TAX POLICY CENTER
URBAN INSTITUTE & BROOKINGS INSTITUTION

THE INTERACTION BETWEEN IRAS AND 401(K) PLANS IN SAVERS' PORTFOLIOS

William Gale, Aaron Krupkin, and Shanthi Ramnath

October 25, 2017

ACKNOWLEDGEMENTS

The opinions represent those of the authors and are not those of the US Department of the Treasury nor any of the institutions with which they are affiliated. The authors thank Alex Gelber for helpful comments.

This publication relies on the analytical capability that was made possible in part by a grant from the Laura and John Arnold Foundation. The findings and conclusions contained in this report are those of the authors and do not necessarily reflect positions or policies of the Tax Policy Center or its funders.

ABSTRACT

Policy makers have long sought to boost households' retirement saving through tax incentives. Little is known, however, about how savers' contributions are linked across different types of tax-preferred accounts. Previous research has concluded that workers who become eligible for a 401(k) plan also see stronger growth in IRA balances. However, the mechanism for this increase – contributions, asset growth, rollovers, etc. – is a puzzle. To examine these issues further, we use a sample of tax returns from 1999-2014. A particularly useful feature of the data is the presence of tax-reported information on IRA balances and 401(k) contributions. Using two different control groups that have stronger and weaker tastes for saving, respectively, than the treatment group, we find virtually no link between new 401(k) contributions and new IRA contributions. Households who start contributing to 401(k) plans do not have higher propensities to start contributing to IRAs, raise IRA contributions, own IRAs, or have higher IRA balances in level or first-differences.

CONTENTS

I. INTRODUCTION

Policy makers have long sought to boost households' preparation for retirement through a variety of tax incentives, including Individual Retirement Accounts (IRAs), 401(k) plans, and other options. The impact of such policies, studied individually, on private and national saving, has led to an extensive literature.[1]

There is little evidence, however, on how the policies interact with each other. To what extent are the retirement programs substitutes or complements? Does eligibility or participation in one such program boost or reduce participation in other similar programs? The programs might logically be thought to be substitutes, since they provide a similar good – a tax incentive for retirement saving. The law essentially treats them as substitutes since the contribution limit of traditional IRAs is lowered by access to an employer-sponsored plan. But it would not be unreasonable, a priori, to consider that they might instead be complements – that is, that eligibility or saving in one form could "crowd in" saving in other forms. This could occur, for example, if eligibility for one form of saving made people more aware of the need to save for retirement and they subsequently responded by saving more in several tax-preferred vehicles.

These issues are of relevance to policy makers because of the perennial focus on ways to raise retirement saving and because of the budgetary costs associated with tax expenditures for saving, with current estimates exceeding $100 billion per year.[2] To the extent that the different tax incentive programs are complements, exposing a worker to one program could raise participation in several programs. To the extent that the programs are substitutes, expansion of one program might cannibalize contributions to the other.

In this paper, we examine the interaction between IRAs and 401(k) plans in savers' portfolios – and in particular, the question of whether the programs act as substitutes or complements – using administrative tax data.

A well-recognized problem in the earlier literature on saving incentives is that needs and tastes for saving are heterogeneous across the population. Households with strong tastes or needs for saving may be more likely to save in many forms than those with weak tastes or needs for saving. Not controlling for this heterogeneity will bias analysis toward finding that different forms of retirement saving are complements even if they are not. To address this problem, we use two different control groups in our analysis. As explained below, one control group plausibly has stronger average needs or tastes for saving than the treatment group, while the other control

[1] See Benjamin (2003); Bernheim (2002); Chetty et al. (2014); Engen, Gale, and Scholz (1996); Engen and Gale (2000); Hubbard and Skinner (1996); Poterba, Venti, and Wise (1996).

[2] The U.S. Department of the Treasury (2016) calculates tax expenditures for retirement programs in two ways. The first estimates current-year revenue losses from all existing accounts. The second examines the present value of revenue loss from all new contributions in a given year. Both procedures yield annual revenue loss estimates above $100 billion in recent years.

group plausibly has weaker average needs or tastes for saving than the treatment group. Our results, however, are not sensitive to which control group is employed. In comparisons of our treatment group with either control group, we find little or no complementarity or substitutability between 401(k) contributions and IRA contributions. As a result, contributions to the two forms of saving appear to be independent.

Section II provides background information on the dataset and regression framework. Section III presents the main results. Section IV compares our results to earlier findings. Section V concludes.

II. DATA AND METHODOLOGY

A. ADMINISTRATIVE TAX DATA

We use administrative tax data that include self-reported personal income tax returns and third-party reported information returns. Our data cover the US population for tax years 1999 through 2014.[3] We begin by drawing a 0.1 percent random sample of individuals ages 18 through 59. For these individuals, we create a merged file that contains information on marital status and household income measures reported on Form 1040; 401(k) contributions and wages reported on Form W2; and IRA contributions and balances reported on Form 5498.

For comparability with previous literature discussed below, we aim to focus on individuals who are in the first full year of a job.[4] Because the tax data do not explicitly report job changes, we create a proxy for people in their first year of a job. We only include individuals who had two jobs in one year (identified through Employer Identification Numbers, or EINs) and then one job (one EIN) in the following year, where the second-year EIN was one of the two first-year EINs. This is intended to capture workers who changed jobs in one year and stayed in that job through the end of the following year.[5] There are two observations for each individual in the dataset, one for each year. Individuals who had more than one job change during our sample period will accordingly appear multiple times in the final dataset.[6]

We define treatment and control groups as summarized in Table 1. Our treatment group consists of individuals who contributed to a 401(k) plan in their first full year on the job (the second year of their observation), but did not contribute to a 401(k) in the first year of their observation (n=13,393). We define two different control groups. Control group 1 consists of individuals who contributed to a 401(k) in both years (n=25,349). Control group 2 consists of individuals who did not contribute to a 401(k) in either year (n=121,212). Thus, the overall sample with control group 1 has 38,742 observations; the overall sample with control group 2 has 134,605 observations.[7]

---

[3] We exclude observations for tax year 2001 due to missing deferred compensation data.

[4] Gelber (2011).

[5] We do not know the exact time pattern of the EIN changes or their causes (for example, it is not unheard of for employers to change their EIN during a year – due to a merger or acquisition or other factors). Nevertheless, these factors would only affect the results if they were distributed in a systematically different way across the treatment and control groups.

[6] We do not cluster the results to account for these individuals.

[7] These figures refer to the number of observations in each group, not the number of individuals. Each individual has two observations, subject to data availability, and it is possible that individuals show up multiple times throughout the dataset if they switch jobs more than once.

B. REGRESSION FRAMEWORK

We are interested, as noted above, in how participation in one saving incentive program affects participation in other saving incentive programs. For comparability with other research, we focus on how 401(k) activity, which is influenced by employer choices, affects individuals' choices regarding individual retirement accounts. We thus employ regression analysis to examine how 401(k) contributions affect (a) the probability of having an IRA (i.e., a positive balance in an IRA),[8] (b) the probability of contributing to an IRA in the current year, (c) the amount contributed to an IRA in the current year, (d) the IRA balance, and (e) the change in the IRA balance from one year to the next.

Our basic regression specification takes the form:

(1) $$I_{it} = \alpha + \beta_1 T_i + \beta_2 Y_t + \beta_3 (T_i * Y_t) + \beta_4 X_{it} + \varepsilon_{it}$$

where the dependent variable *I* is an indicator of IRA activity for individual *i* in period *t* defined in different ways in the various regressions. T is an indicator for treatment status. Y is an indicator for the individual being in the second year of their observation (the first full year of employment at the new job). X is a vector of control variables, including categorical variables for age and income, and indicators for marital status and the presence of Schedule-C income. ε is the robust standard error. We also include year fixed effects.

The coefficient β_3 on $T * Y$ is an interaction term between treatment and the employment year dummy (essentially a difference-in-difference estimator) and shows the increase in IRA activity for the treatment group relative to the control group in the second year of observation relative to the first. If 401(k) contributions boost IRA activity, the coefficient should be positive and significant.

Whether a worker contributes to a 401(k) depends on whether the worker is eligible for a plan and whether the worker makes a contribution given eligibility (either by active choice or via passive enrollment). All workers in the sample are in their first full year on a new job in the second year they are observed. Even if eligibility patterns do differ across the groups, differences in 401(k) contribution behavior are likely to reflect – to at least some extent – workers' heterogeneous needs and tastes for saving. Thus, it is plausible that members of control group 1 – who contribute to a 401(k) plan in both years of the sample – have higher needs or tastes for saving on average than do the treatment group members – who do not contribute in

[8] Individuals are recorded as having an IRA if their IRA has positive fair market value as reported on Form 5498. This form is issued to the IRS each year regardless of whether the account owner made a contribution that year.

the first year but do contribute in the second year. Likewise, it is plausible that treatment group members have, on average, higher needs or tastes for saving than do the members of control group 2, who do not contribute to a 401(k) plan in either year observed. Using two different control groups that have stronger and weaker tastes for saving, respectively, than the treatment group, we show that the results are not sensitive to the choice of control group.

III. EMPIRICAL RESULTS

A. DESCRIPTIVE PATTERNS

Table 2 provides a summary of the characteristics of the treatment group and the two control groups. The groups are similar with respect to age, with a mean age of about 40 years. About 14-15 percent of each group has schedule C income. Relative to the treatment group, members of control group 1 – who contribute to a 401(k) in both years of observation – have higher income, are more likely to be married, are more likely to own and contribute to an IRA, and have higher average balances and contributions than the treatment group. For each of these measures, members of control group 2 – who do not contribute to a 401(k) in either year observed – have lower average values than the treatment group. As noted above, the regressions control for items such as age, income, marital status, and the presence of schedule C income.

Table 3 makes the simple point that IRA ownership patterns and contribution patterns vary significantly. The table shows that, controlling for income, the likelihood of making an IRA contribution does not vary much with age. For example, among individuals with income above $50,000, 7.7 percent of 25-44 year-olds make an IRA contribution in a given year, whereas 8.9 percent of 45-64 year-olds contribute. In contrast, the likelihood of owning an IRA (in effect, the likelihood of ever having made a contribution or a rollover into an IRA) rises substantially with age. Among individuals with income above $50,000, the likelihood of owning an IRA is 29.5 percent for those aged 25-44 and 47.5 percent for those aged 45-64. Thus, variations across groups in IRA contributions are not necessarily a good measure of variation in IRA balances. These differences are an important consideration in understanding previous work.

Before turning to the regression analysis, we provide simple difference-in-difference calculations for the likelihood of contributing to an IRA. Table 4 shows IRA contribution rates by age, income, year of observation, and treatment status. The last column shows the difference-in-difference (the increase in IRA contribution rates for the treatment group relative to the control group from the first observation to the second). If 401(k) participation crowded in IRA contributions, we would expect the difference-in-difference estimates reported in Table 4 to be positive and substantial. Instead, the difference-in-difference calculations are close to zero, and almost all of them are actually negative. These results do not suggest any complementarity between the groups.

B. REGRESSION RESULTS

Table 5 reports the results from estimating Equation (1) for the various dependent variables.[9] The first column reports a linear probability estimate using the likelihood of owning an IRA as the dependent variable. The key estimated coefficient (β_3) shows that the likelihood of owning an IRA *fell* significantly for the treatment group relative to control group 1 in the second period relative to the first. Relative to control group 2, the likelihood of the treatment group owning an IRA in the second period relative to the first also fell, but the estimate was not statistically significant. These results do not suggest any complementarity between 401(k)s and IRAs.

The second column reports linear probability regressions where the dependent variable is an indicator for whether an individual contributes to an IRA. The estimated coefficient on the interaction term is small in absolute value – well below 1 percentage point – and it is negative and insignificant when control group 1 is used. The coefficient is negative and significant when control group 2 is employed. Again, there is no evidence that 401(k)s and IRAs are complements.[10]

The third column reports ordinary least squares regressions using IRA contributions as the dependent variable. The regressions show virtually no impact of 401(k) contribution behavior on the level of IRA contributions. The point estimates suggest that 401(k) participation *reduces* annual IRA contributions, but neither coefficient is precisely estimated.[11] In a more formal analysis of IRA contribution behavior, we use two-limit Tobit models to account for IRA contributions being constrained between zero and a contribution limit (see Appendix Table 1).[12] The results are similar to those in Table 5; in no case was the result both positive and significant.

Column 4 shows ordinary least squares regressions where the dependent variable is the IRA balance. For the first control group, the effect of treatment is statistically significant and *negative* at about $5,800. For control group two, the negative effect is much smaller (less than $300) and not statistically significant. Neither of these results is consistent with 401(k) contributions crowding-in larger IRA balances.

Finally, column 5 reports regression results where the dependent variable is the change in growth of logged IRA balances between periods for each individual (Equation 2).

[9] Table 5 does not display the coefficients on the control variables.

[10] Logit and probit models for IRA ownership and IRA contributions yield similar results to linear probability models and are not reported here.

[11] In other regressions, we looked separately at the effects on contributions to traditional (front-loaded) IRAs and Roth (back-loaded) IRAs. Effects are small for each – approximately -$24 for traditional IRAs and $6 for Roth IRAs. None of the coefficients is precisely estimated.

[12] The lower limit in the tobit model is zero, while the upper limit is the annual IRA contribution limit. Since contribution limits change over time, we run separate regressions for each set of years with the same contribution limit (Appendix Table 1).

$$(2) \qquad I_i = [\ln(A_{i2}) - \ln(A_{i1})] - [\ln(A_{i1}) - \ln(A_{i0})]$$

Period 0 corresponds to the year before the individual's job change, period 1 corresponds to the year when the job change occurs (year one of observation above), and period 2 is the individual's first full year at the new job (year two of observation above). This closely follows the variable construction in previous literature and takes into account all forms of asset accrual over time including contributions, investment returns, rollovers, and withdrawals.[13]

Since this dependent variable represents the change in account balance over time, it is not appropriate to include the employment year dummy (Y) or interaction term $(T * Y)$ from Equation (1) in the regression. Thus, our specification of the estimating equation is:

$$(3) \qquad I_i = \lambda_0 + \lambda_1 T_i + \lambda_2 X_i + \varepsilon_i$$

where *T* and *X* are defined in Equation (1) and the key effect is measured by the coefficient on the treatment variable (λ_1). If 401(k) contributions boost IRA balances, we would observe larger balance increases between periods 1 and 2 versus periods 0 and 1 for the treatment group relative to the control group. That is, λ_1 would be positive and significant. Instead, under both control groups, we find a small negative relationship between 401(k) contribution behavior and the change in IRA balances that is not statistically significant. Similar to our earlier findings, neither result provides evidence of a crowd-in effect of 401(k) contributions on IRAs.

Looking across all of the columns of Table 5 provides further evidence for our view that members of control group 1 (all of whom contribute to a 401(k) in both periods) have on average stronger needs or tastes for saving than members of the treatment group, and that members of control group 2 (each of whom does not contribute to a 401(k) in either period) have on average weaker needs or tastes for saving than members of the treatment group. Specifically, the coefficient on the treatment group dummy is negative and significant for four of the five regressions using control group 1 (in the top panel), and in four out of five cases, is positive and significant for comparisons using control group 2 (in the bottom panel). This means that members of control group 1 are more likely to have an IRA or contribute to an IRA than members of the treatment group. Additionally, it shows that members of control group 2 are less likely to have or contribute to an IRA than the members of the treatment group. This is consistent with the view in the saving incentive literature that groups with higher needs or tastes for saving tend to save more in all forms of saving. In contrast, the coefficients in the first row of each panel show that *changes* in 401(k) contribution status do not induce changes in IRA behavior.

---

[13] Gelber (2011).

IV. DISCUSSION

In our descriptive and regression analyses, we find little evidence supporting the claim that 401(k) participation and IRA participation are complements. Gelber (2011), on the other hand, uses data from the 1996 Survey of Income and Program Participation (SIPP) and finds that 401(k) eligibility raises IRA balances and interprets the results as suggesting that 401(k) eligibility "crowds in" higher IRA contributions.

As discussed further in the Appendix, Gelber's (2011) estimates are hard to explain. His point estimates imply that the impact of 401(k) eligibility on IRA balances is roughly $3,000 over one year. However, the individual annual contribution limit over the period in question was just $2,000, and about the same share of treatment and control group members in Gelber's study had IRAs. It is therefore implausible that changes in contributions could have generated an impact of $3,000 for the treatment group relative to the control group.

Changes in IRA balances could be due to factors other than contributions, such as increases in asset values, withdrawals, or rollovers. However, general increases in asset values cannot explain the differential growth of IRA balances, since Gelber (2011, Table 1) shows that the control group had *higher* initial IRA balances than the treatment group. It is possible that there were differences in rollovers or withdrawals between the two groups, but there is no evidence presented on this.[14] In summary, it is difficult to see the sources of the $3,000 increases in IRA balances for the treatment group relative to the control group.

One possible explanation might be poor asset recall. Whereas the data we employ are based on compulsory administrative filings that are required to be accurate, SIPP responses are voluntary, and asset balances are based on respondents' recall. It is possible that the SIPP data are less than fully reliable in this regard. For example, among individuals in the SIPP with a positive IRA balance at the end of period 0, 18 percent reported a zero IRA balance at the end of year 1. In contrast, among individuals in the administrative data with a positive IRA balance at the end of one year, less than 5 percent had a zero IRA balance in the following year.

There are also other advantages of the administrative data relative to the SIPP. First, the SIPP data are now roughly 20 years old, and there have been many changes in the retirement saving landscape. Second, the SIPP data are based on a much smaller sample size than the administrative data.

The disadvantage of the administrative data is that they do not contain information on 401(k) eligibility, which has been used in previous work to instrument for participation and which

[14] Although the SIPP waves that Gelber (2011) uses do not report contribution data, the data are included in a separate wave set (Copeland 2002).

Gelber (2011) uses to identify policy effects.[15] In contrast, our treatment group consists of individuals who do not contribute to 401(k) in the year they have a job transition, but do contribute to a 401(k) in their first full calendar year on the job. Since 401(k) contributions are a choice (conditional on eligibility), our classification of households into treatment and control groups is not based on exogenous factors. However, we believe our results are of interest, for two key reasons.

First, the approach we use employs two control groups. Members of one control group have on average stronger needs and tastes for saving than the treatment group. Members of the other control group have on average weaker needs or tastes for saving than the treatment group. Thus, our results provide strong evidence that 401(k)s and IRAs are not complements in household portfolios, after controlling for needs and tastes for saving.

Second, as shown in the Appendix (and Appendix Table 2), using the SIPP data, to the extent that there is an impact of 401(k)-status on IRA balances, it is occurring through individuals who actually have positive 401(k) balances, not those who are eligible but don't participate.

[15] Gelber (2011); Heim and Ramnath (2016).

V. CONCLUSION

We examine the relationship between changes in households' 401(k) contribution status and IRA status. If the two savings vehicles were complements, policy makers would obtain a bit of a "free lunch," as they would be able to spur retirement saving through both types of plans merely by encouraging the expansion of one of them. Previous research supports this position.

However, since 401(k)s and IRAs provide similar benefits – tax savings associated with saving for retirement – it would not be surprising if households viewed them as substitutes. This situation would occur if people who contributed to one type of account were also less likely to contribute the other type of account, other things equal.

Our examination of the data suggests an intermediate outcome, as we find virtually no relation between a households' propensity to start contributing to a 401(k) and its propensity to start or continue contributing to an IRA. Our method obtains similar results when using two different control groups: one with stronger saving motives than the treatment group, and one with weaker saving motives than the treatment group. By showing that our results are not sensitive to the presumed heterogeneity in needs and tastes for saving across households, we provide new evidence that policy makers should not expect higher retirement saving in one form to "crowd in" retirement saving in another form.

APPENDIX: ANALYZING SIPP DATA

A. BASIC SPECIFICATION AND RESULTS

Gelber (2011) examines the effect of 401(k) eligibility on Individual Retirement Account saving using data from the 1996 Survey of Income and Program Participation. His sample consists of workers who were younger than 65 and in their first and second years of a job at a firm that offered a 401(k). His treatment group consists of individuals who were not eligible for a 401(k) in their first year because they had not been employed long enough. Members of this group gain eligibility for the 401(k) plan after 12 months of employment. The control group consists of workers whose eligibility did not change over the sample period.[16] The sample consists of 835 individuals, 296 in the treatment group and 539 in the control group.

The year before employment at the job in question roughly corresponds to 1997 (which we will call year 0). The first year of employment roughly corresponds to 1998, and the second year of employment roughly corresponds to 1999.[17] The data presented on asset values refer to balances at the end of each year.

Defining A_{it} as IRA balances (+10, to avoid taking logs of zero) for individual *i* in period *t*, the dependent variable of interest is Y_i, given by Equation (A1) as the change in the natural log of IRA balances from the end of year one to the end of year two.[18]

(A1) $$Y_i = [\ln(A_{i2}) - \ln(A_{i1})] - [\ln(A_{i1}) - \ln(A_{i0})]$$

Gelber's main regression formulation is given in Equation (A2), where T_i is an indicator for treatment status, X_i represents a set of control variables, and E_i is the error term. Robust standard errors are clustered by household.

[16] In the Gelber (2011) analyses that we report, the control group consists of individuals who were eligible for a 401(k) in both periods. In an alternative analysis, Gelber (2011) examines a control group that incorporates those who were never eligible.

[17] Since workers may not start their job on January 1, there is an inherent mismatch between the end of a calendar year and the end of the first year of employment. However, the years presented here provide a rough proxy to organize the sample.

[18] Gelber (2011) identifies these periods using the wave numbers of the SIPP. Our end-of-year 0 refers to data in SIPP Wave 6; end-of-year 1 refers to Wave 9 data; and end-of-year 2 refers to Wave 12 data.

(A2) $$Y_i = B_0 + B_1 T_i + B_2 X_i + E_i$$

The key coefficient, B_1, represents the amount by which IRA balances rose over time for treatment group members versus control group members. By downloading data that Gelber posted, we were able to replicate his regressions exactly.[19] Selected results are shown in Appendix Table 2. In a specification with no controls (row 1), the point estimate for B_1 is 0.56 and is different from zero at usual standards of statistical significance. In dollar terms, the estimate translates into an average increase in IRA balances of $3,319 for the treatment group relative to the control group.[20] Adding control variables, including initial balances, changes the results only slightly (rows 2 and 3). The dollar equivalents are $3,174 and $2,977, respectively.

B. INTERPRETING THE RESULTS

Gelber's (2011, page 112) explanation of these findings is that:

> "If 401(k) eligibility encourages households to overcome the fixed costs of opening accounts with mutual funds or other investment vehicles, or to learn about financial markets, then it may be less costly to put money in IRA accounts... Eligibility often comes with reminders by one's firm to save, pamphlets emphasizing the importance of retirement saving, the necessity of learning about financial markets and the like. Therefore, individuals could be encouraged by 401(k) eligibility to save in IRAs."

We find it implausible that this explanation could account for the estimated coefficients. The change in IRA balances is the result of contributions, withdrawals, changes in the market value of assets, and rollovers. Gelber's interpretation suggests that the coefficient reflects increased IRA contributions by those who become newly eligible. However, this view seems to be hard to reconcile with some simple calculations.

For example, the point estimate of the impact of 401(k) eligibility on IRA balances ranges from $2,977 to $3,319 in the three regressions reported above, but IRA annual contribution limits at that time were $2,000 for individuals. Contributions would not come close to explaining the full difference *even* if (a) *every* individual in the treatment group contributed the *maximum* amount to an IRA in year 2 and (b) *no one* in the control group contributed to an IRA in either

[19] The data and corresponding programming code can be downloaded at https://www.aeaweb.org/aej/pol/data/2009-0165_data.zip.

[20] See Gelber (2011, Table 2). The dollar equivalent can be calculated by taking the natural log of (1 + coefficient) and multiplying it by the mean asset value for the treatment group at the end of year 0.

period 1 or 2 and (c) *no one* in the treatment group contributed to an IRA in period 1.[21] Of course, not everyone in the treatment group contributed the maximum.[22] In addition, IRA ownership was quite similar between the two groups. In both groups, 27 percent of sample members owned an IRA in the second period. IRA ownership rates rose by just 4 percentage points for the treatment group relative to the control group from year 1 to year 2. Thus, it is extremely difficult to explain an increase in the growth of average IRA holdings anywhere close to $3,000 for the treatment group relative to entire control group.

Another way to gauge the plausibility of the results – and the channels that Gelber (2011) describes – is to ask how many people both started contributing to a 401(k) and opened an IRA during the second year. Only 1.5 percent of the treatment group fell into this category, compared to 1.2 percent of the control group. Nor is it plausible that the impact on IRAs was coming from people who did not participate, as we discuss below. For all of these reasons, it seems implausible – indeed, mathematically impossible – that differences in IRAs contributions between the two groups could have accounted for more than a very small share of Gelber's estimated coefficient.

Nor is growth in asset values a likely explanation of the large coefficients. The control group actually held *larger* initial IRA balances than the treatment group, so any broad-based increases in market values would work to reduce the estimated coefficient, not raise it.[23]

That leaves differences in rollovers and withdrawals as potential explanations. Gelber (2011) does not use data on either of those issues, however, which is one reason we turn to administrative data.

Gelber (2011) notes that the confidence intervals in his paper do not rule out responses much smaller than the point estimates. Our calculations here, however, show that the true value *must* be much smaller than his point estimates, and our own analysis raises doubts as to whether the true effects of 401(k)s on IRAs are even positive.

C. EXOGENEITY ISSUES

While the administrative data offer several advantages over the SIPP data (see the main text), the great advantage of Gelber's study relative to ours is his method of finding plausibly exogenous variation in 401(k) eligibility. Specifically, his treatment group consists of individuals

[21] The estimated coefficients in the regressions could, in principle, have been generated by changes in contribution behavior if a large percentage of control group members had IRAs in period 1 but closed them in period 2. There is no reason to suspect that this happened. Incidence of IRA ownership actually increased in the control group from period 1 to period 2.

[22] Although the SIPP waves that Gelber (2011) uses do not report contribution data, other sources indicate that only about two-thirds of IRA contributors contributed the maximum amount in the late 1990s (Copeland 2002).

[23] See Gelber (2011, Table 1).

who are ineligible for a 401(k) in their first year at a job because of the 401(k) eligibility rules at their company. These individuals then become eligible in their second year on the job.

In contrast, our treatment group (described in the main text) consists of individuals who do not contribute to 401(k) in the year they have a job transition, but do contribute to a 401(k) in their first full calendar year on the job. Since 401(k) contributions are a choice (conditional on eligibility), our classification is not based on exogenous factors. However, we believe our results are of interest, for two key reasons.

The first reason – as noted in the main text – is the fact that we generate similar results using different control groups, one with plausibly stronger needs and tastes for saving than the treatment group, and one with plausibly weaker needs or tastes for saving than the treatment group. Thus, heterogeneous needs or tastes for saving are not contaminating the results.

Second, using the SIPP data, we show that the extent there is an effect of 401(k) eligibility on IRAs, the effect occurs through 401(k)-eligible individuals who actually participate in their 401(k). Appendix Table 2, row 4, shows the results of a regression that follows the specification in row 2 exactly, except that the treatment group consists only of the members of the original treatment group who had a positive 401(k) balance in year 2. (That is, it excludes those who are eligible in the second year but choose not to participate). Appendix Table 2, row 5 shows the results of the same regression except that the treatment group consists only of the members of the original Gelber (2011) treatment group who did not have a 401(k) in year 2 (the excluded members in the row 4 regression). As expected, the impact of 401(k) eligibility on IRA balances for second-year 401(k) holders (row 4) is substantially higher – 80 percent – than for second-year 401(k)-eligibles who did not have a 401(k) (row 5). Indeed, the impact for those without a positive balance is not statistically different from zero at conventional significance levels ($p=.127$) even though the treatment group in row 5 is more than twice as large as the treatment group in row 4.

To clarify the impact of contributing even further, we start with the sample in Gelber (2011), but specify the treatment group to be those who owned a 401(k) in period 2 but not in period 1 and therefore must have contributed to a 401(k) in the second period. The control group is those who did not own a 401(k) in either period. Appendix Table 2, row 6, shows that the coefficient in this regression is 0.85, different from zero, and substantially larger than Gelber's estimates in the first several rows. This serves to show that, to the extent that there is an impact of 401(k)-status on IRA balances, it is occurring through individuals who contribute to a 401(k).

TABLE 1

Description of the Treatment and Control Groups

	Contributes to 401(k) in First Year	Contributes to 401(k) in Second Year	Sample Size
Treatment Group	No	Yes	13,393
Control Group 1	Yes	Yes	25,349
Control Group 2	No	No	121,212

TABLE 2

Descriptive Statistics of the Administrative Tax Data by Treatment Status

	Treatment Group		Control Group 1		Control Group 2	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Age	39.0	10.2	41.8	10.2	40.4	10.8
Has Schedule-C Income (%)	13.8	34.5	14.7	35.4	15.0	35.7
Income ($)	80,299	128,707	131,044	395,822	67,767	278,426
Married (%)	55.0	49.7	66.4	47.2	54.7	49.8
IRA Ownership (%)	28.1	44.9	47.4	49.9	20.9	40.7
IRA Contribution (%)	6.5	24.7	9.4	29.2	4.7	21.1
IRA Balance ($)	13,831	61,398	37,198	98,670	11,096	71,869
IRA Contributions ($)	307	1,479	502	1,953	213	1,439
Observations	13,393		25,349		121,212	

Note: Members of the treatment group contribute to a 401(k) in the second year of observation, but not the first. Members of control group 1 contribute to a 401(k) in both years of observation. Members of control group 2 do not contribute to a 401(k) in either year.

Source: Authors' calculation using administrative tax data.

TABLE 3

IRA Behavior by Age and Income

Age Group	25-44	25-44	45-64	45-64
Income Group	$10,000 - $50,000	$50,000+	$10,000 - $50,000	$50,000+
IRA Contribution (%)	2.2	7.7	3.2	8.9
Positive IRA Balance (%)	9.4	29.5	20.0	47.5

TABLE 4

IRA Contribution Rates (%) by Group over Time

Age	Income	Control Group	Control (Year 1)	Control (Year 2)	Treatment (Year 1)	Treatment (Year 2)	Diff-in-Diff
All	All	1	9.5	9.3	6.8	6.3	-0.3
25-44	10k-50k	1	5.9	5.3	3.0	2.5	0.0
25-44	50k+	1	10.2	9.9	8.8	8.3	-0.2
45-64	10k-50k	1	7.9	7.6	4.2	4.1	0.2
45-64	50k+	1	9.8	9.9	10.6	8.6	-2.1
All	All	2	4.6	4.7	6.8	6.3	-0.6
25-44	10k-50k	2	1.9	1.9	3.0	2.5	-0.5
25-44	50k+	2	6.6	6.8	8.8	8.3	-0.7
45-64	10k-50k	2	2.9	2.8	4.2	4.1	0.0
45-64	50k+	2	8.5	8.6	10.6	8.6	-2.1

Note: The difference-in-difference is the increase in IRA contribution rates for the treatment group relative to the control group from the first observation to the second.

TABLE 5

Administrative Data Regression Results

Variables	(1) Have IRA	(2) Contribute to IRA	(3) IRA Contributions	(4) IRA Balance	(5) Natural Log Change in IRA Balance
Control 1					
Coefficient on Interaction Between Treatment and the Second Year	-0.022** (0.009)	-0.006 (0.006)	-16.22 (35.24)	-5,777*** (1,525)	------ ------
Coefficient on Treatment	-0.091*** (0.007)	-0.013*** (0.004)	-83.86*** (26.22)	-4,738*** (1,073)	-0.044 (0.037)
Observations	38,742	38,742	38,742	38,742	37,801
R-squared	0.127	0.013	0.015	0.153	0.023
Control 2					
Coefficient on Interaction Between Treatment and the Second Year	-0.002 (0.008)	-0.008* (0.004)	-19.4 (26.61)	-287.2 (1,074)	------ ------
Coefficient on Treatment	0.053*** (0.005)	0.015*** (0.003)	60.96*** (19.02)	1,675** (770.9)	-0.023 (0.026)
Observations	134,605	134,605	134,605	134,605	129,685
R-squared	0.130	0.025	0.018	0.081	0.019

Note: Robust standard errors in parentheses. * denotes .05 < p ≤ .10. ** denotes .01 < p ≤ .05. *** denotes p ≤ .01.

APPENDIX TABLE 1

Estimated Coefficient on Interaction Term for Two-Limit Tobit Regressions

Year	1999-2001	2002-2004	2005-2007	2008-2012	2013-2014
Control Group 1	2557.6 (1,944.0)	-1,370.0 (1,137.6)	-1064.3 (1042.7)	-630.3 (1,023.6)	818.6 (1,968.8)
Control Group 2	810.4 (1,759.9)	-1,229.3 (877.1)	-1,674.4* (890.6)	-493.7 (883.0)	937.3 (1,739.4)

Note: Robust standard errors in parentheses. * denotes .05 < p ≤ .10. ** denotes .01 < p ≤ .05. *** denotes p ≤ .01.

APPENDIX TABLE 2

SIPP Data IRA Asset Regression Coefficient for the Treatment Variable

Specification	Coefficient
1. Gelber (2011) without Controls	0.56** (0.26)
2. Gelber (2011) with Controls	0.53** (0.25)
3. Gelber (2011) with Controls and Initial Balance	0.49* (0.25)
4. Gelber (2011) Treatment Limited to 401(k) Participants	0.77* (0.40)
5. Gelber (2011) Treatment Limited to 401(k) Non-participants	0.43 (0.28)
6. Treatment Group Owns 401(k) in Period 2, but not Period 1 Control Group Does Not Own 401(k) in either Period	0.85* (0.48)

Note: Robust standard errors are in parentheses. * denotes .05 < p ≤ .10. ** denotes .01 < p ≤ .05. *** denotes p ≤ .01.

REFERENCES

Benjamin, Daniel J. 2003. "Does 401(k) Eligibility Increase Saving? Evidence from Propensity Score Subclassification." *Journal of Public Economics* 87: 1259-90.

Bernheim, B. Douglas. 2002. "Taxation and Saving." In *Handbook of Public Economics*, Vol 3. Edited by Alan J. Auerbach and Martin Feldstein. Amsterdam: Elsevier.

Chetty, Raj, John N. Friedman, Soren Leth-Petersen, Torben Heien Nielsen, and Tore Olsen. 2014. "Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark." *Quarterly Journal of Economics* 129 (3): 1141-219.

Copeland, Craig. 2002. "IRA Assets and Characteristics of IRA Owners." *Employee Benefit Research Institute Notes*, 23 (12): 1-9.

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): 113-38.

Engen, Eric M., and William G. Gale. 2000. "The Effects of 401(k) Plans on Household Wealth: Differences Across Earnings Groups." NBER Working Paper 8032: Cambridge, MA.

Heim, Bradley T., and Shanthi P. Ramnath. 2016. "The Impact of Participation in Employment-Based Retirement Savings Plans on Material Hardship." *Journal of Pension Economics and Finance* 15 (4): 407-28.

Hubbard, R. Glenn, and Jonathan S. Skinner. 1996. "Assessing the Effectiveness of Saving Incentives." *Journal of Economic Perspectives* 10 (4): 73-90.

Gelber, Alexander M. 2011. "How do 401(k)s Affect Saving? Evidence from Changes in 401(k) Eligibility." *American Economic Journal: Economic Policy* 3 (4): 103-22.

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.

U.S. Department of the Treasury. 2016. "Tax Expenditures." Office of Tax Analysis: September 28.

The Tax Policy Center is a joint venture of the
Urban Institute and Brookings Institution.

For more information, visit taxpolicycenter.org
or email info@taxpolicycenter.org