

NBER WORKING PAPER SERIES

SAVING INCENTIVES FOR LOW- AND
MIDDLE-INCOME FAMILIES: EVIDENCE
FROM A FIELD EXPERIMENT WITH H&R BLOCK

Esther Dufo
William Gale
Jeffrey Liebman
Peter Orszag
Emmanuel Saez

Working Paper 11680
<http://www.nber.org/papers/w11680>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2005

We thank H&R Block for the collaboration and resources it has devoted to this experiment. We gratefully acknowledge the help and contributions of the H&R Block team led by Bernie Wilson and 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 are those of the authors alone and do not necessarily reflect the views of H&R Block, any of the funders of this project, or any of the institutions with which the authors are affiliated. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

©2005 by Esther Dufo, William Gale, Jeffrey Liebman, Peter Orszag and Emmanuel Saez. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

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

Esther Dufo, William Gale, Jeffrey Liebman, Peter Orszag and Emmanuel Saez

NBER Working Paper No. 11680

September 2005

JEL No. H0, H3

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 14,000 H&R Block clients, across 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 main findings. First, higher match rates significantly raise IRA participation and contributions. Take-up rates were 3 percent for the control group, 8 percent in the 20 percent match group, and 14 percent in the 50 percent match group. Average IRA contributions (including non-contributors, excluding the match) for the 20 percent and 50 percent match groups were 4 and 7 times higher than in the control group, respectively. Second, several additional findings are inconsistent with the full information, rational-saver model. In particular, we find much 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 contributions to retirement accounts, including among middle- and low-income households.

Esther Dufo
MIT Department of Economics
50 Memorial Drive
Cambridge, MA 02142
and NBER
eduflo@mit.edu

William Gale
Brookings Institution
Retirement Security Project
1775 Massachusetts Avenue
Washington, DC 20036
wgale@brookings.edu

Jeffrey Liebman
Kennedy School of Government
Harvard University
79 JFK Street
Cambridge, MA 02138
and NBER
jeffrey_liebman@harvard.edu

Peter Orszag
Brookings Institution
Retirement Security Project
1775 Massachusetts Avenue
Washington, DC 20036
porszag@brookings.edu

Emmanuel Saez
University of California
549 Evans Hall #3880
Berkeley, CA 94720
and NBER
saez@econ.berkeley.edu

1. 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 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 along with 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 for families with 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 participate in retirement saving accounts. Matching contributions can be independent of the individual's marginal tax rate, and thus provide a significant incentives even for 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 contribute to retirement accounts.

The Saver's Credit offers one example of such matching contributions. Enacted in 2001, the credit provides a federal 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 the full potential of matching, however, in part due to its low income thresholds for eligibility, in part because it is not refundable (taxpayers are eligible for the credit only to the extent of their income tax liability), 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

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). Mills, Gale, and Patterson (2005) evaluate evidence from a randomized experiment examining the effects of Individual Development Accounts on household wealth.

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

401(k) participation and contributions.⁴ A concern with matching studies involving 401(k) plans, however, is that the match rates offered by firms in such plans may not be independent of worker characteristics at the firm, which makes it difficult to disentangle the effect of matching rates on contributions from the selection of workers who are likely to work in firms that offer 401(k) plans. Another concern is that the results apply mainly to relatively affluent households. In particular, automatic payroll deduction and the other workplace features associated with 401(k) plans are more likely to be offered in larger firms with a comparatively better-off workforce. Therefore, the results may not be directly applicable to low- and middle-income families or to policy interventions that occur outside the workplace.

In contrast to the mixed evidence on how matching affects contributions, designs based on behavioral economics, such as default options (Madrian and Shea, 2001) or soft commitments (Thaler and Benartzi, 2004), have been shown to have a substantial impact on retirement contribution decisions, including and even especially among lower- and middle-income households.

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 and transparent 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 the matching contributions in studies of 401(k) plans, the matches studied in this paper are guaranteed (by random assignment) to be independent of individual 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 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 metropolitan area from March 5th to April 5th, 2005. The

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 data from the Health and Retirement Study. 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.

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 of \$300 was maintained in all groups.

The results show that match rates can have large effects on IRA participation and contributions. Take-up rates were 3 percent, 8 percent, and 14 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) were \$765, \$1,100, and \$1,110, respectively. With the match included, average IRA deposits were \$765, \$1,280, and \$1,590 respectively, among contributors. Average IRA deposits (including non-contributors, excluding the match) with the 20 percent and 50 percent matches were 4 and 7 times higher, respectively, than with no match. These effects are substantially larger than those found in the context of 401(k) matches.

The second main result of the paper comes from a comparison between the effect of this match experiment and that of the Saver’s Credit. Using national H&R Block data (not drawn just from St. Louis), we estimate upper bounds of the effect of the Saver’s Credit on contributions, and find that they are almost five times smaller than that of a 50 percent match rate, even though the effective match provided by the Saver’s Credit for eligible lower-income households is substantially larger than 50 percent. Since both programs provide a match for retirement savings contributions, the striking differences in the elasticities highlights the role of information and simplicity in affecting 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. This strategy would not, however, produce contribution rates anywhere near 100 percent.

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.

2. Experimental Design, Implementation, and Data

A. 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 select that particular screen directly.

B. The Express IRA Product

Since tax season 2001, H&R Block has offered a product, called the Express IRA (X-IRA), which allows 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 from a checking or savings account, 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. X-IRA contributions can be either Traditional IRAs, where the contribution is deducted from income for tax purposes and income taxes are paid upon withdrawal, or Roth IRAs, where the contribution is not deductible but no further tax is due upon withdrawal. Roth and Traditional IRAs are economically equivalent if the tax filer faces the same marginal tax rate currently and upon withdrawal.

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 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

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.

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 (HRBFA), 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 monthly 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. To appear, it must be 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 the client asks about it.⁸ Clients who are offered the X-IRA and decline it may change their mind later during the tax preparation process (for example if they obtain a larger-than-expected tax refund). It is simple 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 frequently 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. Despite this advertising, take-up has been modest. 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 new accounts and re-contributions) was approximately 1.4 percent.

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 and this commission structure was in place for our experiment. More generally, a tax professional receives greater compensation for completing a more complicated (and therefore more time consuming) tax return.

The existence of the X-IRA at H&R Block provides an auspicious tool for 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.

C. 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 reach 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. Tax professionals were not informed 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 the X-IRA offer screen. Therefore, the decision of whether or not to offer an X-IRA to the client was independent of treatment status. Because a tax professional's decision regarding whether or not to navigate to the X-IRA 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

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.

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).

The second way that TPS was modified is that, when the X-IRA screen was activated, a special pop-up window would automatically appear presenting the offers corresponding to the group to which the client had been assigned, and asking whether the client wanted to contribute to an X-IRA. In particular, the pop-up window proposed \$500 as an example of a contribution level and presented the associated match. 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. 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 (and many did so). 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-IRAs totaled roughly half a million dollars. They 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.¹² 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. The 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).¹³ The training described the general goal of the experiment and explained all the

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 Withdrawals can be tax-free (and penalty free) if the deposit and withdrawal are within the same tax year, and no deduction is claimed on the deposit.

¹³ A second set of training sessions for 15 offices added a little later in the experiment was organized by Scott McBride on March 8-9.

details of the TPS modification. Tax professionals were instructed to follow the F10 sequence and systematically go through the X-IRA screen and to carefully read and explain to the tax filer the offer on the pop-up window for all clients in the experiment. 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.¹⁴ All tax professionals were provided with abundant documentation about the experiment: goals, 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 are a 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 presented the offer to their clients (as there would be if a match like this were implemented as federal policy).

D. 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. Clients were informed that they were receiving a special X-IRA offer as part of a research project and that they were under no obligation to participate. They also received H&R Block’s standard X-IRA explanation that IRAs are not for everyone and that there can be penalties for early withdrawal. 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. In order to study the Saver’s Credit, we use national H&R Block data for tax season 2005, with a 100 percent sample of all returns contributing to an X-IRA and a 9 percent random sample of all other returns.

¹⁴ 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.

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.¹⁵ We often divide filers into “married” (specifically, married filing jointly) and “non married” (including singles, heads of households, and married filing separately) categories.

E. Descriptive Statistics

Table 1 displays the means of several variables for each of the three experimental groups in the first three columns and tests whether the means are statistically different across groups in the next three columns. The t-statistics of the difference are reported whenever they are 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 on average.

Only two variables, fraction married and fraction homeowners, are significantly different (at the 5 percent level) between the treatment and control groups. The proportion of married tax filers is slightly higher in the 20 percent matching group (37.6 percent) than in the control group and 50 percent matching group (35.4 percent and 35.4 percent).¹⁶ The fraction of homeowners is slightly lower in the 20 percent match group than in the other two groups. Although modest, the differences are statistically significant. Because we are looking at 15 different variables, it is not surprising to find some differences with t-statistics just over 2, even when randomization was successfully implemented.

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 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. Almost half of tax returns report positive investment income (interest, dividends, or rents). About two-thirds reports owning a savings account, and slightly over a quarter have made positive 401(k) contributions.

¹⁵ 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.

¹⁶ This difference could be due to chance. Alternatively, in a few instances, tax professionals could possibly have modified the ordering of the spouses on a joint return when the first draw of the offer generated a zero match. This, however, is an unlikely explanation, because the effect should have been more pronounced for the 50 percent match group. More directly, married filers for whom both SSN endings fell into the same group could not manipulate the offer generated. We find that the effects of matches on take-up for this group is extremely close (well within statistical confidence intervals) to the effects on take-up for married couples on average, further suggesting that tax professionals did not re-order spouses to generate offers.

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 tend to be 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.¹⁷

As we noted above, it was possible for tax professionals to skip the X-IRA screen and never present an offer to the client. TPS recorded which tax returns went through the initial X-IRA screen. When this initial X-IRA screen was reached the pop-up window appeared and the tax professional had to record the client's answer in order to be able to continue. We refer to returns that reached this screen as having received an "offer."¹⁸ It is worth emphasizing that all this dummy variable measures is whether the screen was reached. It does not tell us whether the subject of opening an IRA was discussed with a tax filer. A tax pro could discuss extensively the pros and cons of opening an IRA and then skip the screen after determining that the client was not interested.¹⁹

The first row of Table 1 shows "offer" rates for each of the three groups; the rates were about 75 percent in each group which is consistent with the fact that tax pros had no way to determine which clients would receive a particular offer other than to go to the offer screen (i.e., they did not figure out the randomization algorithm). The second row shows that the 25 percent of cases in which no offer was made do not arise from a few tax professionals systematically avoiding the pop up screen; almost all professionals displayed the screen at some point.

Figure 1 shows that tax professionals became more likely to "offer" the X-IRA as the experiment progressed; the offer rate increased from 55 percent on March 12 to just below 80 percent around March 22, and stayed fairly constant after that date. The increase largely reflects H&R Block's effort to get professionals to systematically

¹⁷ It is also worth noting that tax professionals are less busy in March than they are in February (the main crunch period occurs when people come to get their tax refund as soon as possible after receiving their W-2s from their employers) and this might have implications for tax professionals' willingness to offer the X-IRA product.

¹⁸ 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.

¹⁹ It is also possible for a tax professional to reach the offer screen and click "no" so fast that the client would not be aware of the offer.

²⁰ The statistics for this variable 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.

navigate through the screens using the F10 sequence. In addition, tax pros appear to have become more comfortable offering the X-IRA product after they had done so a few times.

3. Results

A. Take-up Rates and Contributions

Table 2 presents the main results of the experiment. The table shows X-IRA take-up rates and contribution levels 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 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 were 2.9 percent, 7.7 percent and 14.0 percent respectively, in the no match, 20 percent match, and 50 percent match groups. The take-up rate in the no match group was low, and less than one percentage point higher than take-up before the experiment started (2.1 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 can be more effective if contributions are matched.²¹ The no-match results also suggest that the propensity to contribute among those included in 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 were even stronger, since contributions, conditional on take-up, were higher in the two groups receiving the match than in the control group.²² Not surprisingly, when the matching contributions are added, the differences grow even larger: unconditional X-IRA deposits (including the match) for the three groups were respectively \$22, \$99, and \$222. Total IRA deposits were 4.5 times larger with a 20

²¹ 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).

²² The fact that contribution levels were similar across the two match groups could be the result of offsetting income and substitution effects, though we present evidence below that is inconsistent with a simple income and substitution effect model.

percent match than with no match, and 10 times larger with a 50 percent match than with no match.²³

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 $(1+\text{match rate})$, defined as the change in the natural log of the take-up rate divided by the change in the natural log of $(1+\text{match rate})$ were respectively 5.4 for the 20 percent match rate and 3.9 for the 50 percent match rate.²⁴ 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 were about 50 percent larger than for others.²⁵ Take-up rates were 2.4 percent, 9.4 percent, and 17.8 percent for married filers in the three groups, respectively, compared to 3.2 percent, 6.7 percent, and 11.9 percent for other filers. The effects on amounts contributed were also correspondingly much larger for married tax filers than for others.

For married filers, the average unconditional contributions inclusive of match were respectively \$33, \$162, and \$381 in the three groups, and were almost three 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 were 0.4 percent, 4.1 percent, and 8.4 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.²⁶

The cumulative distribution of X-IRA contributions (excluding the matching amounts) is shown in Figure 2A for non married filers and Figure 2B for joint filers. Each graph reports the fraction of tax filers contributing less than a given amount for any level

²³ The contribution amounts inclusive of match presented are calculated on the basis of the match that was slated to have been given. In fact, H&R Block accidentally ended up matching the entire contribution for those who contributed more than \$1,000, so the actual amounts in the X-IRAs are higher than those presented here. Since the matches were deposited only at the end of the season, this did not affect the expectations of anyone who participated in the experiment.

²⁴ For example, in comparing the no-match and the 20-percent match group, $(1+\text{match rate})$ rises from 1 to 1.2 and the take-up rate rises from 2.9 to 7.7. The elasticity is therefore $[\ln(7.7)-\ln(2.9)]/[\ln(1.2)-\ln(1)]$, or 5.4.

²⁵ Our analysis is always done at the tax unit level. In the case of married filing jointly, the tax unit includes both spouses. Each spouse can separately open an X-IRA. In our analysis, a tax unit 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.

²⁶ 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 1.4 percent, 6.7 percent, and 13.1 percent for each of the three groups. This shows that married filers, considered as individuals, are slightly more responsive to match incentives than single filers.

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. About 40 percent of the individuals in the match groups who contribute to an X-IRA contribute exactly \$1,000, the maximum eligible for a match. Some bunching also occurs at \$500 for the match groups (and much less so for the control 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.

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 1 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. Although the increase in take-up between March 5 and March 22 could potentially have been 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.²⁷ It could be that tax professionals became better at presenting the product, or that clients who filed later in the season were savvier or less cash constrained. The effect does not, however, appear to be due to a different propensity to contribute, since the take-up in the no match group did not rise over time.

B. Who took up the match?

The welfare and policy implications of savings incentives depend in part on who is most likely to use them. Table 3 explores the effect of the match on take-up rates by tax filer income and tax refund size. Of particular interest is the response of poorer households, where much of the policy concern is focused. Panel A in Table 3 presents the average X-IRA contributions by quartiles of AGI (to which we have added the traditional IRA contributions, since such contributions are subtracted when computing AGI). Quartiles are defined separately for married joint filers and other filers. Each tax

²⁷ In other words, most of the increase in offer rates over time likely came from tax pros who had already determined that their clients did not want to open an X-IRA becoming more likely to comply with the instruction to always navigate through the offer screen.

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 2000.²⁸

The effect of the match is significant in all income quartiles, although it does increase with income: Relative to the control group, the take-up rate was 5.1 percentage points higher in the 50 percent match group in the first quartile, 9.9 points in the second quartile, 12 points in the third quartile, and 14.6 points in the top quartile. The effects of the 20 and 50 percent match rates remain significant in all groups including the bottom quartile.

Panel B in Table 3 divides tax filers by EITC (Earned Income Tax Credit) receipt and Panel C by size of the federal tax refund. As Table 1 shows, about 17 percent of the sample overall receives the EITC and about two-thirds receive a federal tax refund in excess of \$500. Both of these variables are correlated with income. Given the results in Panel A, it is therefore not surprising that Panel B 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. The effects of matches on take-up rates by EITC status, however, are quite similar: 14.2 percent of those who do not receive the EITC and were offered the 50 percent match contributed, while 12.9 percent of those who do receive the EITC contributed when offered the same 50 percent match. The modest difference across EITC status groups (relative to the larger difference across quartiles) might be due to the size of refund. Indeed, Panel C shows that receiving a refund over \$500 significantly increases take-up rates across all treatment groups. For example, the take-up rate for the 50 percent match group was 12.5 percentage points higher than the control group for those with a large refund and only 8.5 percentage points higher for those with a small refund. This descriptive finding is consistent with the view that the availability of money due to tax refunds makes tax preparation time perhaps the most favorable time of the year for many low- and middle-income families to contribute to retirement accounts outside a payroll deduction system.

Table 4 further explores the effects of the match among various subgroups. Panel A and B show that responses to the match were larger for tax filers owning a savings account or reporting positive investment income (dividends or interest income). Those with a savings account or with positive investment income were presumably less credit constrained and could use other savings to take advantage of the matches. Those attributes might also proxy for higher tastes for savings in general and hence a higher propensity to respond to savings incentives. Panel C shows that tax filers with positive 401(k) contributions were more likely to respond to the match. This suggests that a stronger taste for retirement savings, or more familiarity with retirement savings and investment in general, increased the effects of match rates. Panel D shows that homeowners respond less to matches than non-homeowners. A possible explanation is that homeowners have to service mortgage payments and might be more cash constrained

²⁸ 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.

in the short run than non-homeowners. Finally, Panel E shows that repeat H&R Block customers responded more to match rates than new customers. This could be due to more familiarity with the company and with the tax professional; below, we discuss the important role of the tax professional in savings decisions.

Finally, Figure 3 presents take-up rates by age. It shows that the effects of the match increased sharply from age 18 to age 40, were relatively flat from age 40 to age 60, and then decreased slowly above age 60. Those results may suggest that tax filers perceived the X-IRA as a retirement savings product and were most likely to contribute and respond to incentives at the ages at which it makes most sense to be saving for retirement. Rational individuals age 59.5 and above would have the easiest opportunity to game the system, because there are no penalties for withdrawals at those ages and hence no costs for gaming, but we do not observe disproportionate responses at those ages.²⁹

It is important to emphasize that take-up rates for groups with the 50 percent match are never above 20 percent, even in the groups of people who are not credit constrained and could easily divert tax refunds or other savings to X-IRA contributions. Credit constraints alone thus cannot explain why well under 100 percent of H&R clients took advantage of the match rate (thereby failing to capture the potential “free lunch”), which is what basic rational theory would imply.

Tables 3 and 4 have uncovered interesting patterns across tax filer characteristics but many of those characteristics are correlated. For example, those with investment income tend to have higher overall incomes but also smaller refunds. Table 5 shows simple OLS regressions to evaluate the influence of each characteristic keeping the others constant. These estimates are not estimates of causal effects of the covariates (such as income quartile, or presence of investment income) on X-IRA take-up rates. Instead, they should be interpreted as showing how the effects of matching incentives change across the population when a given characteristic (such as income quartile) changes but all the other characteristics are kept constant.

Columns (1) to (3) report the coefficients of a single OLS regression on a set of covariate dummies (col. (1)) and the same set of dummy covariates interacted with the 20 percent (col. (2)) and 50 percent (col. (3)) treatment group dummies. Therefore, column (1) estimates the effect on take-up of switching a covariate variable from zero to one for those with no match. For example, having a refund over \$500 (and keeping the other covariate constant) increases the likelihood of making a contribution by 3.09 percentage points. Columns (2) and (3) estimate the additional effect of the 20 and 50 percent match rates respectively (in both cases relative to no match) when the corresponding covariate dummy becomes equal to one. For example, having a refund above \$500 increases the effect of the 20 percent match by 3.1 percentage points. Col. (4) to (6) repeat the same regression but with X-IRA unconditional contribution amounts instead of take-up.

²⁹ For related evidence on people over age 60 who do not contribute to their 401(k) even though their employer matches contributions and there is no withdrawal penalty, see Choi, Laibson, and Madrian (2005).

The results confirm our previous findings that having a refund over \$500, having positive investment income, being married, and having a larger income all increase the effects of the match rates on take-up and average contributions. Interestingly, all those covariates, with the important exception of the refund over \$500, have no effect on take-up absent the match. Being married has only a marginal influence with regard to take up (significant at the 10 percent level but not the 5 percent level) but a stronger effect on amounts due to the fact that two individuals have the opportunity to contribute in the case of married filers. In contrast to Table 4, having made positive 401(k) contributions or being a repeat customer no longer has any significant impact on the match rate effects, suggesting that those characteristics do not affect the match rate effects when all the other variables are kept constant. Being a homeowner still shows a negative effect on the 50 percent match rate effect.

C. 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 divide tax professionals according to their experience with X-IRAs during tax season 2005 but before the beginning of the experiment (Table 6, Panel A). Sub-panel A1 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."³⁰ Sub-panel A2 shows the results for clients whose tax professionals had more take-up than the median, whom we label "high experience tax professionals." Unsurprisingly, the take-up in the "no match group" is higher among clients working with a "high experience tax professional," although the difference is small, about 2 percentage points. 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 the experiment began: the effect of a 20 percent match rate on take-up increases from 5.1 percent to 10.4 percent when moving from a "low experience tax professional" to a "high experience tax professional," and the effect of a 50 percent match rate increases from 10.6 percent to 17.4 percent.

Panel B in Table 6 uses the tax professional'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 a "high experience tax pro" when this fraction is above the median (5.4 percent), and we

³⁰ 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 in 2005, but before March 5th.

run the same specification as in panel A. The results are quite striking: the treatment effects are much larger for tax filers served by tax pros with relatively high X-IRA take-up rates. For example, the effect of the 50 percent match rate (relative to no match) is only 5.4 percentage points for tax filers served by a low X-IRA take-up tax pro. In contrast, the effect of the 50 percent match is 16.9 percentage points for clients served by a high X-IRA take-up tax pro. This is an indication that there are very strong tax professional effects.³¹

Column (1) in Table 7 formally shows that the tax professional effects displayed in Table 6 are statistically significant. Column (1) shows the coefficients of a regression of X-IRA take-up on a constant, dummy variables for the two match rates, a dummy variable for whether or not the tax professional was above median (based of pre-experiment experience in Panel A and during the experiment in Panel B), and the interaction of this tax professional experience dummy with the two match rates. The coefficients on the interactions are all statistically significant.

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 the take-up rates when the experiment started. Therefore, we repeat in column (2) the same regression but controlling for a full set of office dummies, as well as all the individual variables included in Table 5, and their interaction with the treatment dummies.³² The coefficients on the interaction variables in column (2) are very close to those reported in column (1), suggesting that the effects we obtained were not due to differences in those observable characteristics, or to unobservable characteristics varying at the level of the office.

To attempt to control for unobserved characteristics, we further restrict the sample to tax filers who did not file their tax return with H&R Block in 2004, since repeat tax filers may choose a tax professional they particularly like. New tax filers are often assigned to the next available tax professional when they arrive at an H&R Block office, so that within an office, the assignment of a new tax filer to a particular tax professional should not be related to the tax pro and the client's characteristics, ameliorating the concern that clients who are more likely to take up the match are working with the same tax pros.³³ Column (3) displays the results with no controls and column (4) shows the results with controls. The drawback of the new customer sub-sample is that it is much smaller. The results nonetheless persist in this sub-sample, especially in Panel B.

³¹ 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).

³² To save space, we do not report these coefficients.

³³ An exception would be when a new client is referred to a tax professional by a friend.

Therefore, taken together, all the results from Table 7 suggest that individual decisions about savings are affected by external cues such as the recommendation of tax preparers. This may reflect a lack of information or understanding, which the tax professional can clarify more or less well, or simply the importance of the tax professional, who is both close to being a peer and in a position to give advice (see Duflo and Saez (2003) for evidence of the role of information and peer effects in financial decisions).

D. 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 accumulation.³⁴ While a full analysis of this effect is beyond the scope of this paper, a first step is to analyze whether X-IRA contributions are withdrawn or whether they stay in the accounts.³⁵ The matching contributions were deposited in the X-IRA accounts on April 15th. As of May 2nd, only 18 of the almost 1,400 X-IRA individual accounts opened during the experiment had experienced any withdrawals. The fact that less than 1.5 percent of contributors made immediate withdrawals just after the match was deposited implies that gaming the program was not the main motivation of contributors.

Table 8 explores withdrawal activity and X-IRA balances as of August 1, 2005, four months after the experiment ended.^{36 37} Panel A reproduces the initial effects on take-up and contributions from Table 2. Panel B displays the same statistics as of August 1, 2005. Four months after the experiment, the differences across groups fully persist. For example, those in the 50 percent match group had contributions (inclusive of match) 10.5 times higher than those in the no match group at the time of tax preparation. As of August 1, those contributions net of withdrawals were still 10.7 higher for the match group.

³⁴ 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).

³⁵ 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 Roth IRA withdrawals. It is therefore important to assess whether withdrawals are important and whether they vary by match rate groups.

³⁶ The amounts contributed inclusive of the match reported in Table 8 differ from those reported in Table 2 because, as we mentioned earlier, H&R Block matched the full contributions instead of the first \$1,000 of contributions. The amounts reported here are the full amounts deposited into the account as of April 15.

³⁷ More precisely, we measure withdrawals as the difference in balances between August 1st, 2005 and January 1st, 2005 less the contributions (inclusive of the match) made during tax season. We also note that it was not possible to obtain balance information in a small number of cases, most likely due to clients or administrative errors in reporting. We assumed no withdrawal activity for observations where balance information could not be obtained.

Panel C reports statistics on the withdrawal activity from April 15th to August 1st. It shows that the fraction of contributors making withdrawals was 10.9, 13.8, and 14.4 for each of the three groups (no match, 20 percent match, and 50 percent match) and that the differences are not statistically significant. While the average amount withdrawn increases with the match rate, the fraction withdrawn (as a percentage of initial contributions and match) does not. In summary, there is no evidence (as of August 2005) that those who made X-IRA contributions because of the match are more likely to withdraw their contributions.

Panel D explores withdrawal activity in more detail. The standard fully informed rational model predicts that three factors should be particularly relevant. First, contributions made before April 15th can be counted as either “2004 tax year” or “2005 tax year” contributions. Contributions for the 2004 tax year are deducted from 2004 income in the calculation of 2004 AGI. Contributions for the 2005 tax year have no immediate tax implication: the net balance of money contributed to an IRA by the end of 2005 will enter in the calculation of 2005 AGI. Therefore, X-IRA contributions made for tax year 2005 can be withdrawn during calendar year 2005 with no tax penalty (the IRS would consider that such a contribution never took place). Thus, the best gaming strategy would have consisted of making 2005 contributions and withdrawing the money, including the match, during the same year. Panel D shows, however, that over 90 percent of tax filers who contributed chose to apply their contributions to tax year 2004, with no significant differences across treatment groups, suggesting that most taxpayers were not aware of this gaming opportunity or did not choose to take advantage of it. Those making withdrawals are indeed more likely to have made 2005 contributions but this effect is not stronger in the matching groups.

Similarly, those aged 59.5 and above can make withdrawals free of tax penalty. We can note from Figure 3 that crossing that age threshold did not increase the effects of take-up significantly. Panel D in Table 8 shows that those aged 60 and above in the matching groups are actually *less* likely to withdraw their contributions even though they face no tax penalty.

Finally, Roth IRAs withdrawals face the 10 percent tax penalty only on the investment returns, and not on the principal (since Roth IRA contributions are not deducted from AGI). After only four months, returns are negligible relative to principal. Using the Roth IRA was thus another easy possibility for gaming the system: Filers could contribute to a Roth, trigger the match, and then withdraw the full amount with only a trivial penalty on the investment returns in the meanwhile. The fraction of Roth contributors, however, was only slightly higher in the treatment groups (60 percent) than in the control group (54 percent). Roth IRA contributors were indeed more likely to make withdrawals in all three groups, but there is no difference across groups.

In summary, tax filers did not appear to game the system at the time of tax preparation, and did not withdraw funds disproportionately in the months following the contribution. It will be important to continue monitoring withdrawals and new

contributions to understand medium-term impacts. Perhaps the most striking point is how poorly the simple fully-informed rational model can explain either contribution or withdrawal patterns: taxpayers were very far from taking full advantage of what could have been perceived as a “free lunch” opportunity.³⁸

E. Lessons from Focus Groups

On May 4th and 5th, 2005 (one month after the end of the experiment), H&R Block arranged focus groups with tax professionals and clients to help generate insights into the processes that generated the experimental results. The focus groups were professionally moderated using a question guide developed in conjunction with our research team. The sessions were observed through a one-way mirror by a member of our research team and by key H&R Block personnel. On May 4th, three two-hour focus groups were held, each with roughly half-a-dozen tax professionals. The three groups were assembled to represent tax professionals with low, high, and mixed X-IRA take-up rates during the experiment. On May 5th six one-hour focus groups were held, each with 3 clients as were two 30 minute one-on-one sessions. Three of the groups sessions and both of the one-on-one sessions were with clients who had been offered the 50 percent match, but declined it. Two other group sessions were with taxpayers who had been offered the 50 percent match and had accepted it. The final session was with control group clients who opened X-IRAs. The sessions were extremely useful for understanding the context of the experiments – such as the skills of tax pros, the different ways in which tax pros described the X-IRA product to the clients, and the ways in which tax pros thought about which clients an X-IRA was appropriate for — and supplemented our direct field observation during the period in which the experiment was in the field. Here we focus on three issues: 1) Why people turned down the 50 percent match; 2) how take-up rates might be different if the federal government established a national program similar to our intervention; and 3) whether contributions represent new savings. The small sample sizes and the possibility of non-random selection in who agreed to participate in the focus groups means that the focus group results should be interpreted as suggestive of issues that need to be investigated in future empirical work.

It appears that an important reason that people did not take up the offer was that they had already decided to use their tax refund for a different purpose. One client reported accumulating credit card debts during the Christmas season with the intention of using the tax refund to pay it off. Another reported letting her property tax bill be delinquent for a few weeks until she could use her tax refund to pay it off. Others had simply decided before they reached the office to allocate it to buying clothes or a new (used) car. It is possible, therefore, that if people were alerted ahead of time that they would have the opportunity to have their saving matched, we might observe higher take-up rates. A second group of clients were suspicious about the offer. One client thought it

³⁸ One should not interpret this as evidence that gaming would be insignificant if a policy like this were implemented nationwide. Our set-up – with taxpayers receiving an unexpected offer during a short period of time – limited opportunities for learning. Moreover, evidence from our focus group suggested that the standard IRA “there may be penalties for early withdrawal” language seems to have a big impression on people; it did not occur to clients that these “penalties” might be trivial or zero. Fear of these “penalties” also seems to discourage some taxpayers from contributing in the first place.

sounded too good to be true; another was puzzled about receiving the offer from a firm he perceived to be in the tax return business, not the financial services industry; a third did not think Block's investment options were very good; and a fourth did not want their savings split among different financial institutions. It is likely that a federal matching program would be more appealing to this group than the offer from Block was. A third group of clients was scared off by the penalties for early withdrawal or the fear that they would end up borrowing against the account and then be in debt to themselves. A fourth group of taxpayers were temporarily poor (one was unemployed and another's husband had just quit work to start graduate school), and thus it was probably not optimal for them to be doing saving. A fifth group of taxpayers consisted of people who were in a rush to get their tax return done as fast as possible (particularly people who brought their children with them to the Block office) and did not want to spend time on anything extraneous. A final group consisted of people who were financially unsophisticated – for example, people who had never heard of an IRA or who were generally uncomfortable with the idea of owning financial assets.

In addition to the lack advance notification about the match opportunity and the suspicion about the offer, there are several other reasons to think a federal program might have larger impacts. First, the limited duration of our program meant that social learning did not occur to any significant extent. Clients reported trying to tell their neighbors and colleagues about the opportunity but finding out that their friends had already filed their tax return. Second, tax pros and clients both said the program would have functioned better if it had been accompanied by a television advertising campaign. Third, tax pros said they were just getting comfortable with the product by the end of our one month experiment and wish they could have offered it through the tax season. Fourth, tax pros said it often takes two or three years of explaining a new product to a client to get them to be comfortable enough with it to try it.

There are, however, reasons a national program could show smaller effects. One is that Block is the only tax preparation firm serving a large low and moderate income client base capable of opening an IRA for a client at the time of tax preparation. Thus the overwhelming majority of taxpayers who have their returns done elsewhere and who, therefore, would not be able to direct their tax refunds into an IRA on the spot might not be as responsive to a tax incentive. A second is that there may have been excitement effects for people who “won the lottery” that mean that our results overstate what would be observed if this program were taken national. We tried to limit these effects by offering people in the control group a fee-free IRA – so that the tax pros could describe everyone as winning something. The tax pros said that people in the control group rarely knew that they had missed out on receiving a match and that tax pros tried hard to make everyone feel like a winner. So the existing evidence points to this having only a small impact, but it is impossible to rule out a larger impact.

In the focus groups, X-IRA contributors were asked what they would have done with their tax refund had they not received the match. While some described scenarios that sounded as if they would have consumed the tax refund instead, there were also several clients who gave answers that made it sound like the IRA contributions were not

net new savings. In particular, one claimed that she would have put the money in an IRA in the bank instead. Another said she contributed rather than paying down credit card and hospital bills. Clearly this is a topic that needs further study in the future.

4. Comparison with the Saver's Credit

A. The Saver's Credit Program

In this section, we compare the effects reported above to those of the Saver's Credit, which provides, over certain income ranges, matching incentives for low- and middle- income tax filers that are even stronger than those in our experiment. Because the nature and magnitude of the economic incentives in the experiment described above have some similarities to those in the Saver's Credit, the comparison between the impact of this experiment and the impact of the Saver's Credit can provide insight into the relative importance of the pure underlying incentives and the details of implementation in affecting take-up rates.

The Saver's Credit was first implemented in tax year 2002 (for tax returns filed in 2003) and is scheduled to expire after 2006 (tax returns filed in 2007 will be the last to benefit from the program unless it is extended). 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.).³⁹ As shown in Table 9, 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.

It is easy to see that a credit at rate t is economically equivalent to a match rate of $t/(1-t)$. For example, a tax filer facing the 50 percent credit rate and contributing \$1,000 would receive a \$500 tax credit, so that her 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 generates effective match rates of 100 percent, 25 percent, and 11 percent. If the impact of these incentives were similar to those in our matching experiment, the Saver's Credit could potentially generate very large effects on retirement account contributions.

Because the Saver's Credit is non-refundable, however, many low-income tax filers who would qualify for the highest credit rate on the basis of their AGI benefit from the credit 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).⁴⁰ The data allow us to define

³⁹ Those contributions are netted of any withdrawals made during the last three years to limit gaming possibilities.

⁴⁰ 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.

precisely whether a tax filer benefits or could potentially benefit from the Saver's Credit.⁴¹ In what follows, we call those filers who are or could be benefiting "eligible" filers. Those who could not benefit are called "ineligible."

At first glance, comparing retirement savings 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. AGI, however, is *net* of retirement contributions (with the exception of the Roth IRA). As a result, even if gross income (defined here as AGI with all tax-deductible retirement contributions added back) is exogenous, AGI itself is not, since rational tax filers just above the boundary have incentives to contribute to tax-deductible IRAs in order to fall below the boundary and benefit from having a higher credit rate that would apply to all of their contributions.

An implication, however, is that we should expect an abnormally large number of taxpayers bunching at AGIs just below the boundary point. Symmetrically, we should expect to see few taxpayers just above the boundary point, 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, not just the marginal contribution). Finding bunching in the share of filers who contribute among filers just to the left of the boundary point (and a lower contribution rate to the right of the boundary) would thus 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.⁴² 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 offers tax filers a way to fine tune their AGI and contributions to take advantage of the Saver's Credit. At the time of tax preparation, the exact AGI amount is revealed and it is straightforward (if sometimes tedious) to estimate the current Saver's Credit rate and whether a tax filer could benefit from a higher rate by making X-IRA contributions.⁴³

⁴¹ More precisely, we define a tax filer as potentially benefiting from the Saver's Credit if, starting from no X-IRA contributions, his/her tax refund would increase or tax liability decrease due to the Saver's Credit should he/she make an X-IRA contribution.

⁴² Indeed, annual salary contracts do not run in general from January 1st to December 31st. 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, and 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.

⁴³ IRA contributions for a particular year can be made until April 15 of the following year, once the AGI information is revealed, and so are clearly an easier tool for precise tax optimizing than automatic monthly pension contributions.

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; such optimization was not flagged as a specific item in the standard F10 sequence. Some tax professionals understand the Saver's Credit well and can experiment 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.

B. Graphical evidence around the Saver's Credit cliffs

To investigate responses to the Saver's Credit, we use national H&R Block retail tax return data for tax season 2005. The data consist of a 100 percent sample of tax returns with positive X-IRA contributions (about 180,000 returns) and a 9 percent random sample selected based on 9 two-digit endings of the Social Security number of the primary taxpayer (about 1,400,000 returns). In what follows, the data are weighted by the inverse of the sampling probabilities. Since the minimum X-IRA contribution is \$300, we focus on the sub-sample of taxpayers with earnings (wages and salary, alimony, and self employment income) above \$300. We also exclude tax filers aged less than 18 (as of January 1st, 2005) or claimed on someone else's return since such taxpayers are ineligible to claim the Saver's Credit. We multiply the 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 9).

Figure 4 plots the percent of tax units making an X-IRA contribution by \$250 bands of normalized AGI, for normalized AGI ranging between \$20,001 to \$60,000. The sample is further limited to "eligible" tax filers defined as those for whom an X-IRA contribution (real or potential) would trigger a positive Saver's Credit not offset by reductions in other credits.⁴⁴ The three Saver's Credit cliffs, where the implicit match falls from 100 percent to 25 percent, 25 percent to 11 percent, and 11 percent to 0 percent, respectively, are indicated by vertical lines. The figure shows clear evidence of a behavioral response of X-IRA contributions to the Saver's Credit: there are three visible spikes in the fraction of taxpayers contributing just below each of the three cliffs.

To further test the hypothesis that the spike is due to the Saver's Credit, we compare "eligible" tax units from Figure 4 to "ineligible" tax units whose X-IRA contribution would not trigger a Saver's Credit because they have no income tax liability (net of other credits) to offset. Ineligible tax units have more children on average, are less likely to be single, and more likely to itemize deductions, each of which reduces tax liability. Figure 5 shows the likelihood of being an X-IRA contributor for eligibles (dark

⁴⁴ Because of the interaction with the Child Tax Credit, about 3 percent of eligible tax filers do not actually benefit from the Saver's Credit even if their return shows a positive Saver's Credit (because the Saver's Credit is offset one for one by a reduction in the Child Tax Credit).

shading) and ineligible (light shading) around the first two cliffs of the Saver's Credit.⁴⁵ The figure appears consistent with the tax explanation. As we saw in Figure 4, the spikes are clearly visible for eligible tax units, and the contribution rates are higher in the 100 percent match bracket than in the 25 percent and 11 percent match brackets. In contrast, there are no spikes for the ineligible and the contribution rate is about the same across the 100 percent, 25 percent, and 11 percent match brackets.

Figures 4 and 5 also show that, while there is a drop in the fraction of contributors to the right of the threshold, a relatively large fraction of taxpayers is contributing in those income ranges. In principle, many of the filers whose income is just above the threshold 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. In particular, taxpayers with 401(k) contributions just above the 100 percent match cliff have very strong incentives to make potentially modest X-IRA contributions in order to increase the credit rate on their pre-existing retirement contributions. In some of those cases, an X-IRA contribution can even increase the tax refund *net* of the contribution. However, the X-IRA take up rate among taxpayers facing this situation is only around 6.2 percent, showing that the vast majority of those fortunate taxpayers actually fail to exploit this opportunity to be paid for saving.

Figure 6 plots separately the percent of tax returns contributing to traditional and Roth IRAs by normalized AGI income bands for eligible tax units. Contributing to a Roth IRA does not change AGI (as a Roth IRA is an after-tax contribution). Consistent with the incentive explanation, we observe only a level effect in Roth IRA take-up above and below the 100 percent cliff with no spike just below the cliff. The fraction contributing to a Roth X-IRA increases from about 0.5 percent in the 25 percent match bracket to about 1 percent in the 100 percent match bracket, consistent with a small response to the incentives. Therefore, the spikes are indeed primarily created by filers making traditional X-IRA contributions in order to take advantage of the higher credit rate.⁴⁶ We also note that some filers contribute to Roth X-IRAs just on the right of the cliff 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 likely not optimizing their tax savings (since it is implausible that 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 7 plots the percent of returns with positive retirement contributions (all IRA types, and other retirement contributions such as 401(k)s, etc.) among all those potentially eligible for the Saver's Credit and then excluding those who made an X-IRA

⁴⁵ Almost all taxpayers around the third cliff have positive tax liability and hence the group of ineligible is too small in that case.

⁴⁶ Indeed, the number of saver's credit eligible tax returns crossing a particular AGI threshold due to their traditional X-IRA contributions peaks precisely at the \$30,000 and \$50,000 cliffs. Furthermore, the fraction of crossers with positive 401(k) contributions also peaks at the \$30,000 and \$50,000 cliffs. This shows that some filers strategically use X-IRAs to make their other retirement contributions qualify for the higher rate, and that such strategic filers account for about half of the cliff spikes.

contribution. The proportion with any retirement contributions (such as 401(k)s, other IRAs, etc.) is much higher than the fraction with X-IRAs (from Figure 4) and increases from about 15 percent to about 20 percent as (normalized) AGI rises from \$25,000 to \$35,000. There is a moderate spike visible at the \$30,000 cliff. However, most of this spike goes away when X-IRA contributors are excluded, and there is no strong evidence that the 100 percent match rate induces more tax filers to make retirement contributions. As we discussed, it is not surprising that tax filers cannot fine tune the Saver's Credit optimization before tax preparation.⁴⁷

Figure 8 shows that the lack of utilization of the credit is not confined to the extensive margin, but that most of those who contribute are far from contributing to the extent that will maximize their refund. Define the fraction of credit used by an X-IRA contributor as the ratio of the actual X-IRA contribution to the amount of potential contribution (before the X-IRA contribution) that would exhaust the Saver's Credit. This maximum amount is attained either when tax liability is exhausted or when total retirement contributions reach \$2,000 (per spouse in case of married filers). Given the significant effective match rate, we would expect to observe a significant fraction of tax filers contributing up to this maximum and having a fraction of credit used equal to 100 percent if households were optimizing their contributions. Figure 8 depicts the density distribution of the fraction of credit used for taxpayers getting the Saver's Credit, contributing to an X-IRA, and having normalized AGI between \$20,000 and \$30,000 (so that all those taxpayers benefit from the 100 percent match rate). The graph shows that there is indeed a spike in the distribution around 100 percent, but this spike is very small, with less than 3 percent of those taxpayers exactly maximizing their Saver's Credit. The vast majority of tax filers actually contribute less than the maximum amount eligible for the credit. This is another important difference between the Saver's Credit and our experiment, where about 50 percent of those who contributed were bunching at the \$1,000 contribution level, which maximized their match amount.

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

The graphical analysis from Figures 4 and 5 suggests that simple comparisons could be made to estimate the effects of the changes in the match rates generated by the Saver's Credit 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. It is important to keep in mind that those estimates will be an upper bound on the pure match rate effect due to the endogenous "piling up" effect below the threshold that we described above: the contributions are artificially high to the left of the threshold because it is precisely by contributing that the taxpayers are reaching this position. We do not attempt to correct for this effect, since, as we will see, the upper bound we obtain is already quite low compared to the results of the experiment.

⁴⁷ Comparing 401(k) and IRA participation rates among low and moderate income earners across years (before and after the saver's credit was implemented) could provide suggestive evidence on the effects of the Saver's Credit on overall retirement contributions. The evidence would only be suggestive because it could be confounded by non saver's credit related time trends.

We consider five normalized AGI groups of interest: (1) AGI between \$27,501 and \$30,000 (entitled to a 100 percent match rate), (2) AGI between \$30,001 and \$32,500 (entitled to a 25 percent match rate), (3) AGI between \$32,501 and \$35,000 (entitled to a 11 percent match rate), (4) AGI between \$45,001 and \$50,000 (entitled to a 11 percent match rate), and (5) AGI between \$50,001 and \$55,000 (entitled to a 0 percent match rate). We consider relatively narrow AGI groups around the cliffs in order to compare tax filers with relatively similar income levels. Columns (1) to (5) in Panel A of Table 10 display the X-IRA take-up rate, average X-IRA contributions, and conditional X-IRA contributions for those five groups. Finally, columns (6) to (8) display the differences between the 100 percent group and the 25 percent, between the 25 percent and the 11 percent group, and between the 11 percent and 0 percent group, respectively. All standard errors are obtained from a weighted OLS regression.

These differences are small but very precisely estimated. Going from the 25 percent to 100 percent match rate increases take up by 1.4 percentage points, from 1.9 percent to 3.3 percent. Going from 11 percent to 25 percent match rates and from 0 to 11 percent match rates increases take-up much more modestly, by 0.3 and 0.4 percentage points respectively. We also find similarly small but significant increases in amounts contributed (unconditional). Amounts contributed conditional on contributions increase modestly (by \$65) around the 25 to 100 percent match rate jump but not around the other two smaller match rate jumps.⁴⁸

Even the differences we find around the largest jump (from the 25 percent to 100 percent match rates) are much smaller than the difference we found between our 20 percent and 50 percent experimental match rates, where we found a 6.3 percentage point difference in take-up and a \$310 difference in amounts contributed (conditional on take-up and inclusive of match). Even though the differences in match rates are smaller in our experiment, and even though the effects of the Saver's Credit we estimate here are upper bounds, the effect of a 50 percent match rate in our experiment on take-up and contribution rates are respectively 4.5 and 4.8 times larger than what we estimate for the Saver's Credit. Furthermore, the simple difference estimates we obtain in Table 10 are likely to overestimate the effect of the match rate due to the artificial "piling up" effect below the cliff. If we exclude boundary tax filers with AGI between \$29,501 and \$30,500, the difference falls further from 1.4 to 1.0 (0.11).

Another reason why the comparisons from Panel A may be an overestimate of the effect of the Saver's Credit 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 filers with different AGIs also have different propensities to save. A simple but admittedly imperfect 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. Such an analysis of ineligible control groups is shown in Panel B of Table 9 for

⁴⁸ The X-IRA amounts contributed are implicitly inclusive of the equivalent match rate because the Saver's Credit is formally a credit rebate instead of a match.

the bottom three groups.⁴⁹ For those groups, there are much smaller (and insignificant) differences in X-IRA take up rates and contribution amounts, suggesting that the differences from Panel A were indeed caused by the Saver's Credit.

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 1.2 percentage points for the 100 percent match rate and very close to zero and insignificant for the 11 percent or 25 percent match groups. 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, eligible tax units tend to have fewer dependent children and are more likely to be single, and hence perhaps have a lower taste for savings. Differences in amounts are equally small and barely significant.

Finally, columns (6) and (7) in Panel C display the differences-in-differences estimates. These are unbiased estimates of the differential impact of the 100 percent rate versus 25 percent rate and the 25 percent versus the 11 percent match rate under the assumption 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. The difference-in-differences estimates for the 100 versus 25 percent match comparison are quite close to the simple difference estimates from Panel A. The difference-in-differences estimates for the 25 versus 11 percent match comparison also remains of similar magnitude but becomes insignificant, again suggesting that the causal effect of the Saver's Credit is at best very small.

All these results suggest modest effects of the Saver's Credit on take-up and even more modest effects on amounts contributed. This contrasts with the results of our experiment, which suggested large effects from a matching program. A first explanation could be that the population on which we estimated the impact of the Saver's Credit is different from the population on which we estimated the impact of the experiment; for example, those who file early in the season may be more likely to be liquidity constrained than those who file in the mid-season as in the experiment. However, if we restrict our analysis of the Saver's Credit to those filing between March 5th and April 5th, and compare take-up in the 100 percent match rate versus 25 percent match rate groups, we still obtain an increase in take-up of only 1.33 (s.e. 0.29) percentage points (relative to 1.29 percentage points for all filers), suggesting that mid-season tax filers' response is actually very similar to the overall response.

⁴⁹ Almost all tax filers around the third cliff of the saver's credit (above which the effective credit rate falls to zero) have positive tax liability and are therefore eligible, making it impossible to obtain control groups large enough for statistical analysis.

⁵⁰ Such a hypothesis could possibly 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.

The pattern of responses to the Saver's Credit by taxfiler characteristics is similar to what we find in the randomized matching experiment. Having a saving account, reporting positive capital income, contributing to a 401(k) plan, being married filing jointly, being served by a tax professional with more experience offering X-IRAs, and especially having a refund larger than \$500, is associated with larger match rate effects around the 100 percent to 25 percent cliff. In all cases, however, the effects of the Saver's Credit effective match rate remain much smaller than the effects of the matching rates in our randomized experiment, implying that the characteristics of tax filers from St. Louis cannot account for the discrepancy. We propose tentative explanations in the conclusion.

5. 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 7 times higher with a match rate of 20 percent and 50 percent (respectively) than with no match. With matches included, IRA deposits were 4.5 and 10 times higher than with no match.

Second tax professionals appear to 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 during the experiment itself.

Third, 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, and relatively few contributors, even among those benefiting from a match rate, made withdrawals in the weeks following the experiment. Four months after the end of the experiment, about 90 percent of the differential effects of match rates on contributions are still present. Note that a permanent, public program could have "anti-gaming" features, which were not in place in this case.

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 very modest. The upper bound of the difference in take-up rates between a 100 percent and a 25 percent effective match rate in the Saver's Credit is only 1.3 percentage points. Therefore, a researcher who would have relied solely on the evidence from the Saver's Credit would have concluded that large matching incentives are not successful in inducing low- and moderate-income families to make retirement contributions, and therefore that matching incentives are unlikely to represent a desirable policy option to improve the financial security of future retirees. Our experiment leads to very different

results and policy conclusions and it is therefore important to understand the reasons for the discrepancy.

We suspect 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.⁵¹ The differential responses thus may represent another piece of evidence suggesting that framing effects are important for understanding behavioral responses and that the very idea of "elasticity" with respect to effective tax rates is somewhat unclear. Indeed, the evidence reported in this paper suggests that the same elasticity, estimated with respect to similar economic incentives but in different contexts, varies by a factor of more than four. As optimal policy making depends sensitively on such behavioral responses, an important task for future empirical work is to go beyond merely estimating the size of behavioral responses in specific contexts and start exploring the factors that shape the size of the behavioral response.⁵²

In line with our findings that framing and cues are a very important dimension of the decision to enroll, the effect of the match remains smaller than the default enrollment effect estimated by Madrian and Shea (2001). After one year, the participation rate in a cohort enrolled in the 401k by default was 60 percentage points higher than for those who were not enrolled by default, and, while the effect diminishes over time, it was still 30 percentage points higher after four years, even though the firm offers a match. Nevertheless, our results suggest that when default enrollment is not possible, a clear, simple matching contribution scheme would be an effective way to increase participation in, and contributions to, retirement savings accounts.

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 middle- and low-income households. This strategy would not, however, produce contribution rates anywhere near 100 percent.

⁵¹ 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. However, a 50 percent credit rate should be at least as effective as a 50 match rate, so this is not the entire explanation.

⁵² It should also be noted that the experiment was a one-time offer while the saver's credit is a program available each year (from 2002 to 2007). Ignoring gaming, rational individuals should respond more to a one-time offer than to a permanent program. Providing more information on saver's credit opportunities would be a way to tell apart the information explanation from the one-time versus permanent program explanation.

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.
- Choi, James J., David Laibson, and Brigitte C. Madrian (2005). "\$100 Bills on the Sidewalk: Suboptimal Saving in 401(k) Plans." NBER working paper no. 11554. August.
- Duflo, Esther and Emmanuel Saez (2003) "The Role of Information and Social Interactions in Retirement Plan Decision: Evidence from a Randomized Experiment" *Quarterly Journal of Economics*, 118(3).
- 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.

Mills, Gregory, William G. Gale, and Rhiannon Patterson (2005). *Effects of Individual Development Accounts on Household Saving Behavior: Evidence from a Controlled Experiment*. Abt Associates and the Brookings Institution. April.

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.

Table 1: Descriptive Statistics

	Mean (Standard error)			Significance of the difference in means: t-stats		
	No match	20% match	50% match	20% vs. no match	50% vs. 20% match	50% vs. no match
	(1)	(2)	(3)	(4)	(5)	(6)
Fraction of returns offered X-IRA	0.748 (0.007)	0.741 (0.007)	0.747 (0.007)			
This tax pro made no offer during experiment	0.005 (0.001)	0.004 (0.001)	0.006 (0.001)			
Fraction of returns with X-IRA by tax pro 01/01-03/04	0.022 (0.000)	0.022 (0.000)	0.022 (0.000)			
Adjusted Gross income (plus Trad. IRA contributions)	42,891 (525)	44,378 (567)	43,433 (572)			
Adjusted Gross income, married taxpayers (filing jointly)	67,961 (986)	68,410 (971)	68,756 (1017)			
Adjusted Gross income, non married taxpayers	29,134 (439)	29,910 (534)	29,539 (542)			
Fraction married, filling jointly	0.354 (0.007)	0.376 (0.007)	0.354 (0.007)	2.145	2.150	
Overpayment amount	1582 (26)	1552 (26)	1561 (26)			
Fraction with overpayment above \$500	0.665 (0.007)	0.662 (0.007)	0.656 (0.007)			
Fraction with investment income	0.448 (0.007)	0.461 (0.007)	0.458 (0.007)			
Fraction homeowners	0.449 (0.007)	0.424 (0.007)	0.448 (0.007)	2.445	2.316	
Fraction filed with H&R Block in the preceding tax year	0.711 (0.007)	0.715 (0.007)	0.714 (0.007)			
Fraction EITC recipients	0.174 (0.006)	0.161 (0.005)	0.173 (0.005)			
Fraction with savings account	0.624 (0.007)	0.631 (0.007)	0.621 (0.007)			
Fraction with 401(k) contributions	0.268 (0.006)	0.269 (0.007)	0.269 (0.006)			
Number of observations	4719	4521	4722			

This table displays descriptive statistics for the three experimental groups (columns 1, 2, 3) and reports t-stats when differences across groups are significant at 5% level (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 3 reports the average over tax pros of the fraction of returns for which an X-IRA was sold 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 X-IRA contributions (as X-IRA contributions are affected by the experiment). Home and savings account ownership data are collected during tax preparation.

Investment income is defined as interest income (taxable and non taxable) and all dividends.

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

	Mean (standard errors)				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 TAX RETURNS							
Opened an X-IRA (%)	2.13 (0.05)	2.90 (0.24)	7.72 (0.40)	13.98 (0.50)	4.82 (0.46)	6.26 (0.65)	11.07 (0.56)
Amount contributed (\$) (unconditional)	8.8 (0.3)	22.2 (3.1)	85.1 (6.1)	154.9 (7.4)	62.9 (6.7)	69.8 (9.6)	132.7 (8.0)
Amount contributed (\$) (positive contributions only)	382.1 (8.2)	765.1 (84.0)	1102.3 (54.9)	1108.2 (34.4)	337.2 (102.2)	5.9 (61.9)	343.1 (84.7)
Amount contributed, inclusive of match (unconditional)		22.2 (3.1)	98.8 (6.9)	222.3 (10.1)	76.6 (7.4)	123.5 (12.3)	200.1 (10.5)
Amount contributed, inclusive of match (positive contributions only)		765.1 (84.0)	1280.3 (59.7)	1590.7 (43.6)	515.2 (108.9)	310.5 (74.0)	825.6 (103.0)
PANEL B: MARRIED TAX RETURNS ONLY							
Opened at least one X-IRA (%)	1.66 (0.10)	2.39 (0.37)	9.36 (0.71)	17.81 (0.94)	6.97 (0.80)	8.45 (1.17)	15.42 (1.01)
Amount contributed (\$) (unconditional)	9.6 (0.8)	32.6 (7.5)	139.9 (14.0)	266.9 (17.9)	107.3 (15.9)	127.0 (22.7)	234.3 (19.4)
Amount contributed (\$) (positive contributions only)	480.7 (30.7)	1364.4 (230.0)	1494.8 (98.4)	1498.4 (62.3)	130.5 (227.4)	3.6 (111.5)	134.1 (189.6)
Amount contributed, inclusive of match (unconditional)		32.6 (7.5)	162.1 (15.8)	381.0 (24.2)	129.5 (17.5)	218.9 (28.8)	348.3 (25.4)
Amount contributed, inclusive of match (positive contributions only)		1364.4 (230.0)	1732.1 (106.2)	2138.7 (76.8)	367.7 (241.1)	406.6 (130.7)	774.3 (225.9)
Both spouses opened an XIRA (%)	0.01 (0.01)	0.36 (0.15)	4.12 (0.48)	8.37 (0.68)	3.76 (0.51)	4.25 (0.83)	8.01 (0.69)
Amount contributed, secondary taxpayer (\$) (unconditional)	1.4 (1.0)	5.4 (2.7)	49.5 (6.9)	90.9 (8.3)	44.1 (7.4)	41.3 (10.8)	85.5 (8.8)
PANEL C: NON MARRIED TAX RETURNS ONLY							
Opened an X-IRA (%)	2.24 (0.05)	3.18 (0.32)	6.73 (0.47)	11.87 (0.59)	3.55 (0.56)	5.14 (0.76)	8.69 (0.67)
Amount contributed (\$) (unconditional)	8.6 (0.3)	16.5 (2.4)	52.1 (4.8)	93.4 (5.6)	35.6 (5.3)	41.3 (7.4)	76.9 (6.1)
Amount contributed (\$) (positive contributions only)	364.3 (7.9)	518.0 (55.6)	773.8 (46.6)	786.9 (26.2)	255.8 (76.4)	13.1 (49.5)	269.0 (58.3)
Amount contributed, inclusive of match (unconditional)		16.5 (2.4)	60.7 (5.4)	135.3 (7.7)	44.3 (5.8)	74.6 (9.6)	118.8 (8.1)
Amount contributed, inclusive of match (positive contributions only)		518.0 (55.6)	902.2 (50.0)	1139.6 (32.8)	384.2 (80.5)	237.5 (58.0)	621.6 (69.6)

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).

Table 3: Effects on X-IRA behavior by income and tax refund size

	Mean (standard errors)			Difference (standard errors)		
	No match	20% match	50% match	50% match		
				20% match vs no match	vs 20% match	50% match vs no match
	(1)	(2)	(3)	(4)	(5)	(6)
A. INCOME QUARTILES						
A.1 Bottom quartile (\$0-\$35,000 married, \$0-\$8,500 others)						
Opened an XIRA	2.49 (0.54)	4.68 (0.75)	7.53 (0.89)	2.19 (0.91)	2.86 (1.18)	5.05 (1.05)
Amount contributed (unconditional)	14.8 (4.9)	49.2 (12.2)	76.2 (10.9)	34.4 (12.9)	27.0 (16.3)	61.4 (12.1)
A.2 Second quartile (\$35,000-\$58,300 married, \$8,500-\$18,400 others)						
Opened an XIRA	2.92 (0.52)	8.69 (0.90)	12.83 (1.05)	5.76 (1.02)	4.14 (1.38)	9.91 (1.16)
Amount contributed (unconditional)	33.8 (9.8)	89.6 (14.0)	120.9 (12.1)	55.9 (16.9)	31.2 (18.5)	87.1 (15.5)
A.3 Third quartile (\$58,300-\$88,200 married, \$18,400-\$32,500 others)						
Opened an XIRA	3.44 (0.50)	7.04 (0.72)	15.41 (1.00)	3.59 (0.87)	8.37 (1.23)	11.97 (1.12)
Amount contributed (unconditional)	26.3 (5.9)	77.4 (10.8)	162.4 (14.2)	51.1 (12.2)	84.9 (17.9)	136.1 (15.4)
A.4 Fourth quartile (\$88,200+ married, \$32,500+ others)						
Opened an XIRA	2.65 (0.41)	9.31 (0.76)	17.24 (0.97)	6.66 (0.86)	7.93 (1.24)	14.59 (1.06)
Amount contributed (unconditional)	14.7 (3.3)	108.2 (11.5)	216.9 (16.3)	93.4 (11.8)	108.7 (20.1)	202.2 (16.7)
B. EITC Status						
B.1 No EITC						
Opened an XIRA	2.44 (0.25)	7.73 (0.43)	14.21 (0.56)	5.29 (0.50)	6.48 (0.71)	11.77 (0.61)
Amount contributed (unconditional)	20.1 (3.4)	92.0 (7.1)	171.2 (8.7)	71.9 (7.8)	79.2 (11.3)	151.0 (9.4)
B.2 Positive EITC						
Opened an XIRA	5.10 (0.77)	7.67 (0.99)	12.88 (1.17)	2.57 (1.24)	5.21 (1.55)	7.78 (1.40)
Amount contributed (unconditional)	32.1 (7.3)	49.3 (8.0)	77.0 (8.6)	17.2 (10.9)	27.7 (11.9)	44.9 (11.3)
C. Refund Status						
C.1 Refund less than \$500						
Opened an XIRA	0.82 (0.23)	3.54 (0.47)	9.35 (0.72)	2.72 (0.52)	5.81 (0.87)	8.53 (0.77)
Amount contributed (unconditional)	15.7 (6.1)	48.1 (7.9)	130.7 (12.9)	32.3 (9.9)	82.6 (15.3)	115.0 (14.3)
C.2 Refund over \$500						
Opened an XIRA	3.95 (0.35)	9.85 (0.54)	16.41 (0.67)	5.90 (0.64)	6.56 (0.86)	12.45 (0.75)
Amount contributed (unconditional)	25.5 (3.5)	104.0 (8.2)	167.6 (9.0)	78.5 (8.8)	63.6 (12.2)	142.1 (9.6)

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 for married filing jointly and all other filers separately.

Table 4: Effects on X-IRA behavior by other tax filers' characteristics

	Mean (standard errors)			Difference (standard errors)		
	No match	20% match	50% match	50% match		
				20% match	vs 20%	50% match
				vs no match	match	vs no match
	(1)	(2)	(3)	(4)	(5)	(6)
A. Savings account ownership						
A.1. No Savings Account						
Opened an XIRA	2.03 (0.33)	5.04 (0.54)	8.38 (0.66)	3.01 (0.62)	3.35 (0.85)	6.36 (0.74)
Amount contributed (unconditional)	13.7 (4.5)	43.1 (6.5)	77.8 (8.4)	29.4 (7.8)	34.7 (10.7)	64.2 (9.6)
A.2 Savings Account owners						
Opened an XIRA	3.43 (0.34)	9.29 (0.54)	17.39 (0.70)	5.86 (0.63)	8.10 (0.89)	13.96 (0.78)
Amount contributed (unconditional)	27.4 (4.1)	109.6 (8.8)	201.9 (10.6)	82.3 (9.6)	92.2 (13.9)	174.5 (11.4)
B. Investment Income						
B.1. No investment income						
Opened an XIRA	3.15 (0.34)	5.45 (0.46)	9.02 (0.57)	2.31 (0.57)	3.57 (0.73)	5.88 (0.66)
Amount contributed (unconditional)	14.3 (2.0)	36.2 (4.7)	66.6 (5.4)	21.9 (5.0)	30.3 (7.2)	52.3 (5.7)
B.2 Positive investment income						
Opened an XIRA	2.60 (0.35)	10.37 (0.67)	19.84 (0.86)	7.77 (0.75)	9.47 (1.09)	17.24 (0.93)
Amount contributed (unconditional)	32.0 (6.4)	142.3 (11.9)	259.5 (14.5)	110.3 (13.4)	117.2 (18.8)	227.5 (16.0)
C. 401(k) contributions						
C.1. No 401(k) contributions						
Opened an XIRA	2.75 (0.28)	7.41 (0.46)	12.32 (0.56)	4.66 (0.53)	4.90 (0.73)	9.56 (0.62)
Amount contributed (unconditional)	20.5 (3.2)	80.3 (7.0)	128.6 (7.5)	59.8 (7.6)	48.3 (10.3)	108.1 (8.2)
C.2 Positive 401(k) contributions						
Opened an XIRA	3.31 (0.50)	8.55 (0.80)	18.49 (1.09)	5.24 (0.94)	9.94 (1.36)	15.17 (1.20)
Amount contributed (unconditional)	26.8 (7.4)	98.0 (12.1)	226.2 (18.1)	71.2 (14.1)	128.1 (22.0)	199.4 (19.6)
D. Home ownership						
D.1. Non homeowners						
Opened an XIRA	2.69 (0.32)	9.02 (0.56)	17.91 (0.75)	6.33 (0.65)	8.89 (0.94)	15.21 (0.82)
Amount contributed (unconditional)	28.4 (5.2)	112.0 (9.3)	225.0 (12.3)	83.6 (10.7)	113.0 (15.4)	196.6 (13.4)
D.2 Homeowners						
Opened an XIRA	3.16 (0.38)	5.95 (0.54)	9.13 (0.63)	2.79 (0.65)	3.18 (0.84)	5.97 (0.73)
Amount contributed (unconditional)	14.6 (2.4)	48.6 (6.7)	68.4 (5.8)	34.0 (6.9)	19.8 (8.8)	53.8 (6.3)
E. New Customer Status						
E.1. New customer						
Opened an XIRA	2.34 (0.41)	5.82 (0.65)	9.76 (0.81)	3.47 (0.76)	3.94 (1.04)	7.42 (0.90)
Amount contributed (unconditional)	17.0 (4.6)	59.0 (10.1)	101.0 (10.2)	42.1 (10.9)	42.0 (14.4)	84.1 (11.1)
E.2 Repeat Customer						
Opened an XIRA	3.13 (0.30)	8.48 (0.49)	15.67 (0.63)	5.35 (0.57)	7.19 (0.80)	12.54 (0.70)
Amount contributed (unconditional)	24.4 (3.9)	95.5 (7.5)	176.5 (9.5)	71.1 (8.4)	81.0 (12.1)	152.2 (10.3)

Notes:

Investment income defined as interest income (taxable and non taxable) and all dividends.

New customer status is based on whether client filed his/her taxes with H&R Block the preceding year.

Table 5: Match Effects on X-IRA take-up and amounts interacted

Dummy variables X	Effect on Take-up (percentage points)			Effects on Amounts (\$)		
	20% match *		50% match *	X	20% match * X	50% match * X
	X	X	X			
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.18 (1.52)	-1.01 (2.15)	1.72 (2.13)	3.83 (21.59)	-29.52 (30.54)	-9.11 (30.26)
Refund above \$500	3.11 (0.83)	3.04 (1.19)	3.10 (1.17)	9.27 (12.02)	44.69 (17.22)	15.21 (16.96)
Married	-0.90 (0.97)	2.31 (1.37)	2.26 (1.37)	4.95 (14.10)	65.95 (19.96)	118.22 (19.96)
Homeowner	0.43 (0.98)	-0.28 (1.40)	-3.21 (1.39)	-6.97 (14.26)	14.38 (20.29)	-19.14 (20.15)
Investment income>0	-0.29 (0.88)	4.54 (1.26)	8.26 (1.25)	15.91 (12.83)	71.66 (18.34)	110.46 (18.10)
Repeat customer	0.92 (0.90)	-0.10 (1.29)	1.12 (1.28)	2.27 (13.12)	-2.69 (18.80)	1.46 (18.52)
401(k) contributions>0	0.66 (0.93)	-1.62 (1.34)	1.45 (1.32)	3.63 (13.49)	-28.93 (19.40)	16.82 (19.13)
Quartile 2	0.19 (1.24)	2.46 (1.79)	2.63 (1.76)	15.19 (18.05)	8.29 (26.04)	1.47 (25.54)
Quartile 3	0.58 (1.24)	0.03 (1.77)	3.60 (1.73)	4.01 (17.96)	2.36 (25.74)	40.00 (25.17)
Quartile 4	-0.38 (1.29)	3.01 (1.85)	4.51 (1.82)	-10.51 (18.77)	49.17 (26.93)	98.30 (26.38)
Adj. R-square	0.0563			0.0569		

Columns (1) to (3) report OLS coefficients of the single regression of an X-IRA take up dummy (normalized to 100) on treatment groups dummies, covariate dummies, and covariate dummies interacted with treatment dummies.

Columns (4) to (6) report OLS coefficients of the single regression of unconditional X-IRA contributions on treatment groups dummies, covariate dummies, and covariate dummies interacted with treatment dummies.

The coefficients in col. (1) can be interpreted as the effect on take-up of switching the corresponding covariate dummy variable from zero to one. The coefficients in columns (2) and (3) can be interpreted as the additional effect of the 20% match and 50% match respectively (always relative to no match) for a tax filer with corresponding covariate dummy equal to one.

Table 6: Effects on X-IRA by tax professional X-IRA experience

	Mean (standard errors)			Difference (standard errors)		
	No match	20% match	50% match	50% match		
				20% match vs no match	vs 20% match	50% match vs no match
	(1)	(2)	(3)	(4)	(5)	(6)
A. PRE-EXPERIMENT TAX PROFESSIONAL X-IRA EXPERIENCE						
A.1 Below median						
Opened an XIRA	1.90 (0.28)	5.09 (0.46)	10.59 (0.63)	3.19 (0.54)	5.50 (0.79)	8.69 (0.69)
Amount contributed (unconditional)	12.3 (2.6)	58.3 (7.2)	119.1 (8.6)	46.0 (7.5)	60.8 (11.3)	106.8 (9.0)
A.2 Above median						
Opened an XIRA	3.91 (0.40)	10.35 (0.64)	17.42 (0.78)	6.44 (0.75)	7.07 (1.02)	13.51 (0.88)
Amount contributed (unconditional)	32.2 (5.6)	111.9 (9.8)	191.3 (12.0)	79.7 (11.2)	79.4 (15.5)	159.1 (13.2)
A. DURING EXPERIMENT TAX PROFESSIONAL X-IRA EXPERIENCE						
B.1 Below median						
Opened an XIRA	1.80 (0.27)	3.63 (0.39)	7.23 (0.53)	1.84 (0.48)	3.59 (0.67)	5.43 (0.60)
Amount contributed (unconditional)	13.7 (3.1)	36.5 (5.8)	69.9 (6.6)	22.8 (6.5)	33.4 (8.8)	56.3 (7.3)
B.1 Above median						
Opened an XIRA	3.99 (0.40)	11.80 (0.68)	20.92 (0.84)	7.81 (0.78)	9.12 (1.09)	16.93 (0.93)
Amount contributed (unconditional)	30.6 (5.3)	133.6 (10.6)	242.3 (13.1)	103.0 (11.7)	108.7 (16.9)	211.6 (14.0)

Panel A divides the sample by tax pro experience in selling X-IRAs before the experiment took place. We divide tax professional according to the fraction of returns with an X-IRA among all the returns they completed before the experiment.

Panel B divides the sample by tax pro experience in selling X-IRAs during the experiment. We divide tax professional according to the fraction of returns with an X-IRA among all the returns completed during the experiment and excluding the current tax filer.

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

	All tax filers		Only new customers	
	No control variables	With control variables	No control variables	With control variables
	(1)	(2)	(3)	(4)
A. TAX PRO XIRA CLIENT TAKE-UP RATES BEFORE MARCH 5, 2005				
20% match	3.19 (0.79)	-2.47 (2.20)	1.72 (1.29)	-0.40 (3.19)
50% match	8.69 (0.78)	-0.45 (2.18)	6.28 (1.26)	1.44 (3.17)
Tax pro above median	2.01 (0.78)	1.53 (0.81)	1.52 (1.27)	0.83 (1.32)
Tax pro above median * 20% match	3.25 (1.12)	3.26 (1.10)	3.41 (1.82)	2.78 (1.80)
Tax pro above median * 50% match	4.82 (1.11)	4.76 (1.09)	2.32 (1.80)	1.65 (1.78)
Observations	13962	13962	4006	4006
B. TAX PRO XIRA CLIENT TAKE-UP RATES DURING THE EXPERIMENT				
20% match	1.84 (0.78)	-3.01 (2.16)	1.98 (1.26)	-0.15 (3.12)
50% match	5.43 (0.77)	-3.11 (2.14)	4.17 (1.24)	-0.48 (3.12)
Tax pro above median	2.10 (0.78)	1.02 (0.79)	2.81 (1.26)	1.15 (1.29)
Tax pro above median * 20% match	5.99 (1.11)	5.69 (1.65)	3.09 (1.82)	3.59 (1.81)
Tax pro above median * 50% match	11.45 (1.10)	8.28 (1.63)	6.81 (1.79)	6.22 (1.77)
Observations	13962	13904	3984	3984

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.

Regressions in cols. (2) and (4) control for all individual variables in table 5, and all the variables interacted with the two match rate dummies; they also controls for a full set (60) of office dummies.

Regressions in (3) and (4) limited to the sample of tax filers who did not file with H&R Block the preceding year.

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).

Table 8: Withdrawals Evidence (4 months after)

				Difference (standard errors)		
	No match	20% match	50% match	20% match	50% match	50% match
				vs no match	vs 20% match	vs no match
	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: Initial Contributions						
Opened an X-IRA (%)	2.90	7.72	13.98	4.82	6.26	11.07
	(0.24)	(0.40)	(0.50)	(0.46)	(0.65)	(0.56)
Initial X-IRA amount inclusive of effective match (unconditional)	\$22	\$102	\$232	\$80	\$130	\$210
	(3.1)	(7.3)	(11.1)	(7.8)	(13.4)	(11.5)
PANEL B: Contributions 4 months after (as of 8/01/2005)						
X-IRA still opened (%)	2.71	7.39	12.88	4.68	5.49	10.17
	(0.24)	(0.39)	(0.49)	(0.45)	(0.63)	(0.54)
Amounts outstanding, inclusive of match (unconditional)	\$19	\$91	\$204	\$72	\$113	\$185
	(2.5)	(6.7)	(10.3)	(7.1)	(12.4)	(10.6)
PANEL C: Withdrawal Activity in 4 months						
Fraction of contributors withdrawing	0.109	0.138	0.144	0.029	0.006	0.035
	(0.027)	(0.018)	(0.014)	(0.034)	(0.023)	(0.032)
Average Amount withdrawn for contributors	\$119	\$143	\$198	\$24	\$55	\$79
	(57.1)	(31.1)	(24.8)	(61.2)	(40.9)	(60.4)
Amount withdrawn/Initial X-IRA balance (percent)	0.156	0.108	0.119	-0.048	0.011	-0.037
	(0.075)	(0.024)	(0.015)	(0.060)	(0.027)	(0.047)
PANEL D: Characteristics of withdrawers						
Fraction contributors making 2005 contributions	0.066	0.092	0.061	0.026	-0.031	-0.005
	(0.021)	(0.015)	(0.009)	(0.028)	(0.017)	(0.023)
Fraction withdrawers with 2005 contributions	0.333	0.125	0.074	-0.208	-0.051	-0.259
	(0.126)	(0.048)	(0.027)	(0.111)	(0.051)	(0.084)
Fraction contributors aged 60 and above	0.109	0.143	0.159	0.034	0.016	0.050
	(0.026)	(0.019)	(0.014)	(0.034)	(0.024)	(0.034)
Fraction withdrawers aged 60 and above	0.267	0.063	0.105	-0.204	0.042	-0.162
	(0.118)	(0.035)	(0.032)	(0.091)	(0.051)	(0.092)
Fraction contributors choosing Roth IRA	0.544	0.602	0.593	0.058	-0.009	0.049
	(0.043)	(0.026)	(0.019)	(0.050)	(0.032)	(0.046)
Fraction withdrawers with Roth IRA	0.633	0.646	0.663	0.013	0.017	0.030
	(0.124)	(0.068)	(0.049)	(0.140)	(0.084)	(0.132)

Notes:

Initial contributions (inclusive of match) differ from those presented in Table 2 because H&R Block mistakenly matched the full contribution instead of only the first \$1,000

Early (non-qualified) withdrawals from Traditional IRAs face a 10% tax penalty (in addition to regular taxes) on principal and return. Early (non-qualified) withdrawals from Roth IRAs face no tax penalty on principal and a 10% tax penalty on return.

Tax filers making a 2005 IRA contribution who withdraw the money during 2005 face no tax penalty at all (not considered as a retirement contribution by IRS). Individuals reaching age 59.5 can withdraw IRA money with no tax penalty.

Table 9: Saver's Credit Parameters

Credit Rate	Equivalent Match Rate	Married Filing Jointly	Head of Household	Single and others
t	t/(1-t)	AGI range	AGI range	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%	\$32,501-\$50,000	\$24,376-\$37,500	\$16,251-\$25,000
0%	0%	\$50,001+	\$37,501+	\$25,001+

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.

Withdrawals from IRAs and 401(k)s (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 married, 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 therefore endogenous.

The savers' 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.

A taxpayer benefits from the saver's credit if its income tax liability net of non refundable credits (excluding the child tax credit)

and net of (full child tax credit-child tax credit which can be recovered with additional refundable children tax credit) is positive.

Taxpayers might have positive saver's credit and not benefit from it because the Saver's Credit might crowd out the non-refundable child tax credit one for one.

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

Table 10: The Effects of the Saver's Credit on X-IRA Behavior

Match rate Credit rate Normalized AGI	Levels					Differences		
	100%	25%	11%	11%	0%	100% vs 25%	25% vs 11%	11% vs 0%
	50%	20%	10%	10%	0%	50% vs 20%	20% vs 10%	10% vs 0%
	\$27.5K-\$30K	\$30K-\$32.5K	\$32.5K-\$35K	\$45K-\$50K	\$50K-\$55K	\$27.5K-\$32.5K	\$30K-\$35K	\$45K-\$55K
	(1)	(2)	(3)	(4)	(5)	(1)-(2)	(2)-(3)	(4)-(5)
A. Savers' credit eligible taxpayers								
X-IRA take up rate (percent)	3.28 (0.083)	1.87 (0.063)	1.55 (0.058)	1.31 (0.043)	0.89 (0.040)	1.41 (0.104)	0.32 (0.086)	0.42 (0.058)
X-IRA average contributions (unconditional)	\$19.2 (0.67)	\$9.8 (0.48)	\$7.9 (0.45)	\$9.0 (0.47)	\$6.4 (0.45)	\$9.5 (0.82)	\$1.8 (0.66)	\$2.6 (0.65)
X-IRA average contributions (conditional on take-up)	\$587 (4.8)	\$522 (6.2)	\$513 (7.0)	\$690 (8.9)	\$722 (12.7)	\$65 (7.9)	\$9 (9.4)	-\$32 (15.1)
Sample size	45,747	45,721	44,508	71,052	61,388	91,468	90,229	132,440
B. Savers' credit ineligible taxpayers (no tax liability)								
X-IRA take up rate (percent)	2.09 (0.110)	2.00 (0.145)	1.86 (1.640)			0.09 (0.182)	0.14 (0.220)	
X-IRA average contributions (unconditional)	\$9.5 (0.64)	\$9.4 (0.90)	\$9.0 (1.05)			\$0.1 (1.09)	\$0.4 (1.38)	
X-IRA average contributions (conditional on take-up)	\$455 (6.3)	\$472 (9.7)	\$485 (12.0)			-\$17 (11.2)	-\$13 (15.4)	
Sample size	17,041	9,301	6,831			26,342	16,132	
C. Difference between eligible and ineligible taxpayers (A-B)						Difference-in-differences		
X-IRA take up rate (percent)	1.19 (0.148)	-0.14 (0.156)	-0.31 (0.164)			1.33 (0.220)	0.18 (0.228)	
X-IRA average contributions (percent)	\$9.7 (1.13)	\$0.3 (1.16)	-\$1.1 (1.23)			\$9.4 (1.67)	\$1.4 (1.70)	
X-IRA average contributions (percent)	\$132 (10.0)	\$51 (14.2)	\$28 (17.3)			\$81 (17.2)	\$23 (22.3)	
Sample size	62,788	55,022	51,339			117,810	106,361	

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

The sample consists of 100% of H&R Block retail returns with X-IRAs, and a 9% random sample of all other returns. All results are re-weighted.

Taxpayers aged below 18 or who are claimed as dependents on somebody else's return are excluded (do not qualify for the savers' credit).

Taxpayers with less than \$300 in earnings are excluded (cannot contribute to an X-IRA with minimum \$300 contribution).

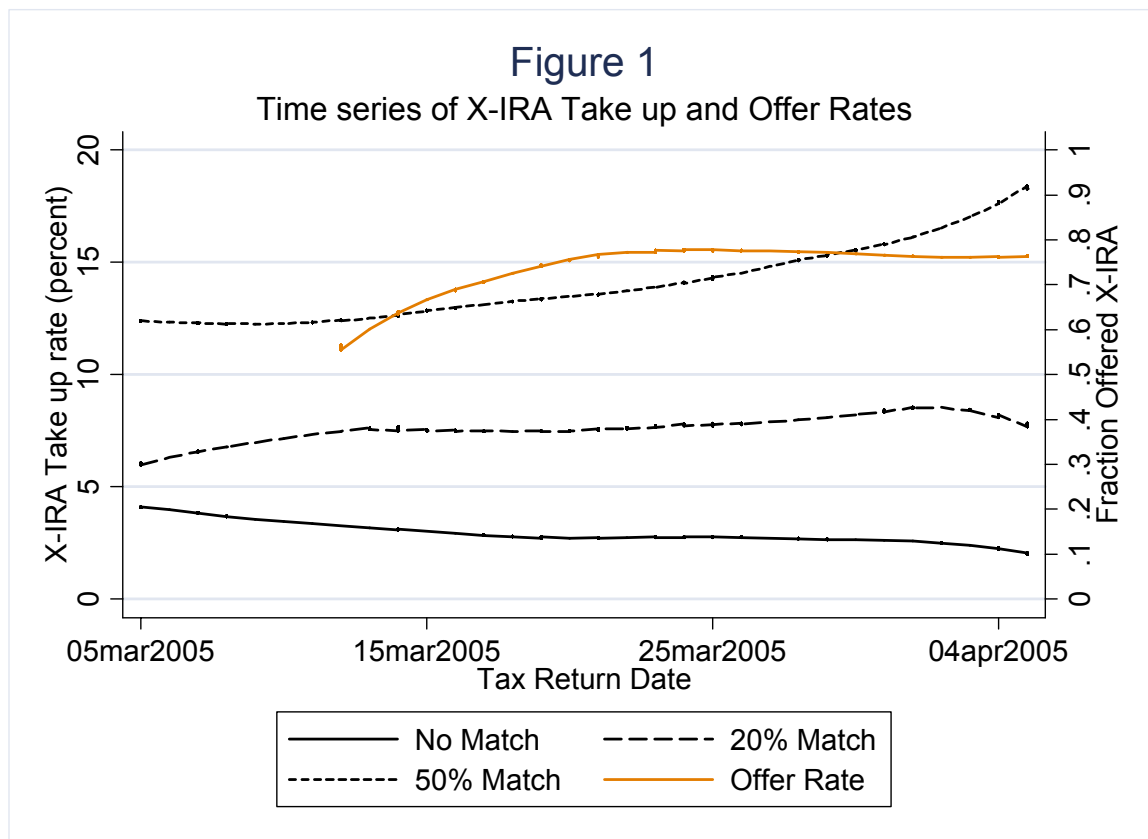
First, groups are defined relative to normalized AGI (Normalized AGI = 100% of AGI for married taxpayers, 133.33% for heads of household, and 200% for singles and others). Columns (1) to (5) display the levels for those five groups and columns (6) to (8) the differences across groups.

Second, groups are defined by savers' credit eligibility (Panel A versus Panel B).

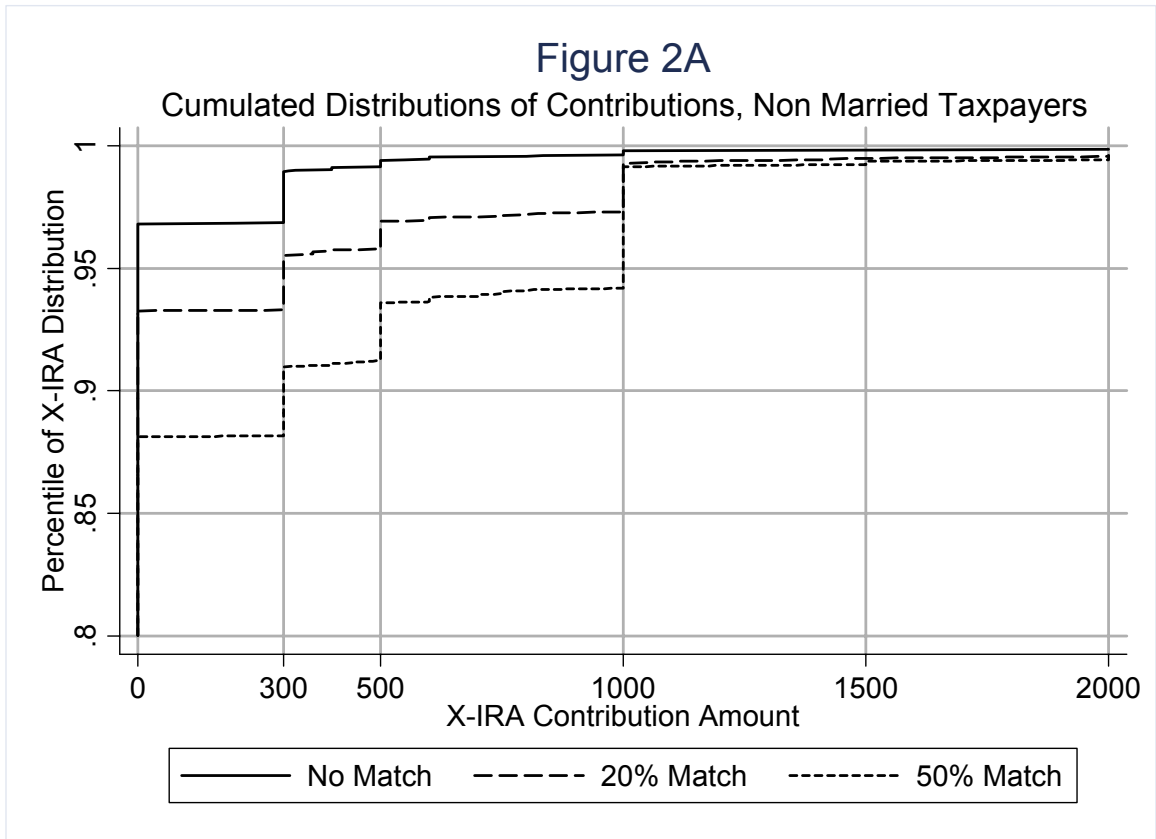
Group A are the eligible, defined as taxpayers whose X-IRA contributions (actual or potential) would benefit from the savers' credit.

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. There is no group B for col. (4) and (5) because almost all taxpayers in that AGI range have positive tax liability.

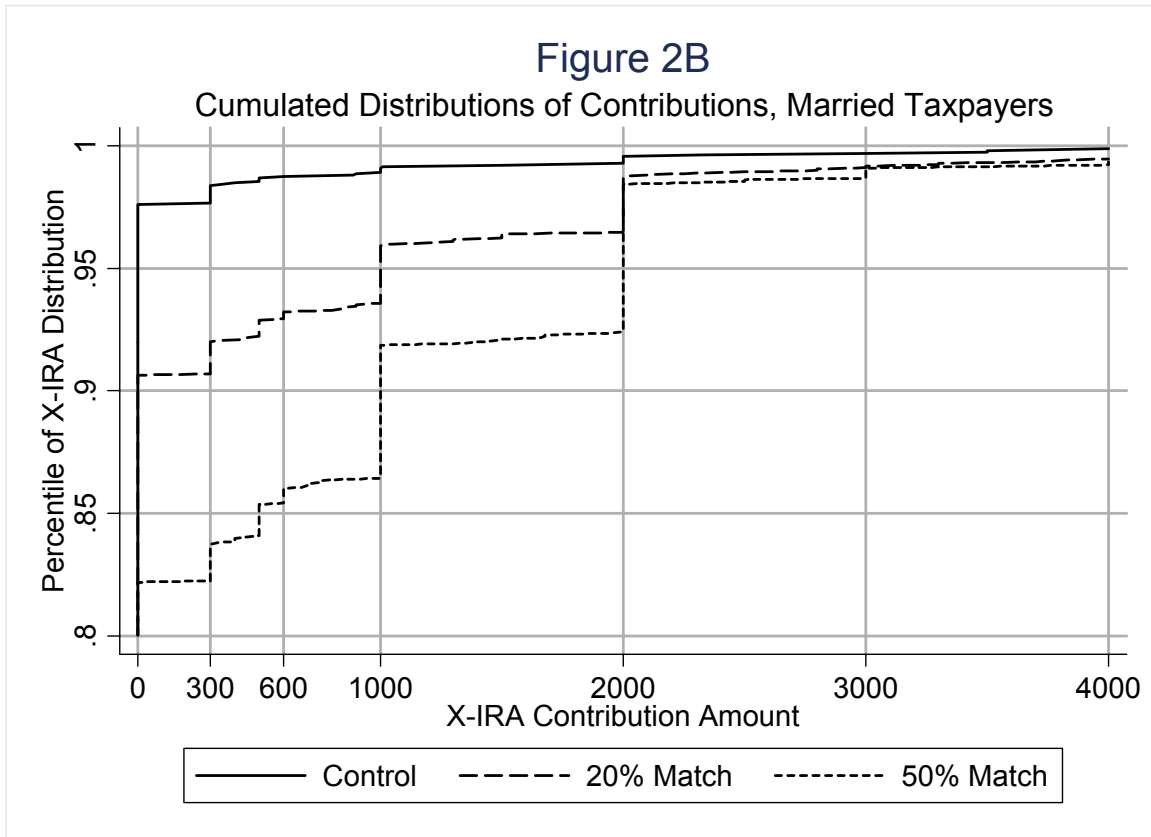
Panel C displays the difference between group A and group B. The bottom right panel displays the difference-in-differences.



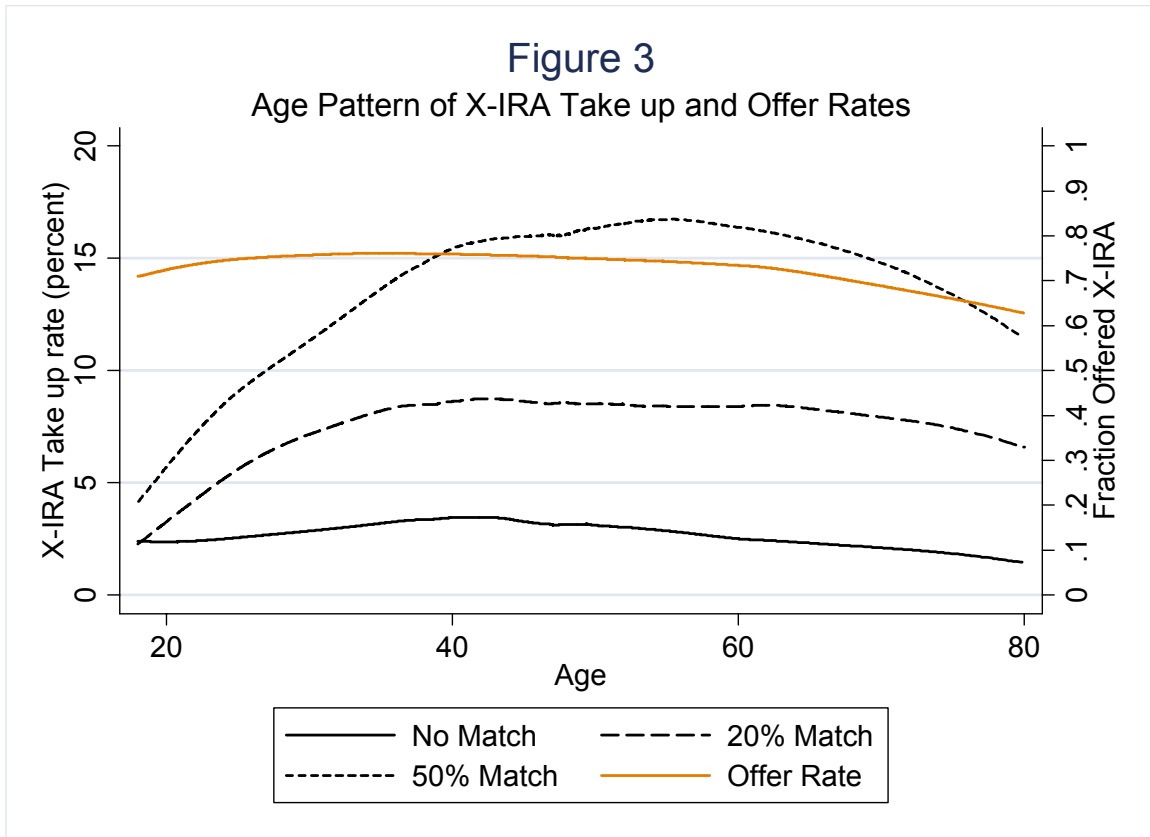
Notes: The figure displays (on the left y-axis scale) the X-IRA take up rate (in percent) for each of the three treatment groups (no match, 20% match, and 50% match) overtime from March 5th (first day of the experiment) to April 5th (last day of the experiment). The figure also displays, on the right y-axis scale, the fraction of tax returns getting an offer. Getting the offer is defined as reaching the X-IRA screen and hence triggering the experiment pop-up screen offer during tax preparation (set to happen by default but could be bypassed manually by tax professionals). If no offer was made, the X-IRA screen was never reached during tax preparation and hence the tax filer could not contribute to an X-IRA and hence not benefit from the offer. All graphs are smoothed using a non-parametric regression.



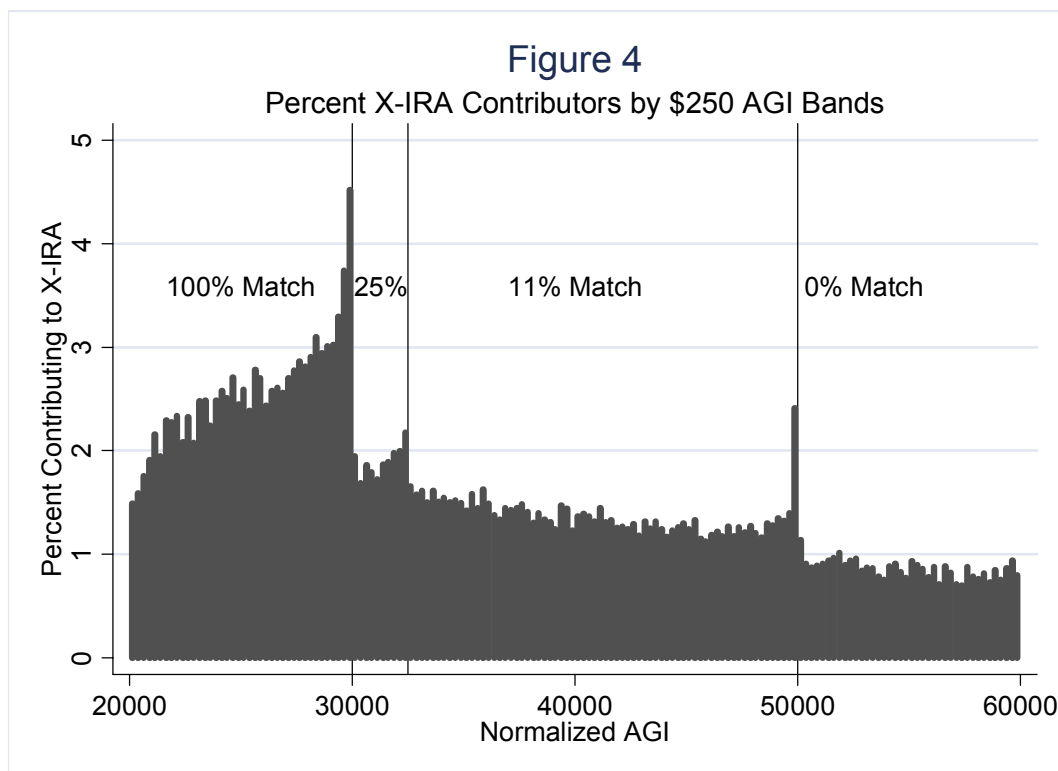
Notes: The figure displays the cumulated distribution of X-IRA contributions in each of the three treatment groups (no match, 20% match, and 50% match) for tax filers not filing jointly. For example, 97% of tax filers in the no match group made no X-IRA contributions, 99% made contributions of \$300 or less, etc. Vertical portions in the graphs are due to bunching of tax filers at those exact contribution levels. The maximum matched contribution is \$1,000.



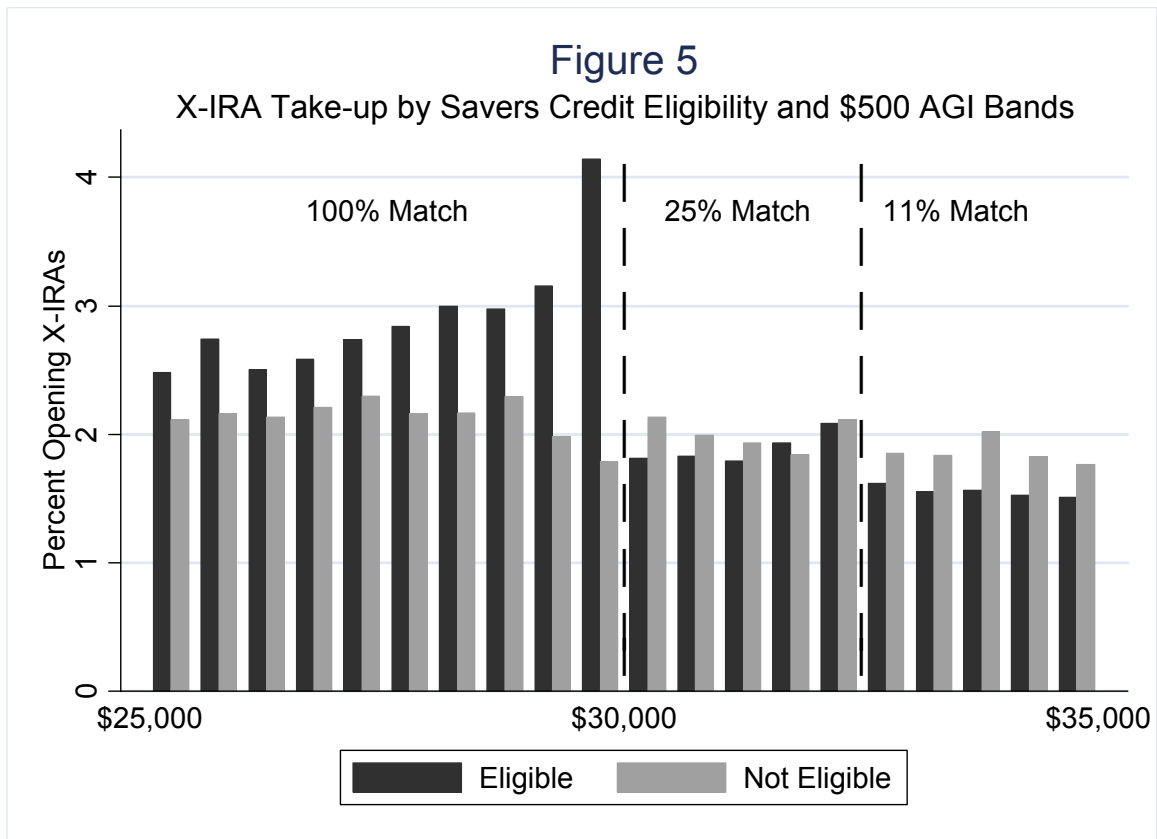
Notes: The figure displays the cumulated distribution of X-IRA contributions in each of the three treatment groups (no match, 20% match, and 50% match) for married tax filers filing jointly. For example, 82% of tax filers in the 50 percent match group made no X-IRA contributions, 84% made contributions of \$300 or less, etc. Vertical portions in the graphs are due to bunching of tax filers at those exact contribution levels. The maximum matched contribution is \$1,000 per spouse and therefore \$2,000 per tax return.



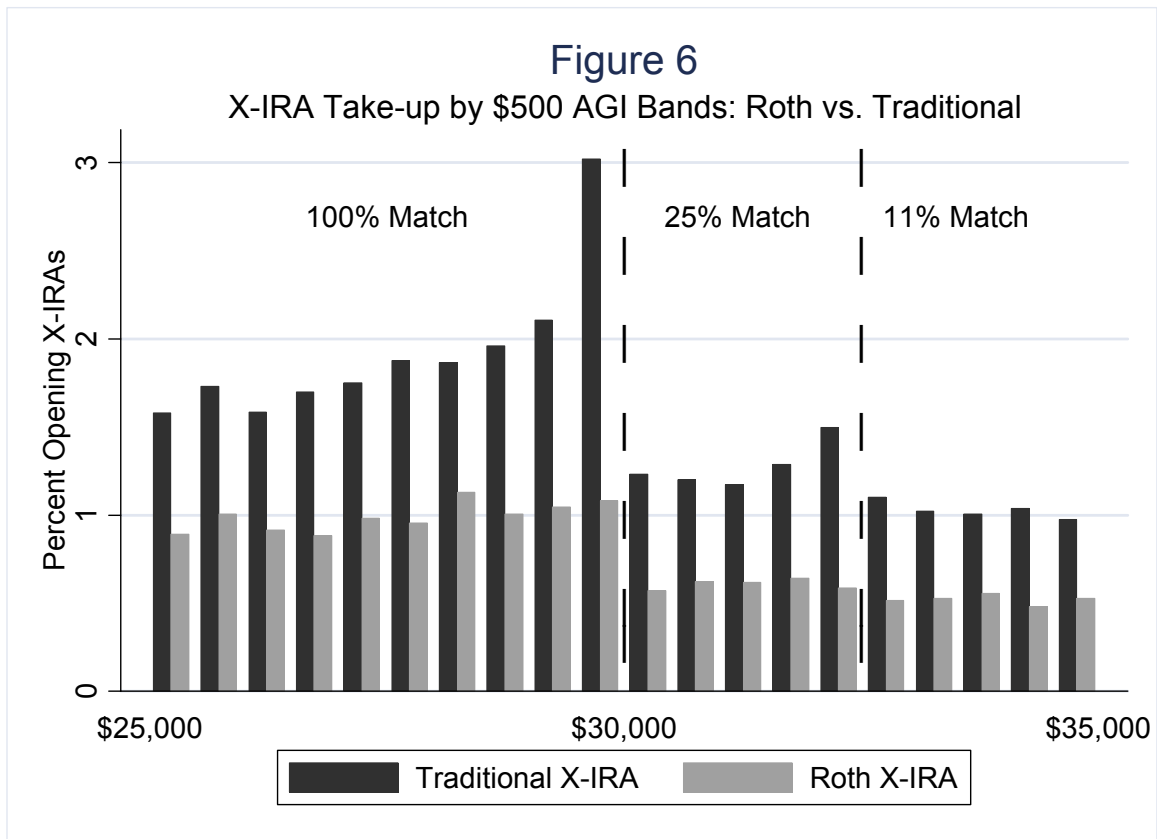
Notes: The figure displays (on the left y-axis scale) the X-IRA take up rate (in percent) for each of the three treatment groups (no match, 20% match, and 50% match) by age of primary tax filer on the tax return. The figure also displays, on the right y-axis scale, the fraction of tax returns getting an offer (same definition as in Figure 1). All graphs are smoothed using a non-parametric regression.



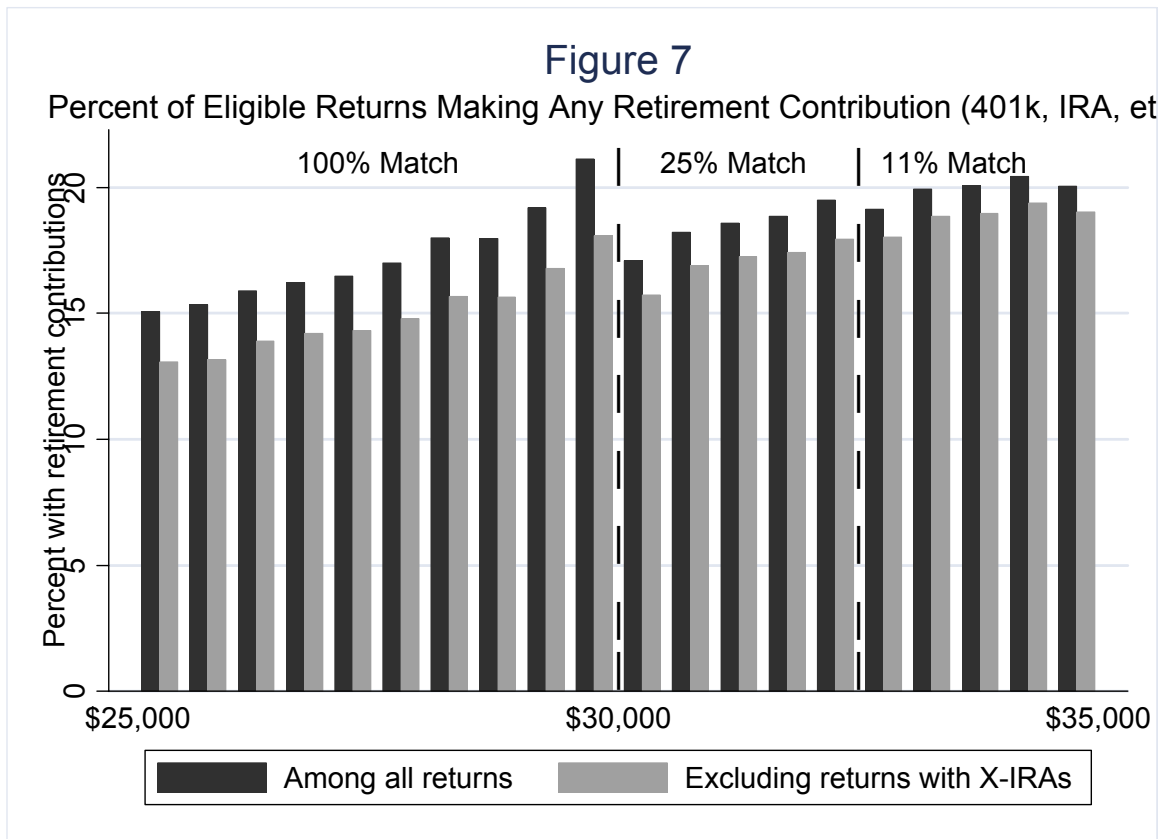
Notes: The figure displays the percentage of tax returns contributing to an X-IRA (for tax year 2004) by \$250 bands of normalized AGI (normalized AGI = AGI for married joint filers, = 133% of AGI for heads of households, = 200% of AGI for singles and others). The sample is restricted to eligible tax returns defined as all tax returns with positive tax liability net of credits (before X-IRA contributions if any), positive earnings, tax filers are aged 18 and above and not claimed as dependents on somebody else's return. Those tax filers benefit from the savers' credit if they make X-IRA contributions and meet the AGI limits. The sample is 100% of H&R Block retail returns with positive X-IRA contributions and a 9% random sample of other returns, filed for year 2004 during tax season 2005. The vertical lines display the location of the AGI cliffs where the savers' credit rate changes.



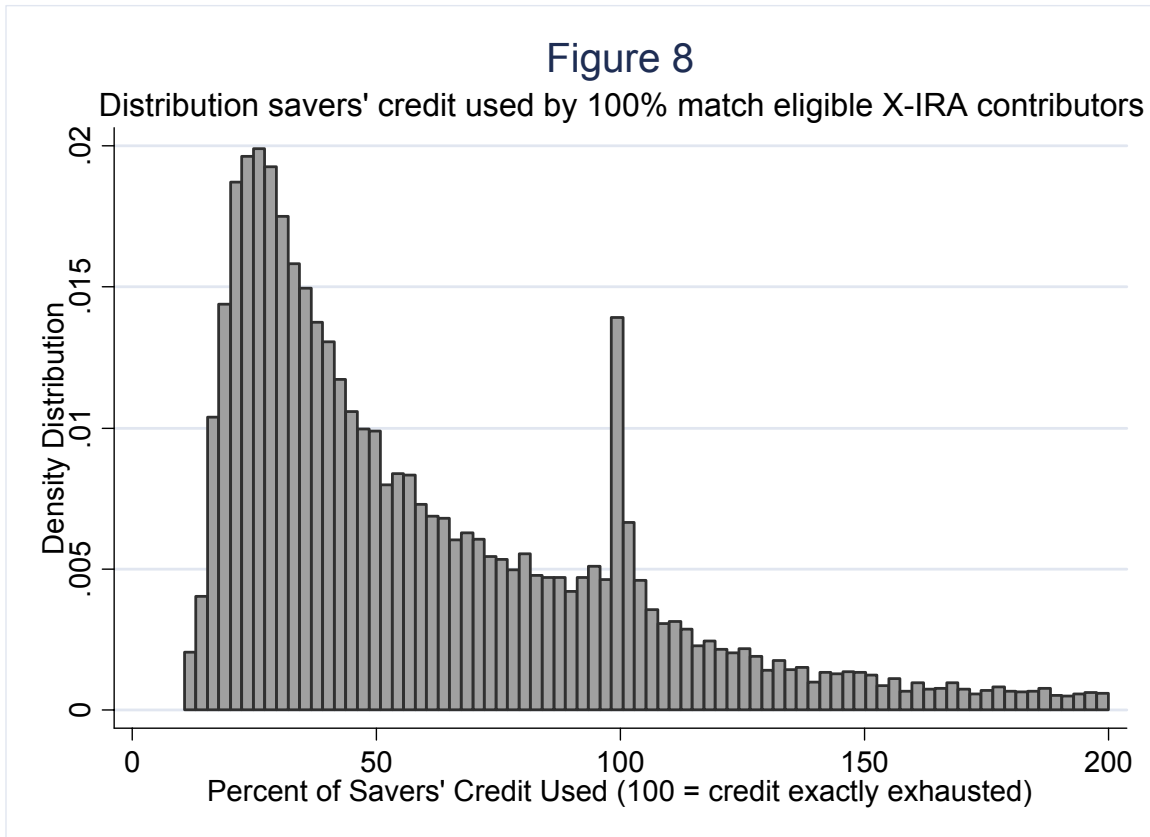
Notes: The figure displays the percentage of tax returns contributing to an X-IRA (for tax year 2004) by \$500 bands of normalized AGI and savers' credit eligibility status. Eligible tax returns defined as in Figure 4. Not eligible returns defined as all tax returns with no positive tax liability net of credits (before X-IRA contributions if any), with positive earnings, and tax filers are aged 18 and above, and not claimed as dependents on somebody else's return. Eligible returns benefit from the savers' credit if they make X-IRA contributions and meet the AGI limits. Not eligible returns do not benefit from the savers' credit because they have no tax liability (net of credits) to offset. The vertical dashed lines display the location of the AGI cliffs where the savers' credit rate changes.



Notes: The figure displays the percentage of tax returns contributing to a traditional X-IRA or to a Roth X-IRA (all for tax year 2004) by \$500 bands of normalized AGI. Sample is limited to eligible returns as in Figure 4. Eligible returns benefit from the savers' credit if they make X-IRA contributions and meet the AGI limits. AGI is net of traditional X-IRA contributions. Roth X-IRA contributions do not affect AGI. The vertical dashed lines display the location of the AGI cliffs where the savers' credit rate changes.



Notes: The figure displays the percentage of tax returns getting a positive “real” savers’ credit by \$500 bands of normalized AGI among all returns and among all returns excluding returns with X-IRAs. All returns sample limited to eligible returns defined as all tax returns with positive tax liability net of credits (before any retirement contributions if any), positive earnings, and tax filers are aged 18 and above, and not claimed as dependents on somebody else’s return. Those tax returns benefit from the savers’ credit if they make a retirement contribution (over and above any retirement contribution withdrawals over the last three years) and meet the AGI limits. Therefore, in this sample, getting the savers’ credit is equivalent to making any retirement contribution (any IRA, 401(k), Keogh, etc.) over and above withdrawals over the last three years. Excluding returns with X-IRAs sample is the same sample but excluding all returns making any X-IRA contribution for year 2004.



Notes: The figure displays the density distribution of the fraction of the savers' credit used among tax returns making a positive X-IRA contribution (for year 2004) and with normalized AGI between \$20,000 and \$30,000 and hence facing the 50% savers' credit rate (equivalent to a 100% match rate). The fraction of savers' credit used is defined as the ratio of actual X-IRA contributions to the maximum amount of X-IRA contributions benefiting from the savers' credit (taking other retirement contributions as fixed). For example, suppose a person has a tax liability (net of credits except the savers' credit) of \$500. The maximum creditable contribution is \$1000 (as a credit rate of 50 percent on \$1000 is \$500 and exactly exhausts the net tax liability). If this person makes an X-IRA contribution of \$300, then the fraction is 0.3 (30 on the graph). If this person makes an X-IRA contribution of \$1,500, then the fraction is 1.5 (150 on the graph).