The Retirement Security Project

Saving Incentives for Low- and Middle-Income Families: Evidence from a Field Experiment with H&R Block

Nº 2005-5

The Retirement Security Project

RSP

We thank H&R Block for organizing and funding this experiment. This experiment was organized by the Outreach & Business Development group at H&R Block, headed by Bernie Wilson. In addition to the key impetus he provided, we also gratefully acknowledge the help and contributions of other H&R Block employees, including Mary Beth Granger, Scott McBride, John McDonald, Andy Olson, Mitch Powers, Arijit Roy, Doris Seyl, John Thompson, Ken White, and Sabrina Wiewel, as well as the district and office managers and the 600 tax professionals who implemented this experiment in St. Louis. We also thank Marc Ferguson (Onesta Software), Yvette Ruiz (YMR), John Marinovich (Group 1), and Laura Bos and Bo Harmon (The Retirement Security Project) for their assistance. We gratefully acknowledge support from the Pew Charitable Trusts, the Sloan Foundation, and NSF Grant SES-0134946. The views expressed in this article are those of the authors alone and do not necessarily reflect the views of H&R Block or any of the institutions with which the authors are affiliated.

Common sense reforms, real world results

www.retirementsecurityproject.org

The Retirement Security Project is supported by The Pew Charitable Trusts in partnership with Georgetown University's Public Policy Institute and the Brookings Institution.

Saving Incentives for Low- and Middle-Income Families: Evidence from a Field Experiment with H&R Block

By Esther Duflo, William Gale, Jeffrey Liebman, Peter Orszag and Emmanuel Saez

Abstract

This paper analyzes the effects of a large randomized field experiment, carried out with H&R Block, offering matching incentives for IRA contributions at the time of tax preparation. About 15,000 H&R Block clients, in 60 offices in predominantly low- and middle-income neighborhoods in St. Louis, were randomly offered a 20 percent match on IRA contributions, a 50 percent match, or no match (the control group). The evaluation generates two broad findings. First, higher match rates significantly raise IRA participation and contributions. Take-up rates were 3 percent for the control group, 10 percent in the 20 percent match group, and 17 percent in the 50 percent match group. Average IRA contributions (excluding the match) for the 20 percent and 50 percent match groups were 4 and 8 times higher than in the control group, respectively. Second, several additional findings are inconsistent with the full-information, rational-saver model, and suggest instead that professional tax assistance, information provision, and ease of saving can play important roles in encouraging IRA contributions among low- and middle-income families. For example, we find more modest effects on take-up and amounts contributed from the existing Saver's Credit, which provides an effective match for retirement saving contributions through the tax code; we suspect that the differences may reflect the complexity of the Saver's Credit as enacted, and the way in which its effective match is presented. Taken together, our results suggest that the combination of a clear and understandable match for saving, easily accessible savings vehicles, the opportunity to use part of an income tax refund to save, and professional assistance could generate a significant increase in retirement saving participation and contributions, even among moderate- and low-income households.

Introduction

A significant share of low- and middle-income American families appears to be saving little, either for retirement or for any other purpose. Families with income below $40,000 have low rates of coverage under employer-provided pensions, are extremely unlikely to contribute to Individual Retirement Arrangements (IRAs), and in 2001 had median net financial wealth outside of retirement accounts of just $2,200.[1]

Researchers and policy makers alike have long considered ways to raise saving among these families.[2] The conventional approach to subsidizing saving through 401(k) plans and traditional IRAs provides tax deductions for contributions and tax deferral on account earnings. This approach has not enticed low- and middle-income families to contribute very much to retirement accounts, in part because the value of tax preferences is modest since these families face low marginal income tax rates.

In contrast, the provision of matching contributions could potentially represent a much more promising way to bolster incentives for low- and middle-income households to save. Matching contributions can be independent of the individual's marginal tax rate, and thus provide a significant incentive for saving even to people in low marginal tax brackets. Little is known, however, about whether such matching contributions would in fact induce a meaningful fraction of low- and middle-income families to save for retirement.

The Saver's Credit offers one example of matching contributions provided by the Federal government. Enacted in 2001, the credit provides an income tax reduction of up to 50 percent of funds contributed to a 401(k) or IRA by qualified filers. The credit as enacted may not accomplish its full potential, however, in part due to its low income thresholds for eligibility, in part because it is not refundable, and in part, perhaps, because it has not been effectively advertised or explained to tax filers.[3]

Matching contributions are also present in many employer-sponsored 401(k) plans. Previous studies have found mixed evidence of the effects of match rates on 401(k) participation and contributions.[4] A significant concern in these studies, furthermore, is that the match rates offered by firms in 401(k) plans may not be independent of worker characteristics at the firm, which makes it difficult to disentangle the independent effect of matching rates on contributions. Another concern is that the results apply mainly to relatively affluent households in the presence of automatic payroll deduction and the other workplace features associated with 401(k) plans, which could affect contributions. Therefore, the results may not be directly applicable to low- and middle-income families or to policy interventions that occur outside the workplace.

This paper reports evidence from the first large-scale, randomized field experiment ever conducted regarding the effects of matching rates on the willingness of low- and middle-income families to contribute to IRAs. By randomizing the matching rate across tax filers, we are able to identify not only the impact of the presence of a match, but also how variations in the matching rate affect both take-up and contribution levels. Unlike the Saver's Credit, the match provided in this experiment is available in full to (virtually) all tax filers, has a simple structure, is explained to potential account holders in a straightforward manner, and is deposited directly into an IRA rather than reducing income tax liability. Unlike 401(k) matching contributions, the match provided is independent of individuals' characteristics and of the workplace environment.

The field experiment was conducted in conjunction with H&R Block, the largest tax preparer in the country, and co-organized with the Outreach & Business Development Group at H&R Block. H&R Block paid the direct costs associated with implementing the experiment, including for the matching contributions, advertising materials, and the training of tax professionals. The experiment was run in 60 H&R Block tax preparation offices in the St. Louis metro area from March 5th to April 5th, 2005. The experiment was built around the Express IRA (X-IRA) product offered by H&R Block, which allows clients to make IRA contributions at the time of tax preparation and to fund those contributions with part or all of their federal income tax refunds or from other sources. In effect, the X-IRA allows the client to "split" their anticipated refund between contributions to a retirement account and other uses.

Each client preparing a tax return in one of the 60 offices during the period was randomly assigned to one of three match rates for X-IRA contributions: zero (the control group), 20 percent, or 50 percent. Contributions were matched up to $1,000, a limit that applied separately for each spouse for married tax filers. Each client, including those in the control

group, received a waiver of the $15 set-up fee for opening an X-IRA. The minimum X-IRA contribution is $300.

The evaluation generates two overarching sets of conclusions. First, match rates can have large effects on IRA participation and contributions. Take-up rates are 3 percent, 10 percent, and 17 percent, respectively, for the control group, the 20-percent match group and the 50-percent match group. Conditional on take-up, average contribution levels (excluding the match) are $860, $1,280, and $1,310, respectively. With the match included, average IRA deposits were $860, $1480, and $1,870 respectively. Taking participation rates into account and including the matching contributions, average IRA deposits with the 20-percent and 50-percent matches were 5 and 11 times higher than with no match. The effects are particularly large for married tax filers, with take-up rates of 3 percent, 13 percent, and 25 percent in the three groups. These effects are substantially larger than those found in the context of 401(k) matches.

The second broad conclusion is that, along a number of dimensions, observed taxpayer behavior is not consistent with simple economic models that posit fully informed and fully rational consumers. First, take-up rates varied significantly across different tax preparers. Second, the IRA take-up response to the experimental match was significantly larger than taxpayers' responses to equivalent changes in incentives embodied in the existing Saver's Credit. Third, take-up rates were significantly less than 100 percent even in the presence of significant matches, which suggests that clients were not attempting to game the system by contributing to the IRA and receiving the matching contribution, and then cashing out the accounts (with minimal penalties) as soon as possible afterwards. These findings are consistent with a view that information and simplicity affect saving choices.

Taken together, the two sets of results suggest that a combination of financial incentives, tax preparer assistance, the opportunity to use part of an income tax refund to save, and easily accessible saving vehicles could generate substantial increases in both the efficacy of federal tax incentives and the willingness of households to contribute to retirement saving accounts.

The remainder of the paper is organized as follows. Section 2 describes the experimental design and data. Section 3 analyzes the effects of the experiment. Section 4 compares the results to evidence from the Saver's Credit. Section 5 concludes.

Experimental Design, Implementation, and Data

Tax Preparation at H&R Block

H&R Block is the largest tax preparer in the United States. It currently prepares over 16 million individual income tax returns (about 12 percent of all individual U.S. tax returns and about 23 percent of tax returns prepared by paid professionals) in 11,000 offices distributed throughout the country. H&R employs over 100,000 professional tax return preparers during the tax season from January to mid April.[5]

Clients come to an H&R Block office with the documents, such as W-2 forms, necessary for tax return preparation. They sit at a desk with a tax professional while the return is prepared. Both the tax professional and the client sit in front of a computer running the H&R Block Tax Preparation Software (TPS). The desk design is such that the clients typically can read and follow the computer screens as the tax professional works through the tax preparation process.

TPS consists of a series of screens corresponding to the various steps in tax return preparation. At each screen, the tax professional asks questions or inputs information from the forms brought by the client. By default, there is a natural ordering of the screens called the "F10 sequence" (the name reflects the fact that tax professionals move from one screen to the next in the sequence using the F10 key). Using the computer mouse, tax professionals can also skip ahead or come back to any screen throughout the tax preparation process. Many screens are not accessed directly through the F10 sequence and appear only when they are relevant (following an input or a particular situation of the tax filer) or if the tax professional uses the mouse to get to this particular screen directly.

The Express IRA Product

Since tax season 2001, H&R Block has offered a product, called the express IRA (X-IRA), to allow tax filers to make IRA contributions at the time of tax preparation. X-IRAs can be funded with a personal check, a one-time direct deposit, or the automatic allocation of part or all of their federal tax refund to the X-IRA (without the need to borrow against the refund). Most contributors fund their X-IRA with part of their refund.

In effect, through the X-IRA, H&R Block offers all of its clients the "split refund" option that has been advocated as a way to significantly increase savings (see Beverly, Schneider, and Tufano, 2004). The logic for this option is that filers are more likely to contribute to an IRA when they have money available in the form of a refund. Indeed, for many low- and middle-income families with children, the federal tax refund is the largest single payment they receive during the year. Federal tax refunds represent over 10 percent of Adjusted Gross Income (AGI) for tax filers with AGI below $50,000 in the 60 offices where this experiment was run. Furthermore, opening or contributing to an IRA at the time of tax return preparation is easy and convenient. In contrast, many low- and middle-income families lack strong ties to banks and other IRA providers, so that opening an IRA outside of tax preparation could be costly in terms of the psychological and information requirements, as well as the transaction fees charged.

The custodian of the X-IRA account is H&R Block Financial Advisors, a wholly-owned subsidiary of H&R Block. The only investment option available within the X-IRA product is a FDIC-insured money market bank account offered by The Reserve Funds. The minimum contribution is $300.[6] There is a $15 set-up fee for opening an account or for re-contributing to an existing account using a portion of the refund at the time of tax preparation. There is also an annual account maintenance fee of $10, which is waived for accounts with balances over $1,000 or for those participating in an automatic savings plan.[7] Once the balance in those accounts reaches $1,000, individuals may transfer assets from their X-IRA into a new HRBFA IRA, which offers access to a menu of investment options.

The X-IRA screen is normally not part of the default F10 sequence, unless the option has been deliberately selected by the tax professional at the beginning of the tax preparation session. Therefore, the X-IRA will not be an option unless the tax professional chooses to raise the topic or clients asks about it.[8] Clients who are offered the X-IRA and decline may change their mind later during the tax preparation process (for example if they obtain a larger tax refund than they expected). It is easy for tax professionals to come back to the X-IRA screen at any point during the tax preparation and change the inputs. In practice, substantial heterogeneity exists among tax professionals in how effectively they offer the X-IRA product.

The X-IRA product is well advertised to tax professionals during pre-tax-season training and to clients in the offices. Many offices display posters that describe the product and encourage clients to think about their future and save. H&R Block opened about 450,000 X-IRA accounts during tax seasons 2001 to 2004. In tax season 2004, the X-IRA take-up rate (including opening and re-contributions) was approximately 1.4 percent.

The existence of the X-IRA at H&R Block provides an ideal tool to use to design and build our experiment. H&R Block is the only large tax preparer offering such a product to such a broad low- and middle-income client base.

Experiment Design

The experiment was run in 60 H&R Block offices in the St. Louis metro area in Missouri and Illinois. The experiment ran for 32 days, from March 5th to April 5th, 2005. During the first week from March 5 to March 11, the experiment ran in 45 offices. It was extended to 60 offices starting March 12. Any client coming to prepare his/her taxes at one of the relevant offices during the experiment was randomly assigned to one of three groups: the control group, the 20 percent match group, or the 50 percent match group. Assignment was based on the last two digits of the Social Security number of the primary filer on the tax return. The probability of assignment was 34 percent, 32 percent, and 34 percent respectively.[9]

Each group received a waiver of the $15 X-IRA set-up fee. The control group received no match. A second group received a 20 percent match on X-IRA contributions up to $1,000 (for a maximum match of $200). A third group received a 50 percent match on any X-IRA contribution up to $1,000 (for a maximum match of $500). X-IRA accounts are individually owned, so the same offer was extended to each spouse for married tax filers filing jointly. For example, if a married couple filing jointly was assigned to the 20 percent match group, both the husband and the wife could simultaneously open X-IRA accounts, contribute to the limit, and receive $200 each in matching contributions, for a total family match of $400.

During the experiment, the Tax Preparation Software (TPS) of H&R Block was modified in the 60 experimental offices along two important dimensions. First, the X-IRA screen was made part of

the default F10 sequence. As a result, any tax professional following the regular F10 flow would hit the screen offering the X-IRA product at some point during tax preparation. However, through active use of the mouse, it was possible for the tax professional to avoid the X-IRA screen altogether and never propose the randomized X-IRA offer to the client. TPS recorded which tax returns went through the X-IRA screen.[10] From now on, we refer to those returns as having received an "offer." Obviously, if a filer was not offered the X-IRA, the filer could not contribute and benefit from the offer. The important point is that the tax professional's decision to offer the X-IRA was made before the treatment assignment was revealed, since the tax professionals were not aware of the algorithm for assigning clients to each of the randomized experimental groups and the match rate was not revealed by the software until after reaching this screen.[11]

Second, when the X-IRA screen was activated, a special pop-up window would automatically appear presenting the offers corresponding to the treatment group in which the client fell, and asking whether the client wanted to contribute to an X-IRA. The tax professionals were instructed to read the text that appeared in the pop-up window, but field observations suggested that they did not systematically adhere to the text and instead presented the offer in a way that they felt would be intuitive for the client. In particular, while the window proposed $500 as an example of a contribution level and presented the associated match, some tax professionals undoubtedly presented other contribution levels and the associated matches. To move on, the tax professional was forced to click "yes" or "no" on the pop-up window. A "yes" click led the tax professional to the standard X-IRA screen where contributions and corresponding matches would be manually inputted by the tax professional. At the standard X-IRA contribution screen, the client could still decide not to contribute. A "no" click on the pop-up window would skip the X-IRA screen and move the tax preparation to the next step. At any point, however, the tax professional could come back to the X-IRA screen and trigger the (same) pop-up offer window and modify the initial X-IRA contribution choice.

Matches to the X-IRA totaled roughly a half million dollars and were funded by H&R Block and deposited into the X-IRA accounts of clients on April 15th. The match was deposited only if the X-IRA contributions had not been withdrawn between tax preparation and April 15th. After the match was deposited on April 15th, the account became a regular X-IRA account and tax filers were completely free to withdraw their contribution and/or the match subject to standard tax rules. In addition to any regular income tax that may be owed, traditional IRAs face a 10 percent penalty for non-qualified withdrawals. Roth IRAs face a 10 percent penalty on interest (and no penalty on the principal) for non-qualified withdrawals.

About 600 tax professionals prepared tax returns in the 60 experimental offices. Those tax professionals were trained in groups during simple one-hour training sessions on March 1st and 2nd. The training sessions were moderated in St. Louis by Scott McBride (implementation manager at H&R Block) and a member of the research team (Emmanuel Saez).[12] The training described the general goal of the experiment, and explained all the details of the TPS modification. Tax professionals were instructed to follow the F10 sequence and systematically go through the X-IRA screen and to always carefully read and explain to the tax filer the offer on the pop-up window. Tax professionals were also instructed to offer a net worth survey at the end of the tax preparation. Finally, tax professionals were given the opportunity to ask questions about the experiment. Importantly, all tax professionals were instructed to present the offers as opportunities for retirement savings and explicitly told not to tell tax filers that they could "game the system" by making contributions and then rescinding the accounts immediately after the match

was deposited on April 15th.[13] All tax professionals were provided with abundant documentation about the experiment: goal, implementation details, and common questions and their answers. Over 90 percent of tax professionals working in the experimental offices attended the training sessions. A number of standard X-IRA refresher training sessions were also offered to tax professionals on a voluntary basis in early March. H&R Block local management and district and office managers made a special effort to ensure that everyone working in the experiment offices was aware of the experiment and complying with the experimental protocol.

Although training sessions were uniform, tax professionals form a very heterogeneous group, in terms of skills, motivations, and interest in the X-IRA product. Therefore, in practice, there was significant heterogeneity in how tax professionals would present the offer to the clients and how convincingly they made the case that those offers were a beneficial opportunity.

Data

H&R Block supplied the data for this analysis. H&R Block treats client privacy with the utmost care and significant safeguards were involved. Throughout the experiment, every measure was undertaken to ensure that clients understood the implications of their participation. All of the data were stripped of any individual identifiers (such as name, Social Security number, phone numbers, addresses, office names, etc.) before being provided to our research team. The data include selected tax return information collected during the tax preparation process, information on X-IRA contributions, selected additional information collected by H&R Block during the tax preparation process that does not get recorded on tax forms, and information about the tax professional who prepared each sample member's tax return. These data are part of the extensive data that H&R Block maintains for its operations and is of extremely high quality. The data we use consist of all clients in the 60 offices during the full tax season from January 1st to April 15th. (We use the non-experimental data from before March 5th to compare the results of our experiment to the Saver's Credit.)

In our analysis, we exclude filers with less than $300 in earned income, since they are not eligible to make the minimum X-IRA contribution. All other filers may open an X-IRA.[14] We often divide filers into "married" (specifically, married filing jointly) and "non married" (including singles, heads of households, and married filing separately) categories.

Descriptive Statistics

Table 1 displays the means of several variables for each of the three experimental groups in the first three columns and tests for whether the means are statistically different across groups in the next three columns. The t-statistics of the difference is reported whenever it is significant at the 5 percent level (i.e., when the t-statistic is above 1.96). Because of randomization, 5 percent of those differences should be statistically significant.

Only two variables, AGI and fraction married, are significantly different (at the 5 percent level) between the treatment and control groups. The overall AGI differences are the result of the difference in the proportion of married tax filers between the treatment and control groups (married taxpayers have higher incomes on average than non-married taxpayers). There are no treatment-control AGI differences among married tax filers or among other tax filers. The proportion of married tax filers is slightly higher in the matching groups (38.6 percent and 37.4 percent) than in the control group (35.5 percent). Although modest, those differences are statistically significant.[15]

The average AGI in our sample ($43,000 in the control group) is similar to the average AGI in the U.S. population. A little less than half of the sample owns a

Table 1: Descriptive Statistics

	Mean (Standard error)			Significance of the difference in means: t-stats		
	No match	20% match	50% match	20% match vs no match	50% match vs 20% match	50% match vs no match
	(1)	(2)	(3)	(4)	(5)	(6)
Fraction of returns offered X-IRA	0.75 (0.007)	0.75 (0.007)	0.76 (0.007)			
This tax pro made no offer during experiment	0.005 (0.001)	0.004 (0.001)	0.005 (0.001)			
Fraction of returns with X-IRA by tax pro 01/01-03/04	0.023 (0.000)	0.022 (0.000)	0.023 (0.000)			
Ajusted Gross income (plus Trad. IRA contributions)	42909 (523)	44789 (561)	44617 (575)	2.45	2.20	
Ajusted Gross income, married taxpayers (filing jointly)	67904 (981)	68487 (943)	69775 (996)			
Ajusted Gross income, non married taxpayers	29168 (438)	29920 (532)	29579 (540)			
Fraction married, filling jointly	0.355 (0.007)	0.386 (0.007)	0.374 (0.007)	3.09	1.98	
Overpayment amount	1584 (26)	1570 (26)	1596 (26)			
Fraction with overpayment above $500	0.67 (0.007)	0.66 (0.007)	0.66 (0.007)			
# Children dependent at home	0.53 (0.013)	0.54 (0.013)	0.53 (0.013)			
Fraction homeowners	0.45 (0.007)	0.42 (0.007)	0.44 (0.007)			
Fraction filed with H&R Block in the preceding tax year	0.63 (0.007)	0.63 (0.007)	0.64 (0.007)			
Fraction EITC recipients	0.17 (0.006)	0.16 (0.005)	0.17 (0.005)			
Fraction with savings account	0.62 (0.007)	0.64 (0.007)	0.63 (0.007)			
# of observations	4739	4617	4894			
# of observations, married	1681	1780	1831			
# of observations, others	3058	2837	3063			

This table displays descriptive statistics for the three experimental groups (columns 1, 2, 3) and reports t-stats when differences across groups are significant (columns 4 and 5). The sample is all experimental returns eligible to contribute to an X-IRA.

Row 1 reports the fraction of offers defined as returns for which the X-IRA offer pop-up screen was reached. This is a necessary step to contribute to an X-IRA.

Row 2 reports the average over tax pros of the fraction of returns for which an X-IRA was taken up among returns completed by the tax pro before the experiment started. The average is taken over all individual experimental returns in each group.

Adjusted gross income is always inclusive of all traditional IRA contributions (as X-IRA contributions are affected by the experiment). Homeownership data are collected during tax preparation and not reported on the tax return.

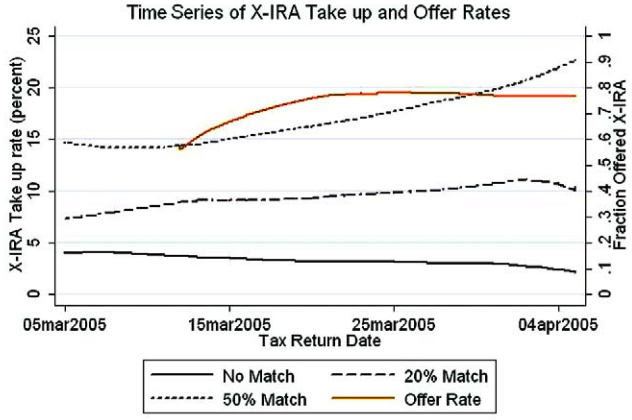
house. Two-thirds have a federal refund larger than $500, which would generally allow them to fund a $300 X-IRA out of their refund even if they owe taxes at the state level.

The experimental sample differs from the large majority of H&R Block clients who file early in the season. In St. Louis, almost 79,000 H&R block clients eligible for an X-IRA filed before March 5, while close to 15,000 filed between March 5 and April 5. Early filers are poorer (their AGI averages $28,418), and less likely to be married (only 21 percent are married filing jointly). Early filers are presumably more impatient to receive their tax refund and more likely to be liquidity constrained than the population subject to this experiment. Our results may therefore be more representative of what would happen in an average American population than among the more traditional H&R Block client base of early filers.

The first row in Table 1, labeled "offer rate," shows the fraction of instances where the pop-up screen appeared and the tax professional had to record the client's answer in order to be able to continue. In each of the three groups, the tax professionals presented the filer with the relevant offer in 75 percent of the cases.[16] The second row shows that the 25 percent of cases where no offer was made does *not* reflect a few tax professionals who systematically avoided the pop up screen. Indeed, almost all professionals displayed the screen at some point.

Figure 0 shows that learning occurred during the experiment. The figure shows how the offer rate varied by day: it increased from 55 percent on March 12 to just below 80 percent around March 22, and stayed fairly constant after that date. The increase early in the experiment partly reflects H&R Block's effort to get professionals to systematically navigate through the screens. The fact that the offer rate does not vary significantly across treatment groups suggests that bypassing the offer was not based on the tax professional's guess as to which match the tax filer would receive. In other words, the tax professionals do not appear to have figured out the randomization algorithm. Because a tax professional's decision regarding whether or not to navigate to the offer screen could have been based on an assessment of how likely the client was to accept the offer, we do not analyze take-up conditional on receiving an offer. Instead, our analysis includes all tax filers regardless of whether they were presented with the offer screen (i.e., we present "intent-to-treat" estimates), and we allocate tax filers to the various experimental groups based on the last two digits of their Social Security number (which is how the randomization algorithm worked).

Figure 0
Time Series of X-IRA Take up and Offer Rates



Results

Take-up Rates and Contributions

Table 2 presents the main results of the experiment. The table shows the X-IRA take-up rates and contributions in the three experimental groups. For comparison, X-IRA take-up and contributions in the period before the beginning of the experiment are also displayed. The last three columns report differences across the experimental groups. Panel A reports statistics for all

Table 2: Effects of the experiment on X-IRA behavior

	Mean (Standard error)				Difference (standard errors)		
	Pre-experiment	No match	20% match	50% match	20% match vs no match	50% match vs 20% match	50% match vs no match
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PANEL A: ALL TAXPAYERS							
Opened an X-IRA (%)	2.15	3.31	9.64	17.00	6.33	7.36	13.69
	(0.048)	(0.260)	(0.434)	(0.537)	(0.503)	(0.695)	(0.603)
Amount contributed ($)	8.8	28.4	123.4	222.8	95.1	99.4	194.4
(unconditional)	(0.27)	(3.74)	(7.72)	(9.36)	(8.51)	(12.21)	(10.20)
Amount contributed ($)	381	856	1281	1310	424	30	454
(positive contributions only)	(8.2)	(91)	(56)	(36)	(108)	(64)	(92)
Amount contributed, inclusive of match		28.4	143.0	318.2	114.6	175.2	289.8
(unconditional)		(3.74)	(8.68)	(12.65)	(9.37)	(15.52)	(13.38)
Amount contributed, inclusive of match		856	1484	1872	627	388	1015
(positive contributions only)		(91)	(60)	(45)	(115)	(76)	(111)
PANEL B: MARRIED TAXPAYERS ONLY							
Opened an X-IRA (%)	1.69	2.91	13.48	24.90	10.57	11.42	21.99
	(0.100)	(0.410)	(0.810)	(1.011)	(0.924)	(1.300)	(1.126)
Amount contributed ($)	9.6	45.2	232.3	433.9	187.1	201.6	388.8
(unconditional)	(0.80)	(9.32)	(18.14)	(22.33)	(20.75)	(28.86)	(24.96)
Amount contributed ($)	470	1549	1723	1742	173	19	193
(positive contributions only)	(30.2)	(236)	(86)	(55)	(218)	(98)	(185)
Amount contributed, inclusive of match		45.2	268.5	616.3	223.4	347.8	571.2
(unconditional)		(9.32)	(20.36)	(29.94)	(22.83)	(36.40)	(32.49)
Amount contributed, inclusive of match		1549	1992	2475	442	483	925
(positive contributions only)		(236)	(92)	(66)	(231)	(113)	(216)
Spouse opened an XIRA as well (%)	0.04	0.83	8.09	15.95	7.26	7.86	15.11
	(0.015)	(0.222)	(0.646)	(0.856)	(0.699)	(1.077)	(0.918)
Amount contributed by spouse ($)	4.4	11.1	96.6	176.5	85.5	79.9	165.5
	(1.83)	(3.74)	(9.29)	(10.95)	(10.22)	(14.39)	(11.97)
PANEL C: NON MARRIED TAXPAYERS ONLY							
Opened an X-IRA (%)	2.25	3.53	7.23	12.28	3.69	5.05	8.74
	(0.054)	(0.334)	(0.486)	(0.593)	(0.583)	(0.774)	(0.681)
Amount contributed ($)	8.7	19.1	55.1	96.6	36.0	41.5	77.4
(unconditional)	(0.28)	(2.70)	(4.90)	(5.61)	(5.49)	(7.50)	(6.23)
Amount contributed ($)	364	542	763	787	221	24	245
(positive contributions only)	(7.9)	(57)	(45)	(25)	(74)	(48)	(56)
Amount contributed, inclusive of match		19.1	64.3	140.0	45.1	75.7	120.9
(unconditional)		(2.70)	(5.53)	(7.81)	(6.02)	(9.70)	(8.27)
Amount contributed, inclusive of match		542	889	1140	347	251	599
(positive contributions only)		(57)	(48)	(32)	(78)	(56)	(67)

This table reports X-IRA statistics for the pre-experiment tax returns (before March 5) in column 1, and the three experimental groups (no match in column 2, 20% match in column 3, 50% match in column 4) for all taxpayers with at least $300 in earned income (IRA eligible). Columns 5, 6, and 7 report the differences across experiment groups.

Unconditional amounts report averages including zeros (those with no X-IRA contributions).

Married taxpayers defined as married filing jointly. Non-married taxpayers defined as all others (singles, head of households, and married filing separately).

tax filers, while the two remaining panels break down the results according to marital status.

The first row of Panel A shows that matching incentives have strong effects on X-IRA take-up rates. The take up rates are 3.3 percent, 9.6 percent and 17.0 percent respectively, in the no-match, 20-percent match, and 50-percent match groups. The take-up rate in the no-match group is quite low, and only one percentage point higher than take-up before the experiment started (2 percent). This suggests that the combined option to have set-up fees waived and to "split" one's refund — to allocate part of it to retirement saving — is not in itself sufficient to induce a large fraction of tax filers to contribute to retirement accounts; the results for the other groups suggest that split refunds may be quite effective if combined with stronger matches for saving.[17] The no-match results also suggest that the propensity to save in the population that was the subject of the experiment is actually not much larger than that of the population of early filers (even with fee waivers), which ameliorates to some extent the concern that the experimental group is not representative of the average H&R Block client.

The results in the first row show that the match rate has a large effect on take up, and the differences between the matches and between a match and no-match are strongly statistically significant (t-statistics above 10). The second row shows that the effects on overall amounts contributed are even stronger, since contributions, conditional on take-up, are higher in the two groups receiving the match than in the control group (though they are quite similar across the two match groups). Not surprisingly, when the matching contributions are added, the differences grow even larger: unconditional X-IRA deposits (including the match) for the three groups are respectively $28, $143, and $318. IRA deposits are 5 times larger with a 20 percent match than with no match, and 11 times larger with a 50 percent match than with no match.

The significant difference in take-up rates between the 20 percent and 50 percent match groups indicates that the level of the match rate, and not only the fact that there is one, influences take up. The elasticity of the take-up rate with respect to (1+match rate), defined as the change in the natural log of the take-up rate divided by the change in the natural log of (1+match rate) are respectively 5.87 for the 20 percent match rate and 4 for the 50 percent match rate.[18] Although the elasticity is higher for low match rates, it remains high even for larger match rates. The fact that tax filers respond to the level of the match suggests that the response is a reasoned calculation, and not simply a case where the match attracted their attention to the existence of the X-IRA (Bernheim, 2003).

The effects of the match on take-up rates for married tax returns are about twice as large as they are for others.[19] Take-up rates are 2.9 percent, 13.5 percent, and 25 percent for married filers in the three groups, respectively, compared to 3.5 percent, 7.2 percent, and 12.3 percent for other filers. Interestingly, married and other tax filers have about the same likelihood of opening an X-IRA without a match. The effects on amounts contributed are also correspondingly much larger for married tax filers than for others. For married filers, the average unconditional contributions inclusive of match are respectively $45, $269 and $616 in the three groups, and are roughly four times as large as contributions for other taxpayers in the 20 percent and 50 percent match groups. The bottom of Panel B shows that take-up rates for a second X-IRA for married filers are 1 percent, 8 percent, and 16 percent for the three groups. This shows that match rates also generate a large response on the spousal (or secondary) X-IRA contribution decision for couples.[20]

The cumulative distribution of X-IRA contributions (excluding the matching amounts) is shown in Figure 1A for married filers and Figure 1B for other filers. Each graph reports the fraction of

Figure 1A

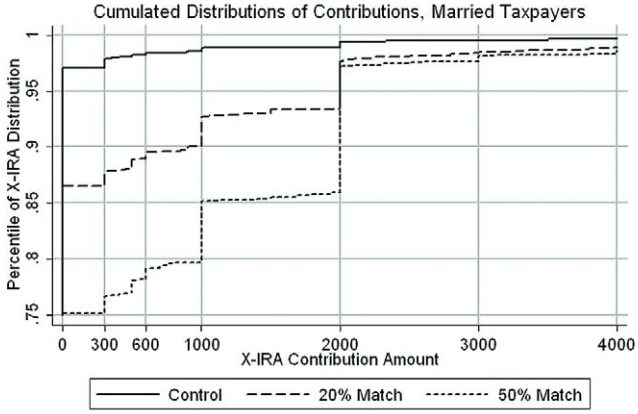
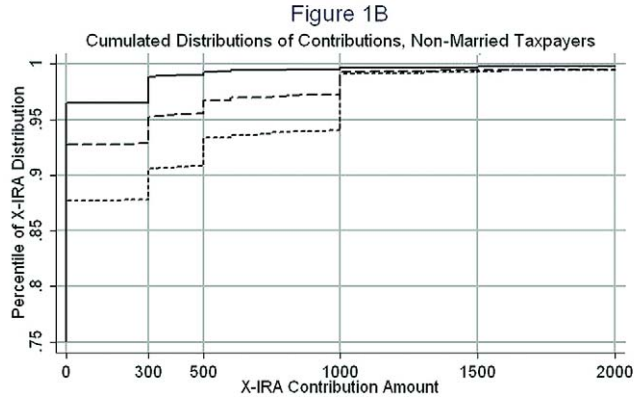


Figure 1B



tax filers contributing less than a given amount for any level of contributions (on the x-axis). The figures show that, for any contribution level, the fraction of filers contributing at least that amount is always higher for the 50 percent match group than for the 20 percent match group, and higher for the 20 percent match group than for the no-match group. The figures also show that contributions tend to be clustered at discrete amounts such as $300 (the minimum contribution) and $1,000 (the maximum contribution eligible for the match). For married tax filers, bunching also occurs at $600 and $2,000 because both spouses may contribute. There is also some bunching at $500 for the match groups (and much less so for the no-match group). This bunching may reflect the fact that the example given on the X-IRA offer pop-up window for the match group was based on a $500 contribution example, which could have created a focal point. We plan to explore these types of focal effects further in future work.

Above the match cap ($2,000 for married, $1,000 for single), the match provides only an income effect, with no substitution effect. Standard theory would therefore predict that the match should reduce contributions above the match cap, and we should see the distributions crossing at that point. This prediction is not what we observe in the data: the fraction of filers contributing in excess of $1,000 (for non married) and $2,000 (for married) remains higher for the match groups than for the control group.

Figure 0 (page 9) shows the take-up rates in the three experimental groups as a function of the date at which the return was filed. The take up of the X-IRA with no match did not increase over time; if anything, it may have been slightly declining. The take-up rate in both the 20 percent group and the 50 percent group increased over time, and increased much faster in the 50 percent group. While the increase in take-up between March 5 and March 22 could potentially be due to the corresponding increase in the offer rate, the take-up rate continued to increase at the same speed afterwards, suggesting that, in practice, the increase in offer rates is likely not responsible for the increase in take-up rates. It could be that tax professionals became better at presenting the product, or that clients who filed later in the season were more savvy or less cash constrained. The effect does not, however, appear to be due to a different propensity to save, since the take up in the no-match group did not rise over time.

Who took up the match?

The welfare and policy implications of savings incentives depend in part on who is most likely to use them. Of particular interest is the response of poorer households, where much of the policy concern is focused. Table 3 (next page) presents the average X-IRA take-up rates

Table 3: Effects on X-IRA behavior by income quartiles

	Mean (Standard error)			Difference (standard errors)		
	No match	20% match	50% match	20% match vs no match	50% match vs 20% match	50% match vs no match
	(1)	(2)	(3)	(4)	(5)	(6)
A. MARRIED TAXPAYERS						
A.1 Bottom quartile ($0-$35,000)						
Opened an XIRA	3.46	11.44	17.21	7.98	5.77	13.75
	(1.026)	(1.822)	(2.059)	(2.071)	(2.772)	(2.343)
Amount contributed	26.1	174.4	235.3	148.3	60.9	209.2
(unconditional)	(12.18)	(41.09)	(32.45)	(42.18)	(51.89)	(35.44)
# observations	318	306	337			
A.2 Second quartile ($35,000-$58,300)						
Opened an XIRA	4.32	14.66	25.39	10.34	10.73	21.08
	(0.996)	(1.656)	(2.066)	(1.978)	(2.641)	(2.342)
Amount contributed	109.2	249.3	343.2	140.2	93.9	234.1
(unconditional)	(32.78)	(38.82)	(32.56)	(51.31)	(50.79)	(46.24)
# observations	417	457	445			
A.3 Third quartile ($58,300-$88,200)						
Opened an XIRA	2.60	12.61	25.13	10.01	12.52	22.53
	(0.686)	(1.381)	(1.814)	(1.577)	(2.276)	(1.985)
Amount contributed	28.2	215.8	424.2	187.6	208.4	396.0
(unconditional)	(9.75)	(29.97)	(38.77)	(32.45)	(48.93)	(41.08)
# observations	539	579	573			
A.4 Fourth quartile ($88,200+)						
Opened an XIRA	1.47	14.84	29.62	13.37	14.78	28.15
	(0.598)	(1.701)	(2.095)	(1.856)	(2.726)	(2.332)
Amount contributed	17.0	276.8	671.0	259.9	394.2	654.1
(unconditional)	(9.25)	(37.42)	(59.69)	(39.83)	(71.84)	(65.12)
# observations	407	438	476			
B. NON MARRIED TAXPAYERS						
B.1 Bottom quartile ($0-$8,500)						
Opened an XIRA	1.90	2.82	4.67	0.92	1.85	2.77
	(0.596)	(0.743)	(0.895)	(0.947)	(1.179)	(1.088)
Amount contributed	8.0	18.3	32.3	10.3	13.9	24.3
(unconditional)	(2.80)	(6.15)	(6.82)	(6.64)	(9.27)	(7.53)
# observations	526	497	557			
B.2 Second quartile ($8,500-$18,400)						
Opened an XIRA	2.93	7.39	8.67	4.46	1.29	5.75
	(0.662)	(1.111)	(1.140)	(1.253)	(1.597)	(1.299)
Amount contributed	12.9	40.6	50.9	27.7	10.3	38.0
(unconditional)	(3.29)	(7.42)	(7.67)	(7.73)	(10.71)	(8.17)
# observations	649	555	611			
B.3 Third quartile ($18,400-$32,500)						
Opened an XIRA	5.03	6.57	14.14	1.54	7.58	9.12
	(0.785)	(0.917)	(1.239)	(1.202)	(1.562)	(1.473)
Amount contributed	34.6	49.1	101.5	14.6	52.3	66.9
(unconditional)	(8.66)	(9.46)	(11.03)	(12.80)	(14.64)	(14.06)
# observations	776	731	792			
B.4 Fourth quartile ($32,500+)						
Opened an XIRA	3.61	9.68	16.77	6.06	7.10	13.16
	(0.561)	(0.911)	(1.125)	(1.059)	(1.456)	(1.256)
Amount contributed	17.2	84.2	150.9	67.0	66.6	133.6
(unconditional)	(3.61)	(10.30)	(12.03)	(10.71)	(15.90)	(12.54)
# observations	1107	1054	1103			

Groups are defined by quartiles of AGI inclusive of IRA contributions. Those quartiles are defined from the universe of all individual US tax returns (with at least $300 in earned income) from the Statistics of Income microfiles for year 2000 (adjusted for income growth to 2004) for married filing jointly and all other filers separately.

and contributions by marital status and by quartiles of AGI (to which we have added the traditional IRA contributions, since such contributions are subtracted when computing AGI). Each tax unit was assigned to its position in the full distribution of tax filers in the United States estimated from the Internal Revenue Service Statistics of Income file for the year 2000.[21]

For married filers, the effect of the match is significant in all income quartiles, although it does increase with income: Relative to the no-match group, the take-up rate was 14 percentage points higher in the 50 percent match group in the first quartile, 21 points in the second quartile, 23 points in the third quartile, and 28 points in the top quartile. For other filers, the effect increases faster with income, and is somewhat less striking in the lower income quartiles. Relative to the control group, the effect of the 50 percent match on X-IRA take-up is respectively 3, 6, 9, and 13 percentage points in the first through fourth quartiles. The effect of the 50 percent match rate remains significant even in the bottom quartile. The difference in take-up rates between the 20 percent match rate and the no-match group is not statistically significant in the bottom quartile, but the difference in contributions remains significant at the 10 percent level in this group.

Table 4 explores the effects of the match on two specific sub-samples: those

Table 4: Effects on X-IRA behavior by EITC status and savings account ownership

	Mean (Standard error)			Difference (standard errors)		
	No match	20% match	50% match	20% match vs no match	50% match vs 20% match	50% match vs no match
	(1)	(2)	(3)	(4)	(5)	(6)
A. EITC Status						
A.1 No EITC						
Opened an XIRA	2.89	9.82	17.64	6.94	7.82	14.75
	(0.268)	(0.478)	(0.598)	(0.546)	(0.770)	(0.663)
Amount contributed	26.9	135.6	249.8	108.6	114.2	222.9
(unconditional)	(4.23)	(9.01)	(11.03)	(9.93)	(14.32)	(11.98)
# observations	3914	3879	4070			
A.2 EITC Recipients						
Opened an XIRA	5.33	8.67	13.83	3.34	5.16	8.50
	(0.783)	(1.037)	(1.204)	(1.283)	(1.606)	(1.435)
Amount contributed	35.2	59.5	89.2	24.4	29.7	54.1
(unconditional)	(7.72)	(9.05)	(9.79)	(11.83)	(13.43)	(12.46)
# observations	825	738	824			
B. Savings Account Ownership						
B.1. No Savings Account						
Opened an XIRA	2.30	5.88	9.80	3.58	3.91	7.49
	(0.355)	(0.574)	(0.698)	(0.667)	(0.911)	(0.788)
Amount contributed	20.3	57.6	105.4	37.3	47.8	85.0
(unconditional)	(6.19)	(7.90)	(10.49)	(9.98)	(13.29)	(12.24)
# observations	1781	1683	1817			
B.2 Savings Account owners						
Opened an XIRA	3.92	11.79	21.25	7.87	9.46	17.33
	(0.357)	(0.596)	(0.738)	(0.693)	(0.953)	(0.830)
Amount contributed	33.2	161.2	292.1	128.0	131.0	258.9
(unconditional)	(4.69)	(11.21)	(13.38)	(12.12)	(17.54)	(14.40)
# observations	2958	2934	3077			

Groups are defined by EITC recipiency status and whether taxpayers own a savings account.

Table 5: OLS regression: determinants of X-IRA take up and treatment effects

	X-IRA take up rate (percentage)					
	Married taxpayers			Non-married taxpayers		
	(1)	(2)	(3)	(4)	(5)	(6)
20% match dummy	10.57	10.57	4.01	3.69	3.68	0.03
	(1.14)	(1.14)	(4.39)	(.69)	(.68)	(2.44)
50% match dummy	21.99	21.92	7.53	8.74	8.87	4.21
	(1.14)	(1.13)	(4.38)	(.67)	(.67)	(2.39)
Second AGI quartile		3.92	0.80		1.39	0.12
		(1.44)	(2.53)		(.91)	(1.54)
Third AGI quartile		1.82	-1.19		2.95	1.97
		(1.40)	(2.46)		(.89)	(1.54)
Fourth AGI quartile		3.92	-2.23		3.29	0.31
		(1.51)	(2.67)		(.90)	(1.56)
Number of Dependents		-0.21	-0.28		1.83	1.05
		(.44)	(.81)		(.42)	(.69)
Home ownership dummy		-2.96	-1.02		-0.76	1.09
		(1.38)	(2.42)		(.63)	(1.08)
Repeat customer dummy		0.31	1.01		1.35	1.19
		(1.06)	(1.92)		(.58)	(.99)
Positive refund dummy		4.72	1.88		5.13	3.71
		(1.04)	(1.84)		(.62)	(1.05)
Bank account dummy		4.88	1.28		3.91	1.90
		(1.12)	(1.99)		(.58)	(.99)
Second AGI quartile *			1.83			2.74
20% match			(3.55)			(2.24)
Third AGI quartile *			1.04			-0.32
20% match			(3.45)			(2.20)
Fourth AGI quartile *			4.53			3.80
20% match			(3.74)			(2.23)
Second AGI quartile *			7.25			1.00
50% match			(3.52)			(2.18)
Third AGI quartile *			7.63			3.19
50% match			(3.41)			(2.14)
Fourth AGI quartile *			12.99			5.06
50% match			(3.68)			(2.18)
Number of dependents *			0.96			0.85
20% match			(1.11)			(1.03)
Number of dependents *			-0.64			1.65
50% match			(1.10)			(1.01)
Home ownership *			-3.56			-0.41
20% match			(3.38)			(1.56)
Home ownership *			-2.27			-5.17
50% match			(3.39)			(1.54)
Repeat Customer *			-2.52			-0.69
20% match			(2.63)			(1.43)
Repeat Customer *			0.48			1.01
50% match			(2.64)			(1.40)
Positive refund *			5.08			1.61
20% match			(2.57)			(1.52)
Positive refund *			3.32			2.59
50% match			(2.54)			(1.49)
Savings account *			3.60			2.05
20% match			(2.73)			(1.43)
Savings account *			6.81			3.91
50% match			(2.77)			(1.40)

This table reports OLS coefficients of the regression of an X-IRA take up dummy (normalized to 100) on treatment groups dummies, covariates, and covariates interacted with treatment dummies.

The positive refund dummy represents taxpayers with a federal income tax refund larger than $500.

receiving the Earned Income Tax Credit (EITC) and those with a savings account. As Table 1 shows, about 17 percent of the sample overall receives the EITC and about 63 percent has a savings account. Both of these variables are correlated with income. Given the results in Table 3, it is therefore not surprising that Table 4 shows that the difference in take-up rates and in contributions across the match rates is more modest for EITC recipients than for non-EITC recipients, and is stronger for those with savings accounts than those without such accounts. For example, the take-up rate for the 50-percent match group is 8.5 percentage points higher than the control group for EITC recipients, and 7.5 percentage points higher for those not owning a savings account. By contrast, the differential take-up rate among non-EITC recipients and those owning a savings account is 14.8 percentage points and 17.3 percentage points respectively. Nonetheless, the increases for EITC recipients and those not owning a savings account still remain statistically significant and meaningful for both match rates.

Table 5 uses OLS regressions to summarize these patterns and compare the effect sizes across different groups. Column 1 for married filers — and column 4 for singles and heads of household — simply reproduces the results in Table 3: the coefficient of the "20 percent (50 percent) match dummy" is the difference between the take-up rate in the 20 percent (50 percent) match and that in the "no match" group. Columns 2 and 5 show that these results are not affected by controlling for income, number of dependents, home ownership status, whether or not the tax filer is an H&R Block repeat customer, and savings account ownership. This is not surprising, since the treatment was randomly assigned and the sample is large.

Columns 3 and 6 investigate whether the effect of the treatment varies according to these characteristics. The coefficient of 1.61 for the variable "second AGI quartile 20 percent match" for married tax filers tells us, for example, that the effect of the 20 percent match (relative to no match) increases take-up by 1.61 percentage points for tax filers in the second quartile, compared to tax filers in the first quartile (and the standard errors lead us to conclude that this difference is not significant). This table confirms the pattern shown in Table 3: the effects of the 50 percent match are significantly larger for higher income tax filers for both married and single tax filers. This income-related pattern is not systematically the case for the 20 percent match. The savings account dummy interacted with treatment groups dummies in Table 5 is statistically significant showing that savings account ownership has a positive effect on the response to matches (especially for the 50 percent match) even controlling for income.

Among the other variables included in this table (homeownership, number of dependents, being a repeat H&R block customer, repeat customer status refund amount), only the refund amount has some impact on X-IRA take up. Married tax filers are respectively 5 and 3.5 percentage points more likely to take up an X-IRA in the 20 percent and 50 percent match groups if their federal refund is larger than $500 than if it is smaller than $500. But these effects are not very large, and even in the group of people who have a refund larger than $500, the take-up is far below 100 percent. This suggests that credit constraints alone cannot explain why less than 100 percent of H&R clients took advantage of the match rate (thereby failing to capture the potential "free lunch").

Tax Professional Effects

Field observations suggested that the way a tax professional presents the program to the tax filer is likely to have an important effect on the client's response. Tax professionals clearly exerted different amounts of effort and enthusiasm in

presenting the product to the client and their understanding of the program was also heterogeneous.

One dimension of heterogeneity is the tax professional's prior experience with X-IRAs. To investigate this issue, we first divided tax professionals according to their experience with X-IRAs before the beginning of the experiment (Table 6). In each panel (married or single), the first sub-panel displays the average take-up rates and amounts contributed for clients whose tax professional had less than the median fraction of X-IRA take-up before the beginning of the experiment (1.5 percent), whom we label "low experience tax professionals";[22] the second panel shows the results for clients whose tax professionals had more take-up than the median, whom we label "high experience tax professionals." Unsurprisingly, for both married and singles, the take-up in the "no match group" is higher among clients working with a "high experience tax professional," although the difference is small. More important, the effect of the match rate (20 percent or 50 percent) is larger for filers working with tax professionals who have relatively higher X-IRA take-up rates before: for married filers for example, the effect of a 20 percent match rate on take-up increases from 7 percent to 12 percent when moving from a "below median experience tax professional" to an "above median experience tax professional," and the effect of a 50 percent match rate increases from 18 percent to 25 percent.

Table 6: Effects on X-IRA by tax pro previous X-IRA experience

	Mean (Standard error)			Difference (standard errors)		
	No match	20% match	50% match	20% match vs no match	50% match vs 20% match	50% match vs no match
	(1)	(2)	(3)	(4)	(5)	(6)
A. MARRIED TAXPAYERS						
A.1 Below median						
Opened an X-IRA	2.56	9.90	20.23	7.34	10.33	17.66
	(0.506)	(0.940)	(1.238)	(1.079)	(1.564)	(1.376)
Amount contributed	23.0	161.3	324.8	138.3	163.4	301.8
(unconditional)	(6.16)	(19.81)	(22.57)	(21.05)	(30.12)	(24.19)
A.2 Above median						
Opened an X-IRA	3.43	15.85	28.29	12.43	12.43	24.86
	(0.561)	(1.088)	(1.329)	(1.251)	(1.721)	(1.490)
Amount contributed	63.1	269.6	493.0	206.5	223.4	429.9
(unconditional)	(14.54)	(24.42)	(30.88)	(28.94)	(39.45)	(35.16)
B. NON-MARRIED TAXPAYERS						
B.1 Below median						
Opened an X-IRA	2.20	4.26	8.96	2.06	4.70	6.76
	(0.344)	(0.495)	(0.668)	(0.594)	(0.844)	(0.752)
Amount contributed	10.3	33.0	72.0	22.7	39.0	61.7
(unconditional)	(2.32)	(4.91)	(6.28)	(5.28)	(8.08)	(6.71)
B.1 Above median						
Opened an X-IRA	4.53	9.04	13.70	4.50	4.66	9.16
	(0.486)	(0.697)	(0.807)	(0.840)	(1.073)	(0.940)
Amount contributed	25.3	68.0	103.8	42.8	35.8	78.6
(unconditional)	(4.00)	(7.02)	(7.44)	(7.93)	(10.26)	(8.43)

This table divides the sample by tax professionals according to the fraction of returns with an X-IRA among all the returns they completed before the experiment.

These results could reflect differences in the tax professional's attitudes and skills, but they may also reflect differences in the mix of clients working with different tax professionals. For example, clients who live in richer neighborhoods may be more likely to respond to the match rate (the discussion on the previous page underscores that take-up increased with income) and also to take up X-IRAs in the absence of any special program. This would induce a correlation between the X-IRA take-up rates across tax professionals in the offices in these areas before the experiment and take up of the incentives when the experiment started.

To investigate this further, we regress, in the first panel of Table 7, the take-up of X-IRAs on dummy variables for the two

Table 7: OLS Regressions: Tax pro effects on X-IRA take up

	Married		Non Married	
	All customers	New customers	All customers	New customers
	(1)	(2)	(3)	(4)
A. TAX PRO X-IRA CLIENT TAKE-UP RATES BEFORE MARCH 5, 2005				
20% match	6.50	6.71	2.44	2.18
	(3.43)	(5.70)	(1.97)	(2.72)
50% match	10.82	19.28	7.93	3.26
	(3.33)	(5.95)	(1.93)	(2.80)
Tax pro above median # of IRAs taken up	0.17	1.93	2.53	1.09
before March 5, 2005	(1.70)	(3.40)	(.99)	(1.33)
Tax pro above median	6.56	9.72	2.36	1.63
* 20% match	(2.25)	(4.40)	(1.36)	(1.82)
Tax pro above median	9.42	-0.26	2.82	3.99
* 50% match	(2.23)	(4.44)	(1.33)	(1.80)
B. TAX PRO X-IRA CLIENT TAKE-UP RATES DURING THE EXPERIMENT				
20% match	3.09	3.78	1.09	2.42
	(2.86)	(5.56)	(1.66)	(2.65)
50% match	2.29	5.87	4.68	1.98
	(3.26)	(5.76)	(1.84)	(2.76)
Tax pro above median in fraction with	1.24	-0.66	1.82	2.27
X-IRA take-up during the experiment	(1.31)	(3.22)	(1.00)	(1.32)
Tax pro above median	14.16	16.66	5.32	2.05
* 20% match	(2.20)	(4.27)	(1.36)	(1.83)
Tax pro above median	26.49	28.25	8.58	7.02
* 50% match	(2.85)	(4.31)	(1.67)	(1.79)
C. FRACTION OF 50% AND 20% MATCH RECEIVED BY THE TAX PRO				
20% match	7.63	7.69	3.06	2.20
	(3.38)	(5.63)	(1.95)	(2.69)
50% match	12.00	15.42	8.78	4.42
	(3.28)	(5.70)	(1.90)	(2.74)
Tax pro received more than median	-0.71	-2.77	-0.91	-1.33
high offers	(1.69)	(3.44)	(1.01)	(1.34)
Tax pro above median	4.20	7.22	1.22	1.86
*20% match	(2.28)	(4.49)	(1.39)	(1.87)
Tax pro above median	6.41	9.52	1.22	2.26
*50% match	(2.27)	(4.57)	(1.36)	(1.83)
Observations	5292	1380	8958	3827
Tax pros	651	493	723	662

X-IRA dummy (normalized to 100) is regressed on treatment dummies, tax pro experience dummy with X-IRAs and tax pro dummy interacted with treatment dummies.

All regressions control for all individual variables in table 5, and all the variables interacted with the two match rate dummies; they also controls for a dummies indicating the office location.

In Panel A, the tax pro experience dummy is equal to one if the tax pro is above median in the fraction of returns with X-IRA taken up before the experiment.

In Panel B, the tax pro experience dummy is equal to one if the tax pro is above median in the fraction of returns with X-IRAs taken up during the experiment (excluding current return observation).

In Panel C, the tax pro experience dummy is equal to one if the tax pro is above median in the fraction of returns with a match offer (20% or 50%) during the experiment (irrespective of whether offer is made).

match rates, a dummy variable for whether or not the tax professional had more than the median number of X-IRAs taken up by clients before the beginning of the experiment, and the interaction of this dummy variable with the two match rates. In addition, we control for a full set of dummies for the office in which the tax pro works (60 office dummies), as well as all the individual variables included in Table 4, and their interaction with the treatment dummies.[23] The coefficient of the interaction "tax pro above median 20 percent match" is therefore equal to the difference between the effect of the 20 percent match (column 4) in panel 1 and panel 2 of Table 6, adjusted for the possible effects of the control variables. For married tax filers, the effect of the 20 percent and the 50 percent match are both larger when they work with a tax professional who had higher X-IRA take up before the experiment started. For example, for married filers receiving the 20 percent match, being served by a low experience tax professional increases X-IRA take-up by 6.50 percentage points relative to no match. Being served by a high experience tax professional increases X-IRA take-up rate by a further 6.56 percentage points. Therefore, receiving a 20 percent match and being served by a high experience tax pro increased the take-up rate by 13.06 percentage points relative to no match and being served by a low experience tax professional. This shows that a high experience tax professional doubles the effect of the 20 percent match. The effects are similar for the 50 percent match for married taxpayers. For non-married tax filers, the corresponding effects are positive as well, but smaller and not significant at the 95 percent level.

The difference thus persists even when controlling for observable characteristics of the tax filer that affect take-up. To attempt to control for unobserved characteristics, we restrict the sample to new tax filers, since repeat tax filers may choose a tax professional they particularly like. In contrast, new tax filers are often assigned to the next available tax professional, so that within an office, the assignment of a tax filer to a particular tax professional should be close to random.[24] The drawback is that this sub-sample is much smaller. The results nonetheless persist in this sub-sample, except for the interaction between the high-experience tax professional and the 50 percent match, which disappears. On balance, these results seem to indicate that the identity of the tax professional and his/her previous experience does matter for the take up of the incentives.

Panel B in Table 7 investigates this result further by using the tax pro's performance in offering X-IRAs during the experiment, rather than before. Because there is a mechanical correlation between take-up of a particular individual and the mean X-IRA take-up for a tax pro, we first compute for each tax filer the fraction of the *other returns* prepared by the same tax pro during the experiment with X-IRA take-up. We then assign a dummy equal to one for "high experience tax pro," when this fraction is above the median (5.4 percent), and we run the same specification as in panel A.

The results are quite striking: *almost the entire treatment effect appears to be due to tax pros with relatively high X-IRA take-up rates.* When a tax filer is matched with a tax pro who, in the course of the experiment, had relatively fewer X-IRAs taken up by other tax filers, she appears to be completely unresponsive to the match incentives. For example, the effect of the 50 percent match rate (relative to no match) is only 2.29 percentage points for married taxpayers served by a lower experience tax pro. Conversely, the treatment effects are quite large for those who were assigned to a tax-pro with take up of more than 5.4 percent of the other returns during the duration of the experiment; for married tax filers, the effect of the 50 percent match in this population is 29 percentage points (26.49+2.29), and the effect of the 20 percent match is 17 percentage points (14.16+3.09). This is an indication that there are very strong tax professional effects.[25]

The result is robust to adding control variables, as well as to restricting the sample to only new customers, a set within which the assignment is unlikely to be related to the propensity to take up the X-IRA (if anything, the effects are even slightly larger in this sample). However, in both panel A and panel B, a concern could remain that the correlation with prior or current tax-professional behavior does not indicate a causal effect. The most robust evidence of a tax professional effect would be obtained if we could randomly assign different levels of training or exposure to the product to a tax professional. It turns out that our experiment generates something akin to this ideal set-up: since most tax professionals did not prepare a very large number of returns during the experimental period, the tax filer randomization into the three experimental groups generated random variation in the level of exposure to one of the match groups for each tax professional. For some tax pros, very few of their clients received the 20 percent match or 50 percent match offer, while for other tax pros, many more did. A tax pro who had many of the matched offers among his clients acquired more experience in dealing with the X-IRA (since his clients were more likely to open an X-IRA or at least discuss this option seriously) and may therefore have been more likely to encourage his other clients to do this.

We therefore construct a dummy equal to one if the tax pro received more than the median proportion of either 20 percent or 50 percent match offers (i.e., less than 33 percent "no match"), and run regressions using the same specification as in panels A and B.[26] Married tax filers served by tax professionals whose caseload included many match offers are no more likely to take up an X-IRA when there is no treatment, but their treatment effect for the 20 percent and 50 percent matches are significantly larger than if they had a tax professional whose caseload included less than the median fraction of matches. This is very robust evidence of the role of the tax professional in influencing the response to the match rates because assignment is fully random.

These results therefore suggest that, in contrast to the standard model of decision making, individual preferences about savings can be shaped or affected by external cues such as the recommendation of tax preparers. These findings are in line with the recent findings in the behavioral economics literature showing that non-standard features such as default options (Madrian and Shea, 2001) or soft commitments (Thaler and Benartzi, 2004) can have a substantial impact on retirement contributions decisions. Our results show that the responsiveness of both participation and contributions to incentives can be affected by such external influences.

Cashing Out Contributions

An important question is whether the extra X-IRA contributions due to matching incentives will translate into higher net worth or are simply substitutes for other forms of wealth.[27] A first step is to analyze whether X-IRA contributions are withdrawn or whether they stay in the accounts.[28] The matching contributions were deposited in the X-IRA accounts on April 15th. As of May 2nd, only 18 of the 1,500 X-IRA accounts opened during the experiment had experienced any withdrawals. If withdrawals continue to be negligible, the results will represent a challenge to the fully rational model (in which all individuals would take advantage of the match in order to cash out distributions as soon as possible). Furthermore, the fraction of contributors using Roth IRAs rather than traditional IRAs was 54 percent in the control group, 60 percent in the 20 percent match group, and 60 percent in the 50 percent match group. Since early withdrawal tax penalties are less severe for Roth IRAs (the 10 percent early withdrawal penalty does not apply to withdrawals of principal from a Roth IRA), those wanting to game the system and pull the match out as soon as possible would be attracted to the Roth rather than the traditional IRA.

The fact that Roth IRAs are only slightly more prevalent in the 20 percent and 50 percent match groups compared to the control group provides further suggestive evidence that gaming was not a significant factor in the results. In any case, it will clearly be critical to continue monitoring withdrawals and new contributions in the future.

Comparison with the Saver's Credit

The Saver's Credit Program

In this section, we compare the effects reported above to those created by the Saver's Credit, which provides similar matching incentives for low- and moderate income tax filers. Because the nature and magnitude of the formal economic incentives in the experiment described above have some similarities to those in the Saver's Credit, the existence of substantial differences in their effects would suggest strongly that information, tax professional assistance, and simplicity can play key roles in encouraging low- and middle-income households to make contributions to retirement saving accounts.

The Saver's Credit was implemented in tax year 2002 (tax returns filed in early 2003) and is scheduled to expire after 2006 (tax returns filed in early 2007). The Saver's Credit is a non-refundable tax credit on the first $2,000 (for each spouse) contributed to IRAs (Roth and Traditional) or voluntary pension plans (Keogh, 401(k), 403(b), SIMPLE IRA, etc.).[29] As shown in Table 8, the credit rate decreases with AGI and is zero above an AGI threshold that depends on filing status. The credit rate is 50 percent at the bottom, 20 percent in a narrow AGI band, and then 10 percent for a relatively broad range. A credit at rate t is economically equivalent to a match rate of t/(1-t) so that the effective match rates generated by the Saver's Credit are

Table 8: Saver's Credit Parameters

Credit Rate t	Equivalent Match Rate t/(1-t)	Married Filing Jointly AGI range	Head of Household AGI range	Single and others AGI range
(1)	(2)	(3)	(4)	(5)
50%	100%	$0-$30,000	$0-$22,500	$0-$15,000
20%	25%	$30,001-$32,500	$22,501-$24,375	$15,001-$16,250
10%	11.1%	$32,501-$50,000	$24,376-$37,500	$16,251-$22,500
0%	0%	$50,001+	$37,501+	$22,501+

Notes: The Saver's Credit is a non-refundable federal income tax credit proportional to the sum of all elective retirement contributions (all IRAs, 401(k)s, etc.) up to $2,000 of contributions ($2,000 for each spouse for married taxpayers).

Full time students, individuals claimed as dependents by other taxpayers, and individuals aged under 18 are not eligible. Early withdrawals (within the last three years) are netted out of annual retirement contributions to estimate eligible contributions.

As shown in the columns (3) to (5), the credit rate varies by AGI range and marital status. The bracket length for heads of household and singles are 75% and 50% of the bracket length for joint filers, respectively. A credit rate of t (col. (1)) is equivalent to a match rate of t/(1-t) (col. (2)).

AGI used to compute the credit rate t is net of most retirement contributions with the exception of Roth IRAs and is endogenous.

The Saver's Credit is non refundable and can be used only to offset tax liability net of other non refundable credits. As a result, the Saver's Credit is independent of the EITC but interacts with the partially refundable child tax credit.

See IRS Form 8880 and IRS Publication 590 for more details.

actually 100 percent, 25 percent, and 11.1 percent. For example, a tax filer facing the 50 percent credit rate and contributing $1,000 would receive a $500 tax credit, so that its out-of-pocket cost for a $1,000 contribution is only $500, which is effectively a 100 percent match rate. Therefore, the Saver's Credit could potentially generate very large incentives for contributing.

Because the Saver's Credit is non-refundable, many low income tax filers who would qualify for the highest credit rate on the basis of their AGI benefit from it only to a very limited extent or even not at all, because they have little or no tax liability due to standard or itemized deductions, personal exemptions, and use of other non-refundable credits (in particular, the child tax credit).[30] Furthermore, about 20 percent of tax filers showing a positive Saver's Credit on their 1040 tax form actually do not benefit from it, on net, because the Saver's Credit crowds out the Child Tax Credit one for one. The data allow us to define precisely whether a tax filer benefits or could potentially benefit from the Saver's Credit.[31] In what follows, we call those filers who are or could be benefiting "eligible" filers. Those who do not or could not benefit are called "ineligible."

At first glance, comparing pension contributions on each side of the AGI boundary points defining the credit rate brackets might seem like a promising way to analyze the effects of the credit (or match equivalent) rates, since there is a discontinuity in the match rate at that point. This would, however, not be a valid methodology, because AGI is net of retirement contributions, with the exception of the Roth IRA. As a result, even taking gross income (defined as AGI with all retirement contributions added back) as exogenous, rational tax filers just above the boundary have incentives to contribute in order to fall below the boundary and benefit from the higher credit rate. AGI is therefore endogenous.

An implication is that if tax filers were rational, we should expect an abnormally large number of pension contributors bunching just below the boundary point. Symmetrically, we should expect to see few retirement contributors just above the boundary, since increasing contributions even a little would increase the credit by a discrete percentage (and the higher credit rate would apply to the entire contribution). Finding bunching below the boundary point (and fewer contributors above the boundary) would constitute convincing evidence that individuals understand and respond to the incentives provided by the Saver's Credit.

In practice, tax filers may not be aware of the precise dollar amount of their annual incomes and pension contributions.[32] Furthermore, it might be difficult for tax filers to tailor their pension contributions (often specified as a percentage of salary) during the year so that their AGI falls precisely below the boundary points. However, the X-IRA is the ideal product to take advantage of the Saver's Credit. At the time of tax preparation, the exact AGI amount is revealed and it is easy to estimate the current Saver's Credit rate and whether a tax filer could benefit from a higher rate by making X-IRA contributions.[33]

In the H&R Block tax return preparation process, optimizing choices regarding the Saver's Credit is left to the judgment and skill of the client or the tax professional; it is not flagged as a specific item in the standard F10 sequence. Some tax professionals understand the Saver's Credit well and can play with numbers in the X-IRA screen to flag Saver's Credit opportunities to clients. The lack of systematic software support, however, means that we should expect that tax filers will be much less informed about the potential benefits of the Saver's Credit than about our simple and very salient matching experiment.[34]

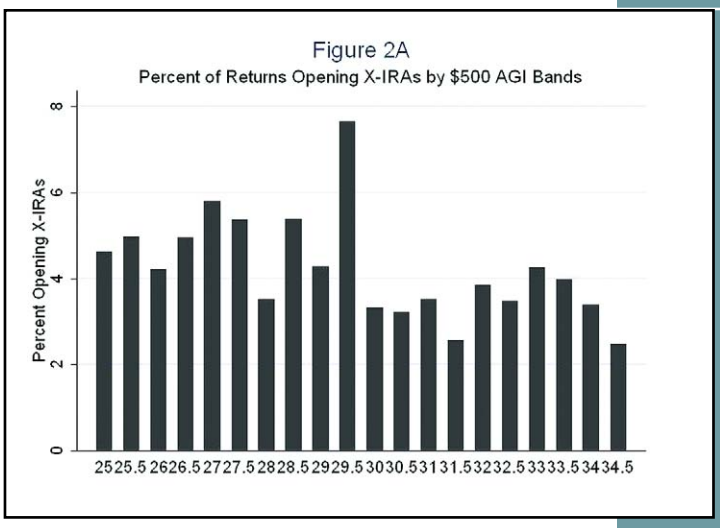
Graphical evidence around the 50 percent-20 percent credit rate boundary

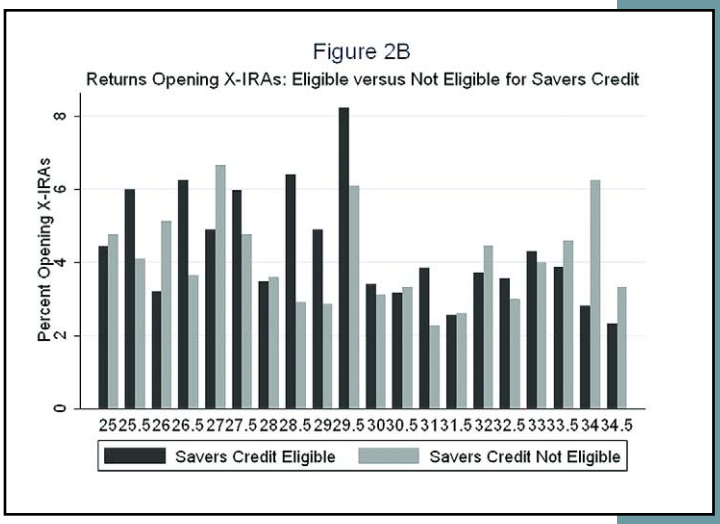
To investigate responses to the Saver's Credit, we use pre-experiment data consisting of tax returns filed in the 60

offices before March 5th. There are 85,000 returns in our data with about 2.3 percent of all returns making an X-IRA contribution.[35] We discard about 7,000 returns of filers ineligible to make IRA contributions (these are mostly retirees with no earned income). We multiply incomes of single and married filing separately tax filers by 2 and incomes of head of households by 1.3333 so that the boundary points, where the Saver's Credit rate changes, are aligned for all types of tax filers. We called these income figures normalized AGI (see Table 8). We concentrate on the first boundary, at a normalized income of $30,000, where the equivalent match rate falls from 100 percent to 25 percent since this is by far the largest discontinuity in the match rate.

Figure 2A plots the percent of tax returns making an X-IRA contribution by $500 AGI bands from normalized AGI $25,001 to $35,000. For example, the group $25,001 to $25,500 is denoted by 25 on the graph. There is a clear spike in the fraction contributing for the band $29,501 to $30,000 (denoted by 29.5 on the graph), which is exactly the band below the $30,000 boundary. The fraction contributing is the highest in that very band (among all 20 bands depicted in the figure) suggesting that the spike is not due to chance alone.

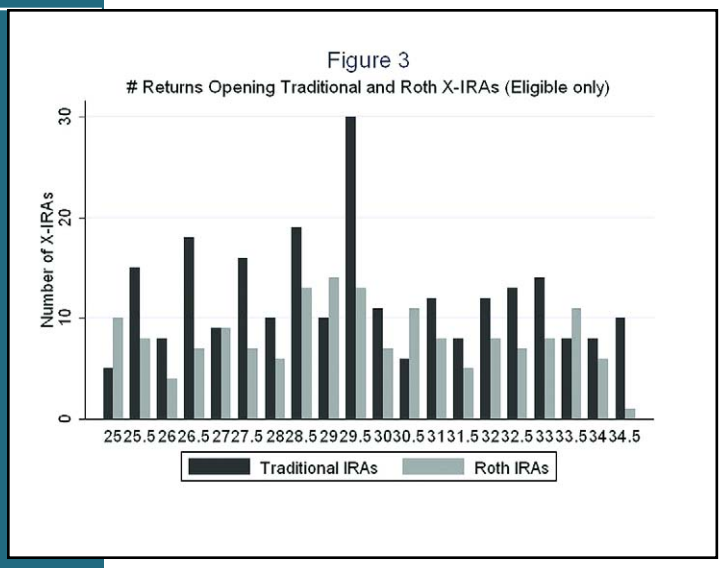
To further test the hypothesis that the spike is due to the Saver's Credit, we divide tax filers into "eligibles," those whose X-IRA contribution (real or potential) would trigger a positive Saver's Credit not offset by reductions in other credits, and "ineligibles."[36] Figure 2B shows the likelihood of being an X-IRA contributor for eligibles (dark shading) and ineligibles (light shading). The figure appears broadly consistent with the tax explanation. The spike is sharp, with over 8 percent of eligible filers in the $29,501-$30,000 band making a contribution. For ineligibles, the spike is less pronounced (two bands among the twenty bands displayed have higher contribution rates) and contribution rates on the left and

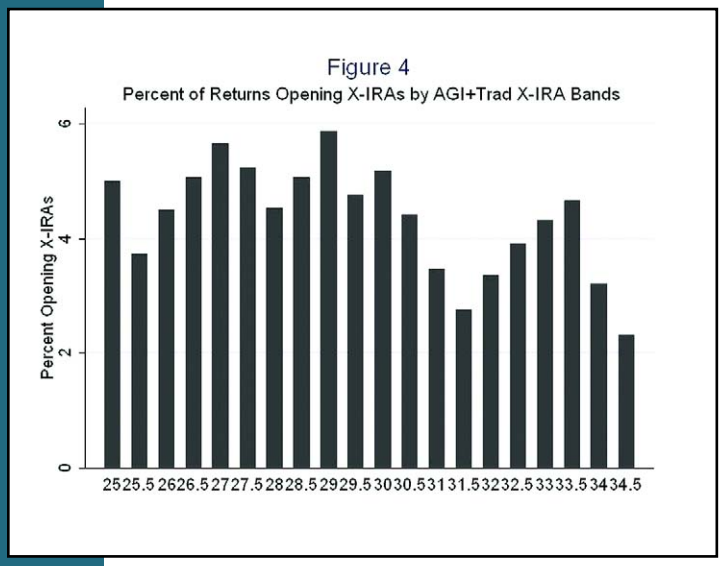


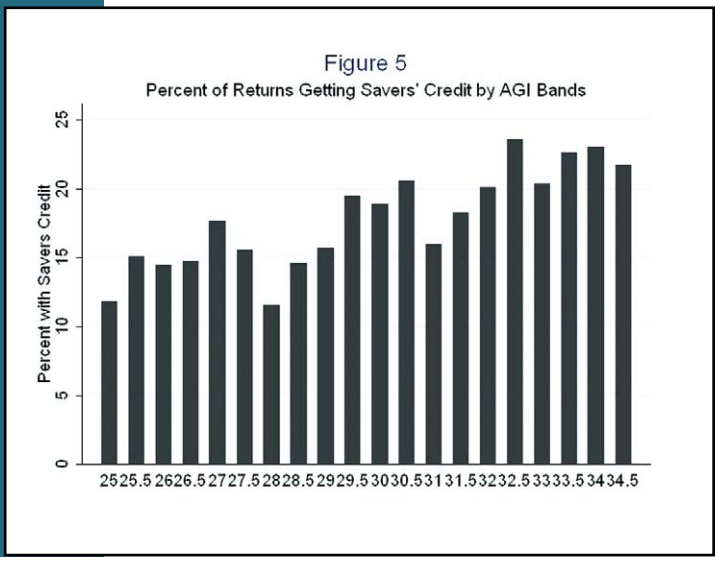


right of the boundary which appear roughly similar in size. It is possible that the small spike in the number of contributors just below the boundary for those ineligible is due to clients and tax professionals (or clients themselves) incorrectly believing that the Saver's Credit will provide a net financial benefit.[37]

In contrast to theoretical expectations, however, we do not observe a clear dip in contribution rates on the right of the boundary for those eligible. In principle, many of those filers could significantly increase their tax refund by contributing more to the X-IRA and hence possibly increase both their X-IRA savings and their after-tax current income net of X-IRA contributions.[38] We also note that, on average, contribution rates are only







slightly higher on the left of the boundary (around 5.5 percent) than on the right (around 3.5 percent), suggesting that the 100 percent match rate has at best small incentive effects relative to the 25 percent match rate.

Figure 3 plots separately the number of Traditional and Roth IRAs opened by AGI income bands for the eligible. Contributing to a Roth IRA does not change AGI (as a Roth IRA is an after-tax contribution). Consistent with the incentive explanation, we see no spike in the Roth IRAs at the boundary. We note that the numbers of Traditional and Roth X-IRAs are both slightly higher on the left than on the right of the boundary, consistent with a small response to the incentives. We also note that some filers contribute to Roth X-IRAs just on the right on the boundary when contributing the same amount to a Traditional X-IRA would have pushed them below the boundary and given them the higher credit rate, suggesting that they are not optimizing their tax savings (since it is implausible that any differences in the tax treatment of Roth and Traditional IRAs would overwhelm the effects of the higher credit rate under the Saver's Credit).

Figure 4 plots the percent of returns contributing to an X-IRA by bands of AGI inclusive of traditional X-IRA contributions (denoted by AGI cum X-IRA). In that figure, the spike at the boundary disappears, providing further evidence that the spike is indeed created by filers making Traditional X-IRA contributions in order to take advantage of the higher credit rate.[39]

Figure 5 plots the percent of returns with positive retirement contributions (all IRA types, and other retirement contributions) among all those potentially eligible for the Saver's Credit. The percent with any retirement contributions (such as 401(k)s, other IRAs, etc.) is much higher than the fraction with X-IRAs (from Figure 2, Panel A). However, no spike at the boundary is visible (the X-IRA spike effect has been diluted). There appears to be a drop in the fraction contributing but it occurs $1,000 above the boundary. This might be simply

noise due to small sample size or might be evidence of failure to optimize (something we will investigate with a larger data set).

Finally, Figure 6 displays the mean X-IRA contributions per contributor for eligible and ineligible tax filers by AGI bands. The mean contributions are quite close to the minimum $300 contribution for X-IRAs and do not display a clear pattern across groups, suggesting that the higher credit rate has little impact on the amount of contributions (note that the amounts are implicitly inclusive of the match equivalent rate as the Saver's Credit is a tax rebate which does not change contributions). About 75 percent of X-IRA contributors make the minimum $300 contribution. Furthermore, this proportion of minimum contributors does not seem to go down when the maximum X-IRA amount eligible for the credit is above $300. Therefore, it seems unlikely that a substantial fraction of X-IRA contributors eligible for the Saver's Credit choose their contribution amounts to maximize their Saver's Credit.[40] This seems to be another important difference between the Saver's Credit and our experiment, where there was significant bunching of contributions at $1,000.

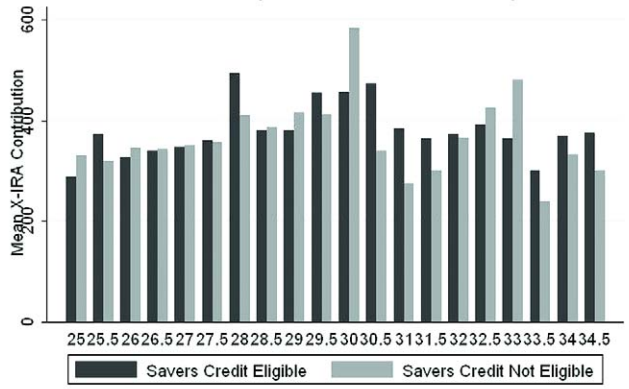
Estimates of the Saver's Credit effects on X-IRA take-up and amounts

The graphical analysis suggests that simple difference-in-differences comparisons could be made to estimate the effects of the large 100 percent match-equivalent credit rate relative to the much lower 25 percent or 11.1 percent equivalent rates generated by the Saver's Credit based on AGI, using the fact that both the AGI and the amount of other credits determine whether a particular tax filer may benefit from the Saver's Credit.

We first compare Saver's Credit eligible tax filers with normalized AGI between $25,001 and $30,000 (entitled to a 100 percent equivalent match rate from the Saver's Credit) to those with AGI between $30,001 and $35,000 (entitled to a 25 percent equivalent match rate or 11.1 percent equivalent rate, see Table 8).

Figure 6

Mean X-IRA Contributions for Eligible Contributors versus Not Eligible Contributors



Columns (1) and (2) in Panel A of Table 9 display the X-IRA contribution rate, average X-IRA contributions, and conditional X-IRA contributions for those two groups respectively. Finally, column (3) displays the differences between the 100 percent group and the 25 percent or 11 percent group.

These differences are small: the X-IRA take up rate difference is only 2.2 percentage points for a match rate increase from 11 or 25 percent to 100 percent. This difference is much smaller than the 7.4 percentage points difference we found between our 20 percent and 50 percent experimental match rates. Furthermore, this simple difference estimate is likely to *overestimate* the effect of the match rate due to the artificial "piling up" effect below the boundary that we described above. Indeed, if we exclude boundary tax filers with AGI between $29,001 and $31,000, the difference becomes even smaller (1.7 percentage points). As we expected from the graphical analysis, there is a $18 insignificant difference in amounts conditional on contribution (the unconditional difference in X-IRA amounts is $9.70).

Another reason why the figure may be an overestimate is that the differences in take-up rates might not be due to the causal effects of differences in the Saver's Credit rate but rather the fact that tax

Table 9: The Effects of the Saver's Credit on X-IRA Behavior

	100% match rate group Normalized AGI $25,001-$30,000	25% or 11% match rate group Normalized AGI $30,001-$35,000	Difference between 100% and 25% or 11% match groups (1)-(2)
	(1)	(2)	(3)
A. Savers' credit eligible taxpayers			
X-IRA take up rate (percent)	5.54 (0.35)	3.32 (0.25)	2.22 (0.42)
X-IRA average contributions (unconditional)	$24.2 (2.03)	$14.5 (1.31)	$9.7 (2.3)
X-IRA average contributions (conditional on take-up)	$398.4 (23.3)	$380.7 (22.1)	$17.7 (32.9)
Sample size	4,227	5,154	9,381
B. Savers' credit ineligible taxpayers			
X-IRA take up rate (percent)	4.56 (0.36)	3.60 (0.55)	0.96 (0.70)
X-IRA average contributions (unconditional)	$18.0 (1.53)	$15.7 (2.66)	$2.3 (3.0)
X-IRA average contributions (conditional on take-up)	$347.7 (14.6)	$372.9 (35.3)	-$25.2 (29.4)
Sample size	3,311	1,140	4,451
C. Difference between eligible and ineligible taxpayers (A-B)			
DD estimates			
X-IRA take up rate (percent)	0.98 (0.51)	-0.28 (0.59)	1.26 (0.81)
X-IRA average contributions (unconditional)	$6.2 (2.67)	-$1.2 (3.06)	$7.4 (4.24)
X-IRA average contributions (conditional on take-up)	$50.7 (30.9)	$7.8 (47.9)	$42.9 (58.1)
Sample size	7,538	6,294	13,832

Notes: This table shows X-IRA take up rates, average contributions (unconditional including zeros and conditional on take-up, excluding zeros) for various groups as well as differences and difference-in-differences. Standard errors are reported in parenthesis.

First, groups are defined relative to normalized AGI (100% of AGI for married taxpayers, 133.33% for Heads of household, and 200% for singles and others).

Group (1) with AGI between $25,001 and $30,000 faces a savers' credit rate of 50% or equivalent match rate of 100% when eligible.

Group (2) with AGI between $30,001 and $35,000 faces a savers' credit rate of 20% or 10% or equivalent match rate of 25% or 11% when eligible.

Column (3) displays the difference between group (1) and group (2).

Second, groups are defined by Saver's Credit eligibility.

Group A are the eligible, defined as taxpayers whose X-IRA contributions (actual or potential) would benefit from the savers' credit.

Group B are the ineligible, defined as taxpayers whose X-IRA contributions (actual or potential) would not benefit from the savers' credit because their tax liability net of other non refundable tax credits (excluding the child tax credit) and net of (full tax credit less potentially refundable additional tax credit) is zero or negative.

Panel C displays the difference between group A and group B. The bottom right panel displays the difference-in-differences (1A-1B)-(2A-2B).

filers with different AGI also have different propensities to save. A simple way to control for this is to consider the same AGI groups for tax filers ineligible for the Saver's Credit because they lack income tax liability before refundable credits. Panel B of Table 9 displays take-up rates and amounts for those two control groups along with the corresponding difference in column (3). For those groups, there are small but insignificant differences in X-IRA take up rates and contribution amounts.

Panel C displays the difference between Panel A (the eligible) and Panel B (the ineligible). If those two groups were identical in their savings tastes, such differences could also potentially capture the Saver's Credit effect. Those differences are even smaller than the results in Panel A. In column (1), the difference in contribution rates is just 0.98 percentage points for the 100 percent match rate and –0.28 percentage points for the 11 percent or 25 percent match group. These negligible effects could be due either to the fact that the Saver's Credit has almost no impact (and the differences in Panel A were biased upward), or that eligible and non-eligible are systematically different. Indeed, the eligible tend to have fewer dependent kids and are more likely to be single, and hence perhaps have a lower taste for savings.[41] Differences in amounts are equally small and barely significant.

Finally, column (3) in Panel C displays the differences-in-differences estimates. These are unbiased estimates of the differential impact of the 100 percent rate relative to the 11 percent or 25 percent match rate if one is willing to assume that, absent the Saver's Credit, the difference in X-IRA behavior between eligible and ineligible filers would be the same across the two AGI groups.[42] All difference-in-differences are small (and insignificant) suggesting again small effects from the Saver's Credit. For example, the difference-in-differences take-up rate estimate is only 1.26 percentage points and insignificant. These results suggest modest effects of the Saver's Credit on take-up and especially on amounts contributed. This contrasts with the results of our experiment, which suggested large effects. We propose explanations for these striking differences and future tests in the conclusion.

Conclusions

This experiment has uncovered a number of intriguing findings. First, simple and saliently presented matching incentives for IRA contributions at the time of tax preparation can have a sizeable impact on IRA take-up rates and on the amounts contributed. IRA contributions (excluding matches) were 4 and 8 times higher with a match rate of 20 percent and 50 percent (respectively) than with no match. With matches included, IRA deposits are 5 and 11 times higher than with no match.

Second, the effects are largest for married tax filers: 25 percent of low- and middle-income married filers contributed to an X-IRA when offered a 50 percent match. For married tax filers, furthermore, the effects are relatively constant across AGI groups, suggesting that even low- and middle-income married families would raise their saving in response to a higher match rate.

Third, tax professionals play a key role in the savings decisions of their clients. Tax professionals who had relatively high client X-IRA take-up rates in the tax season before the experiment started generated much higher take-up rates than those who did not.

Fourth, tax filers apparently did not think about gaming the system by contributing and withdrawing the money very quickly afterwards. Indeed, the take-up rates are too low to be consistent with a systematic "gaming" of the system, only about 60 percent of those contributing chose a Roth IRA (which has much more generous early withdrawal rules for

principal), and virtually no contributors, even among those benefiting from a match rate, made withdrawals in the weeks following the experiment.

The third and fourth findings suggest that tax filers may not be fully informed about savings decisions and hence may rely to a large extent on the advice and trust they place on their tax professional.

We compared the experimental results with those generated by the existing Saver's Credit, which provides an effective match for retirement savings contributions through the tax code. The graphical analysis shows a clear effect of the differential effective matching rates in the Saver's Credit, but simple group analysis indicates that the quantitative effects of the Saver's Credit rates on X-IRA behavior are at best modest. The upper bound of the difference in take-up rates between a 100 percent and a 25 percent equivalent match rate in the Saver's Credit is only 2.5 percentage points. Potentially, larger data samples could be used to tackle this question in a much more precise way. Such an analysis would also require a careful and detailed analysis of the effects of the Saver's Credit, as currently designed, on the inter-temporal budget sets of tax filers and the likely effects on X-IRA behavior. For example, the extent of bunching below the boundary contains information on the size of the behavioral response. Such work, both theoretical and empirical, is beyond the scope of the present paper.

More importantly, the dramatic difference in the size of the responses between our experiment and the Saver's Credit experience needs to be understood. One possibility is that the population on which we estimated the impact of the Saver's Credit is very different from the population on which we estimated the impact of the experiment (those who file early in the season may be more likely to be liquidity constrained than those who file later in the season). This will need to be assessed by estimating the impact of the Saver's Credit among later filers (using another year of data or another city). We suspect, however, that a large part of the difference may be due to the different ways in which the match is presented in our experiment compared to the Saver's Credit. With the Saver's Credit, as currently designed, both the equivalent match rate and the maximum eligible contribution are not easy to decipher.[43] The differential responses thus may represent another piece of evidence suggesting that framing effects are important for understanding behavioral responses. As optimal policy making depends closely on those behavioral responses, an important task for future empirical work is to go beyond merely estimating the size of behavioral responses and start exploring which factors shape the size of the behavioral response.

In summary, the results from a large-scale randomized experiment suggest that the combination of a significant and readily understandable match for saving, easily accessible savings vehicles, the opportunity to use part of an income tax refund to save, and professional assistance could generate a significant increase in retirement saving participation and contributions, even among moderate- and low-income households.

References

Bernheim, B. Douglas (2003). "Taxation and Saving," in Alan Auerbach and Martin Feldstein (eds.), Handbook of Public Economics, Volume 3, North-Holland, 1173-1249.

Boshara, Ray (2005). "Individual Development Accounts: Policies to Build Savings and Assets for the Poor," Brookings Institution Policy Brief Welfare Reform and Beyond #32, March.

Burman, Leonard E., William G. Gale, Matthew Hall, and Peter R. Orszag (2004). "Distributional Effects of Defined Contribution Plans and Individual Retirement Arrangements." National Tax Journal. 57:3. September. 671-701.

Beverly, Sondra, Daniel Schneider and Peter Tufano (2005). "Splitting Tax Refunds and Building Savings: An Empirical Test" Working Paper, Harvard Business School.

Eckel, Catherine and Philip J. Grossman (2003). "Rebate versus Matching: Does How We Subsidize Charitable Contributions Matter?" Journal of Public Economics, 87(3-4), 681-701.

Engelhardt, Gary V. and Anil Kumar (2004). "Employer Matching and 401(k) Saving: Evidence from the Health and Retirement Study", Center for Retirement Research Working Paper 2004-18, Boston College.

Engen, Eric M., William G. Gale, and John Karl Scholz (1996). "The Illusory Effects of Saving Incentives on Saving," Journal of Economic Perspectives 10:4: 113-138.

Kusko, Andrea, James M. Poterba, David W. Wilcox (1998). "Employee Decisions with Respect to 401(k) Plans: Evidence From Individual-Level Data", in Olivia Mitchell and Sylvester Schaefer, eds., Living with Defined Contribution Pensions: Remaking Responsibility for Retirement. (University of Pennsylvania Press: Philadelphia).

Gale, William G., J. Mark Iwry and Peter R. Orszag (2005). "The Saver's Credit: Expanding Retirement Savings for Middle-and Lower-Income Americans." Retirement Security Project Policy Brief No. 2005-2. March 2005.

General Accounting Office (1997). "401(k) Pension Plans: Loan Provisions Enhance Participation But May Affect Income Security for Some." GAO/HEHS-98-5.

Madrian, Brigitte, and Dennis F. Shea (2001). "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior," Quarterly Journal of Economics 116, 1149-1187.

Papke, Leslie E. and James M. Poterba (1995). "Survey Evidence on Employer Match Rates and Employee Saving Behavior in 401(k) Plans." Economics Letters 49, 313-317.

Poterba, James M., Steven F. Venti and David A. Wise (1996). "How Retirement Saving Programs Increase Saving," Journal of Economic Perspectives 10:4: 91-112.

Sherraden, Michael (1991). Assets and the Poor: A New American Welfare Policy. Armonk, NY: M.E. Sharpe, Inc.

Thaler, Richard, and Shlomo Benartzi (2004). "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving," Journal of Political Economy, 112, S164-S187.

Endnotes

[1] See Burman et al (2004) for data on defined contribution pension coverage rates by income group. Calculations from the 2001 SCF imply that only one quarter of households with income below $40,000 have defined benefit coverage. Burman et al (2004, Appendix Table 6) report that among households with cash income below $40,000, less than 2.1 percent contributed to either a Roth or traditional IRA in 2004. Median net financial wealth in the text is calculated from the 2001 Survey of Consumer Finances.

[2] For a discussion of efforts to increase saving among low-income households through Individual Development Accounts and other accounts, see Sherraden (1991) and Boshara (2005).

[3] Gale, Iwry, and Orszag (2005) examine the Saver's Credit in more detail.

[4] Some studies (e.g., Kusko, Poterba, and Wilcox 1998) find that the existence of a match raises 401(k) saving, but that a higher match rate itself does not. Other studies (e.g., General Accounting Office 1997) have even found that a higher match rate, conditional on the existence of a match, actually reduces 401(k) saving because the income effect dominates the substitution effect. Still other studies (e.g., Papke and Poterba 1995) find that higher match rates do increase 401(k) saving. One recent study, Engelhardt and Kumar (2004), found a positive but modest effect of match rates using individual level Health and Retirement Survey data. Their results suggest that introducing a 20 percent or 50 percent match rate should increase dollar contributions by about 10 percent and 25 percent respectively. Nonetheless, the effects of 401(k) match rates, conditional on a match existing, remain unclear. Bernheim (2003) identifies this as an important and unresolved empirical issue.

[5] H&R Block also provides the tax preparation software "Tax Cut" which is the second largest tax software program in terms of market share and is used by millions of tax filers to prepare their own tax returns.

[6] Since 2002, tax filers have the option to set up automatic monthly contributions to their X-IRA account from their bank account. The minimum monthly contribution is $25.

[7] There is also a $25 account termination fee. Contributions or withdrawals by mail are free.

[8] Tax professionals are paid $5.50 for each X-IRA account opened or re-contributed to by their clients. More generally, a tax professional receives greater compensation for completing a more complicated tax return.

[9] In the case of married tax filers, modifying the order to the tax filers listed on the returns could have possibly increased the odds of getting a more generous offer. Tax professionals did not know that assignment was based on Social Security number and therefore presumably very few tax professionals did this reordering systematically. We discuss this point further in the descriptive statistics section.

[10] During the first week of the experiment, from March 5th to March 12th, single or head of household tax filers who declined to contribute to the X-IRA were not properly recorded by the TPS software as being offered the X-IRA even if their tax professional hit the X-IRA screen. We exclude those observations when we study offer rates.

[11] It is possible for some tax professionals to reach the offer screen and decline the offer without presenting the offer to the clients. In our data this will still appear as a client having been offered the X-IRA.

[12] A second set of training sessions for the 15 additional offices was organized by Scott McBride on March 8-9.

[13] H&R Block headquarters is capable of monitoring cases in which an IRA deposit was immediately rescinded and hence discover systematic patterns of behavior. We show later in the paper that to date there have been only a handful of instances in which tax filers made withdrawals.

[14] There is no age limit to make IRA contributions as long as tax filers have earned income. Almost no tax filer had reached the maximum IRA contributions for both 2004 and 2005 at the time of tax preparation. Tax filers with high AGI can still make non-deductible traditional IRA contributions which qualified for the experimental match.

[15] This difference could be due to chance. Alternatively, in a few instances, tax professionals may have modified the ordering of the spouses on a joint return when the first draw of the offer generated a zero match, to see if this change randomized the couple into a better offer. In a future revision to this paper, we plan to correct for this possibility by using the ordering of tax filers in the previous tax season to assign to the tax filer to an "intended treatment" for the 74 percent of the married tax filers in our sample who filed their taxes with H&R Block the previous year. Since the difference in assignment of married couples across groups is minimal, our results should not be greatly affected.

[16] These data exclude the first week of the experiment (from March 5th to March 11th) when, due to a software glitch, the occurrence of the pop-up screen was not properly recorded.

[17] Beverly, Schneider, and Tufano (2005) obtain much larger take-up rates (around 15 percent) for a split tax refund option in an experiment in Tulsa Oklahoma with a non-profit tax preparer. In the Tulsa experiment, the refund was split into a regular savings account set up at the time of tax preparation. Demand for split refunds in Tulsa may possibly have been higher than X-IRA demand in the absence of any additional match because the Tulsa version allowed tax filers to set up savings accounts for free and the money in the savings account could be withdrawn at any time with no penalty (95 percent of initial contributions had been withdrawn from the Tulsa saving accounts 6-8 months after set-up).

[18] For example, in comparing the no-match and the 20-percent match group, (1+match rate) rises from 1 to 1.2 and the take-up rate rises from 3.31 to 9.64. The elasticity is therefore [ln(9.64)-ln(3.31)]/[ln(1.2)-ln(1)], or 5.9.

[19] Our analysis is always done at the tax return level. In the case of married filing jointly, the tax return includes both spouses. Each spouse can separately open an X-IRA. In our analysis, a tax return for married joint filer is defined as having contributed to an X-IRA if at least one of the spouses contributes and the contribution amounts are defined as the sum of contributions for the two spouses. We also analyze subsequently the spousal decision to open a second X-IRA.

[20] The take-up rate for married filers (considered as individuals) is simply the average of the married taxpayer take up rate and the spousal take up rate from Table 2, Panel B. Those married individual take up rates are 1.9 percent, 10.8 percent, and 20.4 percent for each of the three groups. This shows that married filers, considered as individuals, are somewhat more responsive to match incentives than single filers.

[21] We have not done any adjustment for income growth because the nominal income growth between 2000 and 2004 will be very close to zero. Internal Revenue Statistics show that average nominal income per tax return in 2003 was still 3 percent lower than in 2000.

[22] More precisely, these are the tax professionals who had client X-IRA take-up rates of less than 1.5 percent of tax returns they prepared this season before March 5th.

[23] That is, we control for all the variables in specification 3 and 6 in Table 5. To save space, we do not report these coefficients.

[24] An exception would be when a new client is referred to a tax professional by a friend.

[25] The results are essentially identical when we separate the tax professional into those with no X-IRAs taken up by the other clients that they served during the experiment (one third of observations), and those with at least one X-IRA taken up by their other clients (two thirds of observations).

[26] The dummy variable is defined by the Social Security number assignment rule and irrespective of whether the tax pro clicked on the X-IRA window and made the corresponding offer to the client.

[27] There is a large and controversial academic literature on the effects of 401(k) plans on the net worth of households (see Bernheim, 2003, Engen, Gale and Scholz (1996) and Poterba, Venti and Wise (1996) for surveys).

[28] As we discussed earlier, if all tax filers were rational and not severely credit constrained, take up rates for the 20 percent and especially the 50 percent match rate should have been much higher as tax filers are able to cash out contributions (including the match) after the match is deposited on their account on April 15th subject to the relatively small IRS penalties for early IRA withdrawals. It is therefore important to assess whether withdrawals are important and whether they vary by match rate groups.

[29] Those contributions are netted of any withdrawals made during the last three years so that tax filers do not game the program.

[30] The Saver's Credit is determined before refundable credits. Therefore, the Earned Income Tax Credit and the refundable portion of the Child Tax Credit do not reduce the Saver's Credit.

[31] More precisely, we define a tax filer as potentially benefiting from the Saver's Credit if, starting from no retirement contributions, his/her tax refund would increase or tax liability decrease due to the Saver's Credit should he/she make a retirement contribution.

[32] Indeed, annual salary contracts rarely run from January 1st to December 1st. Actual payment dates for work in December might fall in January. There are a number of additional factors, such as Social Security and Medicare taxes, pre-tax parking and health care contributions, which make it difficult to evaluate precisely taxable wages and salaries before the W2 form arrives in January. At that time, it is too late to modify employer pension plan contributions.

[33] IRA contributions are in general made during the tax season once the AGI information is revealed and are clearly an easier tool for tax optimizing than employer's automatic monthly pension contributions.

[34] H&R Block routinely updates the TPS software through the tax season depending on tax law changes and tax professional feedback.

[35] In future work, we plan on using a much larger sample to do a comprehensive and more precise analysis of the Saver's Credit.

[36] Because of the complex interaction with the Child Tax Credit, many tax filers do not benefit from the Saver's Credit even if their return shows a positive Saver's Credit.

[37] Our dataset is not large enough to investigate this issue precisely. We will be able to analyze this effect by separating naïve eligible from truly eligible when we have access to a larger dataset.

[38] Analyzing more precisely whether many tax filers are in that dominated situation would require a larger dataset and is left for a future revision.

[39] Indeed, the largest number of filers crossing a particular AGI threshold due to their Traditional X-IRA contributions occurs precisely at $30,000. Furthermore, about 40 percent of those crossers had non X-IRA contributions (a much larger percentage than for the average X-IRA contributors). This suggests that some filers strategically use X-IRAs to make their other retirement contributions qualify for the higher rate.

[40] We will analyze this issue in more detail when we have access to more data.

[41] With a larger data sample, the same analysis could be repeated within marital status group and within households with the same number of dependents (exploiting for example the fact that the child tax credit is only for children up to age 16 and that families with children aged 16 or 17 should be quite comparable).

[42] Such a hypothesis could be tested using 2001 tax data (2002 tax season) when the X-IRA was fully deployed and the Saver's Credit did not exist yet.

[43] Furthermore, experimental work has shown that credit rates are much less effective than equivalent match rates to induce people to contribute to charities (see e.g. Eckel and Grossman, 2003). It is possible that presenting the Saver's Credit as a 100 percent match rather than a 50 percent credit rate could have a large effect on take-up. This is something we plan to test in the future.

Esther Duflo, Professor of Economics, Massachusetts Institute of Technology and Director, Poverty Action Lab

William G. Gale is Deputy Director of the Economic Studies Program at the Brookings Institution, the Arjay and Frances Fearing Miller Chair in Federal Economic Policy at Brookings, and Co-Director of the Urban-Brookings Tax Policy Center.

Jeffrey Liebman, Associate Professor of Public Policy, Kennedy School of Government, Harvard University and Faculty Research Fellow, National Bureau of Economic Research.

Peter R. Orszag is Director of The Retirement Security Project, the Joseph A. Pechman Senior Fellow in Tax and Fiscal Policy at the Brookings Institution, and Co-Director of the Urban-Brookings Tax Policy Center.

Emmanuel Saez is Associate Professor of Economics at the University of California, Berkeley and Research Associate at the NBER.

Mission Statement

The Retirement Security Project is dedicated to promoting common sense solutions to improve the retirement income prospects of millions of American workers.

The goal of The Retirement Security Project is to work on a nonpartisan basis to make it easier and increase incentives for middle- and lower-income Americans to save for a financially secure retirement.

1755 Massachusetts Ave., NW, Suite 550
Washington, DC 20036
p: 202.483.1370 f: 202.483.1460
www.retirementsecurityproject.org

RSP