio (x_j/e_j), to measure the overall saving tendency:

$$x_j = \begin{cases} \underbrace{e_j + r(a_j + b_j) - T(e_j, ra_j + rb_j, 0)}_{\text{after tax income}} - c_j & \text{w/o TDA} \\ \underbrace{e_j + ra_j - q_j - T(e_j, ra_j, q_j)}_{\text{after tax income}} - c_j + \underbrace{q_j}_{\text{contribution}} - \underbrace{(T(e_j, ra_j, 0) - T(e_j, ra_j, q_j))}_{\text{tax savings on contribution}} & \text{w TDA} \end{cases} \tag{5}$$

If an increase in the contribution does not coincide with an increase in the total savings, it indicates that the introduction of TDA does not lead to a reduction in consumption or the emergence of new savings.

Third, we define a revised TDA balance (b^*_{j+1}) by deducting all TDA tax savings:

$$b^*_{j+1} = \begin{cases} q_j - (T(e_j, ra_j, 0) - T(e_j, ra_j, q_j)) & \text{if } j = 1 \\ (1+r)b^*_j + q_j - (T(e_j, r(a_j + b^*_j), 0) - T(e_j, ra_j, q_j)) & j \geq 2 \end{cases} \tag{6}$$

The comparison between $a_{j+1} + b_{j+1}$ and $a_{j+1} + b^*_{j+1}$ in the TDA treatment reveals the mechanical effect of TDA tax incentives, while the comparison between $a_{j+1} + b^*_{j+1}$ in the TDA treatment and $a_{j+1} + b_{j+1}$ in the non-TDA treatment T2 reveals the behavioral effect of TDA tax incentives.

4 Experimental Design

In this section we first discuss the parameterization of the model that we used in the experiment. Next, we present the rational choice solution for our different treatments, which we use to formulate theoretical predictions for evaluating our experimental data. Finally, we discuss the experimental procedures that we followed in collecting the data.

4.1 Parameterization

Assuming that each period in the model represents 2 years, we set $Jr - 1 = 15$, which corresponds to a working phase ranging from age 30 to age 59. All numbers presented below are for each model period unless otherwise noted. Starting from the first period of retirement, agents face a survival probability of $s = 0.9$, and so the average life expectancy is around ages 78-79, which is close to U.S. life expectancy. Appendix Figure C1.(a) displays the unconditional survival rates, which are 1 from periods 1 to 16 and then decline. The retirement phase was implemented using the block random termination design of Fréchette and Yuksel (2017), so that the random numbers that determine the continuation of the experiment will only be revealed at the end of each 10-period block in the retirement phase.

Within each block, subjects know that there is a constant 10% probability that the current period is the final period of the sequence, and any future decisions will not matter for their experimental payoffs. However, subjects do not learn whether and when the sequence ends until the end of that 10 period block; if the sequence does not end within a 10 period block, then another 10 period block is played, etc. This design allows us to collect the decisions of at least 10 post-retirement periods for every subject. In practice, to control the length of the experiment, we pre-drew 20 different sequences of random numbers from 1 to 10 where each draw was made with equal probability. A sequence ended the first time that a 1 was drawn. Subjects were randomly assigned to one of these 20 possible sequences for the post-retirement phase. The aim here was to allow for some heterogeneity in lifetimes across subjects. The maximum number of periods of the 20 post-retirement sequences was 35 (equivalent to ages 98-99). The average termination period of the 20 post-retirement sequences is period 25 (equivalent to ages 78-79) which, as previously noted, is also the empirical U.S. life expectancy.

We set the endowment process, the tax function, and the statuary contribution limit for the TDA to match those of the 2015 U.S. economy. To make the environment easier for subjects to understand, we round up the endowment income and the tax bracket cutoff to the nearest thousand dollar amount. We construct the lifecycle endowment income using the sum of total personal earned income and social security income in the 2015 American Community Survey. We approximate the average income profile for workers using a fourth order polynomial function of age, and the value for retirees is set to the average income for those aged 60 and above. Appendix Figure C1.(b) displays the estimated endowment income for every model period (two years per period). Notice that the endowment/income profile is "hump-shaped" over the first 15 working periods and is then flat during the retirement phase. We use this same endowment/income profile in all six of our experimental treatments.

We set the tax function to match the U.S. federal income tax schedule for the year 2015, as displayed in Appendix Figure C1.(c). Income below $21,000 per model period is exempt from taxation to capture the provision of a personal exemption and standard deduction. We restrict the highest marginal tax rate to be 25 percent to simplify the environment, as the income brackets relevant for a 28 percent or greater tax rate (applicable for ordinary income greater than $202,000 per model period) are not very relevant for a person who receives the average income over the life cycle as in our framework. In the TDA treatment, income taxes are levied on the sum of endowment income, interest income, and TDA withdrawals while TDA contributions are exempted from income taxes.

The per period utility function was set to $u(c) = 1 - e^{-(c-47400)/20000}$. This function served to convert subjects' consumption choices each period into dollar earnings; this was our way of inducing a concave utility function needed to ensure that we have a unique, optimal solution. We set 47.4k as the minimum consumption amount in every period. This minimum amount is equal to the after-tax income of retirees with zero savings (they receive $51k in endowment income in each period of retirement but pay taxes on this amount of $3.6k). The addition of a minimum consumption requirement was intended to encourage subjects to save for retirement, since if they had no retirement savings, their utility from consuming the after tax endowment amount would be zero in every period of the retirement phase. Appendix Figure C1.(d) illustrates the utility function or how subjects' consumption choices map into dollar payoffs. The pre-tax interest rate, is 10 percent per period, matching the annual return of 5 percent on long-term investments, as summarized in Cooley (1995). The discount factor β is set to 1, as it is hard to induce discounting over the short-time frame of our experimental study. On the other hand, as already noted, we do have survival or mortality risk in the post-retirement phase of life. Any assets held after death are declared to be worthless and have no utility/redemption value to the subject.

4.2 Treatments and Predictions

We consider six different treatments, aiming to understand both the overall effect of TDAs and the various underlying channels by which they may affect behavior. A summary of the different features of the six treatments was provided earlier in Table 1, but here we go into greater detail. In all treatments subjects make consumption/saving decisions over an indefinitely-long lifecycle and always have access to a regular savings account (RA). In treatment T1, subjects also have access to a TDA, the features of which are calibrated to match the current U.S. system. Since T1 represents the status quo, it serves as our baseline treatment. Treatment T2 eliminates the TDA but is otherwise identical to treatment T1; it taxes income in the same manner as T1, so the only difference between treatments T1 and T2 is the absence of the TDA in the latter. Because no income is exempt from taxation, in theory, the government should collect more taxes in the T2 environment as compared with the T1 environment, due to the absence of TDAs. To make the policy change of introducing TDAs *revenue neutral* (in expected terms) for the government, in Treatment T3, we also remove TDAs but we introduce a flat tax on interest income at the rate of 6.82%; that is, in T3, interest income is now taxed differently from endowment income, which remains subject to the same progressive income tax schedule as in all other treatments. By design, if agents make choices according to the rational choice model predictions, then treatment T3 should generate the *same* expected government revenue as treatment T1.[3]

The first three treatments enable three important pair-wise comparisons: the difference between treatments T2 and T1 reveals the overall effect of eliminating the TDA; the difference between treatments T3 and T1 reveals the effect of replacing TDAs with a revenue neutral tax reform that reduces taxes on interest income for all income brackets. Finally, the difference between treatments T3 and T2 reveals the effect of lowering the tax rate on interest income.

To further understand the channels that contribute to the overall effect of a TDA on consumption and savings behavior, specifically the difference between treatments T1 and T2, we employ three additional treatments. The first of these, treatment T4, considers the role that education may play. In the baseline T1 treatment, subjects receive information about the tax benefits of the TDA in both the instruction section and on the decision screen of each period. In particular, the TDA decision screen in each period includes a "tax calculator", showing how the tax liability and after-tax income would change based on TDA contribution or withdrawal choices. In the new treatment T4 we eliminated this tax calculator from that same decision screen. Otherwise, treatment T4 is identical to treatment T1, that is, subjects can save using a TDA. Further Treatments T5 and T6 are modified versions of the non-TDA treatment T2, with the inclusion of a second savings vehicle, a commitment account (CA). The CA offers the same return and tax treatment as the RA but like the TDA, contributions (withdrawals) are limited to the individual's working periods (retirement periods). Thus, treatments T5-T6 explore the commitment-device role played by a TDA. The distinction between T5 and T6 is that in T5, subjects make their CA saving choice on the first decision screen and their consumption choice on the second decision screen (with the residual going to the RA account) just as in TDA treatments T1 and T4, while in treatment T6, consumption is chosen first and then on a second decision screen, subjects choose how much to save in their CA account (the residual again goes to the RA).[4] While the timing of consumption/saving decisions should not matter, according to rational choice theory with standard preferences (no self-control problems), behaviorally it may make a difference.

By implementing treatments T4-T6, we can decompose the overall effect of adding a TDA, which is the only difference between treatments T1 and T2, into five distinct channels. First, we can gauge the effect of informing subjects about the precise tax incentives of TDA contributions or withdrawals on the first decision screen by comparing T1 with T4, the treatment without a tax calculator. This comparison addresses the "educational effect". Second, by creating a counterfactual treatment T4*, which duplicates the decisions of T4 but eliminates the effect of TDA tax savings on inflated TDA balances (i.e., replacing b_j with b_j^* – see Equation (6)), we can determine the "mechanical effect" of TDA tax incentives by comparing T4 to T4*. It should be noted that in T4*, the TDA balance can become negative during the retirement phase, so we focus on analyzing asset accumulation during the working phase. Third, by examining the difference between T4* and T5 (which has no TDA) we can identify behavioral responses to the tax advantages of TDAs, or what we call the "behavioral effect". The mechanical and behavioral effects are the *only* two channels through which TDAs affect retirement savings in a rational choice model. Fourth, the effect of requiring subjects to make a CA saving decision *before* making a consumption choice can be determined by comparing treatments T5 and T6, i.e., the "order effect". Lastly, the commitment device aspect of the CA, a savings vehicle that is not liquid until retirement but also carries no tax benefit relative to the RA can be identified by comparing treatment T6 to T2, i.e., the "commitment effect".

Subjects were randomly assigned to one and only one of the six treatments (that is, we employ a "between subjects" design). Each treatment was repeated for two "sequences" or lifetimes to allow for subject learning. The termination period was randomly selected for each of the two sequences so subjects can live for different lifetimes in each sequence, though as noted earlier we observe decisions for a minimum of 10 post-retirement periods from the random block termination design. The final experimental payment is composed of a show-up fee and the payoff from one randomly chosen sequence and one randomly chosen risk elicitation question.[5]

In Figure 1, we present the rational choice theory predictions for consumption and total net worth across the different treatments and by time periods. The figure includes a counterfactual net worth prediction $(a_{j+1} + b^*_{j+1})$ that deducts the mechanical effect of TDA incentives, which is referred to as the T1*/T4* model. The graph covers the first 25 periods only, as the experiment could be randomly terminated after the first retirement phase.

Consumption is highest in T2 at the very beginning of the life cycle, because the after-tax return from saving is lowest in this treatment. This reflects the dominance of the substitution effect over the income effect for young households with small net worth. Tax incentives enable households in T2 and T3 to accumulate net worth at a faster rate. In most subsequent periods, the income effect dominates the substitution effect, and consumption follows the order of $T1 > T3 > T2$. Net worth is highest in T1 and lowest in T2 in every period. In the final period before retirement (period 15), between T1 and T2, the overall difference in the ending period net worth, what we call the "retirement wealth", is 99.5k. The mechanical effect, the difference between T1 and T1*, accounts for an increase of 124.7k, and the behavioral effect, the difference between T1* and T2, accounts for a *decrease* of 25.2k. This means that in the experimental environment, the income effect dominates the substitution effect of TDA incentives, and TDA actually reduces the amount of total savings.

4.3 Experimental Procedures

Subjects were recruited to participate in our experiment from the UC Irvine Experimental Social Science subject pool using the Sona systems recruitment software. The subjects were undergraduate students at UC Irvine from many different majors with no prior experience with our experiment. No subject was allowed to participate in more than one treatment (experimental session). Subjects were randomly assigned to the different treatment conditions.

The experiment was computerized and programmed in oTree (Chen et al. (2016)). The experimental program involved an instruction phase, with tests for comprehension of the instructions, followed by the play of two lifecycle "sequences" each having unknown lengths. Subjects were instructed that one of the two sequences would be randomly selected for payment, but since they did not know which of the two sequences would be chosen in advance, they were incentivized to do their best in both sequences. Appendix A provides the experimental instructions and computer screenshots for the TDA treatment T1, and the consumption before CA treatment T6. Instructions for the other four treatments are minor variants of these two sets of instructions. Appendix B provides the various post-experiment tasks that subjects had to complete in all six treatments.

T1, T4, T5, and T6 differ from T2-T3 by having two decision screens. In T1 and T4, subjects make decisions on how much of their available income to set aside on a pre-tax basis for their TDA at the start of each working period (1-15) or how much to withdraw from their TDA account at the start of each period of the retirement phase. They then decide how much to consume on the second decision screen. Any residual amount goes into (or is taken out of) their RA. The only difference between T1 and T4 is that T1 provides a tax calculator on the first decision screen which aims to educate subjects about TDA tax incentives – see the screenshot for T1 in the online Appendix A.2 where this calculator is clearly identified; in T4, this calculator is removed. Note that for both T1 and T4, the realized tax liability and after-tax income are always reported on the summary screen at the end of each period. In T5-T6, the two treatments with a CA, subjects also have two decision screens: one to decide how much of their available income they want to contribute to the CA during the working phase or how much they want to withdraw from CA in the retirement phase, and a second screen to decide how much to consume; the residual amount again goes to (or is taken out of) their RA. The difference is that in T5, the first screen is about the CA saving choice, while in T6, the first screen is about the consumption decision and the CA saving choice comes on the next screen – see the screenshot for T6 in online Appendix A.4. In T2-T3, subjects only have one decision screen to choose how much of their available income to consume and the program then calculates the necessary deposits into or withdrawals from their RA. It is always possible for subjects to consume out of RA balances accumulated in previous periods; only the treatments with TDAs (T1 and T4) or CAs (T5 and T6) have lifecycle restrictions on withdrawals from TDAs or CAs.

Subjects in the experiment had the option to make their consumption and contribution/withdrawal decisions using sliders or by directly inputting numbers if they chose to. The slider/input box was used as a *calculator* to display the utility/monetary value of the consumption choice on the consumption decision screen, and for T1 only, it also served as a tax calculator on the TDA decision screen. Additionally, in all treatments, the slider/input box on the final decision screen always calculated the ending period RA balance. Thus, subjects did not have to refer to figures or tables (though we provided these to them as well) and they could consider many different possible choices before clicking a submit button that finalized their decision. The positioning of the sliders was dynamic in nature; in each new period, the slider was positioned at the choice made by the subject in the prior period; for the first period the slider position was at 0.

At the end of the session, we collected additional demographic data on each subject including their age, gender, education, and income. We further asked subjects to complete a risk preference elicitation (RE) task designed to reveal their degree of risk aversion and answer 4 cognitive reflection test (CRT) questions (Frederick (2005), Toplak et al. (2014)) and three financial literacy (FL) questions (Lusardi and Mitchell (2011)) – see Appendix B for details. Only the risk elicitation task was incentivized with a small additional monetary payment depending on the choices subjects made.

5 Findings

We have data for 185 subjects, roughly 30 subjects for each of the six treatments. As all of our treatments are individual-choice, we treat each subject as a single observation. Table 2 reports on some mean characteristics of subjects across the six treatments. Out of the 75 pairwise comparisons of observed characteristics, only the difference in RE between subjects in T2 and T4 was significant at the 5% level.

We conduct our data analysis using all observed behavior from both sequences. Thus, we have $185 \times 2 = 370$ subject-sequence pairs for the working phase and the first 10-period block of the retirement phase. The number of observations falls to 180 for the second 10-period block, because 51% of all subject-sequence pairs end in the first 10-period block. This rather large reduction in the number observations led us to focus on analyzing the working phase and the first retirement block only.

Figure 2 shows the mean consumption (left panel) and mean net worth (right panel) for each treatment and compares those statistics with theoretical predictions; there are just three theoretical benchmarks: the TDA model (T1/T4 model), the no TDA model (T2/T5/T6 model), and the revenue neutral no TDA model (T3 model). Similar patterns for median consumption and net worth across treatments are displayed in the Appendix Figure C3.(a).

Consistent with previous laboratory experiments on life-cycle consumption and saving decisions, (see Arifovic and Duffy (2018) for a review), we find that in treatments T2-T3, that do not have a TDA or CA, there is over-consumption relative to theoretical predictions in the working periods of life, which leads to under-consumption in the retirement periods of life due to reduced wealth accumulation. By contrast, in the two TDA treatments, T1 and T4, consumption levels are quite close to theoretical predictions during the early working periods of life and there is little difference in consumption between these two treatments. It is notable that subjects in T1 and T4 accumulate "retirement wealth", defined by their net worth at the end of their final working period 15, that is greater than predicted for the TDA treatments. Subjects in T4 decumulate their assets at a faster rate than those in T1 during the retirement phase most likely because they do not receive immediate feedback about the tax consequences of TDA withdrawals on their first decision screen (recall there is no tax calculator in T4). However, the quicker decumulation of mean net worth in T4 does not generalize to the median value (see Figure C3.(a)).

We note however that the replacement of the TDA with the CA in treatments T5-T6 also mitigates the tendency to over-consume in the early periods. In both treatments, subjects are able to accumulate more retirement wealth relative to those in the no TDA treatments T2 and T3. In T5, where subjects are required to make a CA decision before deciding on consumption, there is even more of a correction leading to a dramatic under-consumption during the first 15 working periods. It is worth noting that even without tax incentives, subjects in T5 are still able to accumulate retirement wealth at levels similar to those in the two TDA treatments, affording them similar levels of consumption during retirement. The same does not hold for T6, where subjects first choose consumption and then make a CA decision.

To summarize, the observations from Figure 2 yield four main messages. First, the tendency for over-consumption in the early stages of the life cycle, a common finding in prior laboratory experiments, is muted by the addition of a CA (T6 compared to T2-T3). Second, the order of decision-making is important, as subjects tend to under-consume during the working phase if they are required to make a saving decision first (T5 compared to T6). Third, subjects respond to the TDA tax incentives by increasing their consumption and reducing total savings (T4 compared to T5), in line with theoretical predictions. Finally, while providing a tax calculator has little impact on contribution amounts, it does slightly reduce withdrawals in the retirement phase (T1 compared to T4).

To formalize these differences across treatments, we estimate the following regression model

$$y_{is} = \text{Const.} + \sum_{k=2}^{6} \delta_k Tk_i + X_i\theta + \sigma S2_s + \epsilon_{is} \tag{7}$$

where y_{is} is the outcome for subject i in sequence s. The Tk_i is a 0-1 indicator variable that equals 1 if subject i is assigned to treatment Tk. The X_i represents four individual level controls: female, CRT, FL, RE. The $S2_s$ is an indicator variable that equals 1 for sequence 2. Data from the two sequences are pooled together in this analysis due to limited evidence of learning. Standard errors are clustered at the subject level. The constant term captures the level in T1 (baseline), and the coefficient δ_k measures the difference of treatment Tk from T1.

Table 3 presents estimates on how treatment conditions affect asset accumulation during the working phase by using two measures: the *contribution rate*, defined as the ratio of TDA or CA contributions to the endowment, as shown in Panel A; and the *saving rate*, defined as the ratio of total saving (see equation (5)) to the endowment, as show in Panel B. Note that the first measure is not applicable for T2 and T3 as they don't have TDAs or CAs and the second measure includes RA savings in addition to TDAs or CAs (where applicable). The constant terms in column (4) of Panel A indicate that, on average, subjects in T1 contribute 19% of their endowment to the TDA during the working phase. The near-zero coefficients associated with the T4 indicator variable suggest that the elimination of the tax calculator has a negligible effect on TDA contributions during the working phase. The T5 coefficients show that the TDA tax incentive alone has little impact on the contribution rate in first 10 periods (equivalent to 20 years). This observation is consistent with the empirical evidence that in the short run, TDA contributions respond little to tax incentives (Beshears et al. (2017)) or employer contributions (Chetty et al. (2014)). In periods 11-15, the contribution rate decreases from 19% in T1 to 12% in T5 due to the substantial tax liabilities generated by accumulated CA assets.[6] Notably, in T6, where subjects are required to make a consumption decision first, there is a significant decrease in contribution rates. On average, the contribution rate in T6 is 42% (= (16.2% − 9.4%)/16.2%) lower than that found in T5, where CA saving decisions are made first. Our findings contradict the rational choice theory, as the order of decision-making appears to play an important rule.

In addition to studying the impact on contributions, our laboratory environment makes it easier to analyze the effect of TDAs on total savings, an outcome variable hard to observe in field experiments. As discussed earlier in section 3, our measure of total savings excludes the tax savings associated with TDA contributions, and can be interpreted as after-tax income minus consumption plus pre-tax TDA contributions, i.e., TDA contributions minus their tax savings. For instance, if contributing $1000 to a TDA reduces taxes by $250, then only $750 is counted towards total savings. Panel B of Table 3 reveals that in the TDA treatments (T1 or T4), the total saving rates are slightly lower than the contribution rates, reflecting that subjects consume a portion of the tax savings generated by their TDA contributions. The removal of the tax calculator in T4 has little effect on saving decisions during the working periods, and on average, the saving rate is around 15% in both treatments. The elimination of TDA tax incentives in treatment T5 (where the CA saving choice is made first) causes the average saving rate to increase by 17% relative to T1, whereas the rational choice theory predicts a 12% increase. However, requiring subjects to make a consumption decision first as in T6 significantly dampens this increase in the saving rate; the saving rate in T6 is similar to that in T1. The removal of illiquid retirement accounts (TDAs or CAs) in treatments T2 and T3 causes the saving rate for these two treatments to fall into the single digits; on average, savings rates in these two treatments are 7-9 percentage points less than the savings rate in the baseline T1. Conforming to rational choice theory predictions, subjects assigned to the flat tax treatment T3 save slightly more than subjects assigned to T2, although the difference is not statistically significant. Among the control variables, FL is marginally positively linked to the average saving rate, and, on average, females save 5% less than males. Both of these results find support in the empirical data (Lusardi and Mitchelli, 2007; Hasler and Lusardi, 2017).

To analyze the impact of various channels on retirement wealth, Table 4 decomposes the overall effect of TDAs on retirement wealth, i.e., the difference in period 15 ending net worth between treatments T1 and T2 (T1-T2), into five channels. As previously discussed in section 4.2, only two channels, the mechanical effect and the behavioral effect of TDA tax incentives should matter according to rational choice models. Table 4 reveals that retirement wealth in T1 is 429k greater than in T2, surpassing the level predicted by the rational choice theory by more than fourfold. The decomposition exercise reveals that the educational effect (T1-T4) is small and insignificant. The mechanical effect (T4-T4*) is similar in levels to that predicted by the theory, accounting for a 149k increase in retirement wealth and explaining 35% of the gap. The behavioral effect (T4*-T5) generates a 110k reduction in retirement wealth. In percentage terms, the behavioral effect accounts for a 26% decrease in the gap, almost identical to the theoretical prediction of a 25% reduction. The order effect (T5-T6) of requiring subjects to make a saving decision first, generates a 209k increase in retirement wealth, explaining 49% of the gap. And finally, the commitment effect yields a 152k increase in retirement wealth, explaining 36% of the gap.

To summarize, when required to make a saving decision first, subjects tend to save a constant fraction of their endowment income to a TDA or a CA in the first few periods of a life cycle when the difference in tax liabilities is small. This order effect (saving in a TDA or CA before choosing consumption) is absent in the rational choice model and is the most important mechanism accounting for the increased retirement wealth seen in our lab sample. The commitment effect, which is also absent from the rational choice model, is the second most important mechanism. The absence of both mechanisms in the comparison between T2 and T3 is potentially the reason why there is no significant difference in retirement wealth between these two treatments.

We acknowledge that the TDA balances observed in our experiment are much greater than those typically seen in the real world. According to the 2019 Survey of Consumer Finance (SCF), the median balance of quasi-liquid retirement accounts (e.g., IRA and 401(k)s) for some one aged 58-59, corresponding to period 15 in the experiment, is 6k and the average balance is 125.8k. (Balances for married couples are split equally). By comparison, our experiment finds that the median and mean values for the same age group are significantly higher, with T1 (T4) showing values of 576.3k (499.0k) and 527.4k (512.4k), respectively This disparity in the level of retirement wealth between our baseline T1 and the real world contributes to the large increase in consumption following retirement in our experimental data. There are several possible explanations for this discrepancy. Firstly, our experiment does not account for intertemporal discounting or other lifecycle expenditures, such as for education and housing. Housing is the primary form of savings for many households. Secondly, we incentivize savings for retirement by imposing a utility function that strongly penalizes those with no retirement savings, whereas in reality, relying solely on Social Security benefits for retirement could generate a smaller reduction in utility. Thirdly, our experiment greatly simplifies the decision-making environment and does not take into account the many uncertainties individuals may actually face, such as earnings and health risks. These risks would discourage contributions to an illiquid account in favor of more liquid accounts. Lastly, it is possible that in the real world, some individuals may tend to prioritize consumption over savings. Data from T5 and T6 show that when we change the order of decisions from saving first to consumption first, the median (mean) CA balance by the end of period 15 falls from 456.9k to 193.0k (459.2k to 264.9k), a reduction of 58% based on the median values.

Despite the significant differences in balances across our different treatments and between our experimental and SCF data, it is still worth exploring whether individuals have the same tendency to contribute to or withdraw from an illiquid account in relative terms. To construct relative changes, we normalize both the SCF data and the experimental data, to 100 for age group 60-61. Figure 3 compares the relative changes of illiquid account balances over the life cycle, with the SCF data series showing the 45th-70th percentiles of quasi-liquid retirement account balances in the cross-sectional distribution. Surprisingly, in relative terms, our four treatments and the SCF data share similar patterns of accumulation during the working phase and decumulation in the retirement phase.

Table 5 presents the effects of our six treatments on average consumption and expected lifetime utility. Expected lifetime utility is calculating by adding up the discounted period utility, which is the product of period utility and the ex-ante probability of being alive at that period. For subjects whose sequence ended after the first 10-period retirement block, their lifetime utility was adjusted by adding the average discounted utility from the second 10-period retirement block, computed using data collected from subjects who completed both blocks. This measure, unlike the realized utility, is not affected by variations in the termination period across treatments, making it a more suitable indicator for welfare analysis.

The consumption estimates indicate that throughout the working phase (periods 1-15), subjects assigned to treatments T2 and T3 consumed about 20% more on average than those assigned to T1, which is in contradiction with the rational choice theory prediction that they should consume less. Due to the lack of retirement wealth, average consumption in the retirement phase is about 30-40% lower in treatments T2-T3 as compared with T1, which is significantly greater than the levels predicted by the rational choice theory. No significant difference between T1 and T4 is observed, suggesting the minimal role of education. The provision of the CA in T6 helps subjects to cut back on consumption during the working phase and eliminates more than half of the difference in retirement consumption between T2 and T1. In T5, the requirement of making a saving decision first reduces consumption even further, causing the average consumption during the working phase to be less than that in T1, and also less than theoretical predictions. Subjects in T5 accumulate substantial wealth for retirement, allowing them to maintain a consumption level close to that in the two TDA treatments during the retirement phase. While the treatment effects on consumption differ between the working and retirement phases, they mostly balance each other out. As a result, there is no significant difference in lifetime utility between treatments, with the only exception that subjects assigned to T5 experience a considerable reduction in lifetime utility compared to those in the other five treatments. The difference in lifetime utility between T2 and T5 is marginally significant at the 10% level, and the difference between T2 and the other four treatments are significant at the 5% level. While the saving-first design helps subjects accumulate more retirement wealth in T5, this design leads subjects to under-consume during the working phase, ultimately reducing their lifetime utility. However, we caution this finding may not generalize to the real world, as households face other temptations, e.g., present biased preferences to over-consume during the working phase, that our model does not capture. Among all controls, greater experience (S2) slightly improves lifetime utility.

Although the treatment effects on expected lifetime utility are informative, these mean effects mask distributional differences between subjects who ended a sequence early versus those who ended a sequence later. Figure 4 presents the treatment effects on cumulative utility across different periods. It shows that subjects assigned to T2-T3 earn about $2 more if the sequence ends early in periods 16-18, and this advantage diminishes as the termination of the sequence is delayed. By period 21, the difference is no longer significant. Comparing T6 to T1 reveals a similar pattern, but the magnitude is smaller, reflecting the lower consumption levels during the working phase in T6 as compared with T2 and T3. There is no significant difference between T1 and T4, whereas T5 is associated with significantly lower cumulative utility regardless of the ending period.[7]

6 Generalizability

6.1 Results using the MTurk sample

An often heard criticism with using the convenience sample of student subjects is that such subjects lack the appropriate experience of making consumption and saving decisions and planning for retirement so that findings from such a sample may not generalize outside of the laboratory. To address this issue, we repeated our experimental treatments using a separate sample recruited from MTurk. Here we follow other researchers who have used data from MTurk workers as a means of assessing the generalizability of their experimental findings using undergraduate subjects, see, e.g., Arechar et al. (2018), Snowberg and Yariv (2021).

We randomly assigned 180 MTurk subjects to each of our six treatments. MTurk subjects used the same interface and faced the same monetary incentives as the student subjects. The only difference is that the show-up payment (guaranteed minimum) for the MTurk sample was set at $1, reflecting online norms; the variable payment amounts were the same for both the MTurk and student samples. We gathered data on the choices made by 32 subjects for T1, 28 subjects for T3, and 30 subjects for the remaining four treatments.

Table 6 presents the summary statistics of the two samples. The MTurk sample exhibits a female share that is 7 percentage points lower compared to the lab sample. Additionally, the MTurk subjects are, on average, approximately 15 years older than the lab subjects. Regarding racial composition, the lab sample is predominantly composed of Asian students, which aligns with the demographic makeup of UC Irvine undergraduates. On the other hand, the MTurk sample more closely resembles the racial distribution across the United States. As anticipated, the lab sample temporally has lower income and lower education achievements than the MTurk sample, though this likely reflects the age differences between the two samples. We also find that the lab sample scores higher, on average, on both the CRT and the FL test. The lab sample is more risk averse and has higher earnings from the decision stage of the experiment.

Table 7 provides formal estimates of the treatment effects on contribution and total saving rates using the MTurk sample, which are comparable to the results presented in Table 3. The point estimates of the treatment effects from the MTurk sample closely align with those of the lab sample with one exception: the contribution and saving rates in T6 are lower in the MTurk sample as compared to the lab sample. Unlike the lab sample, the addition of a CA in T6 has only a minimal impact on the accumulation of retirement savings in the MTurk sample. This difference reflects that MTurk subjects, consisting of older individuals with more experience in making consumption and saving decisions, are less likely to utilize the CA and be subject to the associated liquidity constraint when the CA saving decision is made following the consumption decision. Out of the 30 MTurk subjects assigned to T6, 7 never contributed to the CA in at least one experimental sequence, whereas all 29 lab subjects assigned to T6 made positive contributions in both sequences. By comparison, in T5 where subjects are first asked to make a CA saving decision, only one subject in both the lab and MTurk samples chose to never contribute to the CA.

Table 8 decomposes the overall effect of TDAs on retirement wealth into five channels, which can be compared to Table 4. We consistently find that the order effect remains the primary contributor to the increase in retirement wealth, explaining 71% of the gap between T1 and T2, and the educational effect is very limited. However, the commitment effect and behavioral effect no longer play a significant role in explaining the gap in retirement wealth within the MTurk sample.

Several less important findings regarding the MTurk sample are relegated to the Appendix. As reported in Appendix Tables C2 and C3, the MTurk sample exhibits more noisy behavior as compared to the lab sample, with average mean square errors (MSEs) nearly double that of the lab sample. Figures C2 and C3.(b) display average and median behavior of the MTurk sample, which closely resembles that of the lab sample. The only notable difference is that unlike lab subjects, MTurk subjects in T6 have consumption and saving behavior similar to those in T2 during the entire working phase. Finally, Appendix Table C4 considers treatment effects on consumption and expected lifetime utility for the Mturk sample. There are two notable differences from the lab sample. First, consumption levels in T6 of the MTurk sample closely resemble those in T2. Second, expected lifetime utility in T4 is significantly lower than what was found in T1, as the absence of a tax calculator leads to a more pronounced degree of over-consumption during the early periods of the retirement phase in the MTurk sample.

In summary, our treatments have similar effects in both the MTurk and lab samples with the exception that the commitment-device aspect of TDAs no longer plays a significant role in the MTurk sample. Consistently across both samples, we observe an order effect, whereby subjects tend to save more if they are required to make a saving decision first.

6.2 Generalizability beyond the MTurk sample

The lifecycle theoretical predictions that are being tested in our experiment emphasize that it is *incentives* that matter for intertemporal consumption/saving decisions. This theory does not speak directly to the *composition* or to the *experience* of the subject population. Still, we have shown that most of our results using student subjects are replicated in the MTurk sample. We think this provides some evidence for the generalizability of our results beyond the lab sample.

Indeed, there is evidence from other laboratory studies that the behavior of student subjects is often no different from that of the general public, operating in more naturalistic settings, see, for example the work of Alm et al. (2015) on the incidence of tax evasion in the lab and in the field, which are found to be quite similar. Further, there is evidence that in many settings, so-called "professional" subjects (to the extent that they can be identified as such) do not perform any better than student subjects see, e.g., Fréchette (2015). Finally, we note that even in more naturalistic field studies assessing interventions for retirement savings, e.g. Beshears et al. (2017), the researchers are only studying how treatment interventions affect a *single* population of individuals and so the generalizability of those results to other populations is also not so clear. Moreover, as discussed earlier in section 5, the findings presented in Figure 3 regarding the balances of quasi-liquid retirement accounts provide some support for the generalization of our experimental findings. Although the mean retirement account balance in our experimental data is significantly higher, there are notable similarities in the relative patterns of accumulation during the working phase and decumulation in the retirement phase between individuals in the field and those in our experiment.

Still, as with any study, much caution is warranted in extrapolating from our experimental findings to lifecycle decision making in the "real world". In particular, we study intertemporal decision-making in a very compressed period of time. We induce preferences upon subjects so that we can test rational choice model predictions. However, we abstract from preference specifications that might give rise to present-bias or other behavioral anomalies. Additionally, the mortality risk in our setup remains a constant 10% throughout the retirement phase, whereas in reality, it should be age-adjusted. This discrepancy contributes to the observed increase in consumption following retirement in our experimental data.

Our model also abstracts from a number of important components in lifecycle planning, including investments in education, housing, children and other non-financial assets, which may affect agents' abilities to save for retirement. Since we abstract from these important lifecycle events and real households may prioritize consumption over savings, we may regard our analysis of the effects of TDAs on retirement savings as providing an upper bound on the benefits that TDAs can provide. We emphasize, however that the magnitude of the treatment effects is not the primary focus of our experiment. Rather, our main contribution lies in providing insights into the *channels* through which TDAs might affect retirement savings. In this regard, in both the lab sample and the MTurk sample, we consistently find that the order effect, specifically the requirement that TDA or CA saving decisions are made first, emerges as the primary driver behind the observed increase in retirement wealth. We note that requiring the saving decisions to be made first is akin to the widely recognized "pay yourself first" strategy, often recommended by financial advisers (For example, see Chilton and Ganser (1998); Classon (2020)).

Despite these caveats, we view our methodology as constructive for future research on retirement policies. We are cautiously optimistic that our results are generalizable as we can replicate our main findings using a more diverse cross-section of participants, beyond the typical convenience sample of undergraduate students.

7 Conclusions

Understanding the value added of mechanisms that might improve retirement savings is complicated by the fact that many variables that matter for such an evaluation such as income, asset returns and retirement horizons are not perfectly known or observable. In our incentivized laboratory experiment we are able to exert control over such variables enabling clearer inferences to be made. We find that subjects in our baseline TDA treatment T1 accumulate significantly larger retirement wealth than those in the non-TDA treatments T2-T3, despite facing the *exact same lifecycle income profile* and without any differences in rates of pre-tax returns. We interpret this as causal evidence that TDAs improve retirement wealth. Our exercise decomposing the channels by which TDAs improve retirement wealth indicates that the primary behavioral factor influencing subjects' choices is the order in which they make decisions, specifically whether they make retirement saving decisions first or second. This amounts to a classic framing effect where individuals make different decisions based on the way in which problems are presented to them, even though theory says there should be no difference (Tversky and Kahneman, 1981). In practice, many households may make retirement saving decisions before consumption choices in each pay period, for example, those who have automatic payroll deductions for employer provided 401(k) plans are effectively doing this. In this case, our experiment shows that TDAs lead to large increases in retirement wealth relative to their absence. However, the timing of contributions to other tax deferred savings vehicles such as IRAs is less clear and our research suggests that encouraging households to make contributions to such accounts prior to making consumption choices could be beneficial. Our decision interface also requires participants to revisit and possibly revise their TDA/CA choices in every period, which is calibrated to be two years. Given the benefits of this interface when saving decisions are made first, this technology might also be considered for adoption by administrators of TDA plans; in future research it would be of interest to compare settings where participants make a one-time TDA decision with our setting where they revisit the decision each period.

Other possible extensions using our laboratory approach would include providing default options, financial advice or social information about the choices made by others; since all participants face the same income profile, this would amount to a setting where other subjects were *true peers* whose choices might carry great weight. Future laboratory experiments can also implement age-adjusted mortality risks and employ alternative utility functions which impose lesser penalties on individuals with zero retirement savings. These revisions would yield a more realistic consumption path that aligns better with empirical regularities, although at the cost of reduced incentives for consumption smoothing. While we find all of these extensions very interesting, we leave them to future research.

Acknowledgment: We gratefully acknowledge funding for this project from the National Science Foundation under grant numbers SES-1918571 and SES-1918687. We thank Patrick Julius and Nishtha Sharma for help with programming the experiment. For helpful comments, we thank our editor Kareen Rozen, three anonymous referees, and seminar participants at 2022 AEA annual meeting and the 2022 ESA international meeting. The experimental protocol governing data collection in our study was approved by UC Irvine's Institutional Review Board, HS#2019-5209.

Table 1: Features of the six treatments

Feature	T1	T2	T3	T4	T5	T6
	Treatment					
Tax Deferred Account (TDA)	✓			✓		
Flat tax on interest income			✓			
Commitment Account (CA)					✓	✓
TDA tax calculator	✓					
Saving choice first	✓			✓	✓	
Consumption choice first		✓	✓			✓

Table 2: Summary statistics

	All (1)	T1 (2)	T2 (3)	T3 (4)	T4 (5)	T5 (6)	T6 (7)
Female	0.55	0.52	0.47	0.68	0.45	0.58	0.59
Cognitive Reflection Test (CRT)	2.55	2.61	2.43	2.94	2.52	2.68	2.07
Financial Literacy (FL)	2.42	2.29	2.40	2.45	2.45	2.58	2.34
Risk aversion (RE)	6.15	6.32	6.93	6.16	5.30	6.19	6.03
Payoff	15.08	15.89	15.02	15.22	14.46	14.64	15.27
Obs	185	31	30	31	33	31	29

Notes: Line 2 reports the number of correct answers to the four cognitive reflection test questions. Line 3 reports the number of correct answers to the three financial literacy questions. Line 4 reports the number of safe choices among 10 paired lottery risk elicitation (RE) choices and line 5 reports experimental payments earned which includes the RE task payment but excludes the show-up payment.

Table 3: Treatment effects on contribution rates and saving rates

	A: Contribution rate				B: Saving rate			
	1-5	6-10	11-15	1-15	1-5	6-10	11-15	1-15
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cons	0.20***	0.19***	0.19***	0.19***	0.19***	0.17***	0.14***	0.17***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.05)	(0.03)
T2					-0.06***	-0.11***	-0.11***	-0.09***
					(0.02)	(0.03)	(0.03)	(0.02)
T3					-0.05***	-0.08***	-0.10***	-0.07***
					(0.02)	(0.02)	(0.03)	(0.02)
T4	-0.01	0.00	-0.01	-0.01	-0.02	-0.02	-0.02	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
T5	-0.03	0.00	-0.07***	-0.03*	0.02	0.16***	0.34***	0.17***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.03)	(0.04)	(0.02)
T6	-0.09***	-0.08***	-0.12***	-0.10***	-0.05***	0.01	0.11**	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)
CRT	0.00	-0.00	-0.01	-0.00	-0.00	-0.00	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
FL	-0.00	0.01	0.01	0.01	0.02**	0.02	0.02	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
RE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female	-0.02*	-0.03	-0.02	-0.02	-0.04***	-0.05***	-0.04*	-0.05***
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
S2	-0.00	-0.00	-0.01	-0.01	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
R^2	0.18	0.11	0.19	0.19	0.16	0.34	0.47	0.42
N	248	248	248	248	370	370	370	370

Notes: This table presents estimates from Equation (7). The dependent variables are the average contribution rate (Panel A) and the average saving rate (Panel B) over 5 or 15 period intervals. Standard errors are clustered at the subject level. *** p<0.01, ** p<0.05, and * p<0.10.

Table 4: Decomposition of the difference in retirement wealth

	Interpretation	Data Level (k) (1)	SE (2)	% (3)	Theory Level (k) (4)	% (5)
T1-T2	overall effect	429.3***	42.7	100.0	99.5	100.0
T1-T4	educational effect	29.4	50.0	6.9	0.0	0.0
T4-T4*	mechanical effect	148.6***	9.2	34.6	124.7	125.4
T4*-T5	behavioral effect	-110.2***	35.7	-25.7	-25.2	-25.4
T5-T6	order effect	209.2***	46.2	48.7	0.0	0.0
T6-T2	commitment effect	152.3***	42.5	35.5	0.0	0.0

Notes: This table reports differences in retirement wealth between treatments. Retirement wealth is defined as period 15 ending net worth, $a_{16} + b_{16}$ for all treatments but T4* where we use $a_{16} + b^*_{16}$. Column 1 (Level in thousands, k) reports the differences in mean retirement wealth for selected treatment pairs and its significance using a two-sided t-test. *** p<0.01, ** p<0.05, and * p<0.10. Standard errors are clustered at the subject level and reported in column 2. Columns 3 and 5 report how the overall difference between T1 and T2, normalized to 100%, is decomposed across the various channels we consider in both the data and in the theory.

Table 5: Treatment effects on consumption and expected lifetime utility

	Consumption (k)					Lifetime
	1-5	6-10	11-15	16-20	21-25	Utility
	(1)	(2)	(3)	(4)	(5)	(6)
Cons	56.23***	70.52***	70.06***	105.81***	94.57***	12.58***
	(1.99)	(2.98)	(3.08)	(9.52)	(8.03)	(0.61)
T2	5.75***	14.90***	16.11***	-42.85***	-38.34***	-0.20
	(1.40)	(2.06)	(2.27)	(5.94)	(4.54)	(0.44)
T3	5.46***	13.78***	18.37***	-39.74***	-31.24***	0.08
	(1.35)	(1.96)	(2.19)	(6.38)	(4.84)	(0.42)
T4	1.54	1.01	0.51	3.17	-2.56	0.65
	(1.19)	(1.87)	(2.25)	(6.81)	(6.23)	(0.46)
T5	0.25	-5.82***	-7.06***	-0.41	4.36	-1.06**
	(1.09)	(2.15)	(2.39)	(7.95)	(6.90)	(0.54)
T6	5.33***	4.71**	3.37	-24.17***	-19.63***	0.00
	(1.46)	(2.34)	(2.39)	(7.41)	(6.12)	(0.51)
CRT	0.31	0.09	0.62	-0.59	0.06	-0.00
	(0.38)	(0.62)	(0.57)	(1.74)	(1.53)	(0.12)
FL	-1.10**	-0.84	-0.54	2.42	1.81	0.21
	(0.54)	(0.95)	(0.88)	(2.56)	(2.70)	(0.17)
RE	-0.09	-0.24	0.06	0.32	0.50	-0.01
	(0.18)	(0.28)	(0.29)	(0.75)	(0.75)	(0.06)
Female	3.12***	2.87**	-0.20	-7.43*	-16.10***	-0.05
	(0.86)	(1.31)	(1.52)	(3.96)	(3.58)	(0.27)
S2	-0.56	-0.06	0.09	4.20**	0.24	0.39**
	(0.44)	(0.72)	(0.85)	(1.93)	(2.27)	(0.15)
R^2	0.19	0.38	0.41	0.32	0.30	0.05
N	370	370	370	370	370	370

Notes: This table presents estimates from Equation (7). The dependent variables are the average consumption over each 5-period interval and expected lifetime utility. Standard errors are clustered at the subject level. *** p<0.01, ** p<0.05, and * p<0.10.

Table 6: Comparison between lab sample and MTurk sample

	Lab sample		MTurk sample	
	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)
Female	0.546	0.499	0.472	0.501
Age	21.081	2.407	36.389	11.840
White	0.162	0.370	0.867	0.341
Asian	0.568	0.497	0.044	0.207
Income: 0-25k	0.832	0.374	0.167	0.374
Income: 25-50k	0.070	0.256	0.222	0.417
College or above	0.259	0.440	0.828	0.379
Cognitive Reflection Test	2.546	1.264	2.244	1.360
Financial Literacy	2.422	0.798	2.122	1.122
Risk aversion	6.146	2.440	5.694	2.891
Payoff (ex. fixed amount)	15.077	3.710	12.880	3.991

Note: For definitions of the last four items, see the notes to Table 2.

Table 7: Treatment effects on contribution rates and saving rates in the MTurk sample

	Panel A: Contribution/Endowment				B: Total saving/endowment			
	1-5	6-10	11-15	1-15	1-5	6-10	11-15	1-15
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cons	0.17***	0.22***	0.29***	0.23***	0.21***	0.23***	0.20***	0.21***
	(0.03)	(0.04)	(0.03)	(0.03)	(0.02)	(0.03)	(0.05)	(0.03)
T2					-0.11***	-0.15***	-0.10**	-0.12***
					(0.02)	(0.03)	(0.04)	(0.03)
T3					-0.10***	-0.14***	-0.14***	-0.12***
					(0.02)	(0.03)	(0.04)	(0.02)
T4	-0.03	-0.02	-0.01	-0.02	-0.00	0.01	0.07*	0.03
	(0.02)	(0.03)	(0.03)	(0.03)	(0.01)	(0.03)	(0.04)	(0.02)
T5	-0.05***	-0.02	-0.09***	-0.05***	-0.00	0.12***	0.34***	0.15***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.01)	(0.03)	(0.04)	(0.03)
T6	-0.15***	-0.17***	-0.23***	-0.18***	-0.15***	-0.14***	-0.06	-0.11***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)
CRT	0.00	-0.00	-0.00	0.00	0.00	0.00	-0.00	0.00
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
FL	0.01	-0.00	-0.01	-0.00	-0.00	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
RE	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female	0.02	0.00	-0.00	0.01	-0.02	-0.01	-0.02	-0.01
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.03)	(0.02)
S2	-0.00	-0.00	-0.02**	-0.01	0.01	0.01	0.03*	0.01*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
R^2	0.37	0.29	0.40	0.39	0.35	0.32	0.37	0.39
N	244	244	244	244	360	360	360	360

Notes: This table presents estimates from Equation (7). The dependent variables are the average contribution rate (Panel A) and the average saving rate (Panel B) over 5 or 15 period intervals. Standard errors are clustered at the subject level. *** p<0.01, ** p<0.05, and * p<0.10.

Table 8: Decomposition of the difference in retirement wealth in the MTurk sample

	Interpretation	Level (k) (1)	SE (2)	% (3)
T1-T2	overall effect	502.1***	49.5	100.0
T1-T4	educational effect	-2.9	53.0	-0.6
T4-T4*	mechanical effect	153.1***	14.4	30.5
T4* -T5	behavioral effect	-16.3	40.8	-3.2
T5-T6	order effect	357.4***	42.8	71.2
T6-T2	commitment effect	10.7	44.6	2.1

Notes: This table reports differences in retirement wealth between treatments in the MTurk sample. Retirement wealth is defined as period 15 ending net worth, $a_{16} + b_{16}$ for all treatments but T4* where we use $a_{16} + b^*_{16}$. Column 1 (Level in thousands, k) reports the differences in mean retirement wealth for selected treatment pairs and its significance using a two-sided t-test. *** p<0.01, ** p<0.05, and * p<0.10. Standard errors are clustered at the subject level and reported in column 2. Columns 3 reports how the overall difference between T1 and T2, normalized to 100%, is decomposed across the various channels we consider.

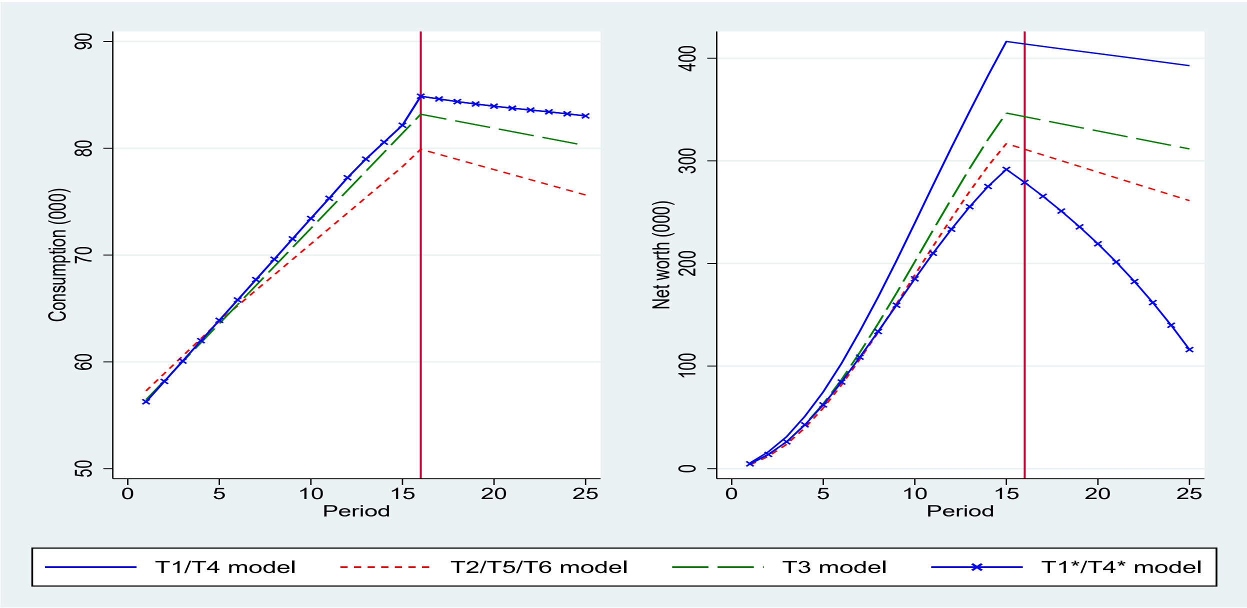


Figure 1: Theory predictions

Notes: T1*/T4* share the same consumption/saving decisions as T1/T4. The right panel displays $a_{j+1} + b_{j+1}$ ($a_{j+1} + b^*_{j+1}$ for the T1*/T4* model).

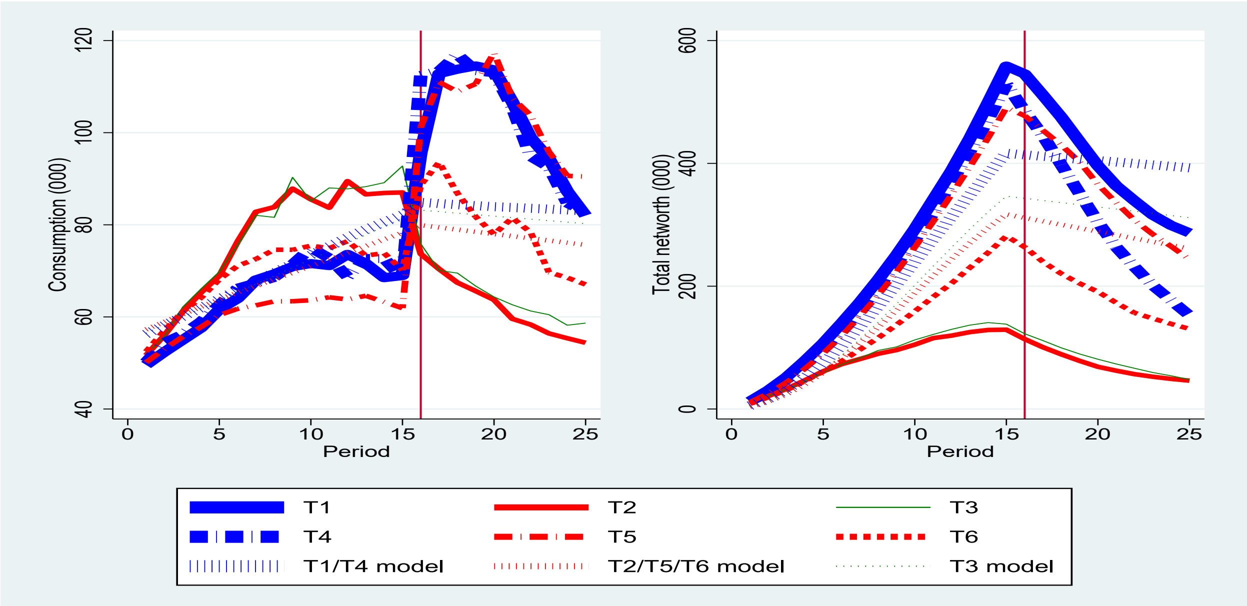


Figure 2: Theory predictions and experimental means

Note: The right panel displays $a_{j+1} + b_{j+1}$.

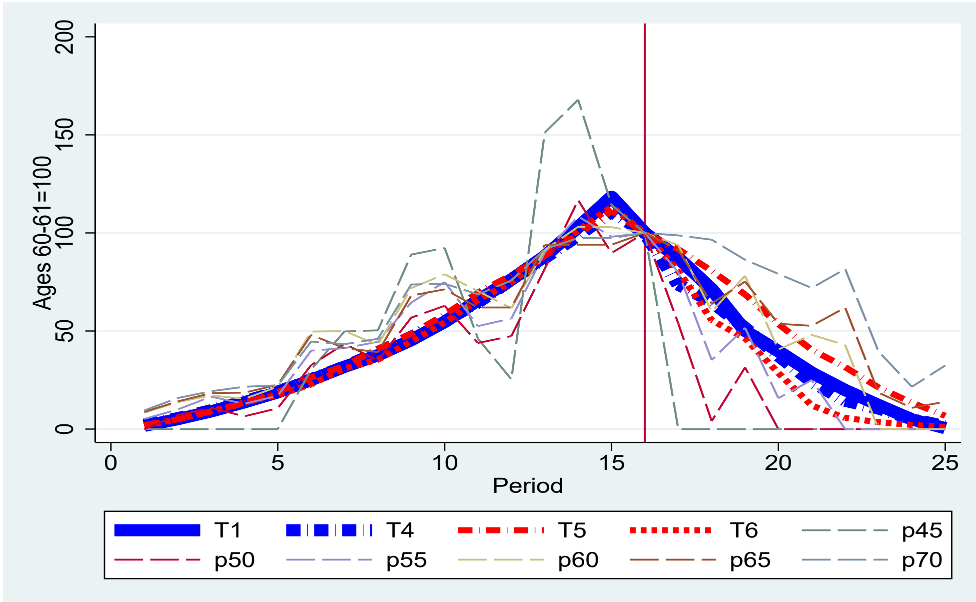
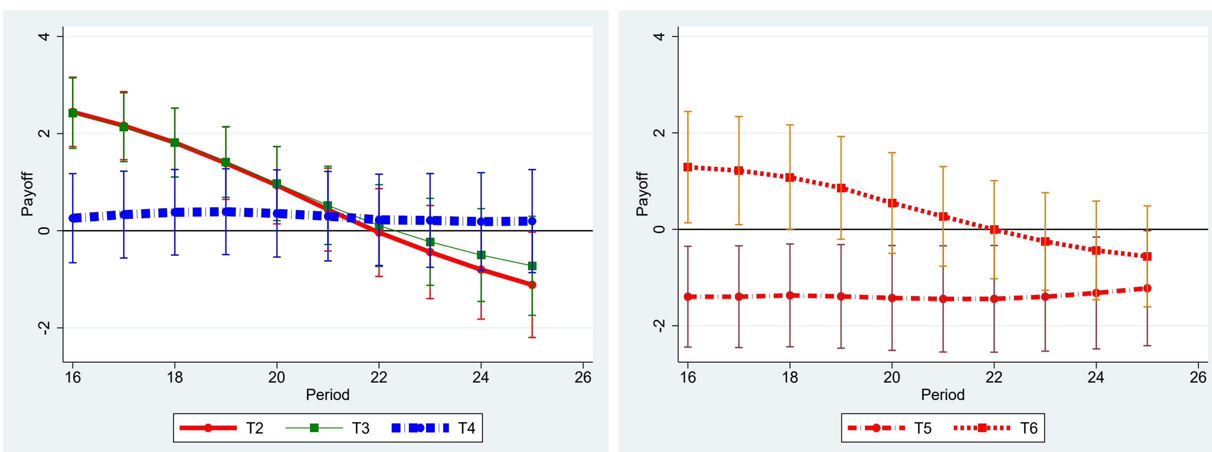


Figure 3: Median TDA or CA balances in the experiment vs. balances in quasi-liquid retirement accounts in the SCF data.

Notes: The dash lines represent different percentiles (e.g., p45 denotes the 45th percentile) of quasi-liquid retirement account balances, derived from the pooled 2013, 2016 and 2019 SCF surveys. For all series, the value for ages 60-61, which corresponds to the first period of retirement phase in the experiment, is normalized to 100.



(a) T2-T4

(b) T5-T6

Figure 4: Treatment effects on cumulative utility across different periods

Notes: This figure presents the estimated coefficients of T2-T6 from equation (7). The dependent variable is the cumulative utility up to the specific period displayed on the x-axis, with each period representing a distinct regression.

Notes

[1]Even in our simplified environment, subjects still face challenges to find the optimal solution. They may not fully understand the experimental rules, or have the perfect ability to optimize. These two factors are not only present within our experimental environment but also extend beyond the lab into real-world scenarios. However, the method of randomly assigning subjects to treatments should to a large extent take care of differences in comprehension and decision-solving ability across treatments.

[2]For simplicity, we focus on traditional TDA, abstract from the early withdrawals and required minimum withdrawal. See Horneff et al. (2023, 2021) for analysis on the implication of Rothification and required minimum withdrawal.

[3]Alternatively, we could keep taxing interest income in the same way as in treatments T1 and T2, and proportionally reduce the progressive income tax rates. But that approach introduces non-integer marginal tax rates, which is harder to comprehend compared to the simpler T3 design.

[4]We maintain the order of consumption/saving decisions for the two TDA treatments for two reasons. First, choosing consumption first leads to varying maximum TDA contribution and minimum TDA withdrawal, complicating the comprehension of rules alongside TDA tax incentives. Second, we wanted subjects to make consumption choices with knowledge of their after-tax income across all treatments. Allowing consumption choices first in the TDA treatment would mean that subjects would choose consumption without knowing their after-tax income. This problem does not arise with the non-tax-advantaged CA.

[5]The show-up fee increased from \$7 to \$10 (for the last 3 treatments) due to inflation, while variable payoff incentives were the same across all treatments.

[6]Appendix Table C1 shows that the average difference in tax to endowment ratios between T5 and T1 is 0.03, 0.06, and 0.12, for periods 1-5, 6-10, 11-15, respectively.

[7]In the smaller sub-sample of subjects who completed the second retirement block, subjects in T2 have significantly lower cumulative utility (-\$3.1 to -\$1.5) than those in T1 if the sequence ends after period 28, and subjects in T4 have significantly higher cumulative utility if the sequence ends at periods 29-30 (\$1.5 to \$1.6). Detailed estimates are available upon request.

References

Alm, James, Kim M Bloomquist, and Michael McKee, "On the external validity of laboratory tax compliance experiments," *Economic Inquiry*, 2015, *53* (2), 1170–1186.

Arechar, Antonio A, Simon Gächter, and Lucas Molleman, "Conducting interactive experiments online," *Experimental Economics*, 2018, *21* (1), 99–131.

Arifovic, Jasmina and John Duffy, "Heterogeneous agent modeling: experimental evidence," in C. Hommes and B LeBaron, eds., *Handbook of Computational Economics*, Vol. 4, Elsevier, 2018, pp. 491–540.

Beshears, John, James Choi, David Laibson, and Brigitte Madrian, "The limitations of defaults," NBER Working Paper RRC NB10-02 2010.

_ , James J Choi, Christopher Harris, David Laibson, Brigitte C Madrian, and Jung Sakong, "Which early withdrawal penalty attracts the most deposits to a commitment savings account?," *Journal of Public Economics*, 2020, *183*, 104144.

_ , _ , David Laibson, and Brigitte C Madrian, "Does front-loading taxation increase savings? Evidence from Roth 401 (k) introductions," *Journal of Public Economics*, 2017, *151*, 84–95.

Blaufus, Kay and Michael Milde, "Tax misperceptions and the effect of informational tax nudges on retirement savings," *Management Science*, 2021, *67* (8), 5011–5031.

_ , _ , and Alexandra Spaeth, "Does a decrease in pension taxes increase retirement savings? An experimental analysis," Working paper 2023.

Bohr, Clement E, Charles A Holt, and Alexandra V Schubert, "A behavioral study of retirement savings accounts: Roth versus traditional," Working Paper 2023.

Chen, Daniel L, Martin Schonger, and Chris Wickens, "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9*, 88–97.

Chetty, Raj, John N Friedman, Søren Leth-Petersen, Torben Heien Nielsen, and Tore Olsen, "Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from Denmark," *Quarterly Journal of Economics*, 2014, *129* (3), 1141–1219.

Chilton, David and LJ Ganser, *The Wealthy Barber: Everyone's Commonsense Guide to Becoming Financially Independent*, Prima Pub., 1998.

Choi, James J, David Laibson, and Brigitte C Madrian, "$100 bills on the sidewalk: Suboptimal investment in 401 (k) plans," *Review of Economics and Statistics*, 2011, *93* (3), 748–763.

_ , _ , _ , and Andrew Metrick, "For better or for worse: Default effects and 401 (k) savings behavior," in D.A. Wise, ed., *Perspectives on the Economics of Aging*, University of Chicago Press, 2004, pp. 81–126.

_ , Emily Haisley, Jennifer Kurkoski, and Cade Massey, "Small cues change savings choices," *Journal of Economic Behavior and Organization*, 2017, *142*, 378–395.

Clark, Robert L, Jennifer A Maki, and Melinda Sandler Morrill, "Can simple informational nudges increase employee participation in a 401 (k) plan?," *Southern Economic Journal*, 2014, *80* (3), 677–701.

Classon, George S, *The richest man in Babylon*, Sristhi Publishers & Distributors, 2020.

Cooley, Thomas F., *Frontiers of Business Cycle Research*, Princeton University Press, 1995.

Cuccia, Andrew D, Marcus M Doxey, and Shane R Stinson, "The impact of tax incentive structure on taxpayers' retirement savings decisions," *The Journal of the American Taxation Association*, 2022, *44* (1), 23–47.

Duffy, John and Yue Li, "Lifecycle consumption under different income profiles: Evidence and theory," *Journal of Economic Dynamics and Control*, 2019, *104*, 74–94.

Duflo, Esther and Emmanuel Saez, "The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment," *Quarterly Journal of Economics*, 2003, *118* (3), 815–842.

_ , William Gale, Jeffrey Liebman, Peter Orszag, and Emmanuel Saez, "Saving incentives for low-and middle-income families: Evidence from a field experiment with H&R Block," *Quarterly Journal of Economics*, 2006, *121* (4), 1311–1346.

Fréchette, Guillaume R, "Laboratory experiments: Professionals versus students," in G.R. Fréchette and A. Schotter, eds., *Handbook of Experimental Methodology*, Oxford University Press, 2015, pp. 360–390.

Fréchette, Guillaume R and Sevgi Yuksel, "Infinitely repeated games in the laboratory: Four perspectives on discounting and random termination," *Experimental Economics*, 2017, *20* (2), 279–308.

Frederick, Shane, "Cognitive reflection and decision making," *Journal of Economic Perspectives*, 2005, *19* (4), 25–42.

Goda, Gopi Shah, Colleen Flaherty Manchester, and Aaron J Sojourner, "What will my account really be worth? Experimental evidence on how retirement income projections affect saving," *Journal of Public Economics*, 2014, *119*, 80–92.

_ , Matthew R Levy, Colleen Flaherty Manchester, Aaron Sojourner, and Joshua Tasoff, "Who is a passive saver under opt-in and auto-enrollment?," *Journal of Economic Behavior and Organization*, 2020, *173*, 301–321.

Hasler, Andrea and Annamaria Lusardi, "The gender gap in financial literacy: A global perspective," *Global Financial Literacy Excellence Center, The George Washington University School of Business*, 2017.

Ho, Anson TY, "Tax-deferred saving accounts: Heterogeneity and policy reforms," *European Economic Review*, 2017, *97*, 26–41.

Horneff, Vanya, Raimond Maurer, and Olivia S Mitchell, "Do required minimum distribution 401 (k) rules matter, and for whom? Insights from a lifecycle model," NBER Working Paper No. 28490 2021.

_ , _ , and _ , "How would 401 (k) 'Rothification' alter saving, retirement security, and inequality?," *Journal of Pension Economics & Finance*, 2023, *22*, 265–283.

Hu, Wei-Yin, Olivia S Mitchell, Cynthia A Pagliaro, and Stephen P Utkus, "Evaluating Web-based Savings Interventions: A Preliminary Assessment," Michigan Retirement Research Center Research Paper 2013-299 2013.

Lusardi, Annamaria and Olivia S Mitchell, "Financial literacy around the world: an overview," *Journal of Pension Economics & Finance*, 2011, *10* (4), 497–508.

— **and Olivia S Mitchelli**, "Financial literacy and retirement preparedness: Evidence and implications for financial education," *Business Economics*, 2007, *42*, 35–44.

Madrian, Brigitte C and Dennis F Shea, "The power of suggestion: Inertia in 401 (k) participation and savings behavior," *Quarterly Journal of Economics*, 2001, *116* (4), 1149–1187.

Nishiyama, Shinichi, "The budgetary and welfare effects of tax-deferred retirement saving accounts," *Journal of Public Economics*, 2011, *95* (11-12), 1561–1578.

Snowberg, Erik and Leeat Yariv, "Testing the waters: Behavior across participant pools," *American Economic Review*, 2021, *111* (2), 687–719.

Toplak, Maggie E, Richard F West, and Keith E Stanovich, "Assessing miserly information processing: An expansion of the Cognitive Reflection Test," *Thinking & Reasoning*, 2014, *20* (2), 147–168.

Tversky, Amos and Daniel Kahneman, "The framing of decisions and the psychology of choice," *Science*, 1981, *211* (4481), 453–458.

Yoo, Kwang-Yeol and Alain De Serres, "Tax treatment of private pension savings in OECD countries and the net tax cost per unit of contribution to tax-favoured schemes," OECD Economics Department Working Paper, No. 406 2004.

Online Appendix to: "Do Tax Deferred Accounts Improve Lifecycle Savings? Experimental Evidence

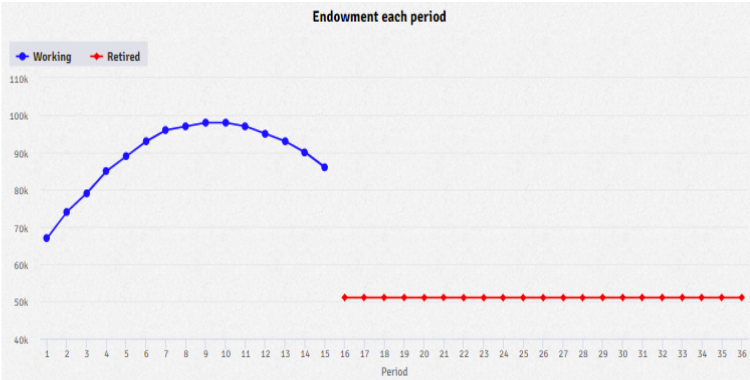
A Experimental Instructions and Screenshots

A.1 Instructions for TDA Treatment T1

Welcome to this study in the economics of individual decision-making. You are guaranteed \$7.00 for completing this study. These instructions explain how you can earn additional amounts of money from the decisions that you make. After completing the main decision task– the two sequences described below– you will be asked to complete a questionnaire. After you complete the questionnaire, you will be paid your earnings from the main decision task and from the questionnaire, in addition to your \$7.00 participation payment. The experiment can take up to 2 hours to complete, though you may be able to complete it more quickly. The main decision task involves two "sequences", which you can think of as two human lifetimes. Each sequence consists of a working phase of 15 "periods", corresponding to the periods of your life in which you work and receive income. This working phase is followed by a retirement phase, consisting of an unknown number of periods. The number of periods in the retirement phase depends on draws made by a random number generator. When the retirement phase ends, so does the sequence (or lifetime). After you have completed both sequences, one of the two sequences from this main decision task will be randomly chosen for payment. You will be paid your earnings from all periods of the chosen sequence. Since you do not know in advance which of the two sequences will be chosen for payment, you will want to do your best in both sequences. In each period of a sequence, you have a certain number of points, which you can think of as your available income. Your task is to decide how much of these available points to consume in the current period and how much to save for future periods.

Think of each period as representing two years in time. The first 15 periods 1 to 15, comprise the periods of life when you are working and earning income. You can think of these periods as corresponding to ages 30-59. Starting in period 16 (or age 60), you enter the retirement phase of life, which consists of blocks of 10 periods. A random number will be pre-drawn from 1 to 10 for each period of the retirement phase. After you make decisions for all the 10 periods of Block 1 of the retirement phase, you will be shown the list of random numbers generated for each period. If the number "1" was drawn for any period, then your earnings will only be calculated till that period and the rest will become worthless (simulating "death"). You will not move to the next block of retirement phase and that sequence will end there. But, if "1" was not drawn for any period in the first block, then the next block of 10 periods of retirement phase will start and you will make the decisions. For this block also, you will see the list of 10 randomly generated numbers from 1 to 10 for each period and the period with the first draw of "1" will become the last period for which you will get paid. Notice that there is a 1 in 10 or 10% chance that "1" was drawn and each new period reached is the last period of the sequence for which you will get paid.

At the start of each period, you will be endowed with a certain number of "points". The exact number of points given to you in each of the 15 working periods and each retirement period is shown in the graph and table below. Please take a moment to look at this graph and table before continuing.



Period	Endowment
1	67000
2	74000
3	79000
4	85000
5	89000
6	93000
7	96000
8	97000
9	98000
10	98000
11	97000
12	95000
13	93000
14	90000
15	86000
16+	51000

In this sequence, you can use two types of savings accounts—a Regular Account (RA) and a Tax-Deferred Account (TDA)—to save points for use in future periods. Neither account can have a negative balance; that is, you can save, but you cannot borrow. Your remaining balance in each account earns 10.0% interest at the start of each period. Thus, if in this period the starting balance in either account is S points, then you will earn interest income of $0.1 \times$ S which is paid to you in additional points.

The information below summarizes the difference between RA and TDA:

1. RA:
    - Deposits to and withdrawals from your RA account can be made in any period.
    - Interest earned in your RA account is counted immediately toward your taxable income.

- Distributions from your RA account are not counted toward your current taxable income.

2. TDA:

   - Deposits to your TDA account can only be made during the working phase, and cannot exceed a limit of 36000 for the first 10 rounds, and 48000 for the remaining rounds of the working phase.
   - Withdrawals from your TDA account can only be made during the retirement phase.
   - Deposits to your TDA account during the working phase are excluded from your current taxable income.
   - Withdrawals from the TDA account during the retirement phase are counted immediately toward your current taxable income.
   - Interest earned in your TDA account is not taxed until it is withdrawn.

Thus, your taxable income is determined as follows:

If it is a working period: (Taxable income) = (endowment) + (interest earned in your RA) - (deposit to your TDA)

If it is a retirement period: (Taxable income) = (endowment) + (interest earned in your RA) + (withdrawal from your TDA).

You will be given tables and figures that illustrate the relationship between taxable income, tax payments, and after-tax income.

Your earnings for each period depend on your consumption choice, which is the number of points you choose to convert into money for that period. What you don't consume is saved, and your savings can matter for the number of points that you can convert into money (or consume) in future periods. You will be given tables and figures showing the relationship between consumption choices in each period and money earned, and the slider used to make choices on your decision screen also calculates your period earnings from your consumption choice.

The precise formula is as follows:

$$(\text{Earnings}) = 1 - e^{-[(\text{consumption})-47400]/20000}$$

where e is Euler's number, 2.718281828...

Note the following:

1. The greater the consumption you choose, the greater are your earnings for that period.

2. The money you earn from consumption is proportionally diminishing; the difference in your earnings from consuming 50,000 rather than 55,000 points is larger than the difference from consuming 55,000 rather than 60,000 points, and so on.

3. To avoid negative payoffs, you will never be able to consume less than 47400 in any period. Your choices will be restricted so that you can never consume less than this amount.

After you get your endowment points each period, you will choose how much to deposit (in working periods) or withdraw (in periods after retirement) in your tax-deferred account, and the remainder will be your taxable income. You will then be taxed on that amount according to the tax schedule below.

For your convenience, the tax schedule is shown below as a graph and a table:

Income Threshold	Marginal Tax Rate
0	0.0%
21000	10.0%
39000	15.0%
96000	25.0%

Each period you are asked to submit two decisions. You first choose the amount you wish to deposit to your tax deferred savings account (TDA) if it is a working period, or the amount you want to withdraw from your TDA if it is a retirement period. Simply move the slider to a particular contribution choice or withdrawal choice and the computer will automatically calculate for you what would be your taxable income, your tax payments, and your after-tax income if you commit to that decision regarding your TDA. Once you click the Submit button, you will be prompted to make a second, consumption choice. Please pay careful attention to the after-tax income resulting from your TDA choice, as that amount will affect how much you can consume and save in your RA in the current period.

Your second choice is the amount you wish to consume. Any amount of after-tax income you do not choose to consume is put into your regular savings account (RA). Once you move the slider to a particular consumption choice, the computer will calculate for you what your current period earnings and ending RA balance would be, using the following formula:

If it is a working period:

$$(\text{Taxable income}) = (\text{Interest earned on RA balance}) + (\text{endowment}) - (\text{deposit to TDA})$$

If it is a retirement period:

(Taxable income) = (Interest earned on RA balance)+(endowment)+(withdrawal from TDA).

In any period:

$$\begin{aligned}\text{(After-tax income)} &= \text{(Taxable income)} - \text{(Taxes paid)}.\\ \text{(Ending RA balance)} &= \text{(Beginning RA balance)} + \text{(After-tax income)} - \text{(Consumption)}.\end{aligned}$$

Once you finish the first sequence, you will begin a second sequence. The second sequence will be just like the first sequence.

Following the first period of a sequence, and in every period thereafter, you will be reminded of your income endowment of points for that period, the beginning balance in your Regular savings Account (RA), the beginning balance in your Tax-Deferred Savings Account (TDA), this period's deposit or withdrawal from your TDA, your consumption choice for the period, the ending balance in your RA, the ending balance in your TDA, your taxable income, your tax payments, after-tax income, and money earned for the period, and your cumulative money earned for the sequence.

Note that once you reach the retirement phase, there is only a 90% chance that the next period in the sequence will not be the last period for payment. and there is a 10% chance that each new period reached will be the last period of the sequence for payment (simulating death). Once the sequence terminates any saved amounts of points that you have will become worthless. You will be informed about the termination of the sequence at the end of each block of 10 periods, and the listed cumulative money earned on the information screen, is assuming that a draw of 1 has not yet occurred. If a draw of a 1 has occurred, then all periods after that draw will be excluded from your final payment.

Your cumulative money earnings from each sequence are the sum of all earnings for each working period, and all earnings from the retirement periods before a random draw of the number 1 was reached. After the second sequence has been completed, we will randomly select one of the two sequences, each with equal probability. Your cumulative money earnings from that one chosen sequence will comprise your earnings for the main decision task of this experiment. You can never earn less than $0 from any part of the experiment. Upon completing the two sequences of the main decision task, you will be prompted on the screen to answer a number of questions (the questionnaire part) and you will be able to earn additional money from completing some of those questions, as will be carefully explained in that part of the study.

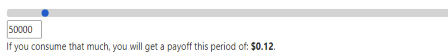
Quiz

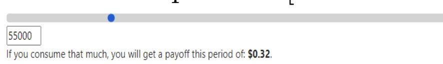
Before continuing on to the experiment, we ask that you successfully complete the following quiz. In answering these quiz questions, you may consult the instructions, tables and figures provided. If you get an answer wrong, you will be told so, and will have the opportunity to correct your answer. Your performance on this quiz does not affect your payoff in any way, but you will not be able to proceed to the experiment until you have correctly answered all of the quiz questions.

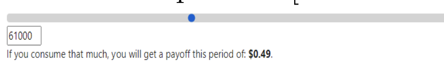
How many sequences will you participate in? [Answer: 2]

How many working periods will be in each sequence? [Answer: 15]

For the next questions, you can move the slider or enter your choice in the box below to show what your payoff would be at different levels of chosen consumption.

Suppose you consume $50,000 in a period. What would be your earnings from consumption for that period? [Answer: Moving the slider reveals the answer as illustrated here] 

Suppose you consume $55,000 in a period. What would be your earnings from consumption for that period? [Answer: Moving the slider reveals the answer as illustrated here] 

Suppose you consume $61,000 in a period. What would be your earnings from consumption for that period? [Answer: Moving the slider reveals the answer as illustrated here] 

True or False: You can contribute to the TDA in any period. [Answer: False]

True or False: You cannot withdraw from the TDA until entering the retirement phase. [Answer: True]

True or False: The length of the retirement phase depends on a sequence of random draws. Once a number 1 is drawn, you will cease to accumulate earnings, but you will continue to make decisions until the end of the current block of 10 periods is complete. [Answer: True]

True or False: In the retirement phase, after a 1 is drawn, the balance remaining in your saving account will still matter for your earnings. [Answer: False]

True or False: Your earnings will depend on your cumulative money earnings from one of the two sequences you play, but you will not know which sequence will be chosen until the end of the session. [Answer: True]

.

A.2 Illustration of the sequence of screens, TDA treatment T1

Sequence 1, period 1

You are currently *working*. You are eligible to *deposit to* the Tax-Deferred Account.

Your income this period is **67,000.0**.

Your starting balance in your **Regular Account** was 0.

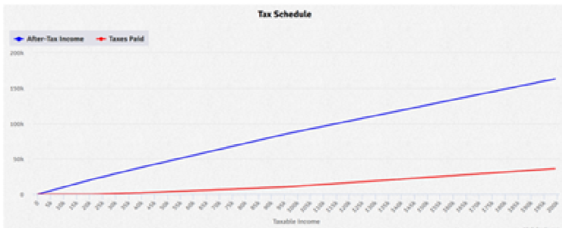
Your starting balance in your **Tax-Deferred Account** was 0.0.

You earned 0.0 on your **Regular Account**; this interest *is* counted toward your current taxable income.

You earned 0.0 on your **Tax-Deferred Account**; this interest *is not* counted toward your current taxable income.

For your convenience, the tax schedule is shown below as a table and as a graph:

Income Threshold	Marginal Tax Rate
96000	25.0%
39000	15.0%
21000	10.0%
0	0.0%



If you do not move the slider or enter your choice in the box below, the default value of 0 will be used. This default will be updated each period to match your choice in the previous period (unless this is not a feasible option).

How much do you want to deposit to your Tax-Deferred Account this period? **7,752**

7752

If you do so, your new balance in the Tax-Deferred Account will be: **$7,752.00**.

Your taxable income will be: **59,248.00**

At this taxable income, you will pay **4,837.20** in taxes, and your after-tax income will be **54,410.80**.

Tax Calculator (T1 Only)

Next

The table below is to remind you what your endowments will be in each period.

Period	Endowment
1	67000
2	74000
3	79000
4	85000
5	89000
6	93000
7	96000
8	97000
9	98000
10	98000
11	97000
12	95000
13	93000
14	90000
15	86000
16+	51000

Sequence 1, period 1

You earned 0.0 in interest on your **Regular Account**.

This interest was counted toward your taxable income.

Your taxable income was 59,248.0.

You paid 4,837.2 in taxes on this income.

Your after-tax income this period is 54,410.8.

Next

Sequence 1, period 1

Your after-tax income this period is 54,410.8.

You have 0.0 saved in your Regular Account.

If you do not move the slider or enter your choice in the box below, the default value of 50000 will be used. This default will be updated each period to match your choice in the previous period (unless this is not a feasible option).

How much do you want to consume this period? **53,340**

53340

If you consume that much, you will get a payoff this period of: **$0.26**.

You will have *deposited* **1,070.80** in your Regular Account. Next period, you will have a total of **1,070.80** in this Regular Account.

The graph and table below show how much money you can earn for different levels of consumption.

The highest payoff you could get from consumption this period is: $0.30

Next

The table below is to remind you what your endowments will be in each period.

Period	Endowment
1	67000
2	74000
3	79000
4	85000
5	89000
6	93000
7	96000
8	97000
9	98000
10	98000
11	97000
12	95000
13	93000
14	90000
15	86000
16+	51000

Results of period 1

Your income this period: 67,000.0

Your starting balance in the **Tax-Deferred Account**: 0

Interest earned in the **Tax-Deferred Account**: 0.0

Your contribution to the Tax-Deferred Account: 7,752.0

Your current balance in the **Tax-Deferred Account**: 7,752.0

Your starting balance in the **Regular Account**: 0

Interest earned in the **Regular Account**: 0.0

Your taxable income this period: 59,248.0

Taxes paid this period: 4,837.2

After-tax income this period: 54,410.8

Your consumption this period: 53,340.0

Your current balance in the **Regular Account**: 1,070.8

Money earned this period: $0.26

Money earned this sequence: $0.26

History

Sequence	Period	Income	Contribution to TDA account	Ending TDA Balance	Ending RA Balance	Consumption	Period Payoff	Cumulative Payoff
1	1	67000.0 points	7752.0 points	7752.0 points	1070.8 points	53340.0 points	**$0.26**	**$0.26**

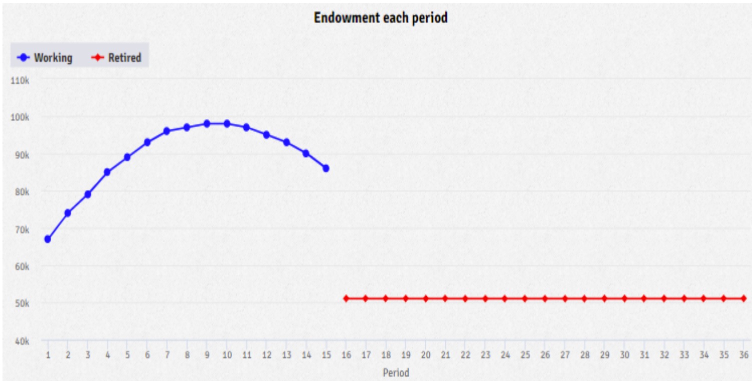
Next

A.3 Instructions for Consumption before CA choice, Treatment T6

Welcome to this study in the economics of individual decision-making. You are guaranteed $10.00 for completing this study. These instructions explain how you can earn additional amounts of money from the decisions that you make. After completing the main decision task– the two sequences described below– you will be asked to complete a questionnaire. After you complete the questionnaire, you will be paid your earnings from the main decision task and from the questionnaire, in addition to your $10.00 participation payment. The experiment can take up to 2 hours to complete, though you may be able to complete it more quickly. The main decision task involves two "sequences", which you can think of as two human lifetimes. Each sequence consists of a working phase of 15 "periods", corresponding to the periods of your life in which you work and receive income. This working phase is followed by a retirement phase, consisting of an unknown number of periods. The number of periods in the retirement phase depends on draws made by a random number generator. When the retirement phase ends, so does the sequence (or lifetime). After you have completed both sequences, one of the two sequences from this main decision task will be randomly chosen for payment. You will be paid your earnings from all periods of the chosen sequence. Since you do not know in advance which of the two sequences will be chosen for payment, you will want to do your best in both sequences. In each period of a sequence, you have a certain number of points, which you can think of as your available income. Your task is to decide how much of these available points to consume in the current period and how much to save for future periods.

Think of each period as representing two years in time. The first 15 periods 1 to 15, comprise the periods of life when you are working and earning income. You can think of these periods as corresponding to ages 30-59. Starting in period 16 (or age 60), you enter the retirement phase of life, which consists of blocks of 10 periods. A random number will be pre-drawn from 1 to 10 for each period of the retirement phase. After you make decisions for all the 10 periods of Block 1 of the retirement phase, you will be shown the list of random numbers generated for each period. If the number '1' was drawn for any period, then your earnings will only be calculated till that period and the rest will become worthless (simulating "death"). You will not move to the next block of retirement phase and that sequence will end there. But, if the number "1" was not drawn for any period in the first block, then another block of 10 periods of retirement phase will begin and you will continue making decisions. At the end of this next block you will also see the list of 10 randomly generated numbers from 1 to 10 for each period and the period with the first draw of "1" will become the last period for which you will get paid. Notice that there is a 1 in 10 or 10% chance that the number "1" was drawn and each new period reached is the last period of the sequence for which you will get paid.

At the start of each period, you will be endowed with a certain number of "points". The exact number of points given to you in each of the 15 working periods and each retirement period is shown in a table and graph below. Please take a moment to look at this table and graph before continuing.



Period	Endowment
1	67000
2	74000
3	79000
4	85000
5	89000
6	93000
7	96000
8	97000
9	98000
10	98000
11	97000
12	95000
13	93000
14	90000
15	86000
16+	51000

In this sequence, you can use two types of savings accounts—a Regular Account (RA) and a Commitment Account (CA)—to save points for use in future periods. Neither account can have a negative balance; that is, you can save, but you cannot borrow. Your remaining balance in each account earns 10.0% interest at the start of each period. Thus, if in this period the starting balance in either account is S points, then you will earn interest income of 0.1 x S which is paid to you in additional points.

Interest earned in both accounts is counted immediately toward your taxable income. The only difference is that for your Commitment Account, deposits can only be made during the working phase and withdrawals can only be made during the retirement phase; for your Regular Account, deposits and withdrawals can be made in any period.

Thus, your taxable income is determined as follows:

(Taxable income) = (points endowed in current period) + (interest earnings in points from both RA and CA)

You will be given tables and figures that illustrate the relationship between taxable income, tax payments, and after-tax income.

Your earnings for each period depend on your consumption choice, which is the number of points you choose to convert into money for that period. What you don't consume is saved, and your savings can matter for the number of points that you can convert into money (or consume) in future periods. You will be given tables and figures showing the relationship between consumption choices in each period and money earned, and the slider used to make choices on your decision screen also calculates your period earnings from your consumption choice.

The precise formula is as follows:

$$(\text{Earnings}) = 1 - e^{-[(\text{consumption})-47400]/20000}$$

where e is Euler's number, 2.718281828...

Note the following:

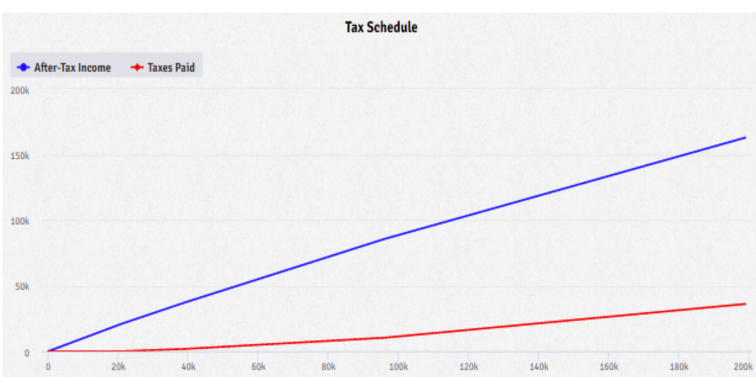
1. The greater the consumption you choose, the greater are your earnings for that period.

2. The money you earn from consumption is proportionally diminishing; the difference in your earnings from consuming 50,000 rather than 55,000 points is larger than the difference from consuming 55,000 rather than 60,000 points, and so on.

3. To avoid negative payoffs, you will never be able to consume less than 47400 in any period. Your choices will be restricted so that you can never consume less than this amount.

After you get your endowment points each period, you will be taxed on the sum of your endowment income and interest earnings according to the tax schedule below.

For your convenience, the tax schedule is shown below as a table and as a graph:



Income Threshold	Marginal Tax Rate
0	0.0%
21000	10.0%
39000	15.0%
96000	25.0%

Each period you are asked to submit two decisions. You first make a consumption choice. Once you move the slider to a particular consumption choice, the computer will calculate for you what your current period earnings would be. Once you click the Submit button, you will be prompted to choose the amount you wish to deposit to your commitment account (CA) if it is a working period, or the amount you want to withdraw from your CA if it is a retirement period. Any amount of after-tax income you do not choose to consume or deposit to your CA is put into your regular savings account (RA). The computer program will calculate for you your ending RA balance using the following formula:

In any period:

(After-tax income) = (Endowment) + (Interest earned from both RA and CA) - (Taxes Paid)

If it is a working period:

(Ending RA balance) = (Beginning RA balance) + (After-tax income) - (Consumption) - (Deposit to CA) - (Interest earned from CA).

If it is a retirement period:

(Ending RA balance) = (Beginning RA balance) + (After-tax income) - (Consumption) + (Withdraw from CA) - (Interest earned from CA).

Once you finish the first sequence, you will begin a second sequence. The second sequence will be just like the first sequence.

Following the first period of a sequence, and in every period thereafter, you will be reminded of your income endowment of points for that period, the beginning balance in your Regular savings Account (RA), the beginning balance in your Commitment Account (CA), this period's deposit or withdrawal from your CA, your consumption choice for the period, the ending balance in your RA, the ending balance in your CA, your taxable income, your tax payments, after-tax income, and money earned for the period, and your cumulative money earned for the sequence.

Note that once you reach the retirement phase, there is only a 90% chance that the next period in the sequence will not be the last period for payment. and there is a 10% chance that each new period reached will be the last period of the sequence for payment (simulating death). Once the sequence terminates any saved amounts of points that you have will become worthless. You will be informed about the termination of the sequence at the end of each block of 10 periods, and the listed cumulative money earned on the information screen, is assuming that a draw of 1 has not yet occurred. If a draw of a 1 has occurred, then all periods after that draw will be excluded from your final payment.

Your cumulative money earnings from each sequence are the sum of all earnings for each working period, and all earnings from the retirement periods before a random draw of the number 1 was reached. After the second sequence has been completed, we will randomly select one of the two sequences, each with equal probability. Your cumulative money earnings from that one chosen sequence will comprise your earnings for the main decision task of this experiment. You can never earn less than $0 from any part of the experiment. Upon completing the two sequences of the main decision task, you will be prompted on the screen to answer a number of questions (the questionnaire part) and you will be able to earn additional money from completing some of those questions, as will be carefully explained in that part of the study.

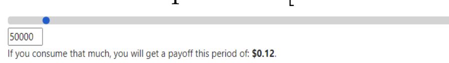
Quiz

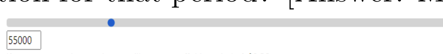
Before continuing on to the experiment, we ask that you successfully complete the following quiz. In answering these quiz questions, you may consult the instructions, tables and figures provided. If you get an answer wrong, you will be told so, and will have the opportunity to correct your answer. Your performance on this quiz does not affect your payoff in any way, but you will not be able to proceed to the experiment until you have correctly answered all of the quiz questions.

How many sequences will you participate in? [Answer: 2]

How many working periods will be in each sequence? [Answer: 15]

For the next questions, you can move the slider or enter your choice in the box below to show what your payoff would be at different levels of chosen consumption.

Suppose you consume $50,000 in a period. What would be your earnings from consumption for that period? [Answer: Moving the slider reveals the answer as illustrated here] 

Suppose you consume $55,000 in a period. What would be your earnings from consumption for that period? [Answer: Moving the slider reveals the answer as illustrated here] 

Suppose you consume $61,000 in a period. What would be your earnings from consumption for that period? [Answer: Moving the slider reveals the answer as illustrated here] 

The interest rate on the RA is 10.0%. Suppose your ending balance in the RA is $60,000. What amount of interest would you earn on this amount for the next period? [Answer: $6,000]

True or False: You cannot withdraw from the RA until entering the retirement phase. [Answer: False]

True or False: You can deposit to your CA in both the working and retirement phase. [Answer: False]

True or False: Interest income from both the CA and RA are taxed immediately. [Answer: True]

True or False: The length of the retirement phase depends on a sequence of random draws. Once a number 1 is drawn, you will cease to accumulate earnings, but you will continue to make decisions until the end of the current block of 10 periods is complete. [Answer: True]

True or False: In the retirement phase, after a 1 is drawn, the balance remaining in your saving account will still matter for your earnings. [Answer: False]

True or False: Your earnings will depend on your cumulative money earnings from one of the two sequences you play, but you will not know which sequence will be chosen until the end of the session. [Answer: True]

.

A.4 Illustration of the sequence of screens, Consumption before CA choice, Treatment T6

Sequence 1, period 1

Your income this period is **67,000.0**.

Your starting balance in your **Regular Account** was 0.

Your starting balance in your **Commitment Account** was 0.

You earned 0.0 in interest on your **Regular Account**.

You earned 0.0 in interest on your **Commitment Account**.

Interest payments on both accounts are counted toward your current taxable income.

Your after-tax income this period is 61,000.0.

You have 0 saved in your Regular Account.

If you do not move the slider or enter your choice in the box below, the default value of 50000 will be used. This default will be updated each period to match your choice in the previous period (unless this is not a feasible option).

How much do you want to consume this period? **53,340**

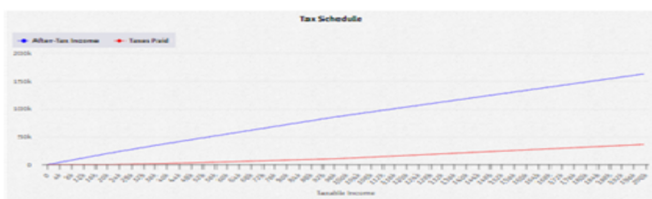
53340

If you consume that much, you will get a payoff this period of: **$0.26**.

The graph and table below show how much money you can earn for different levels of consumption, and provide a reminder about what your future tax payments will be based on your taxable income.

The highest payoff you could get from consumption this period is: $0.49

Next



The table below is to remind you what your endowments will be in each period.

Period	Endowment
1	67000
2	74000
3	79000
4	85000
5	89000
6	93000
7	96000
8	97000
9	98000
10	98000
11	97000
12	95000
13	93000
14	90000
15	86000
16+	51000

Sequence 1, period 1

You are currently *working*. You are eligible to *deposit to* the Commitment Account.

If you do not move the slider or enter your choice in the box below, the default value of 0 will be used. This default will be updated each period to match your choice in the previous period (unless this is not a feasible option).

How much do you want to deposit to your Commitment Account this period? **5,500**

5500

If you do so, your new balance in the Commitment Account will be: **$5,500.00**.

You will have *deposited* **2,160.00** in your Regular Account. Next period, you will have a total of **2,160.00** in this Regular Account.

Next

The table below is to remind you what your endowments will be in each period.

Period	Endowment
1	67000
2	74000
3	79000
4	85000
5	89000
6	93000
7	96000
8	97000
9	98000
10	98000
11	97000
12	95000
13	93000
14	90000
15	86000
16+	51000

Results of period 1

Your income this period: 67,000.0

Your starting balance in the **Commitment Account**: 0

Interest earned in the **Commitment Account**: 0.0

Your contribution to the Commitment Account: 5,500.0

Your current balance in the **Commitment Account**: 5,500.0

Your starting balance in the **Regular Account**: 0

Interest earned in the **Regular Account**: 0.0

Your taxable income this period: 67,000.0

Taxes paid this period: 6,000.0

After-tax income this period: 61,000.0

Your consumption this period: 53,340.0

Your current balance in the **Regular Account**: 2,160.0

Money earned this period: $0.26

Money earned this sequence: $0.26

History

Sequence	Period	Income	Deposit to CA	Ending CA Balance	Ending RA Balance	Consumption	Period Payoff	Cumulative Payoff
1	1	67000.0 points	5500.0 points	5500.0 points	2160.0 points	53340.0 points	**$0.26**	**$0.26**

Next

B End of Experiment Tasks (All Treatments)

The experiment concluded with four additional tasks, reproduced here.

Task 1: Lotteries

In this task you make 10 choices over pairs of "lotteries". Each pair of lotteries is labeled lottery A and lottery B. Each lottery describes two different possible dollar amounts that you could earn with an equal (50%) probability. For each pair of lotteries, you must choose whether you prefer lottery A or lottery B for possible payment. You may choose lottery A for some choices and lottery B for other choices and you may change your decisions or make them in any order that you want. Once you have made a lottery choice for all 10 pairs and clicked the Next button, the computer program will randomly select one of the 10 lottery pairs and implement your choice for that pair. It will then randomly determine whether you got the lower amount or the higher amount for that lottery, each with a 50% chance.

1. Lottery A: 50% chance of $0.75, 50% chance of $0.83; Lottery B: 50% chance of $0.15, 50% chance of $1.16: Choose: A, B
2. Lottery A: 50% chance of $0.75, 50% chance of $0.91; Lottery B: 50% chance of $0.15, 50% chance of $1.32: Choose: A, B
3. Lottery A: 50% chance of $0.75, 50% chance of $0.99; Lottery B: 50% chance of $0.15, 50% chance of $1.47: Choose: A, B
4. Lottery A: 50% chance of $0.75, 50% chance of $1.07; Lottery B: 50% chance of $0.15, 50% chance of $1.63: Choose: A, B
5. Lottery A: 50% chance of $0.75, 50% chance of $1.15; Lottery B: 50% chance of $0.15, 50% chance of $1.8: Choose: A, B
6. Lottery A: 50% chance of $0.75, 50% chance of $1.23; Lottery B: 50% chance of $0.15, 50% chance of $2.01: Choose A, B
7. Lottery A: 50% chance of $0.75, 50% chance of $1.31; Lottery B: 50% chance of $0.15, 50% chance of $2.29: Choose A, B
8. Lottery A: 50% chance of $0.75, 50% chance of $1.39; Lottery B: 50% chance of $0.15, 50% chance of $2.69: Choose A, B
9. Lottery A: 50% chance of $0.75, 50% chance of $1.47; Lottery B: 50% chance of $0.15, 50% chance of $3.65: Choose A, B
10. Lottery A: 50 chance of $0.75, 50% chance of $1.55; Lottery B: 50% chance of $0.15, 50% chance of $3.85: Choose A, B

Task 2 Quiz [Cognitive Reflection Test] Questions [with Answers not provided to subjects.]

Please answer the following questions.

1. If John can drink one barrel of water in 6 days, and Mary can drink one barrel of water in 12 days, how many days would it take them to drink one barrel of water together? [Answer: 4 days]
2. Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class? [Answer: 29]
3. A man buys a pig for $60, sells it for $70, buys it back for $80, and sells it finally for $90. How many dollars has he made? [Answer: $20]
4. Simon decided to invest $8,000 in the stock market one day early in 2008. Six months after he invested, on July 17, the stocks he had purchased were down 50%. Fortunately for Simon, from July 17 to October 17, the stocks he had purchased went up 75%. As of October 17, Simon has: [Answer: a]

   a lost money in the stock market

   b broken even in the stock market

   c made money in the stock market

Task 3 [Financial Literacy] Quiz Questions [With answers not shown to subjects]

1. Suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow?

   a More than $102

   b Exactly $102

   c Less than $102

   d Do not know

   e Refuse to answer.

   [Answer: a]
2. Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account?

   a More than today

   b Exactly the same

c Less than today

d Do not know

e Refuse to answer.

[Answer: c]

3. Please consider whether this statement is true or false. “Buying a single company’s stock usually provides a safer return than a stock mutual fund.”

a True

b False

c Do not know

d Refuse to answer.

[Answer: False]

Task 4: Questionnaire

What is your age?

What year were you born?

What is your race?

What is your personal annual income?

What is your highest level of education completed?

What is your ZIP or postal code?

Please enter any comments about the experiment, or suggestions for improvement.

C Additional Tables and Figures

Table C1: Treatment effect on tax liability

	A: Tax (k)			B: Tax/Endowment		
	(1)	(2)	(3)	(4)	(5)	(6)
	1-5	6-10	11-15	1-5	6-10	11-15
Cons	5.57***	8.04***	7.79***	0.07***	0.08***	0.08***
	(0.26)	(0.50)	(0.84)	(0.00)	(0.01)	(0.01)
T2	2.50***	4.36***	4.80***	0.03***	0.05***	0.05***
	(0.17)	(0.35)	(0.49)	(0.00)	(0.00)	(0.01)
T3	2.33***	3.08***	3.24***	0.03***	0.03***	0.04***
	(0.17)	(0.28)	(0.34)	(0.00)	(0.00)	(0.00)
T4	0.11	-0.13	-0.19	0.00	-0.00	-0.00
	(0.22)	(0.35)	(0.40)	(0.00)	(0.00)	(0.00)
T5	2.62***	6.24***	10.61***	0.03***	0.06***	0.12***
	(0.17)	(0.31)	(0.55)	(0.00)	(0.00)	(0.01)
T6	2.48***	4.84***	6.99***	0.03***	0.05***	0.08***
	(0.17)	(0.39)	(0.74)	(0.00)	(0.00)	(0.01)
CRT	-0.03	-0.06	-0.11	-0.00	-0.00	-0.00
	(0.04)	(0.09)	(0.16)	(0.00)	(0.00)	(0.00)
FL	0.05	0.09	0.11	0.00	0.00	0.00
	(0.06)	(0.14)	(0.25)	(0.00)	(0.00)	(0.00)
RE	-0.00	0.01	0.01	-0.00	0.00	0.00
	(0.02)	(0.04)	(0.08)	(0.00)	(0.00)	(0.00)
Female	-0.04	-0.41**	-0.83**	-0.00	-0.00**	-0.01**
	(0.08)	(0.19)	(0.34)	(0.00)	(0.00)	(0.00)
S2	0.10***	0.23**	0.38**	0.00***	0.00**	0.00**
	(0.04)	(0.09)	(0.17)	(0.00)	(0.00)	(0.00)
R^2	0.80	0.76	0.72	0.81	0.76	0.72
N	370	370	370	370	370	370

Notes: The table presents estimates from Equation (7). The dependent variables are the average tax liability (Panel A) and the average ratio of tax liability to endowment (Panel B) over each 5-period interval. Standard errors are clustered at the subject level. *** p<0.01, ** p<0.05, and * p<0.10.

Table C2: Treatment effect on MSE (k) in lab sample

	1-5 (1)	6-10 (2)	11-15 (3)	16-20 (4)	21-25 (5)	1-25 (6)
Cons	95.07***	151.92*	277.13**	5755.51**	2381.82***	1732.29***
	(19.74)	(90.92)	(119.00)	(2309.02)	(587.69)	(497.15)
T2	15.90	431.36***	160.18	-1818.31***	-1066.13***	-455.40***
	(12.08)	(109.77)	(101.87)	(507.74)	(246.91)	(133.30)
T3	13.84	330.98***	174.72**	-1705.22***	-711.21***	-379.38***
	(13.70)	(74.48)	(79.33)	(501.93)	(247.05)	(126.21)
T4	-14.83	75.49	131.50	974.75	-129.25	207.53
	(12.72)	(60.84)	(85.68)	(1298.48)	(327.67)	(279.42)
T5	4.18	51.35*	46.20	2682.98	870.84	731.11
	(12.32)	(29.88)	(50.06)	(2204.84)	(640.70)	(471.19)
T6	9.78	49.59*	-62.00	-1393.67**	-335.46	-346.35**
	(13.06)	(29.34)	(50.72)	(538.67)	(340.02)	(156.11)
CRT	2.95	-30.25	-11.29	-163.16	73.70	-25.61
	(3.35)	(21.33)	(18.42)	(196.79)	(134.57)	(54.75)
FL	2.66	21.60	-4.23	-539.82	-306.88	-165.33
	(4.64)	(30.43)	(29.09)	(435.72)	(257.49)	(107.95)
RE	-2.83	4.32	15.83	-281.16	7.95	-51.18
	(1.72)	(10.77)	(11.46)	(249.41)	(47.80)	(50.98)
Female	-15.05*	-63.22	-104.99*	214.55	-508.22*	-95.39
	(7.88)	(56.52)	(59.00)	(793.14)	(265.66)	(180.54)
S2	-10.21**	5.27	-67.13*	299.52	-77.66	29.96
	(5.04)	(41.57)	(39.66)	(502.07)	(182.03)	(112.78)
R^2	0.04	0.10	0.03	0.04	0.08	0.06
N	370	370	370	370	370	370

Notes: The table presents estimates from Equation (7). The dependent variable is the mean square error (MSE) of experimental data compared to theory predictions over 5 or 25 period intervals. Standard errors are clustered at the subject level. *** p<0.01, ** p<0.05, and * p<0.10.

Table C3: Treatment effect on MSE (k) in MTurk sample

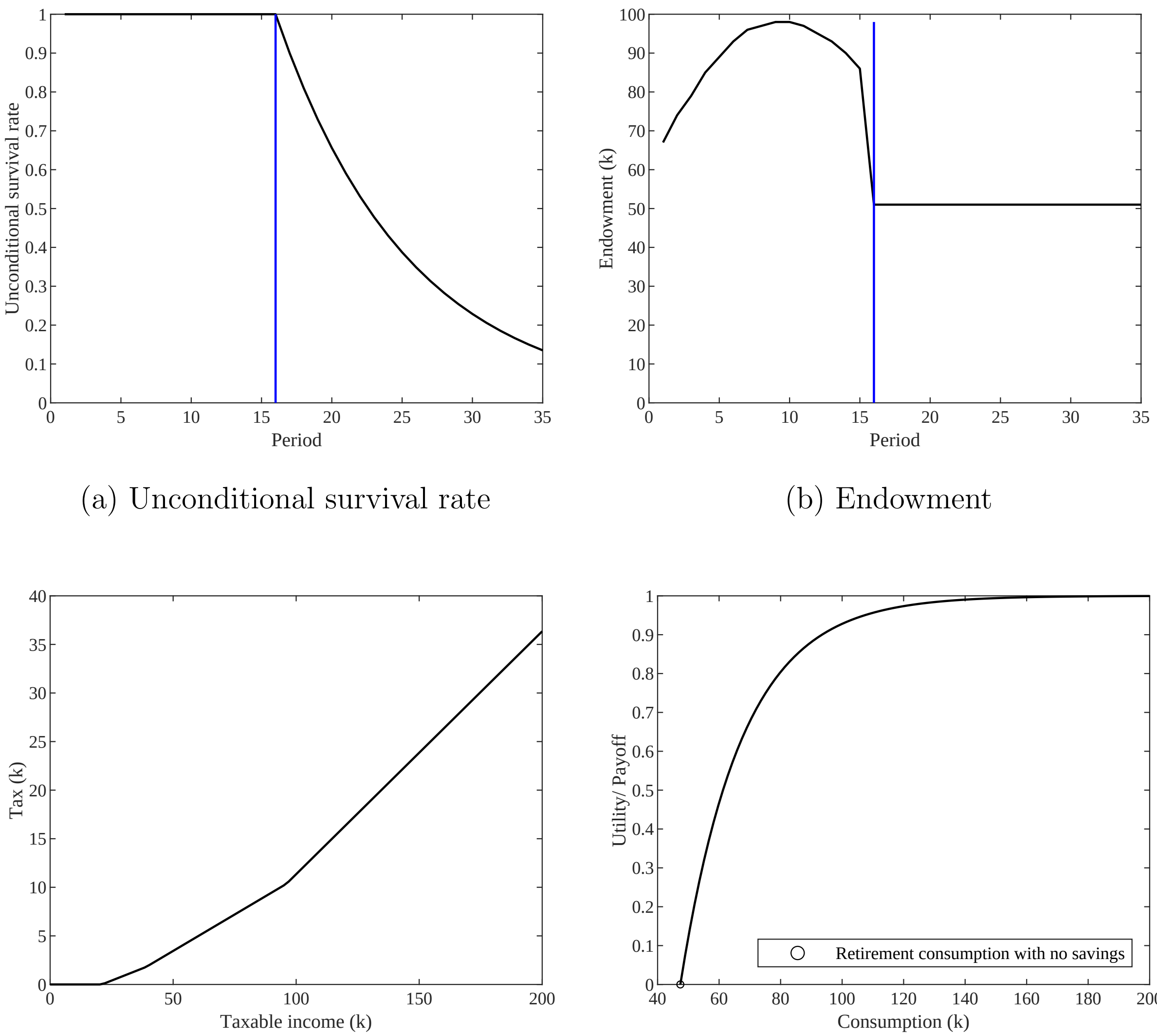
	1-5 (1)	6-10 (2)	11-15 (3)	16-20 (4)	21-25 (5)	1-25 (6)
Cons	85.05***	399.00***	1112.53***	10436.01**	3919.52***	3190.42***
	(12.65)	(143.30)	(326.48)	(5197.44)	(1066.79)	(1045.51)
T2	68.86***	631.08***	-213.06	-5675.97***	-526.98	-1143.21**
	(19.66)	(223.97)	(207.53)	(2176.76)	(551.02)	(453.30)
T3	25.79*	279.35**	-238.77	-7485.88***	-359.52	-1555.81***
	(14.86)	(120.32)	(204.15)	(2173.16)	(561.16)	(438.43)
T4	-2.20	29.33	25.67	12853.10**	461.96	2673.57**
	(10.65)	(86.26)	(257.05)	(6438.56)	(872.23)	(1278.55)
T5	-9.45	41.89	-268.99	7227.01*	2195.02	1837.10**
	(11.00)	(68.39)	(225.78)	(3668.21)	(1460.84)	(771.04)
T6	25.36**	39.13	-645.60***	-5671.11**	-2256.83***	-1701.81***
	(12.60)	(66.66)	(242.54)	(2794.07)	(695.64)	(559.50)
CRT	0.66	-59.35	46.52	-1522.26	956.84**	-115.52
	(3.86)	(49.03)	(41.55)	(1013.48)	(403.71)	(222.42)
FL	-1.57	3.85	-133.85**	-443.74	-1841.13**	-483.29
	(4.20)	(44.80)	(61.41)	(1392.73)	(773.19)	(318.83)
RE	-1.07	2.15	-27.08	-30.21	38.81	-3.48
	(1.80)	(17.74)	(17.48)	(398.50)	(70.08)	(78.82)
Female	7.37	-100.01	-6.06	3614.69	848.77	872.95
	(10.05)	(88.36)	(103.91)	(2821.12)	(732.85)	(580.56)
S2	-17.15**	-118.37	-244.35**	346.96	-1410.09**	-288.60
	(7.87)	(90.62)	(97.25)	(1144.67)	(625.39)	(252.73)
R^2	0.08	0.05	0.06	0.15	0.11	0.19
N	360	360	360	360	360	360

Notes: The table presents estimates from Equation (7). The dependent variable is the mean square error (MSE) of experimental data compared to theory predictions over 5 or 25 period intervals. Standard errors are clustered at the subject level. *** p<0.01, ** p<0.05, and * p<0.10.

Table C4: Treatment effects on consumption and expected lifetime utility in the MTurk sample

	Consumption					Lifetime
	1-5	6-10	11-15	16-20	21-25	Utility
	(1)	(2)	(3)	(4)	(5)	(6)
Cons	54.77***	66.56***	67.45***	112.40***	85.44***	9.81***
	(1.39)	(2.85)	(3.86)	(11.52)	(8.12)	(0.69)
T2	9.33***	16.52***	14.42***	-59.57***	-36.82***	-0.18
	(1.42)	(2.38)	(2.83)	(6.53)	(5.90)	(0.63)
T3	8.80***	17.50***	19.68***	-57.87***	-35.33***	1.04*
	(1.25)	(2.08)	(2.94)	(6.78)	(5.87)	(0.55)
T4	0.43	-0.04	-4.47	8.44	-21.78***	-1.79**
	(1.11)	(2.18)	(3.11)	(11.60)	(7.69)	(0.79)
T5	1.74	-2.54	-7.03**	16.21	-17.74**	-2.02***
	(1.08)	(2.20)	(2.96)	(10.73)	(8.46)	(0.71)
T6	12.12***	14.20***	9.88***	-46.96***	-44.61***	0.76
	(1.24)	(2.07)	(2.78)	(8.03)	(6.02)	(0.59)
CRT	-0.13	-0.29	0.26	-2.71	6.72***	0.11
	(0.34)	(0.57)	(0.64)	(2.25)	(1.71)	(0.17)
FL	0.07	0.38	0.34	0.71	-4.44	0.49**
	(0.40)	(0.76)	(0.88)	(3.12)	(3.18)	(0.23)
RE	-0.14	-0.14	-0.27	0.97	1.25**	0.09
	(0.14)	(0.23)	(0.28)	(0.84)	(0.62)	(0.06)
Female	1.09	-0.44	0.07	4.62	-3.27	-0.40
	(0.79)	(1.37)	(1.55)	(5.49)	(3.81)	(0.39)
S2	-0.55	-0.54	-1.20	8.69***	-3.42	0.58***
	(0.45)	(0.91)	(1.14)	(2.89)	(2.65)	(0.20)
R^2	0.40	0.38	0.36	0.38	0.23	0.18
N	360	360	360	360	360	360

Notes: The table presents estimates from Equation (7). The dependent variables are the average consumption over each 5-period interval and expected lifetime utility. Standard errors are clustered at the subject level. *** p<0.01, ** p<0.05, and * p<0.10.



(a) Unconditional survival rate

(b) Endowment

(c) Tax function

(d) Utility function

Figure C1: Environmental Variables

Note: The vertical line in Panels A and B indicates the first period of retirement.

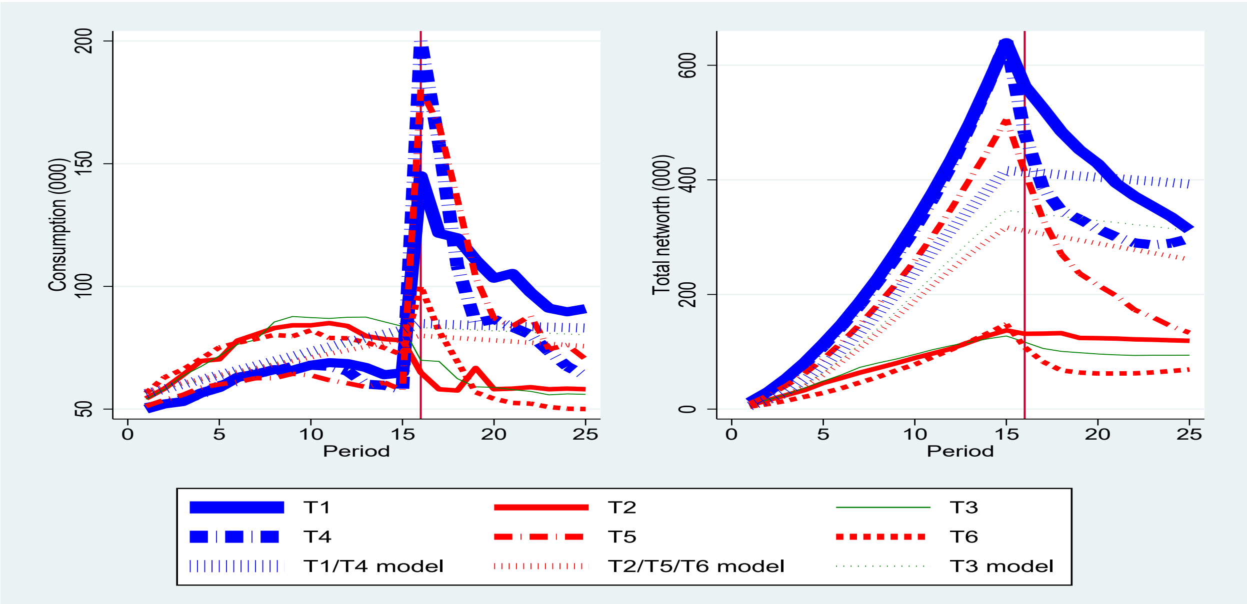
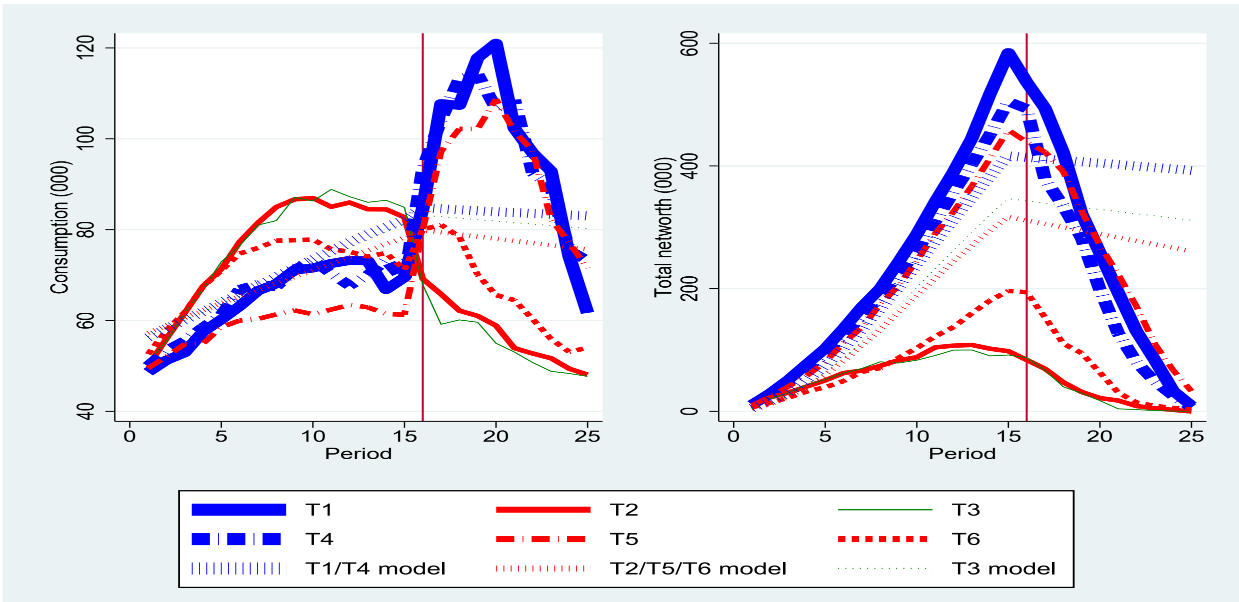
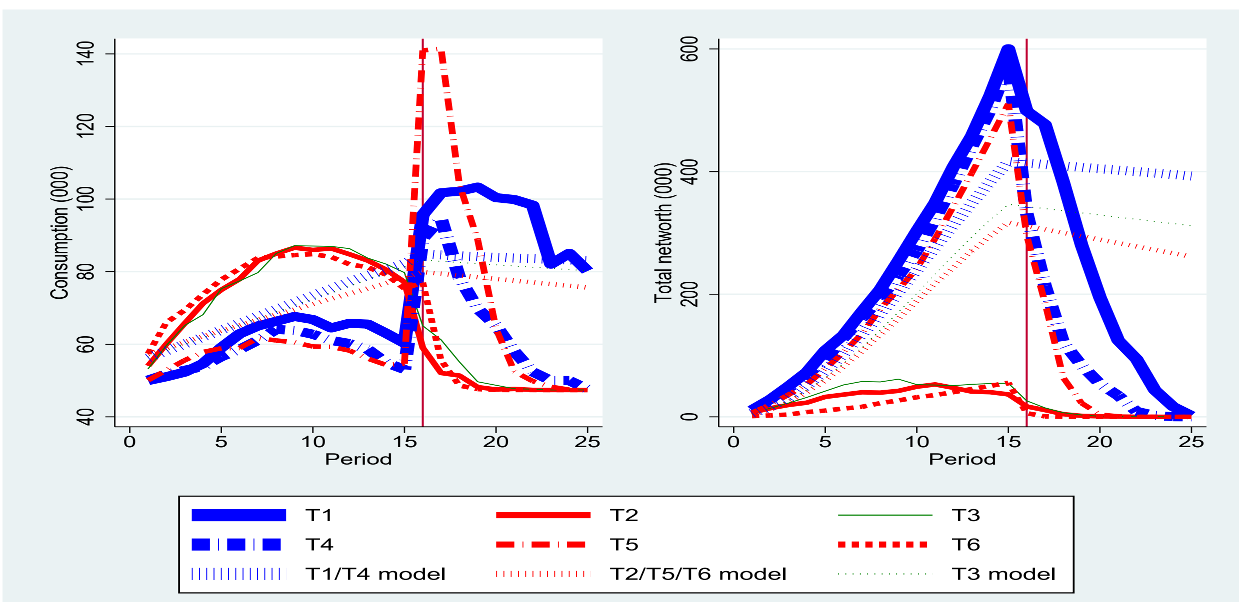


Figure C2: Theory predictions and experimental means of MTurk sample

Note: The right panel displays $a_{j+1} + b_{j+1}$.



(a) Lab sample



(b) MTurk sample

Figure C3: Theory predictions and experimental medians

Note: The right panels display $a_{j+1} + b_{j+1}$.

D Numerical method for solving the theoretical model

Given the complexity of the model, obtaining an analytical solution is not feasible. Hence, the optimal decisions are solved using numerical methods, as outlined below. The Matlab codes used for solving the model are provided in the online data appendix.

1. Discretize the continuous state variables as follows:
   - For the TDA treatment, utilize 50 grid points for the RA balance, and 100 grid points for the TDA balance.
   - For the non-TDA treatments, utilize 100 grid points for the RA balance.
2. Solve the retirees' problem through value function iteration
3. Solve the workers' problem backwards
4. Find the optimal decisions using the value functions obtained in steps 2 and 3