



**University of
Zurich** ^{UZH}

University of Zurich
Department of Economics

Working Paper Series

ISSN 1664-7041 (print)
ISSN 1664-705X (online)

Working Paper No. 440

A Behavioral Study of Roth versus Traditional Retirement Savings Accounts

Clement E. Bohr, Charles A. Holt and Alexandra V. Schubert

November 2023

A Behavioral Study of Roth versus Traditional Retirement Savings Accounts

Clement E. Bohr
Northwestern University

Charles A. Holt
University of Virginia

Alexandra V. Schubert*
University of Zurich

November 2023

Abstract: Motivated by a popular perception that Roth accounts are welfare-improving for most people, this paper compares the effects of mandated Traditional (tax-deferred) or Roth (tax-prepaid) retirement policies in a controlled laboratory setting. Selection effects, which complicate analyses of observational data, are avoided by random assignment to policies. Subjects receive exogenous incomes during “working” periods, followed by no-income “retirement” periods. In each period, subjects decide how many lab dollars to convert into “take-home pay,” akin to consumption with diminishing returns. Subjects’ decisions determine retirement savings and tax payments. Flat income and tax-rate profiles facilitate the analysis of behavioral factors like present-period tax avoidance, while optimal consumption and after-tax savings are identical for both treatments. Our results show that observed savings are suboptimal in both treatments and are influenced by gender, patience, and risk aversion measures. In contrast to conventional wisdom, there are no significant differences between policies; if anything, the Traditional treatment leads to marginally higher post-retirement consumption.

JEL Codes: C91, D84, D91, E21, H24 J32

Keywords: retirement, tax-deferred savings, Roth, IRA, compound interest bias, laboratory experiments.

*Affiliations: Clement E. Bohr, Northwestern University, Department of Economics, 2211 Campus Drive, Evanston, IL 60208, USA, cebohr@u.northwestern.edu; Charles A. Holt, University of Virginia, Department of Economics, 248 McCormick Rd, Charlottesville, VA 22903, USA, cah2k@virginia.edu; Alexandra V. Schubert, University of Zurich, Department of Economics, Schönberggasse 1, 8001 Zurich, Switzerland. lexi.schubert@econ.uzh.ch

We gratefully acknowledge support from the University of Virginia Bankard Faculty Research Grant (2019-20-21). We would like to thank Madeleine Green, David Li, Juliette Sellgren, and Natasha Swindle for research assistance. This paper has benefitted from discussions with John Duffy, Daniel Harper, Charles Noussair, Olivier Armantier, and participants at the Theory and Experiments in Monetary Economics (TEME) Conference 2022 and the ESA North America Meeting 2023.

1. Introduction

How to save for retirement is one of the most important financial decisions individuals have to make during their lifetime as it is a major determinant of income and consumption later in life. The stakes of these savings decisions are high, and rational-agent models predict adequate and careful planning for retirement years. Nonetheless, most people lack sufficient savings for retirement.^{1 2} Hence, it is crucial to better understand policies that encourage savings. Mandated or nudged retirement plans have, for instance, been associated with higher savings rates (e.g., Bernatzi and Thaler 2007; Bohr, Holt, and Schubert 2019). Besides studying the effect of policies that mandate retirement plans, it is important to understand which policy characteristics map to higher savings rates, consumption during retirement, and overall welfare.

Tax incentives are an example of these plan characteristics that have been used to encourage saving. In the US, two popular types of accounts are “traditional” individual retirement savings accounts (IRAs) and Roth accounts. Taxes on income that is saved in traditional accounts are deferred. That is, income put into a traditional retirement account will not be taxed at the point in time when it is saved. Instead, income taxes are paid when savings are withdrawn for consumption during retirement. In contrast, income put into Roth accounts is taxed at usual rates when earned, and no taxes are deducted from withdrawals during retirement. If the proportional income tax rate is unchanging, the permanent income hypothesis would predict that the timing of the taxes does not matter for consumption behavior across the life cycle. Taking behavioral biases into account, however, it is plausible that one of these retirement accounts would induce higher levels of savings than the other.

This paper experimentally compares the implications of mandated Traditional or Roth accounts on saving and consumption over a simulated lifetime. We conduct a lab experiment with two treatments that represent the two main types of retirement accounts: Traditional (tax-deferred) and Roth (tax-prepaid). Both accounts can be used to smooth consumption over the

¹ For example, Bond and Porell (2020) report that about 40 percent of retirees in the US rely exclusively on social security, which was never intended to be the sole source of support. Social Security typically replaces only about 40 percent of pre-retirement earnings. As a result, a significant fraction of retired Americans fall below poverty income levels.

² Moreover, gender pension gaps emerge. The 2022 TIAA Financial Wellness Survey reports that, compared to men, women in the US are saving less for retirement and feel less confident that they can enjoy a comfortable retirement without running out of money (Teachers Insurance and Annuity Association of America, 2022). Some, but not all, of these gaps can be explained by institutional factors like the child penalty (Kleven et al., 2019), leaving room to explore what else might be driving this.

simulated lifetime; however, they differ in when taxes are paid. Subjects are randomly assigned to one of two treatment conditions. In the Traditional treatment, no income taxes are paid on money that is saved; taxes are applied when money is withdrawn for consumption. In the Roth treatment, income is taxed upon receipt; no taxes are paid later. Every period, subjects receive an exogenous income and decide how much of this income to save or consume. Money invested in their retirement account earns interest, which can introduce present-value considerations. Additionally, as in the real world, choices might be biased by compound-interest bias. In addition to interest on savings, subjects receive exogenous incomes every period up until “retirement,” when it drops to zero. Every period, subjects must decide how much of their available lab earnings to convert to take-home pay (“consumption”). The marginal returns to consumption are decreasing such that the first lab dollar converted yields more marginal take-home pay than later lab dollars converted within a period. The setup also features a constant income tax, which is a known proportion of the stationary, exogenous work income. This avoids tax-planning motives for saving and allows us to isolate behavioral factors like loss (tax) aversion and a failure to anticipate future consumption needs or tax liabilities. We abstract from capital gains taxes since we are comparing two different types of retirement accounts, which by design are both exempt from such taxes.

We find that subjects in both tax treatments, tax-deferred (Traditional) and tax-prepaid (Roth), did not save enough for retirement relative to optimal levels determined by the consumption conversion process. The observation of sub-optimal savings and consumption during retirement is consistent with our prior findings in an experiment that did not consider tax effects (Bohr, Holt, and Schubert, 2019).³ This prior experiment, along with that of Duffy and Li (2023), found that life cycle consumption patterns were improved by programs that automatically transfer working-period incomes into retirement income without altering the present value of the income stream.

Our paper introduces income tax considerations in a neutral context intended to permit a clear focus on behavioral biases. As mentioned, experts typically recommend Roth accounts

³ Retirement savings were about half of optimal levels in a baseline treatment with a 5 percent interest rate. In contrast, savings were close to optimal in a parallel treatment with a 1 percent interest rate. This difference was consistent with the presence of a compound-interest bias that has been documented elsewhere (Levy and Tasoff, 2016).

without tax deferral.⁴ In contrast to this conventional wisdom, subjects who were exogenously assigned to the Traditional treatment performed marginally better by under-consuming qualitatively less during retirement than those in the Roth account treatment. We hypothesize that this difference could be due to an aversion towards paying present taxes, which may augment savings in the tax-deferred treatment relative to the Roth treatment. Additionally, “future neglect,” the complement of present bias, could explain a failure to fully anticipate the negative effects of required tax payments for post-retirement withdrawals from a tax-deferred account but not from a Roth account. However, we cannot discern statistically significant differences between the accounts with respect to optimal lifecycle savings and spending patterns.

Comparing Roth and Traditional accounts in the laboratory allowed us to circumvent several constraints present in observational data comparisons of users of both types of accounts. First, Roth accounts have only been around since 2006 for 401(k) retirement accounts. This implies that the oldest of these accounts are at most 18 years old—much less than a full working age life cycle of 40 years or more.⁵ Second, individuals who opt into Roth accounts are likely different from those who do not, which introduces selection bias. For example, it is well known that many individuals do not open a retirement account unless accounts are established as a default (Benartzi and Thaler, 2007). This suggests that there might be a tendency for most individuals to stick with the default traditional account and that those who switch to Roth accounts are, on average, more cognizant of saving for retirement. Indeed, expert advice also tends to favor Roth accounts, which suggests that individuals who adopted Roth accounts may have been those who sought expert advice, which is likely a different demographic from those who did not seek advice. Third, mechanically, Roth accounts are disproportionately beneficial for individuals with an increasing income trajectory and are therefore preferred for some income groups relative to others⁶, aggravating the selection bias and making it tricky to tease out the causal effect of Roth tax incentives on savings. Given the limited data availability and selection biases associated with Roth accounts, observational studies of actual retirement decisions can be

⁴ Of course, this advice is generic and may be geared toward people who benefit from locking in lower tax rates at earlier stages.

⁵ Roth accounts first came into effect for Individual Retirement Accounts in 1998 making the longest feasible duration for these accounts 25 years to date.

⁶ Roth accounts allow savers to “lock in” low tax rates at the time that the income is saved. In contrast, traditional accounts force savers to pay taxes when withdrawing the money. These withdrawal tax rates may fall into a higher tax bracket than the individual belonged to when saving the money.

complemented with a laboratory experiment. Moreover, an experiment can assess the gender gap in retirement savings. A laboratory environment allows for identical work and income profiles to be implemented exogenously for both male and female subjects. Any gaps that emerge between men and women in the experiment may then be attributed to innate or acquired gender-related characteristics.

Our study contributes to multiple strands of literature. At a high level, we contribute to research on behaviorally-motivated policies aimed at increasing savings as summarized in Bernatzi and Thaler (2007). Within this field, we contribute to research about the intensive margin of retirement savings. In particular, we aim to shed light on the implications of specific plan characteristics—in our case the timing of tax payments—for savings and consumption behavior. Existing empirical evidence on whether tax-prepaid Roth accounts induce higher or lower savings than traditional retirement accounts has been mixed.

In a natural experiment, Beshears et al. (2017) exploit quasi-random variation around the rollout of Roth plans at different firms. When employees are given the choice between either savings vehicle, contribution rates are no different than when employees only have the option to save via tax-deferred accounts. These identical contribution rates translate into consuming less before retirement and saving more under the Roth treatment. The authors identify employee confusion or neglect of the tax properties of Roth accounts as a potential reason for this behavior. On the other hand, a lab-in-the-field experiment by Beshears et al. (2020) results in higher savings rates in accounts with commitment features (such as early-withdrawal penalties that are present in Traditional accounts). The latter finding is also in line with the seminal lab-in-the-field experiment by Ashraf et al. (2006) on the role of commitment savings products in increasing short-term savings of low-income households in the Philippines.

In the laboratory, a recent growing body of work has started to shed light on the consequences of tax incentives in retirement savings as well as the behavioral mechanisms. On the extensive margin, Cuccia, Doxey, and Stinson (2022) document that attitudes and preferences govern choices between tax-deferred and tax-prepaid retirement plans more so than awareness of the economic benefits. On the intensive margin, Blaufus and Milde (2020) find that tax-deferred accounts lower savings relative to prepaid accounts in a lab experiment that is relatively more framed and incorporates a real-effort task and filing tax returns, for example. Their paper focuses on savings during the work period and does not include consumption

decisions during retirement. Additional treatments suggest that information, experience, and government-matching frames reduce the differences they find. As posited in the field experiment by Beshears et al. (2017), Blaufus, Milde, and Spaeth (2023) find that myopia and confirmation bias may explain reactions to different tax deduction rates and a lack of reactions to different income taxes. Recent work by Duffy and Li (2022) is the most closely related to our experiment as they investigate the impact of adding access to a tax-deferred savings account in a life cycle consumption experiment. Where our study probes the differential impact of two retirement savings vehicles, their study probes the impact of introducing a retirement savings vehicle with tax benefits. The rich experiment by Duffy and Li (2022) focuses on painting a realistic picture of retirement decisions, which include multiple treatments with varying account combinations, marginal tax rates, death probabilities, and retirement calculators, among others. On average, they find that tax-deferred accounts increase wealth during retirement.⁷ These results are somewhat in line with our suggestive findings that tax-deferred accounts are associated with higher savings, although we do not find a statistically significant effect.⁸

We add to the existing work by abstracting away from many real-world features and framings and adding early withdrawals. This allows us to isolate any behavioral biases that go along with suboptimal consumption and tax deferral, which is the main difference between tax-prepaid and tax-deferred accounts.

The rest of the paper is organized as follows. Section 2 provides a theoretical framework for Traditional and Roth retirement accounts in a finite lifecycle consumption model; section 3 outlines the experimental procedures and design; section 4 discusses the results; and section 5 concludes.

2. Optimal Consumption and Savings Profiles

The experimental setup and associated theory are closely aligned with that used in our previous retirement study (Bohr et al. 2019). A subject is put into an environment with an incentive to maximize utility over T periods by choosing how much to convert into take-home

⁷ While their tax-deferred accounts track theoretical predictions quite closely, their other treatments feature overconsumption early on and underconsumption in retirement.

⁸ Additionally, low power may prevent us from finding statistically significant results in our experiment.

pay (consumption) and how much to save each period. There is no initial wealth endowment but a pre-tax income of y_t is received during periods 1 through $R-1$. Income is reduced to 0 in the remaining “retirement” periods R through T . Hence, any cash conversions during retirement will be based on prior savings balances.

There is a flat income tax rate, τ , and the two treatments affect how the subjects are exposed to this tax. The Roth (RA) treatment requires saving from after-tax income, but then no taxes will be paid during retirement. In the tax-deferred (TDA) treatment, subjects do not pay income taxes on the portion of their income that is saved. Instead, they must pay income taxes on the amount saved when it is withdrawn (including on the compounded interest gains). Moreover, any savings withdrawn before retirement are subject to an additional penalty (at a rate π).

Formally, the maximization of take-home earnings amounts can be modeled as choosing consumption c_t and additional savings, s_t , for each period, $t \in \{1, \dots, T\}$:

$$\max_{\{c_t, s_t\}} \sum_{t=1}^T u(c_t)$$

where $u(c_t)$ is increasing in its argument and globally concave. The total stock of savings allocated to retirement savings in period t will be denoted by b_t . This savings stock earns interest at a fixed rate r that is received at the end of every period. Therefore b_t will be the sum of the beginning of period savings $(1+r)b_{t-1}$ and the portion of a person’s income diverted to new savings s_t . Going into debt is prohibited, so the stock of savings must be weakly positive in every period. By choosing the amount of current income to save, s_t , the subject implicitly chooses the total stock of savings b_t that will earn interest going into the next period.

In our simple setup, current income, y_t , can only be used for current consumption, savings, and tax payments, denoted by X_t , as shown in the top line of the budget constraint list that follows:

$$\begin{aligned} y_t &= c_t + s_t + X_t, \\ b_t &= s_t + (1+r)b_{t-1}, \\ b_0 &= b_T = 0, b_t \geq 0, \text{ and } s_t \leq y_t. \end{aligned}$$

Due to the differential tax treatment, the budget constraint differs across treatments. The treatment-specific tax variable, X_t , for the Roth treatment is simply the tax paid on all income:

$$X_t = \tau y_t.$$

For the tax-deferred treatment, tax payments are determined by:

$$X_t = \tau (y_t - s_t) - \pi s_t \mathbb{1}[s_t < 0, t < R]$$

where π is the penalty rate imposed on pre-retirement withdrawals (before period R) when the additional savings amount, s_t , is negative as specified by the indicator function. Therefore, the penalty payment, $-\pi s_t$, is positive when savings are negative. In this case the penalty is added to the tax payment to determine the tax burden X_t for this period.

Given the flat income process and parameterizations used in the experiment, it is never optimal to withdraw savings before retirement. Optimality for both treatments requires the equality of consecutive marginal utilities over time, adjusted for interest earnings:

$$(1) \quad u'(c_t) = (1 + r)u'(c_{t+1})$$

By spending an additional dollar today, a person misses out on spending that dollar tomorrow plus the interest it accumulated, everything else equal. A person would be indifferent between such adjustments across periods if the marginal utility of the extra dollar today is equal to the marginal value of the dollar plus the associated interest in the next period.

The differential tax treatments do not affect the optimal consumption path. In the RA treatment, the taxes do not affect any of the choice variables and do not show up in the optimality conditions. In the TDA treatment, the effect of the tax rate on savings choices is symmetric across periods and therefore cancels out. Intuitively, if a person saves a dollar more in period t , the after-tax marginal effect on current utility is $-(1 - \tau)u'(c_t)$. The dollar saved earns interest at rate r , which could be used to reduce savings and increase consumption in the following period. Consequently, the after-tax effect of the increased utility in the next period is $(1 + r)(1 - \tau)u'(c_{t+1})$. Equating these effects, the $(1 - \tau)$ terms cancel out, which produces the geometric decrease in marginal utilities in (1).⁹ The optimal consumption and savings paths are identical since the total after-tax income (barring any early withdrawals) is the same across

⁹ Equation (1) can be derived formally by using the budget constraints to express $u(c_t)$ in the utility sum as a function of both the beginning-of-period total cash $(1 + r)b_{t-1}$ and the end-of-period post-saving balance b_t that will earn interest. For the tax-deferred case, it can be shown that $u(c_t) = u((1 - \tau)y_t - (1 - \tau)b_t + (1 - \tau)(1 + r)b_{t-1})$. Note that b_t is present in consecutive utility terms $u(c_t)$ and $u(c_{t+1})$. Therefore, the derivative of the utility sum with respect to b_t will have two terms that can be equated to derive the Euler optimality condition in (1). The analogous equation for the Roth plan is essentially the same, but with no τ tax parameter associated with b_t and b_{t-1} terms.

treatments. Any differences that appear in the experimental data would therefore be due to behavioral biases.

We apply the same utility parameterization as in our previous study, which used a constant-relative-risk-aversion utility function, $u(c_t) = \frac{1}{1-\sigma} c_t^{1-\sigma}$ with the coefficient of relative risk aversion being $\sigma = \frac{1}{3}$. In this case, the ratio of marginal utilities is the ratio of consumptions raised to the power of $1/3$, so the optimal consumption amounts in (1) follow a simple geometric path: $c_t = (1+r)^3 c_{t-1}$. The sequence of optimal consumption levels can be calculated numerically by specifying initial consumption, c_1 , and the geometrically-increasing sequence such that the remaining cash in the final forced-conversion period T just equals $(1+r)^3 c_{T-1}$. This optimal path is the same for both treatments.

3. Experiment Procedures

The experiments were conducted at a public university in the United States. Subjects were recruited and brought to a lab in person. Although each person's earnings were independent of others' decisions, we chose to conduct the sessions in small groups. Each "session" lasted about an hour and a half and began with a reading of the instructions and a short practice sequence with one period mimicking work life and one period of retirement.¹⁰

The main experiment consisted of 19 periods in which subjects received incomes (periods 1-14) or not (periods 15-19). In each period, a subject chose how much of the available balance of lab dollars to convert into take-home pay (with diminishing returns) and how much to deposit into or withdraw from their retirement savings account. Interest was paid at a rate of 5 percent on savings at the end of each period, which would remain in the savings account.¹¹ In the final period 19, any remaining cash and interest payments were automatically converted into take-

¹⁰ "Work" is loosely defined in our context. Subjects do not engage in real-effort tasks, and the income is stationary. Nonetheless, it mimics the period in life where one receives income—usually from working.

¹¹ One issue that has come up in the experimental literature on consumption and savings is how to induce present-value considerations in a dynamic setting. One common method has been to use a random stopping mechanism. There is, however, some evidence that subjects do not understand or accept the independence of random stopping events, or that volatile beliefs might induce "consumption binges" when a subject has a hunch that the process is about to end (Hey and Dardanoni, 1988, and Noussair and Matheny, 2000).

home pay using the same formula that was implemented in all prior periods. The pre-announced exogenous income was \$125 in the first 14 periods, which was reduced to \$0 afterward.

The income tax rate of 20 percent for all periods was pre-announced and stationary. In the Roth treatment, this tax was to be paid on all income at the point when it was received; no tax was paid on withdrawals from savings both before and during retirement. In contrast, Traditional treatment, money deposited in the savings account was not taxed. However, any withdrawals were taxed at the 20 percent rate both before and after retirement. Early withdrawals prior to retirement were permitted in this tax-deferred treatment, but such withdrawals were subject to a 10 percent penalty payment in addition to the 20 percent income tax.

Subjects used a drop-down menu to select a consumption expenditure (“lab dollars” and the corresponding increase in final earnings (“take home dollars”). After submission, a confirmation page explained the tax and savings consequences of the selected conversion amount, and subjects could either confirm or return to the submission page to rechoose. Each market period lasted about 4 minutes and ended when all decisions were submitted and confirmed. The cash conversion was implemented with a utility function $u(c) = c^{2/3}$, which, as noted, yields a geometric optimal consumption path $c_{t+1} = (1+r)^3 c_t$. This function was adjusted with additive and multiplicative constants: $u(c) = 0.04c^{2/3} - 0.2$, where the 20-cent deduction was intended to highlight the suboptimality of consuming small amounts. The 0.04 multiplier was selected to set an appropriate earnings level. This function was implicitly provided by the dropdown conversion menu that showed the earnings implications of a long list of available conversion amounts. This menu was supplemented with a table showing both total and marginal increments in take-home pay for each additional \$100 in lab earnings that could be converted. Lab dollars that were not converted were “saved” and carried over from one period into the next, earning interest at a rate of 5 percent at the end of each period.

The experiment was run with the web-based *VeconLab* software, which generated instructions reproduced in the Appendix. Instructions were read out loud while participants viewed the instructions pages on their screens.¹² Each session began with a two-period practice sequence, in which participants could become familiar with the software, conversion, and tax environment (Roth or Traditional) to be used in the main experiment. An income of \$125 was

¹² The experiment was conducted with the “Base Cash Earnings on Consumption” option for the Leveraged Asset Market program listed on the Finance/Macro Menu of the website: <http://veconlab.econ.virginia.edu/admin.html>

received in the first practice period; there was no income in the second, after which any remaining cash was converted into take-home pay. This practice task was intended to clarify procedures without giving subjects a chance to live their economic lives twice. After the practice sequence, a summary instructions page reviewed the payoff procedures, exogenous income receipts, and final payoff procedures. The main experiment with 18 periods of consumption decisions followed, after which participants were forced to convert the remaining lab cash into take-home earnings in the final 19th period. In total, 100 subjects were recruited from a mixed-gender student pool.¹³

All conversion earnings from the practice and the research experiment were paid in US dollars at the end of the experiment after the subjects had completed a short 7-question survey. In addition to basic demographics (college major and binary self-reported gender), this survey asked subjects for a self-reported, unincentivized measure of patience on a 10-point Likert scale.¹⁴ Also, there was a short risk assessment done at the beginning, with earnings and results not released until after the final period. The risk tolerance measure was done with a 12-box version of the BRET (bomb) risk assessment (Crosetto and Filippin, 2013). Subjects were presented with an ordered list of 12 boxes from which they could “extract” a dollar, but one of the boxes contained a hidden “ink bomb” that would negate all earnings if it was encountered. Subjects would decide which boxes to mark, knowing they would later receive a dollar for each box marked, provided they had not marked the box with the bomb. Results for the bomb task were not released until after the final round 19 of the experiment. These earnings were paid in US dollars (not scaled down) and added to other tasks’ earnings. Final earnings (for all rounds, including practice rounds) ranged between \$20 and \$35, including a \$10 show-up payment.

4. Results

The sudden income drop in retirement periods results in an optimal savings stock that rises during work periods and falls during retirement. Figure 1 shows the optimal savings

¹³ One subject was dropped from the study since they clearly misunderstood the instructions and consumed \$0 in all 18 choice periods, which resulted in very low total earnings due to the diminishing marginal value structure of cash conversions. The remaining 100 subjects were equally divided between Roth and TDA treatments.

¹⁴ The question used was: “In comparison to others, are you a person who is generally willing to give up something today in order to benefit in the future?” on a Likert scale from 1 (completely unwilling) to 10 (absolutely willing) (Falk, et al., 2022).

balances by period for the two treatments (dashed and dot-dashed lines) and the actual savings balances averaged over the 50 subjects in each treatment (connected data markers). However, in stark contrast to optimal savings, observed savings stocks are only about two-thirds of optimal levels for each treatment, which supports our first result:

Result 1: Subjects under-save relative to optimal levels in both treatments.

Support: The perspective provided by the data averages in Figure 1 is also apparent in analyzing individual savings patterns. We compare the savings available after the end of the final work period 14 with the optimal levels. Only 12 of the 50 subjects in the Traditional treatment had sufficient savings by the end of period 14 for optimal consumption during retirement. Similarly, only 11 out of 50 subjects in the Roth treatment had sufficient savings by the end of period 14.

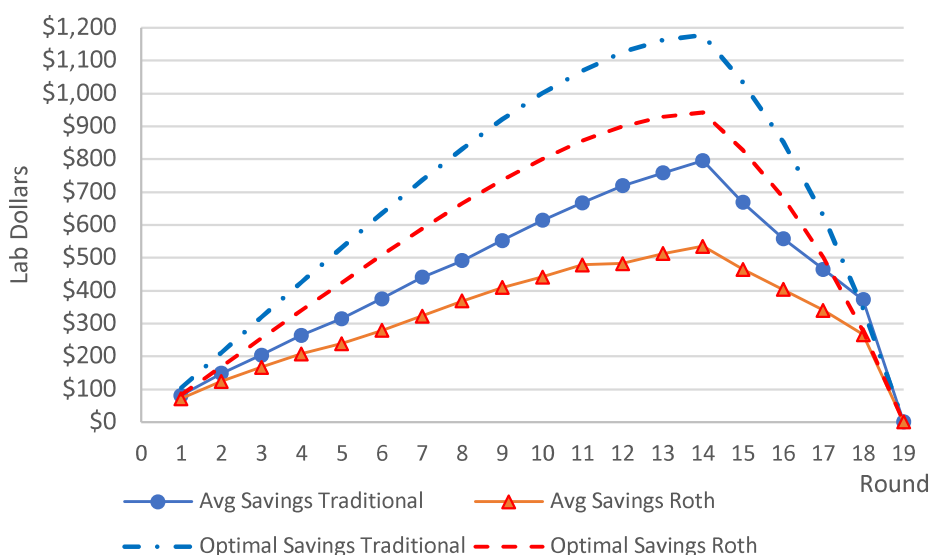


Figure 1. Optimal and Average Retirement Savings Balances by Treatment: Dot-dashed and dashed lines denote optimal savings paths for TDA and RA treatments, respectively. Solid lines with circular and triangular markers denote average after-tax savings paths for TDA and RA treatments, respectively.

While it might look like savings are significantly higher in the Traditional treatment, the visual comparison in Figure 1 is obscured by the fact that the tax must be paid on withdrawals from the tax-deferred retirement account. This adjustment is made in Figure 2, which shows after-tax savings balances for each treatment. This downward adjustment of the TDA savings line reflects the 20 percent income tax that individuals pay when withdrawing their savings balances after retirement.

Discussion: Recall that the optimal after-tax savings path is the same for both treatments, as shown by the dashed line in Figure 2. The observed after-tax savings amounts for the two treatments are roughly comparable and only slightly higher for the Traditional treatment. The dotted line shows the optimal after-tax savings trajectory for a 2.5 percent interest rate, which is half of what was used in the experiment. The observation that savings are near optimal levels for the *wrong* interest rate can be interpreted as indirect evidence that subjects underestimate the power of interest compounding.

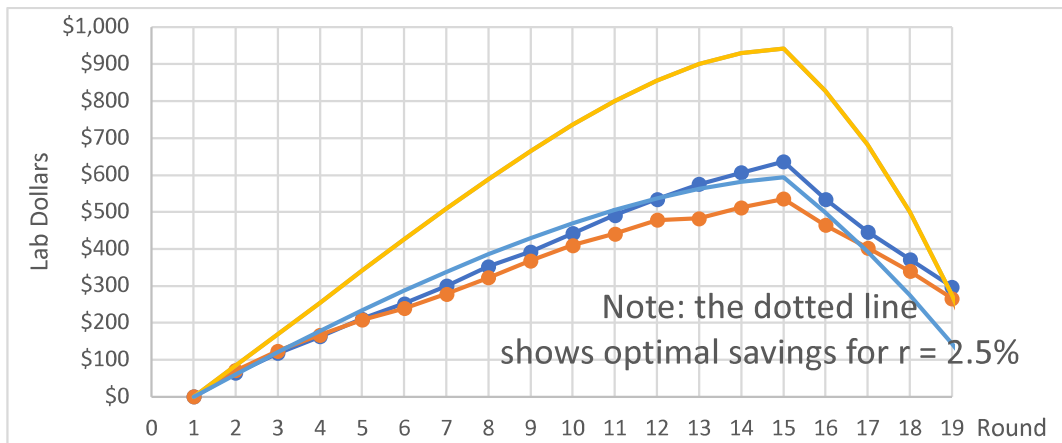


Figure 2. Optimal and Observed Average After-Tax Savings Balances by Treatment: Dashed and dot-dashed lines denote optimal after-tax savings paths. Solid lines with circular and triangular markers denote average after-tax savings paths for TDA and RA treatments, respectively. The dotted line denotes the optimal after-tax savings path if the interest rate had been set to 2.5% interest rate instead of 5%.

The suboptimal savings rates shown in Figures 1 and 2 are the result of overconsumption during working periods, as shown by the average consumption rates by treatment for the first 14 periods on the left side of Figure 3.

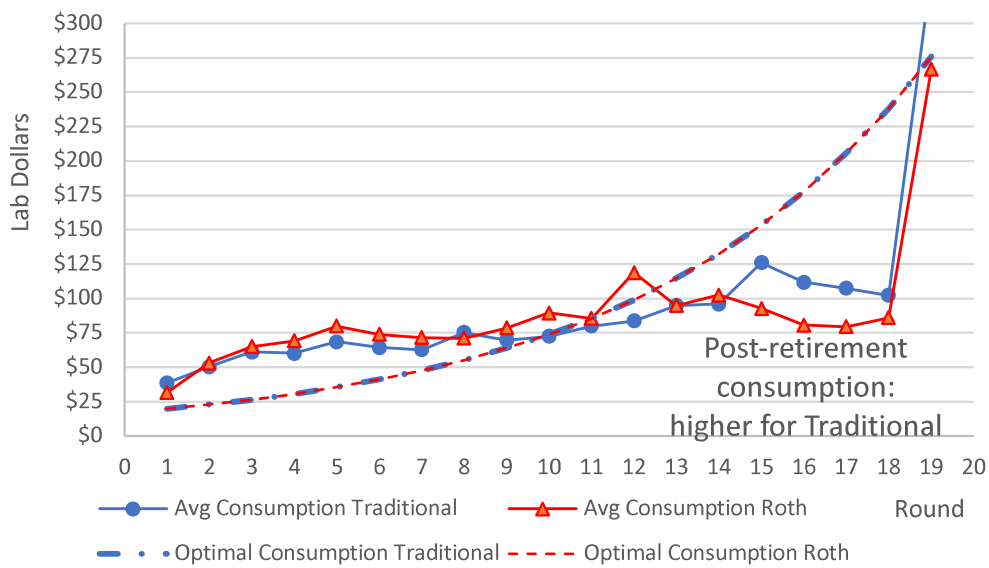


Figure 3. Optimal and Observed Average Consumption by Treatment: Dashed and dot-dashed lines denote optimal consumption paths. Solid lines with circular and triangular markers denote average consumption paths for TDA and RA treatments, respectively.

Note that consumption in early periods is slightly higher in the Roth treatment, with the result that consumption during retirement is, on average, lower, as shown on the right side of the figure. There is no evidence to support the notion that the Roth retirement account improves the suboptimal retirement saving levels in this simulated life cycle experiment, which leads to our second result.

Result 2: Marginally higher savings and consumption rates with Traditional (tax-deferred) retirement accounts than with Roth retirement accounts on average.

Support: As shown on the right side of Figure 3, retirement consumption is about one-third lower for the Roth treatment. However, this difference is not statistically significant at conventional levels ($p = 0.13$).

Table 1. Regression Results for Retirement Consumption and After-tax Savings at Retirement
(Standard Errors in Parentheses)

	(1)	(2)	(3)	(4)
VARIABLES	Avg. Consumption, Periods 15-18		After-tax Saving, Period 15	
Roth Account	- 29.32 (18.22)	-26.97 (16.72)	-105.2 (99.49)	-106.7 (97.52)
Female		-65.99*** (15.08)		-249.4** (103.6)
Risk Aversion		23.88 (13.62)		152.1*** (36.11)
Patience		8.28** (3.25)		75.05*** (21.10)
Constant	114.1*** (11.76)	83.49*** (21.91)	640.5*** (67.19)	194.2 (186.9)
Observations	100	100	100	100
R ²	0.024	0.175	0.012	0.162

Note: Significance levels are indicated: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, based on standard errors clustered at the session level; each session usually had about 9 participants. We used independent observations of measures of individual saving and consumption, which are controlled for gender, risk aversion, and patience measures in regressions shown. The dependent variable in the two left columns is the subject average retirement consumption (over periods 15-18), which is regressed on the tax treatment in column 1, with the covariates gender, risk-aversion, and patience added in column 2. Similarly, the dependent variable for the regression in column 3 is accumulated after-tax savings at the beginning of retirement in period 15, and controls for this regression are added in column 4.

Discussion: This finding is noteworthy and goes against the conventional wisdom that Roth accounts are superior, all else equal. If anything, there is suggestive evidence that subjects assigned to traditional savings accounts fare slightly better than those assigned to Roth accounts. To conclude, we find no evidence of anchoring biases and nominal targeting which tend to enhance subjects' savings contributions in the Roth treatment.

Result 3: Savings at retirement are significantly lower for female subjects and are higher for subjects who are more risk averse or who self-report a higher level of patience.

Support: On average, across treatments, women have \$249 lower savings in lab dollars when they enter retirement. The magnitude is remarkable given that the average savings level for both genders, shown in Figure 2, is about \$500 lab dollars. Next, a one standard-deviation increase in

risk aversion (0.47) amounts to \$71.39 more lab dollars saved across both treatments on average. However, this has no statistically discernible consequences for retirement consumption. Similarly, a one standard-deviation increase in the level of self-reported patience (1.67) is associated with \$125.42 more lab dollars in savings. The higher savings translate into higher consumption levels during retirement; the effect is quantitatively more pronounced for risk-averse individuals but statistically more significant for patient individuals.

Discussion: Our lab design allows us to abstract away from institutional or selection confounds and to attribute observed behavioral differences to demographic variables. The difference in savings and consumption patterns by gender, absent of institutional concerns such as due to discrimination- or maternity-leave-induced lower labor income, is striking. Education (such as more mathematics or finance classes) or internalization of gender norms could still be driving the observed gender differences and warrant further investigation. The observed effects concerning risk aversion and patience are intuitive: increases in both traits are associated with higher pre-retirement savings, reflecting pre-cautionary and delayed-gratification savings behavior.

The Traditional treatment may induce higher savings and, therefore, higher levels of retirement consumption as it discourages subjects from pre-retirement withdrawals, which incur a 10 percent penalty on top of the 20 percent income tax. However, as reported in our final result, we do not strong evidence in favor of this hypothesis.

Result 4: Early withdrawals were relatively infrequent and low in both treatments.

Support: Figure 4 compares average withdrawals for each treatment. As would be optimal in both treatments, zero withdrawal amounts are the norm, especially in early periods. These zero withdrawals cause the average of pre-retirement withdrawals to be close to zero for both treatments. Even in the Roth treatment that does not penalize early withdrawals, there are only small withdrawal amounts before period 15, the onset of retirement.¹⁵ Since early withdrawals are not impeded in the Roth treatment and yet were relatively minor, we conclude that it is unlikely that the early-withdrawal penalty is underlying the marginally higher savings in the Traditional treatment.

¹⁵ Specifically, early withdrawals were observed in 21 percent of the cases in pre-retirement periods with the Roth treatment, and in 12 percent of the cases in pre-retirement periods with the Traditional treatment. In the relatively few periods where early withdrawals did occur, the median withdrawal amounts were \$20 with Roth and about \$29 with Traditional.

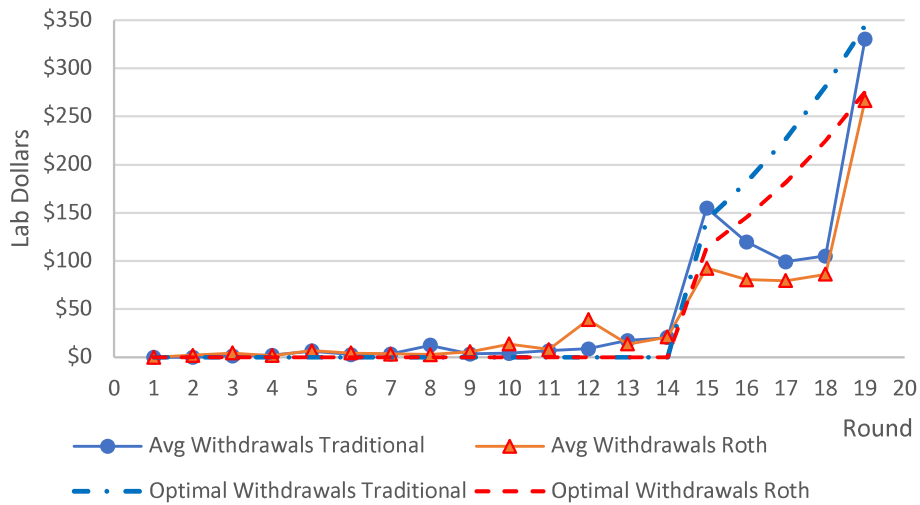


Figure 4. Life Cycle Withdrawal Amounts by Treatment: Dot-dashed and dashed lines denote optimal withdrawal paths for TDA and RA treatments, respectively. Solid lines with circular and triangular markers denote average withdrawal paths for TDA and RA treatments, respectively.

5. Conclusion

It is an active goal of public policy to encourage the average person to save more for retirement. In the US, this effort has resulted in the creation of tax-advantaged retirement savings accounts. The primary options are the traditional (tax-deferred) and the Roth (tax prepaid) retirement accounts. Due to selection effects and the limited duration of time that Roth accounts have been in existence, there is no clear evidence on whether one is more conducive for retirement savings than the other. Moreover, the variation in income trajectories of different people and how that interreacts with the tax treatments likely means that one is not strictly superior to the other. That is, unless there are behavioral factors that may lead one treatment to encourage more savings than the other.

In this paper, we reported the results of a life cycle consumption-savings experiment with treatments that only differed by the type of retirement account. Overall, subjects in both treatments accumulated less-than-optimal savings for retirement and therefore had deficient consumption levels during retirement. Despite this, we found that subjects who were using a tax-deferred account ended up saving more on average and, therefore, also consumed marginally more during retirement. However, the qualitative differences we observe are not statistically significant. One bias that might push behavior in this direction could be an aversion to paying

taxes, which might cause people to save more for retirement to defer paying taxes. We also observe heterogeneity in savings and consumption behavior associated with gender, level of risk aversion, and self-reported patience.

The null result of the insignificant Roth treatment variable surprised us because it is commonly thought that Roth accounts are superior to the traditional tax-deferred accounts. Additionally, the behavioral biases, usually discussed in the context of retirement savings, would also favor the Roth account. For instance, anchoring biases might suggest that individuals aim to hit a nominal target for the amount they save over time regardless of pre- or post-tax deductions. This would induce higher levels of retirement consumption for holders of Roth accounts since they would not have to pay taxes during retirement on that nominal income target.

Knowing that people generally do not save enough for retirement, one might suspect that deferring expenses such as taxes until retirement, as in the tax-deferred savings account, is harmful. In contrast, our experiment provides suggestive evidence that other biases must outweigh this consideration; if anything, individuals assigned to a tax-deferred treatment end up consuming marginally more during retirement. The tax-deferred account may be harnessing an aversion towards paying taxes in order to induce higher levels of savings. By saving for retirement subjects can defer the perceived “losses” they would otherwise have incurred from paying the taxes in that period. If this is the case, then a higher income tax rate may further increase the difference in retirement savings and consumption between the two treatments.

References

- Ashraf, N., Karlan, D. and Yin, W. 2006. Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. *The Quarterly Journal of Economics* 121, no. 2 (2006): 635-672.
- Benartzi, S., and Thaler, R. 2007. Heuristics and biases in retirement savings behavior. *Journal of Economic Perspectives* 21, 81-104.
- Beshears, J., Choi, J.J. Laibson, D. and Madrian, B.C. 2017. Does front-loading taxation increase savings? Evidence from Roth 401(k) introductions. *Journal of Public Economics* 151, 84-95.
- Beshears, J., Choi, J.J., Harris, C., Laibson, D. Madrian, B.C. and Sakong, J. 2020. Which early withdrawal penalty attracts the most deposits to a commitment savings account? *Journal of Public Economics* 183, 104-144.
- Blaufus, K. and Milde, M., 2021. Tax misperceptions and the effect of informational tax nudges on retirement savings. *Management Science*, 67(8), pp.5011-5031.
- Blaufus, K., Milde, M. and Spaeth, A., 2023. Does a Decrease in Pension Taxes Increase Retirement Savings? An Experimental Analysis. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.4372027>
- Bohr, C. E., Holt, C.A. and Schubert, A.V. 2019. Assisted savings for retirement: An experimental analysis, *European Economic Review* 119, 42-54.
- Bond, T. and Porell, F. 2020. Examining the nest egg: The sources of retirement income for older Americans. Report for the National Institute on Retirement Security, <https://www.nirsonline.org/wp-content/uploads/2020/01/Examining-the-Nest-Egg-Final-1.pdf>
- Crosetto, P., and Filippin, A., 2013. The ‘‘bomb’’ elicitation task, *Journal of Risk and Uncertainty* 47, 31-65.
- Cuccia, A.D., Doxey, M.M. and Stinson, S.R., 2022. The impact of tax incentive structure on taxpayers' retirement savings decisions. *The Journal of the American Taxation Association*, 44(1), pp.23-47.
- Duffy, J. and Li, Y., 2019. Lifecycle consumption under different income profiles: Evidence and theory, *Journal of Economic Dynamics and Control* 104, 74-94.
- Duffy, J. and Li, Y., 2023. Do tax deferred accounts improve lifecycle savings? Experimental evidence, Working Paper, UC Irvine.
- Falk, A., Becker, A. Dohmen, T., Huffman, D.B. and Sunde, U., 2022. The preference survey module: A validated instrument for measuring risk, time, and social preferences *Management Science*, <https://pubsonline.informs.org/doi/epdf/10.1287/mnsc.2022.4455>
- Hey, J.D., and Dardanoni, V., 1988 Optimal consumption under uncertainty: An experimental investigation. *Economic Journal* 98, 105-116. <https://doi.org/10.2307/2233308>
- Kleven, H., Landais, C. and Sogaard, J.E., 2019. Children and gender inequality: Evidence from Denmark, *American Economic Journal: Applied Economics* 11, 181-209.
- Levy, M., and Tasoff, J., 2016. Exponential growth bias and lifecycle consumption. *Journal of the European Economic Association* 14, 545-583. <https://doi.org/10.1111/jeea.12149>

Noussair, C., and Matheny, K., (2000) An experimental study of decisions in dynamic optimization problems. *Economic Theory* 19, 389–419.
<https://doi.org/10.1007/s001990050019>

Teachers Insurance and Annuity Association of America, 2022. *TIAA 2022 Financial Wellness Survey*. Retrieved March 2023, from
https://www.tiaa.org/public/pdf/2022_financial_wellness_survey_final_results.pdf.

Instructions Appendix: (see separate file for online publication)

Link to Data and Instructions File: <https://github.com/Veconlab/IRA-Savings-Experiment>