Abandoned Retirement Savings*

Lucas Goodman Anita Mukherjee Shanthi Ramnath

2021-06-10

Abstract

Retirement savings abandonment is a rising concern connected to defined contribution systems and default enrollment. We use tax data on Individual Retirement Accounts (IRAs) to establish that in 2017, 2.7% of 72.5 year-old account-holders in total abandoned $790 million; the median abandoned account held $5,400. Nearly all of these funds remain with plans and are not sent to state unclaimed property. Regression discontinuity estimates show that abandonment is 10 times higher in automatic rollover IRAs, a type of default account. We nest our findings in a model of retirement savings featuring forgetting to derive implications for passive and active savers.

JEL Codes: D83, H24, H31, J32, J14, J63
Keywords: retirement savings, defaults, memory, escheatment

*Goodman: U.S. Department of the Treasury, lucas.goodman@treasury.gov. Mukherjee: University of Wisconsin-Madison, anita.mukherjee@wisc.edu. Ramnath: Federal Reserve Bank of Chicago, sramnath@frbchi.org. We thank Phil Armour, Dani Bauer, Youssef Bezarti, Peter Brady, Taha Choukhmane, James Choi, J. Michael Collins, Keith Ericson, Maria Fitzpatrick, John Friedman, Bill Gale, Gopi Shah Goda, Damon Jones, Benjamin Keys, Olivia Mitchell, John Shoven, Jonathan Skinner, Justin Sydnor, Jeremy Tobacman, and seminar participants at the 2021 NBER Aging Program, 2021 Virtual Public Finance Seminar, 2021 Michigan Retirement and Disability Research Center meeting, 2021 Cherry Blossom Financial Education Institute at The George Washington University School of Business, 2020 Stanford Working Longer and Retirement Conference, 2020 Financial Capability & Asset Building Conference at Washington University in St. Louis, 2019 National Tax Association conference, 2019 and 2018 Social Security Administration conferences, RAND Corporation, Temple University, Northern Illinois University, San Diego State University, University of St. Gallen, and the University of Wisconsin-Madison for helpful comments. Corina Mommaerts provided extraordinary guidance on the state unclaimed property analysis. Lois Miller, Arpita Patnaik, and Alan Yang provided excellent research assistance. Erin Egan provided helpful information about state unclaimed property databases. We are grateful to the U.S. Social Security Administration and Wisconsin Alumni Research Fund for funding. All views are our own and not of the funding agencies. The views expressed in this paper do not necessarily represent the official position of the U.S. Department of the Treasury, The Federal Reserve Bank of Chicago, or the Federal Reserve System.

1 Introduction

A substantial share of private retirement savings in the U.S. is accumulated in tax-preferred retirement savings plans that account-holders must manage. If individuals fail to keep track of these accounts over their lifetime, they risk that their funds will be forgotten (i.e., abandoned or unclaimed). Despite increasing policy concerns about retirement account abandonment (GAO, 2019; Bonamici, 2020), to date even basic facts about the prevalence of abandoned accounts are unknown. This information is important in the context of forces that could contribute to increased retirement account abandonment, specifically the shift to defined contribution plans (which require individual management) and an increased use of default enrollment in retirement saving plans (which increases passive savings that must be remembered later in life).

In this paper, we analyze multiple datasets derived from IRS administrative records and state unclaimed property databases to study abandoned accounts. In the first part of our paper, we establish the share of abandoned accounts among the retirement-age population. We focus on Individual Retirement Accounts (IRAs), which is where retirees hold the vast majority of their account-based retirement wealth (Goodman et al., 2019). We find that 2.7% of 72.5 year-old owners of IRAs in 2017 had an abandoned account, with a total value of $790 million.[1] In the second part of our paper, we study a population of "passive" savers (Chetty et al., 2014) induced into holding an IRA through a default policy known as automatic rollover. Using a treatment-effects framework, we estimate that abandonment is approximately ten times higher among those induced into enrolling in an IRA by a default policy.

To estimate the prevalence of abandoned accounts, we must overcome a measurement challenge. Given that retirement accounts are long-term saving vehicles, an individual might reasonably choose not to interact with the account for a long period of time, while being fully aware of the account's existence. To measure abandonment, we make use of a tax law that mandates individuals take required minimum distributions (RMDs) when they reach age 70.5. In brief, we define an account to be abandoned if the individual fails to take a distribution from the account during any of the first three years in which the RMDs were required to be taken. (We define this more precisely in Section 3.1.1, including an adaptation of this definition for owners with multiple accounts, and a discussion of measurement error related to unawareness of RMD rules.) Because individuals can own multiple IRA accounts,

[1] The analysis to follow uses 2012 USD as the monetary unit; 790 million in 2017 USD is approximately equal to 730 million in 2012 USD.

the abandonment rate at the account level is lower than the rate among account owners: 2.2% versus the aforementioned 2.7%. While abandoned accounts tend to be small in value relative to non-abandoned accounts, they are not economically trivial: conditional on abandonment, the median account value is $5,351, or about 12% of the account holder's income.[2]

A natural question to arise is what factors are correlated with abandonment. We find that abandonment decreases with account balance, but remains substantial at higher balances – e.g., abandonment was 3% for accounts valued near $10,000 in 2017. Holding account balance fixed, we find that measures of financial sophistication observable in the tax data – filing a tax return, earning capital income, and paying estimated tax – are negatively correlated with abandonment, as expected. Our analyses also reveal that abandonment is positively correlated with the non-white share of the population within a zip code, even after controlling for education, income, and population density.

We estimate that IRA balances totaling $1.79 billion were abandoned cumulatively between 2003 and 2008. We then examine the extent to which account owners show interest in reclaiming these abandoned assets as measured by any distribution from the account. We find that among abandoned accounts, about 60% of those valued at $3,000 or more are eventually reclaimed within 10 years. Account value positively predicts reclaiming, but the behavior levels off at about 60% even for high-value accounts, suggesting a substantial level of continued abandonment. In total, about 60.7% (or $1.09 billion) of funds held within abandoned IRAs were reclaimed within ten years (or would have been reclaimed if the individual remained alive for that period).

Next, we turn to state unclaimed property data, which allow us to study retirement accounts that plans have turned over to state governments (i.e., escheated accounts). We find that 36% of these accounts had a balance of less than $100, suggesting that plan fiduciaries escheat accounts that have management costs exceeding returns. There are almost no escheated accounts worth more than $10,000, though there is substantial density in this range in the tax data. We conduct an analysis comparing the escheated accounts to those in the tax data for six states and find that plans escheated only 2.3% of abandoned funds. The unclaimed property data also allow us to examine the effectiveness of state policies in reuniting abandoned funds with their owners. We choose Massachusetts and Wisconsin to focus on as they represent two extremes of effort a state expends in finding account owners, where Wisconsin is a "high" effort state and Massachusetts is a "low" effort state. In Massachusetts, only 3.4% of unclaimed retirement accounts reported in 2016 were claimed

[2] To protect taxpayer privacy, all quantiles in this paper derived from tax data are quasi-percentiles, equal to the mean of the 30 individuals closest to the reported percentile.

within two years, compared to 67% in Wisconsin. The positive relationship between account balance and reclaiming also appears in the Massachusetts data. Overall, these findings suggest that most state unclaimed property portals currently play a limited role in reducing retirement account abandonment.

Having established a set of stylized facts related to abandonment during retirement age, we examine a subset of working-age account holders who might be at higher risk for abandonment. In particular, we study the behavior of individuals induced to engage in an automatic rollover (also known as "forced transfer") of savings left with a former employer – i.e., a subset of individuals who could plausibly be described as "passive savers" (Chetty et al., 2014). The policy allows employers to create a default option for employees leaving their jobs: if the separating employee takes no action, an account with a balance between $1,000 and $5,000 is allowed to be transferred automatically into an IRA designed for this purpose.[3] This policy creates two empirical discontinuities, which we use to impute abandonment in a treatment-effects framework among these passive savers "complying" with this treatment. Under the admittedly strong assumption that any difference in certain behaviors (conditional on amount of the rollover) is due to abandonment, we estimate abandonment rates that are much higher for defaulted savings than for traditional IRAs – ranging from 23% to 45% (versus 2.7%) depending on the specification and threshold used.

This analysis of forced-transfer IRAs serves two purposes. First, because those affected by this default are primarily working-age, it suggests that abandonment could increase in importance as the current working-age population reaches retirement. Second, it informs welfare implications of default policies in general. In the policy discussion section of our paper, we develop a model of retirement saving with forgetting and show the conditions under which the welfare benefits of default policies can be eroded. Forgetting in our model could also represent other mechanisms that lead to abandonment such as hassle costs, information frictions, or rollover difficulties. We show that the conceptual implications of abandonment differ for passive and active savers: specifically, the latter group may under-save out of fear of failing to manage and remember the account. Additionally, the empirical conflation of myopia with forgetting means that traditional lifecycle models may overestimate myopia among active savers. We also discuss the model's connection to naivete versus sophistication in forgetting.

[3]Employees who leave their jobs always have the option to make an active choice to take a cash distribution or to roll their account over into an IRA of their choosing. If the account has a balance less than $1,000, this default option is allowed to be a full distribution in cash. These "force-out" policies are meant to allow firms to unburden themselves of small accounts, which can bear relatively high administrative costs. If the account balance is greater than $5,000 then the employer must offer the option to maintain the account.

This study contributes to several strands of literature. The first relates to passive behavior in retirement saving, which we connect to potential abandonment later in the lifecycle. Default policies such as auto-enrollment in retirement plans, where participants must choose to opt-out of plans rather than opt-in, are shown to substantially increase plan participation (Thaler and Benartzi, 2004; Chetty et al., 2014; Madrian and Shea, 2001; Benartzi and Thaler, 2007). This literature is complemented by research demonstrating inertia in retirement plan choices (Kim et al., 2016) and in other settings (e.g., in tax withholding as studied in Jones 2012). Most of this prior work studies consumer choice and policy design with the purpose of shifting behavior towards an outcome deemed to be more desirable. There is continued debate, however, about the benefits of auto-enrollment in plan design (Bubb and Warren, 2020; Scott et al., 2020; Bernheim and Gastell, 2020). We contribute to this literature by emphasizing a less recognized drawback of default policies: the possibility that such accounts, due to being less salient (Ekerdt and Hackney, 2002), are more vulnerable to becoming abandoned. Plan participants who are defaulted into saving have been shown to have lower financial literacy (Goda et al., 2020; Carroll et al., 2009), and therefore are especially at risk for abandoning accounts. Our results suggest this to be the case.

Our study is also related to literature on forgetting that can lead individuals to "leave money on the table." Compelling evidence of forgetting has been documented in settings such as payment choices for a task (Ericson, 2011) and lapse-based insurance (Gottlieb and Smetters, 2021), where the main takeaway is that people exhibit overconfidence in their prospective memory. Ericson (2017) shows theoretically that ignoring limited memory can lead us to overestimate present bias. Moreover, the paper considers the effects of procrastination, which is likely present in behaviors such as retirement account consolidation. Yet prior work also shows a relative unwillingness of participants to invest in reminder policies that could curb such forgetting (Rogers and Milkman, 2016). Our paper contributes to this literature by incorporating forgetting in the retirement savings context, which has implications for various policies including defaults, reminders, and deadlines to roll funds over. We note that hassle costs could also explain at least some abandoned accounts. These costs have shown to be important in tax filing (Benzarti, 2020), health insurance (Baicker et al., 2012), and unemployment insurance (Ebenstein and Stange, 2010).

Finally, our research is nested within a more general literature on retirement savings adequacy (Poterba, 2014) and consumption smoothing in retirement (Banks et al., 1998). Prior studies have examined whether households are saving enough for retirement (Scholz et al., 2006; Skinner, 2007), while others have studied the role of specific factors such as financial literacy (Lusardi and Mitchell 2014, Clark et al. 2006), information (Mastrobuoni, 2011), and “leakages” (i.e., cash-outs) at job separation (Armour et al., 2016; Clark et al., 2014; Munnell and Webb, 2015) or more generally prior to retirement (Goodman et al., 2019). Our results suggest that even if individuals are saving during their working lives, the risk of account abandonment could have implications for optimal consumption during retirement.

In the next section, we provide background on unclaimed retirement accounts. Section 3 describes the individual level tax data and the state unclaimed property data. Section 4 contains analysis related to the prevalence, trends, and correlates of abandoned accounts. Section 5 presents analysis on a specific type of default retirement saving account, the forced transfer IRA, to show how abandonment changes with default enrollment. Section 6 provides a model-based policy discussion and Section 7 concludes.

2 Institutional Background

We begin with a brief background on the policies related to abandoned accounts. The GAO (2019) defines unclaimed retirement savings as “savings that individuals are entitled to receive, but have not claimed because employers or other entities that maintain their retirement accounts cannot locate the individuals or because the individuals have forgotten about the savings.” We adopt a similar conceptual definition of abandonment, noting that the behavioral forces explaining this behavior are empirically difficult to distinguish. For example, forgetting – a force we highlight in our model in Section 6 – is best considered as the case when the account owner is unaware of the abandoned account. By contrast, a force of rational inattention would imply that the account owner is aware that the account exists, but that the hassle cost of accessing those funds exceeds the value of the account. We do not attempt to isolate these drivers, but consider them together as forces contributing to abandonment.

We focus on forgetting in our model as it is likely an important factor in managing retirement accounts, which are long-term savings vehicles that are designed to remain untouched for decades. In fact, distributions that take place before an individual turns 59.5 are generally penalized. Yet, withdrawals from accounts *must* begin at age 70.5, and there is an individual-specific minimum amount that must be withdrawn each year referred to as the “Required Minimum Distribution” (RMD).[4] The RMD is calculated by dividing the individual’s tax-

[4] The 2020 Setting Every Community Up For Retirement Enhancement (SECURE) Act increased the RMD age to 72 for years after the end of our sample period. If an individual is still working at or after age

deferred retirement account balance as of December 31 in the prior year by that individual's life expectancy factor from the IRS Uniform Life Table. Failure to meet one's RMD triggers a hefty penalty tax equal to 50% of the RMD; additionally, this tax contains no statute of limitations. In practice, however, we observe few individuals reporting payment of this penalty (or requesting a waiver of the penalty) on the necessary tax forms.[5] Congress has occasionally eliminated the RMD requirement during periods of national hardship: in both 2009 and 2020, Congress passed legislation that waived all RMDs so as not to force owners to liquidate funds during a potential downturn.

States use rules based on the RMD to determine whether a retirement account should be considered as an unclaimed asset. In particular, states specify a dormancy period, typically three to five years, after the date for when an RMD should have taken place. If no distribution is taken at any point before that extended period ends, then state policy generally mandates that the participant's assets be remitted to the state, i.e. plans must escheat unclaimed funds. Each state then hosts an unclaimed property database aimed to link individuals with their unclaimed assets. The state unclaimed property databases generate meaningful activity, with owners having claimed $25 million in retirement savings in 2016 according to data from 15 states (GAO 2019).[6] These unclaimed accounts had an average value of $601 from 401(k) plans and $5,817 from traditional IRAs.

Reducing the number of unclaimed retirement accounts is a key policy concern in the U.S. In particular, Congress proposed the Retirement Savings Lost and Found Act of 2020, which is focused entirely on this issue. The Act would expand an existing online database of pension and 401(k) account owners to help facilitate matching owners with their lost plans.[7] The legislation would also clarify the rules for categorizing account owners as "missing" and place a greater burden on employers and plan managers to find these individuals. We note that some state policies could exacerbate the problem of abandonment, however: Pennsylvania in 2016 attempted to categorize retirement accounts as unclaimed if there was inactivity for just three years *even during working age*, but wide criticism of the law kept it from becoming enforced.[8] Determining which types of policies would be most

70.5, he or she can defer the RMD from the DC plan of that employer until he or she retires. The RMD age affects personal financial decision-making in many ways including options to annuitize defined contribution savings, as studied in Horneff et al. (2020), Mortenson et al. (2019), and Brown et al. (2017). RMDs do not apply to Roth IRAs; we do not study Roth IRAs in this paper.

[5]It is possible that the IRS may impose this penalty through an enforcement action that we do not observe.

[6]Note that this report was limited to aggregated state-level data, not account-level data as we study.

[7]This proposal goes far beyond the current Pension Benefit Guaranty Corporation (PBGC) registry, which contains pension plan information (and DC plans as of 2019) for terminating plans.

[8]Hopkins (2018) provides a discussion of Pennsylvania's unclaimed property laws related to retirement.

effective in reducing abandonment requires an improved understanding of abandonment and its predictors, which our study aims to contribute.

3 Data Description

Our analysis makes use of two individual-level administrative data sources: (1) federal tax and information returns and (2) state unclaimed property records. Using both datasets in tandem enables a comprehensive analysis of abandoned retirement accounts in the U.S.

3.1 Administrative Tax Data

We use data drawn from the near-universe of tax returns and information returns. In general, our analysis spans tax years 1999 through 2018, though we restrict to subsets of years for certain analyses.

Information returns are separate from tax returns and include forms sent by a third-party to the IRS and, in some cases, also to the relevant individual. We use two information returns in particular. First is the Form 1099-R, which reports information on distributions made from pensions, IRAs, and similar accounts. This form is used to identify the amount of a distribution and whether it came from an IRA (versus another account). The custodian (i.e., the financial services firm that manages the account) also reports (in Box 7) up to two codes which give further information about the nature of the distribution. In particular, we use the presence of Box 7 codes "G" or "H" to infer that the distribution was part of a direct rollover. The Form 1099-R also includes an identifier for the custodian of retirement funds, which we can link across IRS datasets. We use Form 1099-R to identify whether an individual received a distribution from a given IRA; we also use this form to determine the running variable in our analysis of forced transfer IRAs.

The second information return used is Form 5498, which reports information about IRAs held by an individual. Importantly, this identifies the type of IRA (traditional, Roth, or certain specialized types of small-business IRAs), the value of the IRA as of the end of the year, and any contributions made (rollover or otherwise) during the year. Form 5498 is sent to the IRS every year in which the IRA maintains a positive value, though the account holder typically receives a copy only when a contribution is made. As with Form 1099-R, Form 5498 also includes an identifier for the custodian of the account.

For our analysis, we will focus on IRAs largely because of data availability. We observe IRA account-holders on Form 5498, even for individuals who are not interacting with the account. By contrast, we do not have analogous information on holders of DC plans held through employers. Fortunately, Goodman et al. (2019) estimates that nearly 90% of distributions from the joint DC-IRA system to those age 65 or greater in 2015 were from IRAs, not DC accounts. The prevalence of IRAs among this group suggests that most individuals have rolled their DC plans over into an IRA prior to reaching retirement-age.

We also make use of other components of the IRS database for our analyses. For example, we use the Death Master file, which records each individual's date of birth, date of death, and sex. We use the Form 1040 to obtain a comprehensive measure of income. Additional analyses require controls for Social Security income, which we obtain from Form 1099-SSA (an information return).

3.1.1 Defining Unclaimed Accounts in Tax Data

Unlike in the state unclaimed property data, where accounts will by definition be unclaimed, in the tax data we must infer which accounts have been abandoned. We begin with the universe of traditional IRA accounts held by individuals at age 72.5 (whose date of death, if applicable, is later than the year in question), which are indexed by the combination of individual i and custodian j.[9] We define an account ij to be abandoned if it satisfies each of the three following conditions:

1. The account ij has existed, with positive value, for at least four years (i.e., the 69.5 year, the 70.5 year, and the 71.5 year, in addition to the 72.5 year).

2. There have been no distributions from that account ij (as measured on Form 1099-R) over those four years.

3. The individual i has not satisfied his or her RMD in any year (i.e., the 70.5 year, the 71.5 year, and the 72.5 year).[10]

Using this criterion, in Section 4 we estimate the prevalence of unclaimed accounts in the U.S. as observed in the tax data. It is possible that some accounts identified by this procedure are not truly "abandoned." In particular, some account-holders might be aware of the account but failed to comply with RMD rules, either intentionally or because they are

[9]That is, if an individual receives multiple Forms 5498 from the same custodian, we aggregate to the individual-custodian level, after dropping Forms 5498 that appear to be duplicates on important variables.

[10]If an individual held only one account, then condition (2) would imply condition (3). However, the RMD is calculated with respect to *all* IRA distributions and *all* IRA assets held by the taxpayer; the taxpayer can remain in compliance with RMDs by taking a larger distribution from one IRA (j) and taking no distribution from another (j'). Our algorithm conservatively would *not* consider j' to be abandoned. Furthermore, for this purpose, we measure compliance with RMDs using the larger of 1099-R distributions and taxable IRA distributions reported on Form 1040.

unfamiliar with the RMD rules. While these may be reasons for measured abandonment in our setting, our analyses (to follow) present evidence that the abandonment we observe is inconsistent with either of these forms of non-compliance.

3.2 State Unclaimed Property Data

Our data on escheated retirement accounts comes from state unclaimed property (SUP) databases. These data contain account-level information on each unclaimed property, and include details such as the type of account (IRA or pension, for example), account balance, and names and addresses of the account owner and account custodian. The name and geographic information of the account owner enable inference of owner characteristics such as ethnicity, sex, and neighborhood covariates. The data are collected separately from each state's division of unclaimed property (or related agency). An advantage of the SUP data is that all states use a standard reporting format following guidance from the National Association of Unclaimed Property Administrators (NAUPA). Our sample includes as many states as possible (up to 13) for different analyses.[11]

We use property codes associated with each unclaimed property to identify retirement accounts. The SUP data contain numerous property types – NAUPA lists 123 categories – including uncashed checks, securities, insurance property, mineral proceeds, and trusts. We isolate retirement accounts according to the codes listed in Table B.1, which include pension checks, traditional IRAs, pension and profit-sharing plans, and annuities. We supplement these data with information from the American Community Survey, the U.S. Census, and Social Security names databases to further ascertain the characteristics of these unclaimed account owners. We also analyze information on claims from unclaimed property from two states, Massachusetts and Wisconsin, to better understand the effects of state policy on reducing abandonment.

While much of our primary analysis will focus on results derived from tax data, an advantage of the SUP data is that we can look at abandonment across all types of retirement plans, not just IRAs. We will use this information to construct a national estimate of abandoned retirement accounts of all types that are held in these state portals. For the main analyses, however, we will focus on traditional IRAs in the SUP data to obtain an estimate of abandoned account escheatment.

[11] States vary in their policies for sharing data on unclaimed property with researchers. Most states referenced us to aggregated information on their websites. The 13 states that responded with data containing account-level property type and account balance of unclaimed properties were AK, CA, FL, HI, LA, MA, MN, ND, NV, OH, PA, TX, and WI.

4 Stylized Facts on Abandoned Accounts

We find clear evidence that a small but meaningful share of retirement accounts become abandoned by the time an account owner turns age 72.5. We begin by plotting the abandonment share over time, using administrative tax data, in the left panel of Figure 1. The solid line plots the share of IRAs that are abandoned over time, beginning with 1.0% in 2003 to 2.2% in 2017. The dashed line aggregates the accounts to the individual level and plots the share of individuals who owned at least one abandoned IRA. This share was 1.2% in 2003 and rose to 2.7% in 2017. The right panel of Figure 1 plots the time trend in the dollar value of abandoned accounts (in 2012 dollars). This value increased from $285 million in 2003 to $730 million in 2017. In each of these two years, the aggregate amount abandoned represents approximately 0.35% of the total value of IRAs held by age 72.5 individuals. Note that the figures contain dashed vertical lines separating out the period 2009 through 2011 – we drop observations for individuals turning 72.5 in each of these three years due to a change in RMD policy that affected these cohorts. Specifically, RMDs were waived in 2009 due to the Great Recession, so our definition of abandonment applies differently to these individuals.

Next, we examine the relationship between abandonment and account balance. The reason we do this analysis is to shed light on the mechanisms leading to abandonment. If rational inattention is a key factor for abandonment, we would expect to see abandonment concentrated among low-value accounts. Alternatively, if forgetting plays a significant role, we would expect to see a substantial level of abandonment throughout the account value distribution (though likely forgetting is also a function of account balance). We uncover evidence consistent with a mix of these forces, as abandonment decreases with account balance but remains a problem even for high balance accounts. These results are shown in Figure 2, which plots the abandonment share by value of that account at age 72.5, separately for the early period of our sample (2003-2006) and the later part of our sample (2014-2017). The relationship appears linear on the log-log scale, with abandonment decreasing in account size, with the highest levels of abandonment around 30% (10%) for the later (earlier) sample. Nonetheless, abandonment is substantial for accounts that are large in value: the abandonment share for accounts worth $10,000 is about 1.5% in the early sample and 3% in the later sample.

We provide individual-level context for the abandoned account values in Table 1. Column (1) presents the 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentile of the value of abandoned accounts; column (4) presents the same quantiles for accounts that are not abandoned.

Panel A aggregates from 2003 through 2006, while Panel B aggregates from 2014 through 2017; we show this breakdown to illustrate the trends in abandonment over time. The table shows that, as expected, the median abandoned account was much smaller than the median non-abandoned account in both periods. Furthermore, except at the 90^{th} percentile, each quantile of abandoned accounts shrank in value from the early part of our sample (2003-2006) to the late part of our sample (2014-2017). Columns (2) and (5) report various quantiles of the ratio of account value to income at age 72.5 for abandoned and non-abandoned accounts, respectively.[12] As an alternative measure, in columns (3) and (6), we convert the individual account balances to an annuity value, using the Social Security life expectancy tables and a (conservative) assumption of a 1% interest rate. The median abandoned account was equal to 24.1% of income in the 2003-2006 period and 12.7% of income in the 2014-2017 period; these correspond to annuity values of 2.0% and 1.1% of income, respectively. We also observe that a non-trivial minority of abandoned accounts are much larger. At the 75^{th} percentile, the annuity value of an abandoned IRA represents 6.5% (2003-2006) or 4.8% (2014-2017) of income; at the 90^{th} percentile, these values are 16.7% and 14.5%.

We supplement our results from the tax data with estimates on the prevalence of abandonment using state unclaimed property records.[13] In our sample of 13 states, about 36,500 retirement accounts totaling $18.3 million were escheated in 2016. The summary statistics reveal that abandonment is a problem that extends beyond IRAs – only 25% of the escheated accounts are of this type. The other large categories of escheated retirement accounts are pension and profit sharing plans (42%), uncashed pension checks (30%), and Roth IRAs (2%). Despite its name, the category of pension and profit sharing plans consists of DC accounts such as 401(k)s.[14] Pension checks are a less defined category and could include uncashed defined benefit (DB) distributions as well as required minimum distributions sent by DC account custodians.

Next, for the purposes of producing a national estimate, we assume that these accounts belong to 74 year-olds and extrapolate the account-level data to the entire U.S. using a log-linear model. In this exercise, we estimate that about 70,000 retirement accounts totaling $38 million were escheated in 2016. This amounts to about 3.3% of 74 year-olds having an abandoned retirement account (of any type, not only IRAs) in 2016 with an average value of $547. An immediate takeaway is that the accounts sent to the state are much smaller

[12] Our measure of income is adjusted gross income, plus non-taxable Social Security, minus taxable IRA distributions. The measure is meant to capture disposable income and therefore includes income sources that are not taxed but still available for consumption.

[13] The analysis is in Table B.3.

[14] https://www.dol.gov/general/topic/retirement/typesofplans

compared to the tax data summarized in Table 1, which occurs because plans do not send most abandoned retirement accounts to the state – we will show this in Section 4.3. For that reason, and because the escheated data contain also non-IRA accounts, we highlight that the 3.3% estimate is not directly comparable to the tax-data based analysis.

4.1 Factors Correlated with Abandonment

As shown in the left panel of Figure 1, we find that 2.7% of all IRA account owners subject to an RMD had an unclaimed account in 2017.[15] In this section, we examine individual-level factors that are correlated with this abandonment. Table 2 shows estimates of a regression of abandonment on relevant covariates, conditional on year fixed effects interacted with granular bins of account value; the interactions mean that these results can be interpreted as holding account value fixed. The first column includes four measures of financial sophistication: filing a tax return, paying estimated tax during the year, having non-zero capital gains or dividend income, and having interest income.[16] Across each proxy for sophistication, the results show that financial sophistication is associated with lower abandonment. Additionally, the magnitudes of these coefficients are stable across the columns as controls are added. This suggests that intentional non-compliance is not the main driver of our measure of abandonment, since we would expect knowing non-compliers to tend to be more sophisticated.

The second column adds two demographic variables: sex and race. We examine these variables as they may capture a range of behaviors or vulnerabilities related to abandonment. We directly observe sex for each individual as reported in the tax data, but information on race is not available – instead, we use the zip code of the account owner and calculate the share of population that is white for that zip code (from the American Community Survey) as a proxy. We find that men are slightly (0.1 percentage points) more likely to abandon accounts. More substantially, we find that the white share in a zip code strongly predicts abandonment. Moving from a zip code that is 50% white to a zip code that is 100%

[15] A striking pattern in Figure 1 is the more-than-doubling of abandonment of IRAs between 2003 and 2017. Figure 2 shows that the proliferation of small-balance accounts cannot explain this increase. In particular, the figure shows that the probability of abandonment conditional on IRA value increased through (almost) the entire distribution of IRA values. That is, holding IRA value fixed, the probability of abandonment still increased from 2003 to 2017. To explore other possible factors in this rise, we performed an Oaxaca-Blinder style decomposition of the increase into the amounts explained by changes in observable characteristics such as account value and financial sophistication. These too do not explain the rise in abandonment, however; the analysis is presented in Figure B.1. Further explanations for the increase in abandonment over time is a fruitful area for further research.

[16] For this exercise, we use the sample from 2012 onward because certain variables, most notably the amount of estimated tax paid, are not available to us for earlier years.

white reduces abandonment by 1.3 percentage points – a large effect relative to the mean abandonment of 1.55 percentage points. Column (3) adds a covariate for the population density of the account owner's zip code, which reduces the coefficient on the white share by about 15%. Column (4) adds additional zip code level demographics, including education and poverty status – these have little effect on the estimated coefficients. Column (5) adds fixed effects for year interacted with the custodian (j); this specification reveals the extent to which the other covariates show an effect through the channel of individuals selecting into different financial service firms. The additional controls modestly diminish the partial correlation between abandonment and the presence of capital income, reducing the magnitude of the coefficients on having dividends/capital gains or interest income. Additionally, this set of fixed effects lowers the coefficient on the share white by an additional 18%, suggesting that a modest share of the component of the race correlation is driven by selection into firms. That is, firms with clients in more minority neighborhoods appear to be associated with greater abandonment.[17]

Lastly, column (6) adds a proxy for the awareness of tax rules in a given zip code: the share of 2014-2018 Earned Income Tax Credit (EITC) recipients with children that (1) have self-employment income and (2) have earnings within $500 of the first EITC kink point. As argued by Chetty et al. (2013), this is a proxy for knowledge of the EITC schedule, which is plausibly related to knowledge about other tax rules, such as RMDs, as well. We find that the coefficient on this term is significantly positive, but small in magnitude. In particular, moving from the 25th to 75th percentile of sharp bunching (i.e., from 1.58% to 3.78%) would increase the probability of abandonment by only 0.001 (that is, 0.1 percentage points), relative to the baseline mean of 0.0187. Nevertheless, to the extent that sharp bunching is a reasonable proxy for awareness of other tax rules such as RMDs, this result suggests that (if anything) people more aware of the rules are more likely to abandon their accounts.

To complement our analysis, we also analyze correlates of escheated accounts from the unclaimed property data. As with the tax data, we supplement the dataset with county-level demographic characteristics from the American Community Survey. Since we observe the first and last names of each account holder in this dataset, we can also predict sex and race from this information. Comparing means of escheated retirement account holders with the overall 74 year-old population, we find that individuals with escheated assets are 10 percentage points less likely to be female and 12 percentage points more likely to be

[17]Additionally, a small share (0.37%) of our data include individuals residing outside the United States, as indicated on Form 1040 or 1099-SSA. Conditional on year and IRA value fixed effects, such individuals are 12.7 percentage points more likely to meet our definition of abandonment.

of Hispanic origin. We also find a large difference in average county population: it is 2.1 million for those with escheated accounts versus 158,000 for all 74 year-olds, suggesting that abandonment – or at least escheatment – is more prevalent in urban areas. Other county-level characteristics, such as the age makeup, educational attainment, migration, and median household income are similar between individuals with an unclaimed retirement asset and the population-weighted average county characteristics.[18] As mentioned before, we highlight that the state unclaimed property data and the tax data are not comparable due to the known differences in account type, account value, and plan selection in the escheatment decision.

4.2 Reclaiming abandoned accounts in the tax data

We estimate the "reclaiming rate", i.e., the rate at which abandoned accounts are subsequently claimed by their owners. The reclaiming rate is important for a number of reasons. First, a reclaiming rate substantially less than 100% offers validity to our measurement of abandonment. Second, the welfare consequences of abandonment depend on when – and whether – the accounts are reclaimed. If the accounts are never reclaimed, then the savings within the IRA can be considered lost. On the other hand, if the accounts are reclaimed, then the individual may regain that consumption that would have been otherwise lost, though she may have made suboptimal consumption decisions while the account was abandoned. We measure reclaiming by whether the account owner receives any distribution from the abandoned account. For each year relative to the abandonment year (age 72.5), we compute the hazard of reclaiming as equal to the share of accounts with a distribution in that year, conditional on (1) not yet having been claimed and (2) the individual being alive for the entire year.[19] One caveat to this definition is that the act of escheatment might lead to a distribution, which we would erroneously categorize as reclaiming. We will show, however, that only a small portion of abandoned IRAs undergo escheatment.[20] The left panel of Figure 3 plots the cumulative hazard of reclaiming over a ten-year period. The one-year hazard is just under 20%, and the cumulative hazard increases steadily to approximately 58% over 10 years. Thus, a slight majority of accounts is claimed within ten years (conditional on not dying prior to that point); however, a substantial minority remains abandoned. Note that our restriction to living account owners likely increases our measure of reclaiming relative to the full population of abandoned IRAs.

[18] The unclaimed property analysis is presented in Table B.2.

[19] To have ten years of data, we restrict this analysis to those who were 72.5 years old in 2008 or earlier.

[20] In 2018 (after our sample period), the IRS issued Revenue Ruling 2018-17, which clarified that escheatment events should trigger a Form 1099-R issued in the individual's name. We could also fail to identify true reclaiming if the account custodian changes its Employer Identification Number.

The right panel of Figure 3 plots the 10-year cumulative reclaiming hazard separately by 25 bins of account value. The lowest-value abandoned IRAs, with account balances under $100, have a reclaiming rate of approximately 30%; this increases to 60% for accounts with balances around $3,000, after which point it remains approximately constant.[21] This suggests that one component of failure to reclaim is rational inattention, but such behavior cannot explain the 40% of high-value IRAs that remain unclaimed after 10 years. In sum, IRAs worth $1.79 billion were abandoned cumulatively between 2003 and 2008. Of these, we estimate that 60.7%, or $1.09 billion, were reclaimed within ten years (or would have been reclaimed if the individual remained alive for that period).

4.3 Reclaiming escheated accounts

We now turn to unclaimed property records to help understand what happens to abandoned accounts that are passed on to states. We first examine the extent that plans escheat abandoned retirement assets. While state policies prescribe escheatment after a dormancy period of two to five years, federal guidance causes this process to be unevenly applied. Plans may choose to do so when accounts are too costly to manage, or when they experience a structural change such as a merger or termination.

Figure 4 shows the number of individuals with abandoned IRAs that appear in both the escheatment and tax datasets in 2016. Note that these samples are not subsets of each other; the tax data are for individuals aged 72.5, which is prior to when escheatment would typically occur.[22] An immediate observation is that 36% of individuals in the escheatment data have account balances below $100 – a closer look reveals that all of these accounts are actually below $50 – suggesting that plans may be more likely to escheat accounts that have management costs exceeding returns. (It could also be that custodians act altruistically and attempt to locate abandoned account owners, but decide that the benefit of reuniting accounts below $50 with their owners is too small to justify that altruism.) There are almost no escheated accounts worth over $10,000, though there is substantial density in this range in the tax data. In total, Figure 4 shows that only 2.6% of abandoned IRA dollars are escheated ($4.4 million out of $170 million). A key takeaway from this chart is that state unclaimed property databases are not currently the most effective means of uniting abandoned accounts holders with their lost funds. In part, this is because a large portion of the escheated accounts

[21] Additionally, we examine the reclaiming behavior of individuals with multiple abandoned accounts. We find that, conditional on person fixed effects, an account worth 10 percent more is 0.3 percentage points more likely to be reclaimed within ten years.

[22] We use six states in this analysis: FL, LA, MA, MN, OH, and TX. The reduced sample size is due to keeping only states that report enough (non-missing) addresses for account owners to enable comparison with the tax data.

are likely to be rationally abandoned, as the utility costs of claiming can exceed the small value. For example, most state unclaimed property departments require account holders to provide proof of ownership via pay stubs or bank account statements to claim funds. We also note that from a policy perspective, these low-value accounts are likely to be the least important in terms of enhancing retirement security.

Next, we consider variation in reclaiming by state policy. In general, each state employs a different procedure for reuniting escheated accounts with their owners. We focus on Massachusetts and Wisconsin because they help illustrate the spectrum of such policies. In Massachusetts, like most states, unclaimed account owners must initiate claims to their funds through the state. By contrast, in Wisconsin, the state uses Social Security Number information to match unclaimed funds with their owners. This process started in June 2015 and is rare among states; it is known as the "Wisconsin model". In Wisconsin, it is known as the Department of Revenue (DOR) Auto Match.

Table 3 shows summary statistics on claiming for the two states. In Massachusetts, there were 3,320 retirement accounts reported as unclaimed to the state in 2016, of which only 3.4% were claimed within two years. The average account value of claimed accounts was much higher at $2,110 than unclaimed accounts ($581), consistent with our analysis of tax data. The Wisconsin claim data are available from 2016 to 2018, limiting our ability to study long-term claiming behavior. We observe that in 2016, there were 815 unclaimed retirement accounts reported to the state, of which 67% were claimed within two years. The average account value is higher than the national average at $980, and the ones that remain unclaimed are of higher value ($1,315 versus $812). The state's auto match appears successful, as 54%[23] of accounts are reunited with their owners via the match. Of the accounts that are claimed, 80% are done through this method; an additional 12% of account owners initiate claims on the state's website, and 6.5% of owners are connected to their lost accounts via an online locator service.

As we saw in Figure 3, accounts of higher value are more likely to be reclaimed in a ten year horizon. We investigate whether this relationship also holds for escheated accounts. Table 3 suggests this to be the case, however we provide a formal test below. The Massachusetts data are more representative of the country as most states are unable to use Social Security numbers to match unclaimed funds with their owners. As such, we continue with an analysis of the Massachusetts data to examine whether account value predicts claiming behavior.

[23]This equals the proportion claimed, 0.67, multiplied by the proportion claimed via DOR Auto Match, 0.80, from Table 3.

We take this estimation approach because the parametric analysis is less demanding of the (much) smaller dataset here compared to the tax data. Additionally, due to the varied content of the data, controls for property type are helpful in assessing the impact of account value with reclaiming. We estimate the following regression for each unclaimed account i separately for cutoffs of 1, 2,..., 13 years since the property was reported as unclaimed:

$$\text{Claimed}_i = \beta \ln(\text{Account Value}_i) + \eta_p + \gamma_y, \tag{1}$$

where $Claimed_i$ is equal to 1 if the property was claimed within 1, 2,..., 13 years (13 separate regressions, with decreasing sample size), and $AccountValue$ is in units \$10,000.

The fixed effects η_p and γ_y are for property codes and year reported as unclaimed, respectively. Figure 5 (with corresponding regression results in Table B.4) shows that the relationship between account value and claiming increases with time since the property is reported as unclaimed. In other words, higher value accounts are more likely than lower value accounts to be claimed many years after being reported as unclaimed. This pattern is consistent with rational inattention as a driver of abandonment, as it would predict that individuals knowingly abandon small balance accounts with greater likelihood.

5 Defaults and Account Abandonment

The previous section reported our best estimates of abandonment of IRAs among RMD-aged individuals. Savers in these cohorts tend to be fairly "active" (in the language of Chetty et al. (2014)), since individuals in these cohorts would have had to make an active decision to save. By contrast, over the past twenty years, there has been a policy push to expand participation in retirement plans using interventions, such as automatic enrollment, motivated by behavioral economics, (e.g., Madrian and Shea (2001) and Thaler and Benartzi (2004)). These interventions, by construction, predominantly affect "passive" savers. It is plausible that passive savers are at much higher risk of abandonment, both because the individuals themselves are less sophisticated and because the accounts into which they have been enrolled are not salient. In this section, we test this hypothesis directly. In particular, we study a population of individuals that are especially likely to be passive savers: those who are enrolled in a "forced transfer" IRA by virtue of it being the default, passive option.[24]

[24] Separately, one could directly examine the effect of automatic enrollment on abandonment. Given data constraints – in particular, the lack of data on DC accounts and the lack of widespread automatic enrollment into IRAs – we do not do so. Examining the effect of state-level automatic IRA enrollment – the earliest of which was implemented in 2015 – is a fruitful area for further research.

Our empirical identification comes from exploiting a policy that allows firms to "force out" accounts of separating employers, which are a function of the value of the account. In particular, employers are allowed to force a cash-out of any DB or DC balance under \$1,000 (with separating employees below age 59.5 generally paying a 10% early withdrawal penalty on DC distributions). For balances between \$1,000 and \$5,000, the policy mandates that any force-out distribution must be in the form of an IRA established for this purpose; these IRAs are known as forced-transfer IRAs or automatic rollover IRAs (we will use these terms interchangeably). Typically, the new IRA provider contacts these employees with information on how to roll their funds over prior to the force-out, and they are given 30 to 60 days to take action. If none is taken, the retirement savings are transferred to the automatic rollover IRA. The funds are invested in a default plan designed to protect principal, and the new account holders are mailed paperwork designating their ownership of a new plan.

We note that these rules are a minimum standard: employers are allowed to keep accounts below \$5,000 in their own plan, or use the automatic rollover IRA option for accounts below \$1,000. However, using data on plans maintained by Vanguard, Hung et al. (2015) find that most plans maintain only this minimum standard. Additionally, we show below that the \$1,000 and \$5,000 thresholds appear to be binding for many participants.

5.1 Empirical methodology

As the policies creating the automatic rollover IRA were fully implemented fairly recently, the subpopulation of workers affected remains mostly working-age in our sample. This timing offers the advantage of capturing a cohort who is more subject to default policies in retirement savings, but a drawback is that we cannot use our RMD-based definition of abandonment from Section 3.1.1. Thus, we impute abandonment in this population using regression discontinuities and a treatment-effects framework.

We begin by assembling a dataset of all individuals who make a rollover distribution to an IRA between 2005 and 2015.[25] Figure 6 plots the counts of observations in \$10 bins. There are approximate discontinuities in these counts exactly at each of the two policy thresholds. The counts jump immediately to the right of \$1,000. In the language of treatment effects, individuals to the left of \$1,000 are always-takers (AT): those who would have made an

[25] Specifically, we require the distribution (from Form 1099-R) to have a "direct rollover" code (typically "G") and to arise from a non-IRA, and we include the distribution in the sample only if we observe a Form 5498 with a similar rollover contribution amount (within 20%). We drop distributions that are even multiples of \$500; such distributions predominantly represent partial distributions (i.e., distributions of less than the full account balance), which by construction are not induced by the forced transfer policy. We discuss this point further in Appendix A.

IRA rollover regardless of their location with respect to the \$1,000 threshold. Those to the right of \$1,000 are a mix of always-takers and compliers (C): compliers, our main group of interest, are those who roll over into an IRA if and only if their account value was in excess of \$1,000. Likewise, the counts fall immediately to the right of \$5,000; those to the left of \$5,000 are a mix of compliers and always-takers, while those to the right of \$5,000 are exclusively always-takers. By construction, always-takers tend to be active savers: they made an active choice to roll their funds over into an IRA. On the other hand, compliers tend to be passive savers, since they are taking no action, even when prompted.

We recover the probability of abandonment among auto-rollover compliers using a three-step procedure. The first step is to use regression discontinuity to estimate the share of compliers at each threshold. This is simply equal to the magnitude of each discontinuity in counts, scaled by the counts immediately to the right (\$1,000) or left (\$5,000). For exposition, consider the \$1,000 threshold. Let x_i denote the rollover amount and let T_i denote an indicator that $x_i \geq 1000$. We aggregate to the level of single dollar of rollovers to compute C_x, the count of rollover distributions equal to x (which takes on only integer values in our data). We estimate the following local linear regression, where each observation is a single dollar amount, and "FS" denotes that this is the "first stage" regression:

$$C_x = \gamma^{FS} + \beta^{FS}(x - 1000) + \pi^{FS} T_x + \eta^{FS}(x - 1000) \times T_x + \epsilon_x^{FS} \tag{2}$$

Our estimate for the complier share, which we denote α, is equal to $\frac{\hat{\pi}^{FS}}{\hat{\pi}^{FS} + \hat{\gamma}^{FS}}$.

The second step is to again use a regression discontinuity design to estimate certain specific outcomes for compliers and always-takers. For exposition, we continue in the context of the \$1,000 threshold. We will use two observed binary outcomes that indicate whether the account holder: (1) had any financial interaction (i.e., distribution or contribution) with the IRA over the next five years and (2) updated their address with the IRA custodian over the next five years, conditional on moving at some point during that period. At the \$1,000 threshold, the mean outcome for always-takers y_{AT} is equal to the mean outcome for those to the left of the \$1,000 threshold: y_L. The mean outcome to the right of \$1,000, y_R comprises a mix of always-takers and compliers such that:

$$y_R = \alpha y_C + (1 - \alpha) y_{AT},$$

where α is the share of compliers, discussed above. We can recover y_R and $y_L = y_{AT}$ using standard local linear regression (maintaining the notation from (2)):

$$y_i = \gamma^y + \beta^y(x_i - 1000) + \pi^y T_i + \eta^y(x_i - 1000) * T_i + \epsilon_i^y \tag{3}$$

y_L is equal to $\hat{\gamma}^y$ and y_R is equal to $\hat{\gamma}^y + \hat{\pi}^y$. Using our estimates for α, y_L, and y_R, we can solve for the mean outcome of the compliers.

Once we have estimated these regression discontinuities, the third step algebraically imputes the probability of abandonment for auto-rollover compliers, which is our object of interest. Let A denote a dummy for abandonment. Let Y denote either of our two outcomes measuring interaction with accounts. We define A such that $Pr(Y = 1|A = 1) = 0$; that is, abandonment precludes any further interaction with the account. We also make the simplifying approximation that $A = 0$ for always-takers.[26] We can then express y_C, the outcome for passive savers, as follows:

$$y_C = E(Y|C) = Pr(A = 1|C)E(Y|A = 1, C) + (1 - Pr(A = 1|C))E(Y|A = 0, C). \tag{4}$$

The left-hand-side of this equation is estimated directly, as discussed above. By assumption, $E(Y|A = 1, C)$ is zero, so the first term on the right hand side disappears. Given an estimate for $E(Y|A = 0, C)$, we are left with one equation and one unknown, so we can solve for our object of interest, $Pr(A = 1|C)$. The term $E(Y|A = 0, C)$ represents the probability of Y for compliers, conditional on not abandoning the account. We make a natural, but admittedly strong, assumption in order to solve for that term: that the probability of Y for compliers who do not abandon is the same as it is for always-takers, i.e., $E(Y|A = 0, C) = E(Y|AT)$. This assumption would be violated if passive and active savers are different, even conditional on non-abandonment, in ways that drive Y. The need for such assumptions is not uncommon in these methodologies – for example, Goldin and Reck (2020) use preferences of time consistent choosers to proxy for the preferences of the time inconsistent choosers in a study of framing effects. In our setting, the assumption could be strong because taking a distribution from an IRA in the near future is likely to be correlated with (unobservable) characteristics that differ between compliers and always-takers. Likewise, it may be the case that compliers are less likely to update their address even if they do not abandon the account. If that is true, i.e. if $E(Y|A = 0, C) < E(Y|AT)$, our imputed abandonment shares will be overestimated. In the next section, we will provide evidence on how observed characteristics differ among these groups to lend support to our assumption.

[26] If some individuals to the left of \$1,000 are in fact passive savers – perhaps because their former employer opts to use forced transfers for accounts under \$1,000 as well – then our estimated abandonment shares for compliers will be too low. However, *none* of the 385 plans studied by Hung et al. (2015) follow this approach.

One complication to this empirical approach is we observe the rollover distribution amount (x_i), which contains some measurement error relative to the true balance of the DC account when it is valued (x_i^*). It is this latter quantity that is relevant for determining whether a separating employee is subject to forced transfers. One possible source of measurement error is that there is a short lag between when accounts are valued for determining whether they are eligible for an automatic rollover and when the distributions are actually made, during which period asset prices may change. This measurement error creates the slight "S" shape at each of the thresholds in Figure 6.

In our baseline approach, we address this noise by introducing a donut hole of \$250 (that is, \$875-\$1,125) at the \$1,000 threshold and \$400 (that is, \$4,800-\$5,200) at the \$5,000 threshold. We drop observations within this donut hole when estimating each of Equations (2) and (3). In Appendix Figure B.2, we show that the estimated abandonment shares are robust to changes in the width of the donut hole. As an alternative approach, we replace the binary T_i with a continuous function that (quickly) rises from 0 to 1 near the threshold, but not discretely at it. In particular, at the \$1,000 threshold, we define $T_i = \Phi\left(\frac{x_i - 1000}{\sigma}\right)$, where $\Phi(\cdot)$ indicates the normal distribution and σ is a parameter that represents the standard deviation of the measurement error between x_i^* and x_i. In Appendix Figure B.3, we show that the estimated abandonment shares are robust to this alternative method for various plausible values of σ.[27]

5.2 Results

We begin with the estimated share of compliers for each threshold-outcome combination using equation (2). The first outcome, whether the individual has any interaction with the IRA over the next five years, implies that the share of compliers (α) at the \$1,000 threshold among the full population is 0.555. This is consistent with the graphical evidence in Figure 6 as the counts just to the right of \$1,000 are approximately double of the counts just to the left of \$1,000. Analogously at the \$5,000 threshold, we estimate an α of 0.377. The second outcome on whether individuals who move update their address with the IRA custodian produces estimated complier shares of 0.567 and 0.355 for the respective thresholds.[28]

Next we compute means of both "interaction" outcomes for compliers and always-takers. The top panels of Figure 7 plot the mean probability of having any interaction

[27] In the baseline, we use a \$400 bandwidth at the \$1,000 threshold and a \$800 bandwidth at the \$5,000 threshold. Appendix Figure B.4 shows how our results vary by bandwidth.

[28] For this sample, we require the zip code, as indicated by Form 1040, changes from one valid zip code to another valid zip code at some point within five years of the rollover.

with the account over the next five years. The left panel plots the data around the $1,000 threshold and the right panel plots the data around the $5,000 threshold. This illustrates the clear discontinuity in the share of individuals who subsequently interact with their account, implying that y_C is substantially less than y_{AT}, consistent with greater abandonment among compliers. The bottom panels repeat these plots for the second outcome of updating the IRA address conditional on moving. These bottom panels give similar results as the top panels: there is a clear discontinuity, implying a substantial gap between y_C and y_{AT}, and a fairly large level of estimated abandonment among passive savers.

The corresponding estimates for Figures 6 and 7 are presented in Table 4. The first two columns use any interaction with the account as the outcome. For passive savers, the outcome means are 0.37 at the $1,000 threshold and 0.36 at the $5,000 threshold, relative to 0.60 and 0.65 for active savers, respectively. This leads to large implied abandonment shares for passive savers: 0.38 at the $1,000 threshold and 0.45 at the $5,000 threshold. The second two columns use the address-based approach. In this sample restricted to movers, the outcome means for active savers are 0.70 and 0.75, respectively, relative to 0.45 and 0.58 for the passive savers. This leads to estimated abandonment shares of 0.36 and 0.23 at the two policy thresholds.

The estimated abandonment among working-age default plan compliers is substantially higher than the 2.7% share we estimated for retirees in Section 4. The magnitude of these estimates, however, may be biased by a violation of the assumption that $E(Y|C, A = 0) = E(Y|AT)$. Here, we provide some context for that assumption by showing how the outcomes change across other characteristics observed in this sample. We show that the two outcomes vary for always-takers at the policy thresholds as a function of income and age; i.e., we plot $E(Y|AT, X)$, where X is age or income.[29] For example, the probability of updating one's address varies from 60% to 80% as income increases. This provides context to the 25 percentage-point difference between $E(Y|C)$ and $E(Y|AT)$, as reported in column (1) of Table 4. In the absence of abandonment, this difference could be explained only if the difference between $E(Y|C, A = 0)$ and $E(Y|AT)$ were larger than the difference between $E(Y|AT, X)$ for the lowest and highest incomes in this sample. Likewise, the probability of interacting with the account within five years naturally increases substantially near retirement age, from just under 60% for those age 50 to over 80% for those age 65 and above. In the absence of abandonment, the 23 percentage point gap between active and passive savers (as reported in column (1) of Table 4) would be approximately equivalent to the gap between working-age and retirement-age individuals. While we cannot rule out

[29] These results are in Appendix Figures B.5 and B.6.

the gap between $E(Y|C, A = 0)$ and $E(Y|AT)$ is indeed this large, these analyses support a key role for abandonment (not other differences between non-abandoning compliers and always-takers) in our estimates.[30]

6 Policy Implications

The empirical analysis in Section 4 established that 2.7% of all IRA accounts are abandoned by retirees who are legally required to draw down their assets. Furthermore, we showed in Section 5 that abandonment among a primarily working-age population is substantially higher for savings accounts created by a default policy that forced account transfers to automatic rollover IRAs. If our findings have external validity to abandonment rates of accounts induced by other default policies (e.g., automatic enrollment), then they raise the possibility that default policies could reduce welfare unless myopia is severe.

Our results help inform a broad set of possible policies. For example, tax policy providing preferential treatment for IRAs may require modification as many do not withdraw their savings and pay taxes at the expected ages. Defaults into retirement saving plans may require safeguards to prevent forgetting these accounts. State policy for connecting unclaimed property with their owners could feature greater automation to improve these efforts. We do not take a stand on any one of these policies, but believe that policymakers should pay attention to interventions that could reduce abandoned retirement savings. Here, we continue our discussion by providing a conceptual framework for thinking through the ambiguous welfare effects of default policies and how our findings on abandonment, both among the older population and among passive savers, can provide further insight.

We propose a two-period model of retirement savings and consumption that features active saving, passive saving (by default), and forgetting. Forgetting can be rational or boundedly rational. We begin by showing that incorporating forgetting alone into a consumption-savings model can have different implications for policy design attempting to improve retirement security.

Each individual lives for two periods. In period 1, the individual earns W and saves S. In period 2, the individual consumes S and a Social Security benefit g. Thus, consumption

[30] We note that both outcomes are likely to have a causal effect on abandonment. By failing to update the mailing address, the custodian is less likely to be able to communicate with the account holder in the future. Likewise, interactions with the account signal a lower probability of the owner forgetting that account in the future. Thus, even if these differences in conditional means do not reflect short-term abandonment, they are plausibly linked to future abandonment.

in each period is given by:

$$c_1 = W - S, \qquad c_2 = S + g. \tag{5}$$

Also, suppose individuals are myopic regarding the retirement period and thus undersave as in Laibson (1997). If the individual undervalues retirement utility by a factor of $0 < \beta \leq 1$ at the time of the savings decision, total decision utility is given by:

$$U(S) = u(c_1) + \beta u(c_2) = u(W - S) + \beta u(S + g). \tag{6}$$

Note that β is a *myopia* parameter and not a discount factor, which we omit for simplicity. The optimal (interior) retirement savings for such an individual is $\hat{S}$, which satisfies:

$$\beta = \frac{u'(W - \hat{S})}{u'(\hat{S} + g)}. \tag{7}$$

Given equation (7), we present $\hat{S}$ as a function of β. Thus, $\hat{S}(\beta) < \hat{S}(1)$, $\forall \beta < 1$, i.e., individuals undersave because of myopia.

Default savings in our empirical analysis resulted from automatic rollovers, but conceptually we should consider default savings to be from any type of passive behavior. We use d to represent these savings. The individual remembers to consume active and passive savings with different probabilities at retirement. Specifically, the individual remembers fraction $0 < \gamma_a \leq 1$ of active savings (S) and fraction $0 < \gamma_p \leq 1$ of passive savings (d), thereby allowing for some savings to be forgotten. (A more general model could have γ represent the probability of being unclaimed in retirement, with this possibility emerging from forgetting or other factors such as hassle costs.) The individual does not forget g as these checks are generally sent to the account holder without any action necessary; the purpose of including g is simply to better calibrate individual decisions. The lifetime utility optimization problem with default savings d is given by:

$$U(S, d) = u(W - S - d) + \beta u(\gamma_a S + \gamma_p d + g). \tag{8}$$

Policies to reduce retirement account abandonment have different implications for passive and active savers. As in Section 5, we import this terminology from Chetty et al. (2014), where active savers (estimated to be 15% of the population in that paper) respond to incentives and passive savers (the remaining 85%) default to automatic contributions. In what follows, we derive and discuss insights for those two types of savers. We also discuss a conceptual implication on the estimation of myopia in the presence of forgetting.

6.1 Implications for Passive Savers

Default enrollment in retirement savings is a key policy aimed toward increasing financial security for passive savers. Yet, the lack of salience for these savings may lead to a greater probability of forgetting them in retirement. Here, we examine the conditions under which default contributions increase experienced utility.

To simplify the model, we assume that all of the active savings are remembered, i.e., $\gamma_a = 1$, and that default savings are not salient (as they are intended). Thus, the individual continues to optimize S without considering d.

The reduction in lifetime utility when default savings can be forgotten ($\gamma_p < 1$) is relevant in considering the trade-offs of "nudge" policies.[31] Tension arises because default savings help savers correct for myopic beliefs about retirement needs but also introduce a risk of forgetting these savings. The agent's experienced utility (as a function of d) is given by:

$$V(d) = u(W - \hat{S}(\beta) - d) + u(\hat{S}(\beta) + \gamma_p d + g). \tag{9}$$

The derivative of this experienced utility with respect to d, evaluated at $d = 0$, is:

$$V'(0) = -u'(W - \hat{S}(\beta)) + \gamma_p u'(\hat{S}(\beta) + g). \tag{10}$$

Combining (10) with (7) (which defines $\hat{S}(\beta)$) reveals that $V'(0) > 0$ if and only if $\gamma_p > \beta$. Intuitively, the first dollar of default contributions increases experienced utility when the myopia correction $(1 - \beta)$ is greater than the probability of forgetting $(1 - \gamma_p)$. As such, default savings could prove more beneficial when complemented by efforts in helping account owners remember and locate their savings (at least in retirement).

6.2 Implications for Active Savers

We now turn our attention to active savings. Again, we use (8). To simplify the discussion, we consider the case with $d = 0$; this is (approximately) without loss of generality, as active savers perfectly crowd out any default savings with reduced active savings away from corners. The optimal active savings, $S^*(\beta, \gamma_a)$ satisfies the following equality:

$$\beta\gamma_a = \frac{u'\left(W - S^*(\beta, \gamma_a)\right)}{u'\left(\gamma_a S^*(\beta, \gamma_a) + g\right)}. \tag{11}$$

[31] If $\gamma_p = 1$, i.e., all passive savings are remembered, then the optimal d can fully fix the myopic saving problem by setting $\hat{d} = \hat{S}(1) - \hat{S}(\beta)$.

Based on equation (11), the impact of forgetting on savings is ex ante ambiguous and depends on the individual's utility function. Active savers may fail to take advantage of saving vehicles that arise throughout their working lives out of fear of forgetting them in retirement. (Another empirical signature of such fears could be penalized cash-outs of valuable savings at job transitions.) Or, active savers may lose utility while working (i.e., period 1) as they have to over-save for retirement because a fraction of savings is forgotten. Which behavior dominates depends on the utility function – with log utility, for example, active savers decrease retirement saving ($\hat{S}(\beta) > S^*(\beta, \gamma_a), \ \forall \gamma_a < 1$). These responses from active savers suggest that policies helping individuals to consolidate and keep track of their retirement accounts could impact total savings.

Our model of retirement savings with forgetting raises a new question about the estimation of myopia among active savers. As we have shown, myopia is an important parameter in setting the optimal default rate. It is also a key parameter in lifeycle models of retirement savings, for example in Laibson (1997) and Ericson (2017). If individuals reduce saving because of fear of forgetting as opposed to myopia, policymakers will overestimate the degree of myopia among savers and mistakenly increase default savings. The problem arises because the amount of saving in our model with forgetting is equal to one in a parallel economy with a higher discount rate, so a traditional analysis will conflate these two worlds. The implication is important because addressing myopia versus forgetting requires different policies: while myopia can be aided by default enrollment, forgetting requires reminders or other policies to keep funds with their account owners. Under our model with a log utility function, we determine that myopia is overestimated by one percent for the level of abandonment we estimate in the overall sample ($\beta = 0.71$ if $\gamma_a = 1$, versus $\beta = 0.72$ if $\gamma_a = 0.97$).[32]

Many prior papers starting with O'Donoghue and Rabin (1999) introduce naivete or sophistication in behavior to explain decision-making. What we have described in this section is consistent with sophistication, as individuals are aware they may forget savings in period 2 and thus adjust their period 1 savings. Naive individuals are overconfident in their prospective memory, and behave as if forgetting is not a concern. Thus, their period 1 savings would be unchanged, but they would experience reduced utility in period 2. Such a model reduces to our representation for passive savers in Section 6.1, but without any default savings.

[32] These estimates assume $W = 1$, $g = 0.2$, and $S = 0.3$.

7 Conclusion

The U.S. retirement savings landscape places a great deal of responsibility on individuals: how much to save, which saving vehicles to use, how to allocate investment funds, and how to decumulate savings. Additionally, individuals must keep track of numerous savings accounts accumulated over their working lives. Failing to do so can result in account abandonment, either due to forgetting or hassle costs that likely increase with the age of these accounts. While there is a considerable literature on various aspects of retirement saving, there is a gap in the study of unclaimed accounts.

The present paper fills this gap by providing the first set of stylized facts regarding the extent of unclaimed, or abandoned, retirement assets. We do this by analyzing individual level tax and information returns, along with account-level data from state unclaimed property. The descriptive analysis contains estimates and correlates of unclaimed accounts, including the extent to which they are "reclaimed" after missing their legal withdrawal minimums or being escheated to state unclaimed property. We also explore the apparent abandonment of automated rollover IRAs, a type of default savings account, where we exploit sharp policy thresholds that enable a regression discontinuity analysis. Our contribution also includes an analysis of a theoretical model of retirement saving featuring empirically plausible values of forgetting among passive and active savers.

We find that nearly 2.7% of retirees own an abandoned retirement account, and this percentage is increasing over time. The amounts abandoned in IRAs are substantial at about \$790 million per year, with over 40% remaining abandoned even after a decade. We also find that policies which promote retirement saving, such as auto-enrollment, may unintentionally encourage the accumulation of smaller balance accounts that appear to be about 10 times more likely to be abandoned over the lifecycle. Current policy to mitigate abandonment is focused on the use of escheatment to unclaimed property. Yet plan participation is mostly voluntary, and most accounts are neither escheated nor reclaimed upon escheatment.

Taken together, our results demonstrate that retirement account abandonment is a timely topic of increasing policy attention – our estimated annual flow of \$790 million from IRAs alone is substantial. Forgetting and hassle costs are particularly important as retirees live longer and become vulnerable to cognitive decline (Chandra et al., 2020). We demonstrate the conceptual importance of our results for more general studies of lifecycle saving and consumption, as abandonment can result in suboptimal retirement consumption or decreased retirement saving. Future research on these topics could incorporate frictions in retirement account consolidation and memory so that policymakers can consider the tradeoffs that these frictions create.

References

Armour, P., M. D. Hurd, and S. Rohwedder (2016). Trends in pension cash-out at job change and the effects on long-term outcomes. In *Insights in the Economics of Aging*, pp. 15–39. University of Chicago Press.

Baicker, K., W. J. Congdon, and S. Mullainathan (2012). Health insurance coverage and take-up: Lessons from behavioral economics. *The Milbank Quarterly 90*(1), 107–134.

Banks, J., R. Blundell, and S. Tanner (1998). Is there a retirement-savings puzzle? *American Economic Review*, 769–788.

Benartzi, S. and R. Thaler (2007). Heuristics and biases in retirement savings behavior. *Journal of Economic Perspectives 21*(3), 81–104.

Benzarti, Y. (2020). How taxing is tax filing? using revealed preferences to estimate compliance costs. *American Economic Journal: Economic Policy 12*(4), 38–57.

Bernheim, B. D. and J. M. Gastell (2020). Optimal default options: The case for opt-out minimization. Technical report, National Bureau of Economic Research.

Bonamici, S. (2020, July). Bonamici, Warren, Daines, Banks re-introduce bipartisan bill to upgrade America's retirement saving system. Press Release. `https://bonamici.house.gov/media/press-releases/bonamici-warren-daines-banks-re-introduce-bipartisan-bill-upgrade-americas`.

Brown, J. R., J. Poterba, and D. P. Richardson (2017). Do required minimum distribution rules matter? the effect of the 2009 holiday on retirement plan distributions. *Journal of Public Economics 151*, 96–109.

Bubb, R. and P. L. Warren (2020). An equilibrium theory of retirement plan design. *American Economic Journal: Economic Policy 12*(2), 22–45.

Carroll, G. D., J. J. Choi, D. Laibson, B. C. Madrian, and A. Metrick (2009). Optimal defaults and active decisions. *Quarterly Journal of Economics 124*(4), 1639–1674.

Chandra, A., C. Coile, and C. Mommaerts (2020). What can economics say about Alzheimer's disease? *NBER Working Paper No. w27760*.

Chetty, R., J. N. Friedman, S. Leth-Petersen, T. H. Nielsen, and T. Olsen (2014). Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from denmark. *Quarterly Journal of Economics 129*(3), 1141–1219.

Chetty, R., J. N. Friedman, and E. Saez (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings. *American Economic Review 103*(7), 2683–2781.

Clark, R. L., M. B. d'Ambrosio, A. A. McDermed, and K. Sawant (2006). Retirement plans and saving decisions: The role of information and education. *Journal of Pension Economics and Finance 5*(1), 45–67.

Clark, R. L., M. S. Morrill, and D. Vanderweide (2014). Defined benefit pension plan distribution decisions by public sector employees. *Journal of Public Economics 116*, 73–88.

Ebenstein, A. and K. Stange (2010). Does inconvenience explain low take-up? Evidence from unemployment insurance. *Journal of Policy Analysis and Management 29*(1), 111–136.

Ekerdt, D. J. and J. K. Hackney (2002). Workers' ignorance of retirement benefits. *The Gerontologist 42*(4), 543–551.

Ericson, K. M. (2011). Forgetting we forget: Overconfidence and memory. *Journal of the European Economic Association 9*(1), 43–60.

Ericson, K. M. (2017). On the interaction of memory and procrastination: Implications for reminders, deadlines, and empirical estimation. *Journal of the European Economic Association 15*(3), 692–719.

GAO (2019, January). Retirement accounts: Federal action needed to clarify tax treatment of unclaimed 401(k) plan savings transferred to states. Technical report, United States Government Accountability Office Report.

Goda, G. S., M. R. Levy, C. F. Manchester, A. Sojourner, and J. Tasoff (2020). Who is a passive saver under opt-in and auto-enrollment? *Journal of Economic Behavior & Organization 173*, 301–321.

Goldin, J. and D. Reck (2020). Revealed-preference analysis with framing effects. *Journal of Political Economy 128*(7), 2759–2795.

Goodman, L., K. Mackie, J. Mortenson, and H. Schramm (2019). The evolution of leakage and retirement asset flows in the US. *Available at SSRN 3450866*.

Gottlieb, D. and K. Smetters (2021). Lapse-based insurance. *American Economic Review*.

Hopkins, J. (2018, July). A new retirement risk: Unclaimed property laws. Technical report, Forbes Magazine.

Horneff, V., R. Maurer, and O. S. Mitchell (2020). Putting the pension back in 401 (k) retirement plans: Optimal versus default deferred longevity income annuities. *Journal of Banking & Finance 114*, 105783.

Hung, A. A., J. E. Luoto, and J. Burke (2015). Defaulting in and cashing out? the impact of retirement plan design on the savings accumulation of separating employees. *Working paper*.

Jones, D. (2012). Inertia and overwithholding: explaining the prevalence of income tax refunds. *American Economic Journal: Economic Policy 4*(1), 158–85.

Kim, H. H., R. Maurer, and O. S. Mitchell (2016). Time is money: Rational life cycle inertia and the delegation of investment management. *Journal of Financial Economics 121*(2), 427–447.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics 112*(2), 443–478.

Lusardi, A. and O. S. Mitchell (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature 52*(1), 5–44.

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

Mastrobuoni, G. (2011). The role of information for retirement behavior: Evidence based on the stepwise introduction of the social security statement. *Journal of Public Economics 95*(7-8), 913–925.

Mortenson, J. A., H. R. Schramm, and A. Whitten (2019). The effects of required minimum distribution rules on withdrawals from traditional IRAs. *National Tax Journal 72*(3), 507–542.

Munnell, A. and A. Webb (2015). The impact of leakages from 401(k)s and IRAs. *Center for Retirement Research at Boston College* (15-2).

O'Donoghue, T. and M. Rabin (1999). Doing it now or later. *American Economic Review 89*(1), 103–124.

Poterba, J. M. (2014). Retirement security in an aging population. *American Economic Review 104*(5), 1–30.

Rogers, T. and K. L. Milkman (2016). Reminders through association. *Psychological Science 27*(7), 973–986.

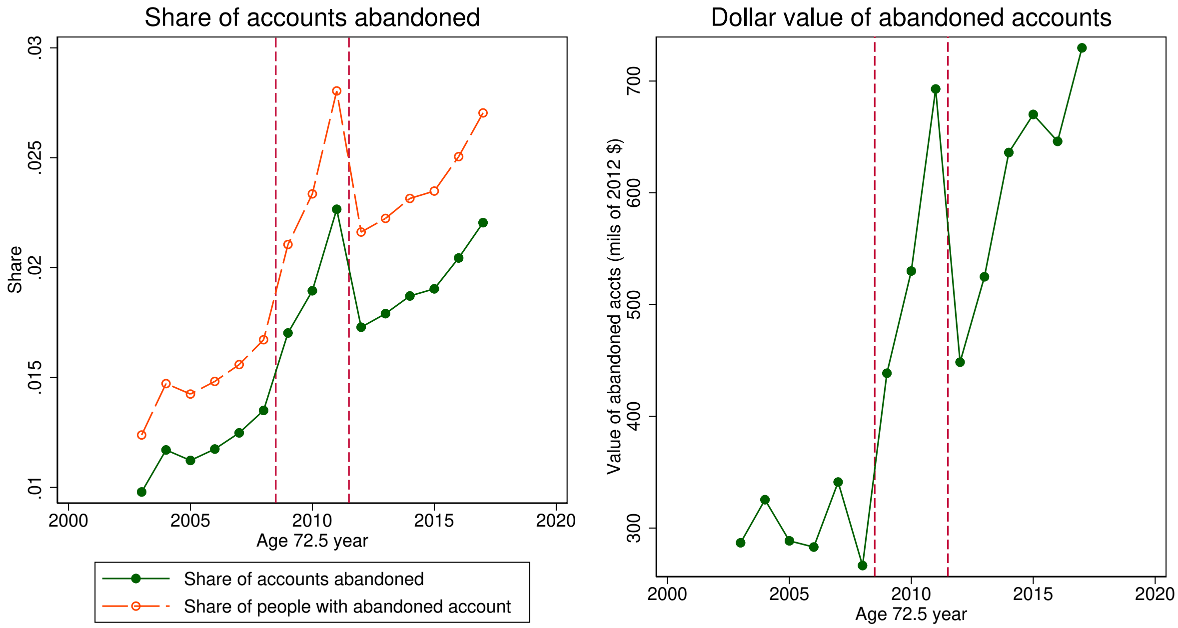
Scholz, J. K., A. Seshadri, and S. Khitatrakun (2006). Are Americans saving "optimally" for retirement? *Journal of Political Economy 114*(4), 607–643.

Scott, J. S., J. B. Shoven, S. N. Slavov, and J. G. Watson (2020). Can low retirement savings be rationalized? *NBER Working Paper No. w26784*.

Skinner, J. (2007). Are you sure you're saving enough for retirement? *Journal of Economic Perspectives 21*(3), 59–80.

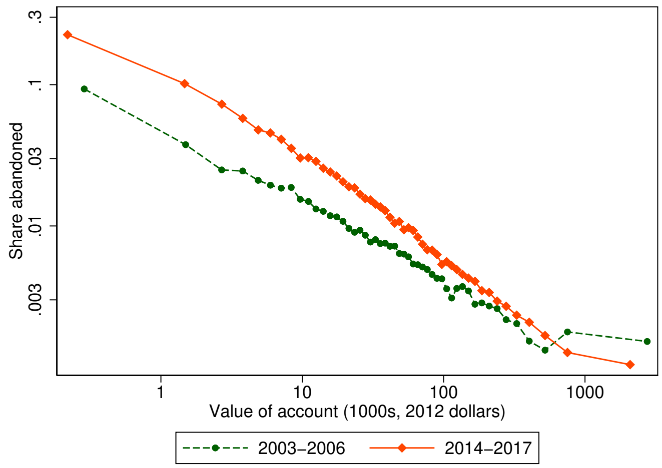
Thaler, R. H. and S. Benartzi (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy 112*(S1), S164–S187.

Figure 1: Time series of abandoned IRAs over time



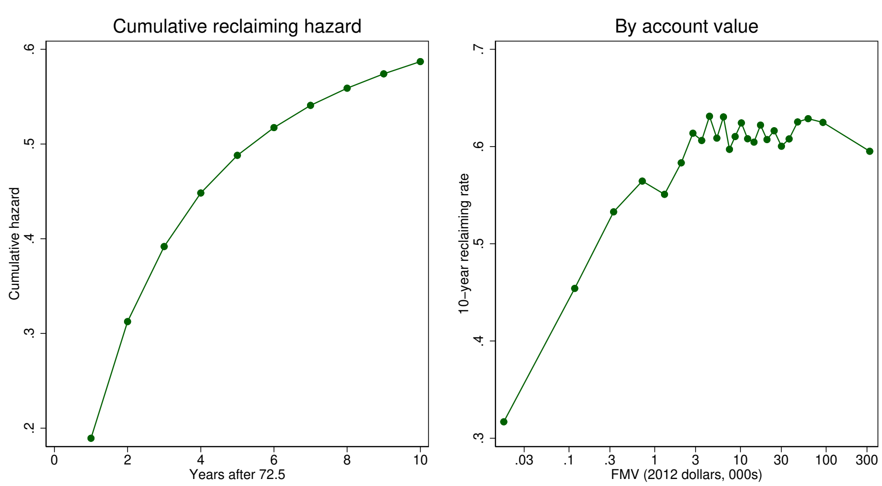
Notes: Left panel: This solid series plots the share of traditional IRAs of 72.5-year-olds that we estimate to be abandoned in each year, where an account is defined as the combination of individual and custodian. The dashed series plots the share of IRA-owners (age 72.5) with at least one abandoned IRA. Right panel: This series plots the aggregate dollar value (measured at the end of the age 72.5 year) of accounts that we estimate to be abandoned at age 72.5. Dollars are adjusted to 2012 dollars via the PCE deflator. Dotted lines bracket the years 2009 through 2011; for each of these three years, the measurement of abandonment is artificially affected by the 2009 suspension of required minimum distributions. Data Source: U.S. tax and information records.

Figure 2: Probability of IRA abandonment conditional on account balance



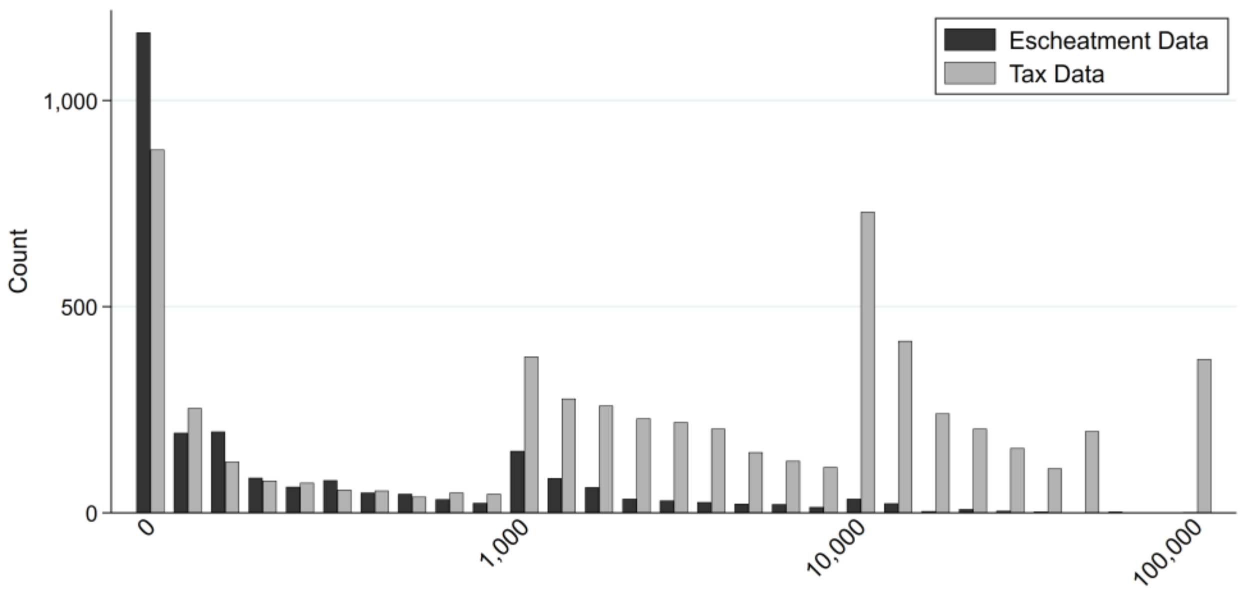
Notes: This figure plots the non-parametrically estimated probability of abandonment at age 72.5 as a function of the account value at age 72.5, measured in 2012 dollars. These series are plotted separately for the period (i.e., cohorts) 2003-2006 and 2014-2017. Data Source: U.S. tax and information records.

Figure 3: Reclaiming of abandoned IRAs



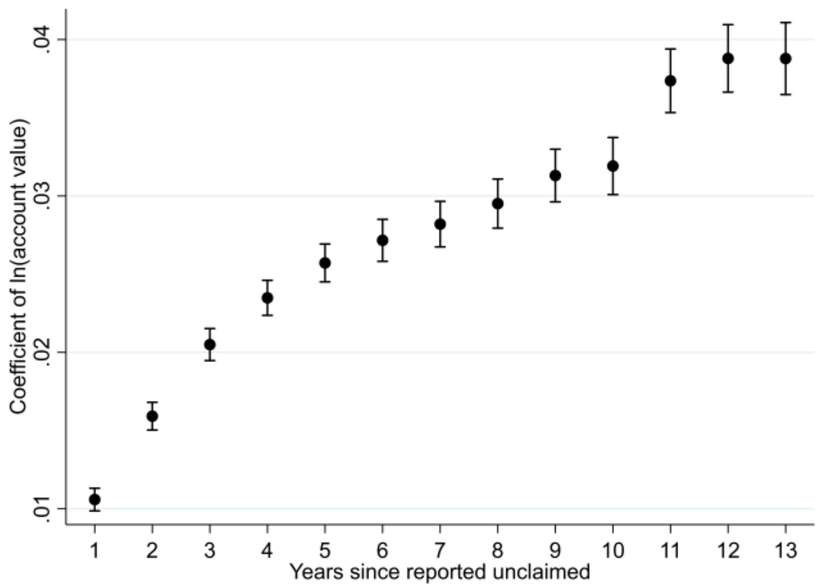
Notes: Left panel: This figure plots the cumulative probability of reclaiming an account in subsequent years, conditional on being estimated to be abandoned at age 72.5 and remaining alive through the year in question. Right panel: This figure plots the probability of reclaiming an account in ten years as a function of the value of the account at age 72.5. Reclaiming is deemed to occur when the individual takes any IRA distribution from the account, as indicated on Form 1099-R. This cumulative probability is computed using year-by-year hazard rates for each year t, which are estimated on the subset of the sample that remains alive through the end of t. This figure includes those who were age 72.5 between 2003 and 2008. Data Source: U.S. tax and information records.

Figure 4: Individuals with abandoned IRAs in escheatment and tax databases



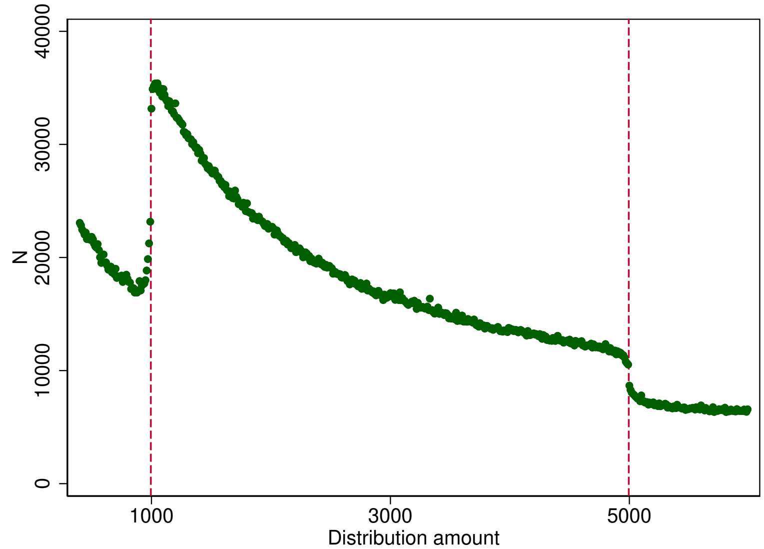
Notes: Figure shows the number of people with abandoned IRA accounts in the escheatment and tax databases in 2016. Note that the horizontal axis is not linear. The chart includes the six states (FL, LA, MA, MN, OH, and TX) that report a sufficient number of names and addresses of the unclaimed property holders to enable comparison with the tax data. Data Source: State unclaimed property records (escheatment data); U.S. tax and information records.

Figure 5: Impact of account value on reclaiming escheated accounts (MA)



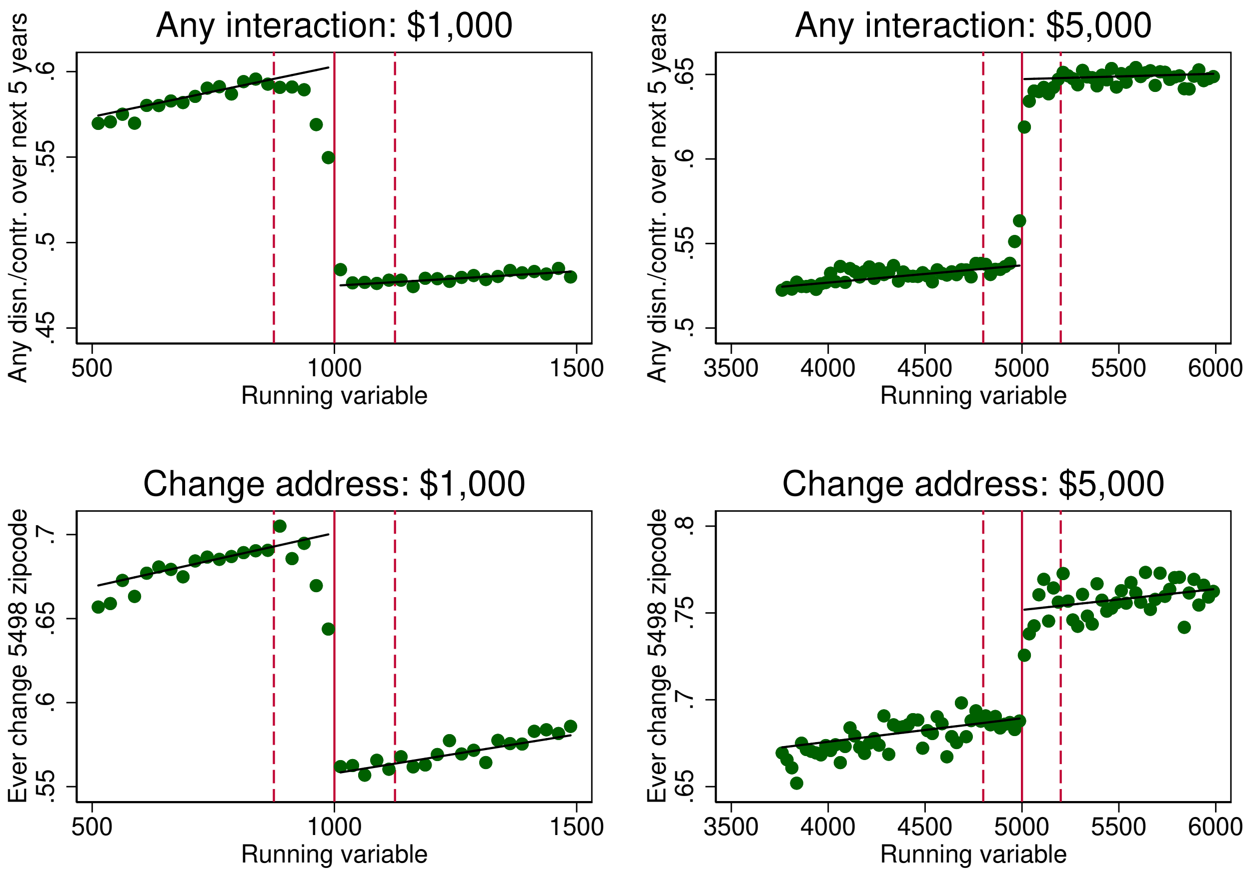
Notes: Figure shows the coefficients, each from a separate regression, on ln(Account Value) as the window of years since reported unclaimed increases from 1 to 13. The coefficients are from regressions of the following form: Claimed $= \beta \ln(\text{Account Value}) + \eta + \gamma$, where *Claimed* is whether the property was claimed within 1, 2,...,13 years (each separate regressions), *Account Value* is measured in units of \$10,000, η represents property type fixed effects, and γ represents year reported as unclaimed fixed effects. The full regression results corresponding to these coefficients are shown in Table B.4. Data Source: Massachusetts unclaimed and claimed property records, 1998-2018.

Figure 6: Automatic rollover IRA counts in $10 distribution bins



Notes: This figure plots raw counts of observations in the automated rollover (forced-transfer) IRA sample with direct rollovers in Form 1099-R as a function of the nominal distribution amount, in $10 bins. Distributions that are exact multiples of $500 are dropped. This figure uses distribution data from 2005 through 2015. Dotted lines indicate the policy thresholds at $1,000 and $5,000. Data Source: U.S. tax and information records.

Figure 7: Outcomes used to impute abandonment of automatic rollover IRA compliers



Notes: Each panel plots the expected mean of some outcome as a function of the nominal distribution amount. Each panel also plots the linear fit on each side of the threshold, weighted using a triangular kernel, dropping observations within the donut hole indicated by the dashed vertical red lines. The left two panels use the $1,000 threshold while the right two panels use the $5,000 threshold. In the top two panels, the outcome is a dummy for having any interaction with the account – distributions or contributions – within five years of the initial distribution. In the bottom two panels, the outcome is a dummy for updating the address with the IRA custodian within five years, as measured using the address reported by the custodian on Form 5498; in these panels, we condition on the individual moving during this period, measured using the addresses reported on Form 1040 and/or Form 1099-SSA. Data Source: U.S. tax and information records.

Table 1: Summary statistics for abandoned and non-abandoned IRA accounts

	Abandoned accounts			Non-abandoned accounts		
	Value (2012 dollars) (1)	Total share of income (2)	Annuity share of income (3)	Value (2012 dollars) (4)	Total share of income (5)	Annuity share of income (6)
Panel A: 2003-2006						
25^{th} percentile	\$ 2,609	0.057	0.005	\$ 18,125	0.447	0.037
50^{th} percentile	\$ 9,015	0.241	0.020	\$ 45,148	1.245	0.104
75^{th} percentile	\$ 27,131	0.774	0.065	\$111,772	3.241	0.272
90^{th} percentile	\$ 61,803	1.988	0.167	\$276,117	7.184	0.603
Observations	33,800	33,800	33,800	2,421,200	2,421,200	2,421,200
Panel B: 2014-2017						
25^{th} percentile	\$ 677	0.012	0.001	\$ 34,300	0.598	0.050
50^{th} percentile	\$ 6,277	0.127	0.011	\$ 97,065	1.797	0.151
75^{th} percentile	\$ 25,640	0.571	0.048	\$250,216	4.604	0.386
90^{th} percentile	\$ 73,182	1.727	0.145	\$562,565	9.856	0.827
Observations	86,600	86,600	86,600	3,518,100	3,518,100	3,518,100

Notes: Column (1) reports quantiles of the values of abandoned IRAs. Column (2) reports quantiles abandoned IRAs to income. Income is defined as adjusted gross income, plus non-taxable Social Security, minus taxable IRA distributions. Columns (3) reports quantiles of the ratio of the value of the abandoned IRA, converted to an annuity stream, to income at age 72.5. The annuity calculation assumes an interest rate of 1% and uses estimated mortality rates from the Social Security Administration, separately by sex. Columns (4) through (6) repeat the same calculation for accounts that are not abandoned. Panel A uses the early part of our sample (2003-2006), while Panel B uses the late part of our sample (2014-2017). To protect taxpayer privacy, all quantiles reported as psuedo-quantiles, equal to the 30 observations nearest the true quantile. Data Source: U.S. tax and information records.

Table 2: The impact of financial sophistication and demographics on abandonment

	(1)	(2)	(3)	(4)	(5)	(6)
Financial sophistication metrics:						
Files tax return	-0.0412	-0.0410	-0.0412	-0.0413	-0.0411	-0.0411
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Pays estimated tax	-0.0014	-0.0013	-0.0013	-0.0018	-0.0019	-0.0019
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Has dividends or capital gains	-0.0044	-0.0037	-0.0039	-0.0042	-0.0040	-0.0040
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Has interest	-0.0023	-0.0021	-0.0021	-0.0021	-0.0017	-0.0017
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Demographics:						
Male		0.0014	0.0016	0.0016	0.0011	0.0011
		(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Zip share white		-0.0267	-0.0227	-0.0227	-0.0183	-0.0168
		(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0004)
Tax awareness:						
Zip share sharp EITC bunching						0.0434
						(0.0039)
Observations	6,760,000	6,757,000	6,757,000	6,757,000	6,747,000	6,747,000
Baseline mean	0.0187	0.0187	0.0187	0.0187	0.0187	0.0187
Year-by-value FE	X	X	X	X	X	X
Control for zip density			X	X	X	X
Control for zip educ. and poverty				X	X	X
Year-by-payer FE					X	X

Notes: This table reports regression estimates for a regression of a dummy for abandonment (at the account-level) on various outcomes using data from 2012 through 2017, restricted to those observations with valid zip codes from Form 1040 and/or Form 1099-SSA. Each column corresponds to a different regression. Each regression includes fixed effects for year interacted with 500 bins of real IRA value. All coefficients are statistically significant ($p < 0.001$). Data Source: U.S. tax and information records, zipcode characteristics from the American Community Survey.

Table 3: Comparison of escheated accounts in MA and WI

	MA	WI
Summary		
# unclaimed retirement accounts in 2016	3,320	815
Proportion claimed within 2 years	0.034	0.666
Avg account value	\$633	\$980
...of claimed accounts	\$2,110	\$812
...of accounts remaining unclaimed	\$581	\$1,315
Proportion of claims by initiation type		
DOR Auto Match	—	0.801
Online (own)	—	0.122
Online (locator service)	—	0.065
Other	—	0.013

Notes: Table shows summary statistics on unclaimed and claimed retirement accounts in Massachusetts and Wisconsin. Statistics are for accounts escheated in 2016. The claim initiation data are not available for Massachusetts. Data Sources: Massachusetts and Wisconsin unclaimed and claimed property records, 2016 to 2018.

Table 4: Imputed abandonment in automatic rollover (forced-transfer) IRAs

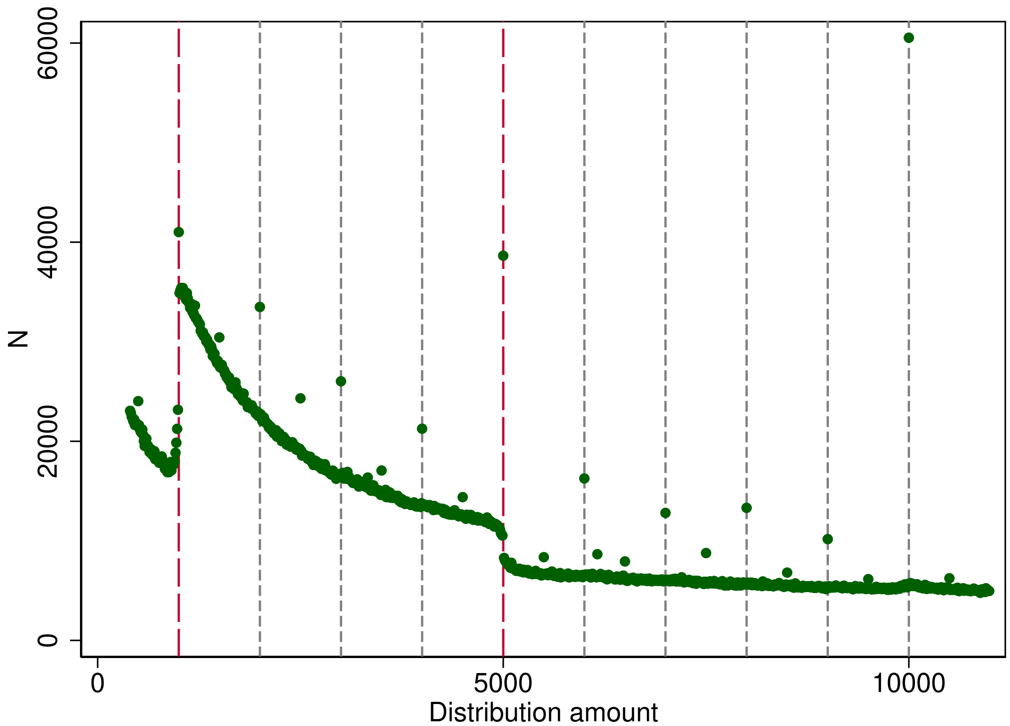
	Any interaction over next five years		Update address over next five years	
	\$1,000 threshold (1)	\$5,000 threshold (2)	\$1,000 threshold (3)	\$5,000 threshold (4)
Share compliers	0.555	0.377	0.567	0.355
	[0.547, 0.562]	[0.367, 0.386]	[0.548, 0.588]	[0.333, 0.377]
$E(Y\|AT)$	0.603	0.647	0.701	0.752
	[0.597, 0.609]	[0.641, 0.653]	[0.687, 0.716]	[0.738, 0.765]
$E(Y\|C)$	0.372	0.355	0.449	0.576
	[0.361, 0.382]	[0.338, 0.372]	[0.423, 0.472]	[0.533, 0.611]
Prob. abandoned	0.383	0.452	0.360	0.234
	[0.362, 0.404]	[0.423, 0.479]	[0.316, 0.403]	[0.178, 0.299]
Observations	1,399,000	1,170,000	219,000	161,000

Notes: This table reports estimated parameters from the calculation of abandonment in the forced transfer sample. In columns (1) and (2), we use as the outcome a binary variable for having any interaction with the account – distributions or contributions – within five years of the initial distribution. In columns (3) and (4), we use the outcome of updating the address with the IRA custodian within five years, as measured using the address reported by the custodian on Form 5498; in these columns, we condition on the individual moving during this period, measured using the addresses reported on Form 1040 and/or Form 1099-SSA. Columns (1) and (3) exploit the \$1,000 threshold, while columns (2) and (4) exploit the \$5,000 threshold. In the first row, we report the estimated share of compliers immediately to the right of the \$1,000 threshold or left of the \$5,000 threshold in the sample in question. The second row reports the estimated mean outcome for always-takers – those who enroll in an IRA voluntarily. The third row reports the estimated mean outcome for compliers – those who are induced to enroll in an IRA due to the forced transfer policy. The fourth row reports the implied share of compliers that abandon, under the assumptions maintained in the text. See text for further discussion about how these objects are calculated. Data Source: U.S. tax and information records.

– Appendices for Online Publication –

A Round-number bunching in rollover IRAs

Figure A.1: Counts of distribution amounts (including round numbers)

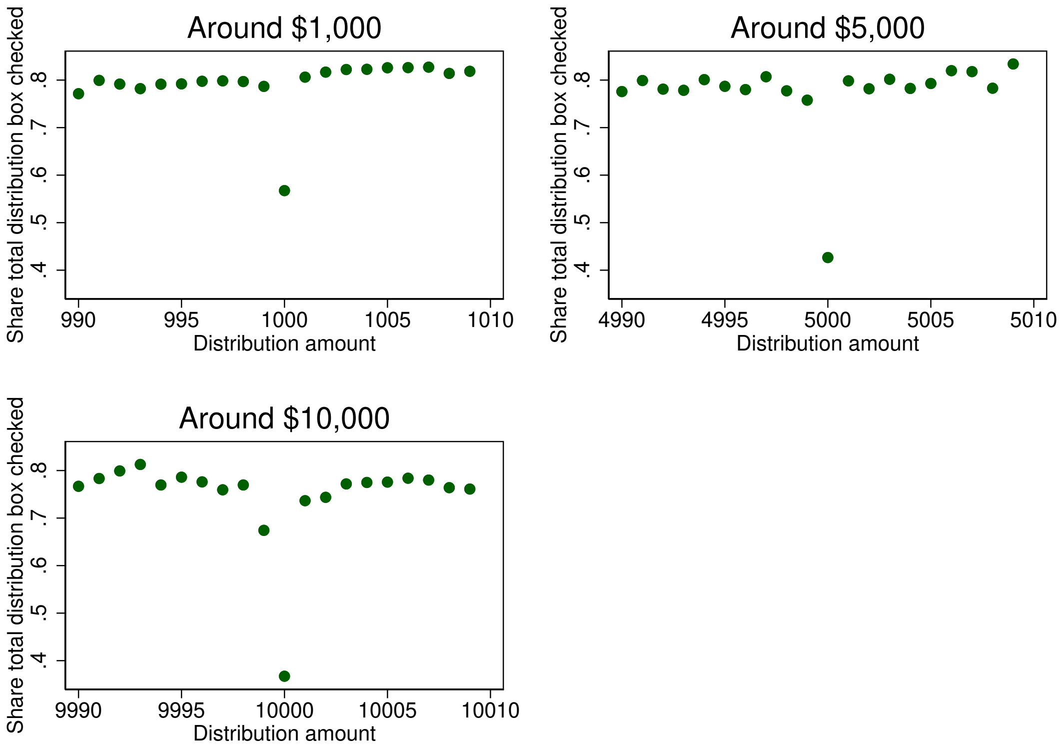


Notes: This figure plots raw counts of observations in the forced transfer sample with direct rollovers (Form 1099-R) as a function of the nominal distribution amount, in \$10 bins, without dropping round numbers. Dotted lines indicate even multiples of \$1,000. Data Source: U.S. tax and information records.

Figure A.1 plots raw counts of 1099-R rollover distributions as a function of distribution amount, in \$10 bins, for distributions between \$400 and \$11,000, between 2005 and 2015. This figure shows the clear round number bunching at multiples of \$1,000 and (to a lesser extent) multiples of \$500. The bunching at \$5,000 is very large; however, the similarly large bunching at \$10,000 is reassuring that the magnitude of the \$5,000 spike is related to the "roundness" of \$5,000 rather than the policy threshold.

The spike of distributions exactly at \$1,000 and \$5,000 (among other places) would be highly problematic in the empirical approaches that we pursue, especially because round number bunchers are likely to be quite different than those with distribution amounts nearby. In particular, round number bunchers are much more likely to be rolling over only a portion of the account balance, rather than the entire balance. Such distributions could very plausibly bunch at round numbers. And, furthermore, such distributions will generally be unaffected by our policy variation, while differing from those taking full distributions in important ways.

Figure A.2: Share of distributions with total distribution box checked

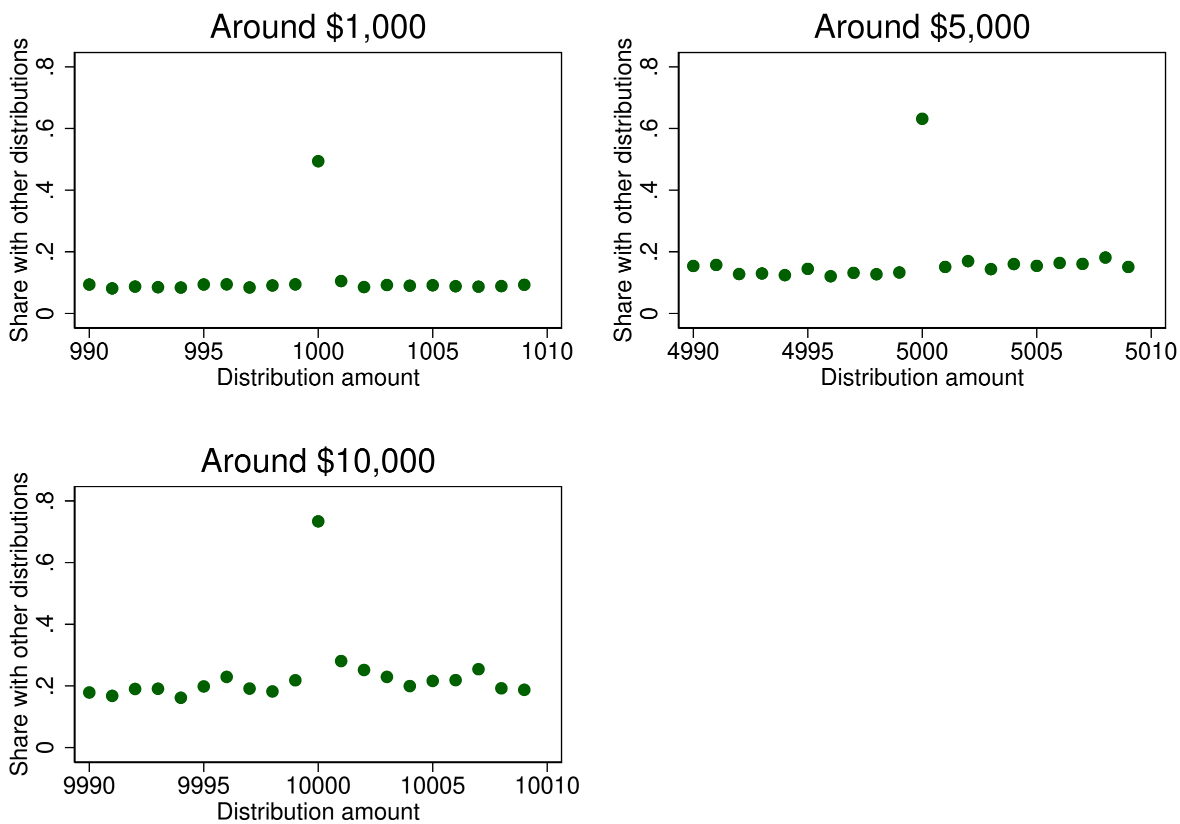


Notes: This figure plots the share of distributions in the forced transfer sample with the "total distribution" check-box checked, by single dollar of distributions. The three panels zoom in on thresholds of $1,000, $5,000, and $10,000 respectively. Data Source: U.S. tax and information records.

To explore this, we plot in Figure A.2 the share of distributions in our data that have the "total distributions" checkbox checked on Form 1099-R, as a function of the distribution amount, by exact dollar of distributions around $5,000 and $10,000. Indeed, the share of distributions with the total distribution box checked declines substantially at round numbers, including round numbers (such as $10,000) that are outside of our policy variation.

One strategy to proceed would be to restrict attention to 1099-R's with the total distribution checkbox checked, since the policy variation affects only total distributions. Unfortunately, there appears to be too much measurement error in the checkbox variable: some portion of distributions with the checkbox checked would in fact not be a total distribution. One common example of mismeasurement is the case when an individual makes two distributions to close out an account; e.g., suppose an individual at separation has a balance of $7,000, chooses to roll over $5,000 and take a cash distribution of $2,000. This

Figure A.3: Share of individuals with same-year regular distribution from same payer, conditional on total distribution box checked

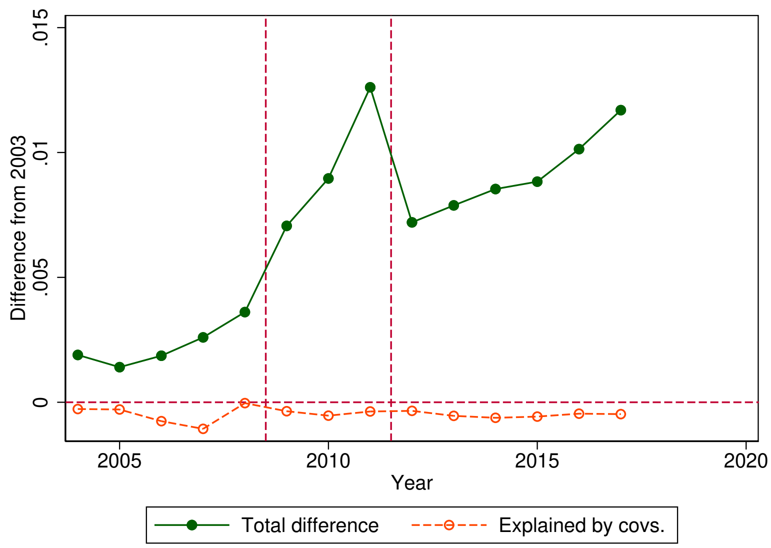


Notes: This figure analyzes the forced transfer sample, restricted to those with the total distributions checkbox checked. The figure plots the share of such observations that have some other distribution (with a Box 7 code of 1, 2, or 7, indicating a non-rollover distribution) from the same payer in the same year, by single dollar of distributions. The three panels zoom in on thresholds of $1,000, $5,000, and $10,000 respectively. Data Source: U.S. tax and information records.

would generate two different Forms 1099-R, which sometimes would both have the total distribution box checked. Figure A.3 explores this further. Among distributions with the total distribution box checked, we determine whether the individual received a different Form 1099-R from the same payer with a distribution code indicating a “normal” distribution (with codes 1, 2, or 7). We indeed find a large spike at both the $5,000 and $10,000 threshold. This suggests that this type of mismeasurement would cause the problem of round number bunching to remain if we used this restriction. For this reason, we take a simpler approach: we drop those with round number distributions (that is, at even multiples of $500) from our data. This conservative approach allows us to focus primarily on those rolling over the full account balance.

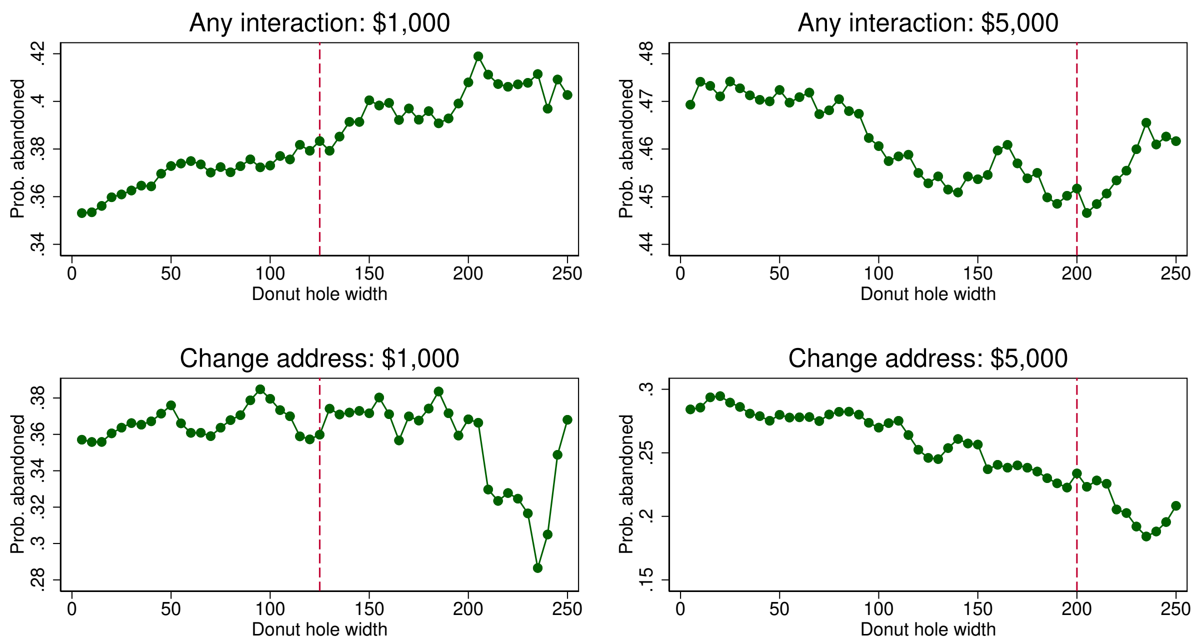
B Additional Figures and Tables

Figure B.1: Oaxaca-Blinder decomposition of abandonment rise since 2003



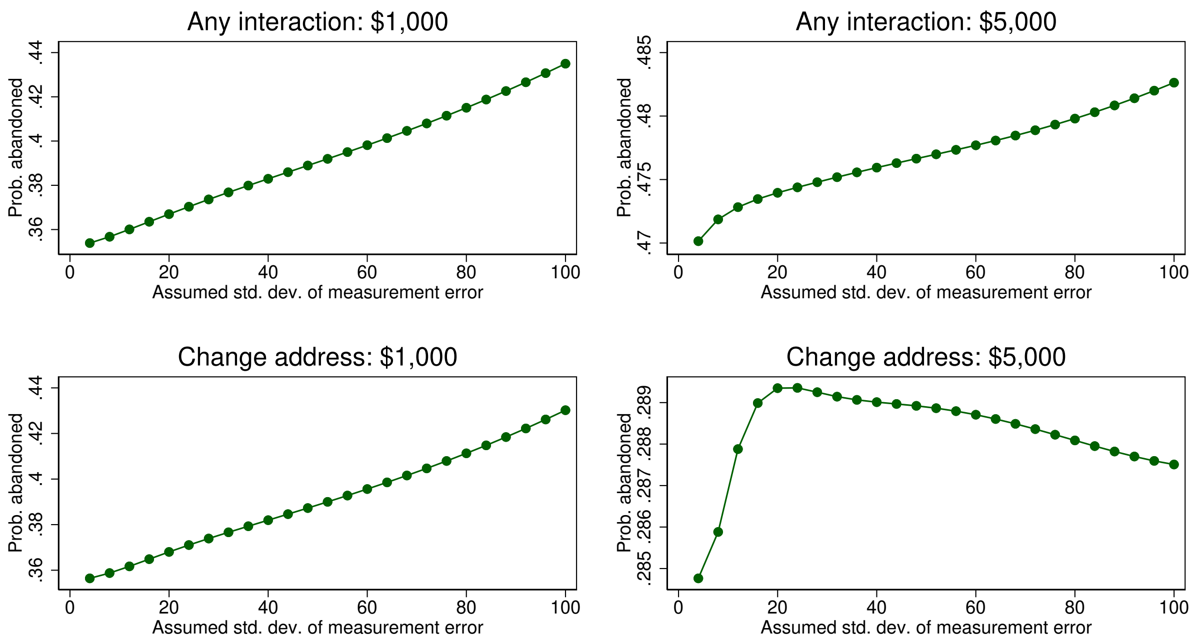
Notes: The solid series plots the share of IRAs abandoned in year t minus the share abandoned in 2003. The dashed series plots the amount of this difference that can be explained by covariates, $(X_t - X_{2003})'\beta_{2003}$. X includes a fully-interacted set of fixed effects for (1) real value of the IRA (in 50 bins), (2) 10 bins of Social Security income relative to the national distribution, (3) 10 bins of the white share of the zip code, (4) a dummy for being male, (5) a dummy for having any interest income, and (6) a dummy for having non-zero capital gains or dividend income. The coefficient β_{2003} is the coefficient from a regression of abandonment on X using observations in 2003. Data Source: U.S. tax and information records.

Figure B.2: Estimated abandonment of forced transfer compliers: robustness by donut hole width



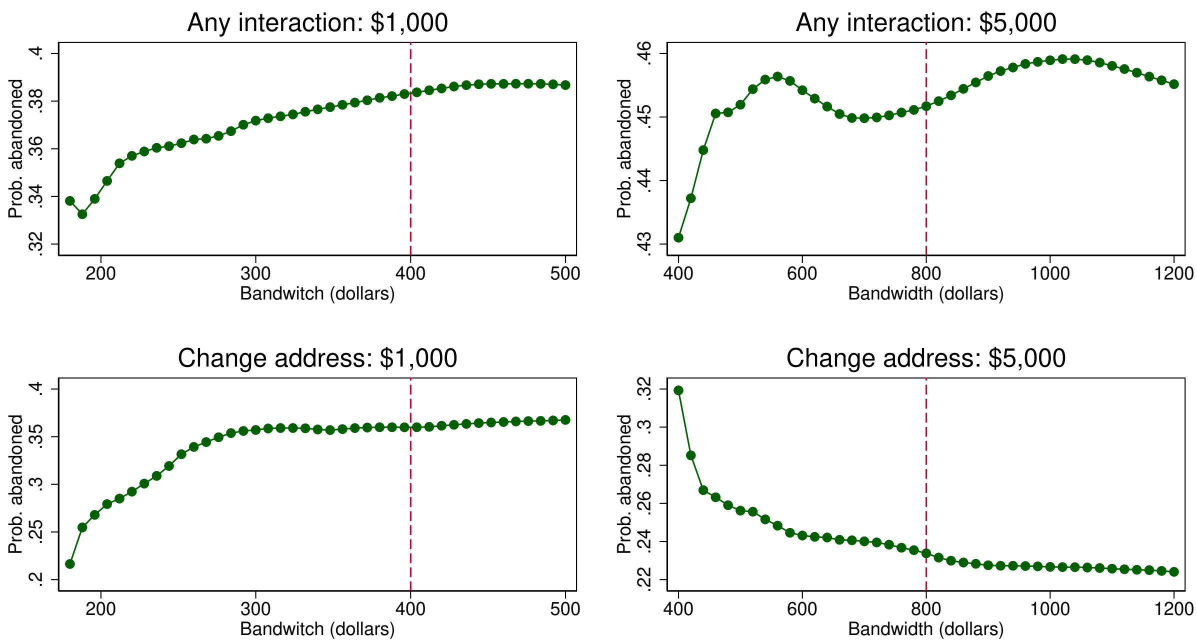
Notes: This figure plots the probability of abandonment for forced transfer compliers, as estimated using the method of Section 5, varying the width of the donut hole. The default bandwidths ($400 at $1,000 and $800 at $5,000) are used. The top two panels use having any interaction with the account over the next five years as the key outcome, while the bottom two panels use changing the address with the IRA custodian as the key outcome (conditional on moving). The left two panels use the $1,000 threshold; the right two panels use the $5,000 threshold. See text for details of the calculation and sample restrictions. Data Source: U.S. tax and information records.

Figure B.3: Estimated abandonment of default compliers: alternative treatment construction, varying σ



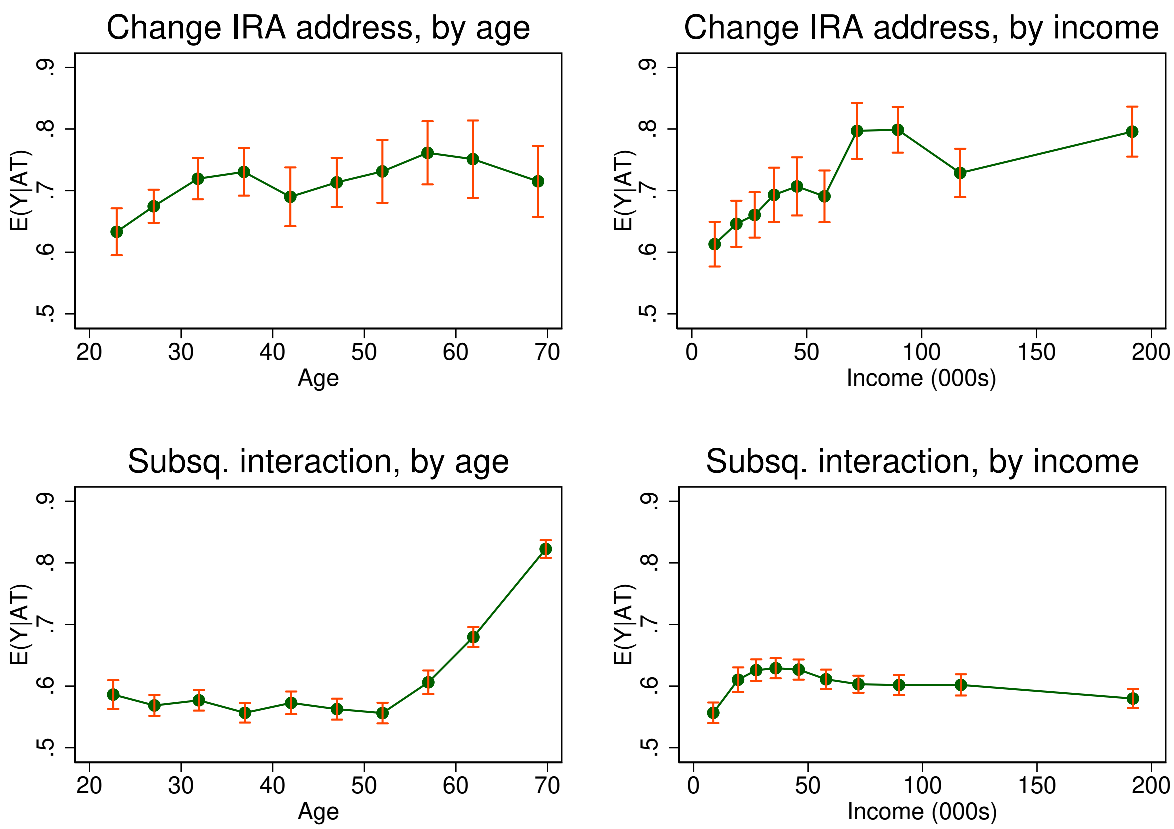
Notes: This figure plots the probability of abandonment for forced transfer compliers, as estimated using the method of Section 5, replacing the binary treatment dummy in Equations (2) and (3) with $\Phi\left(\frac{x_i - c}{\sigma}\right)$, where c is the cutoff in question, $\Phi(\cdot)$ indicates the normal density function, and σ is a parameter that varies along the x-axis. The default bandwidths ($400 at $1,000 and $800 at $5,000) are used. There is no donut hole. The top two panels use having any interaction with the account over the next five years as the key outcome, while the bottom two panels use changing the address with the IRA custodian as the key outcome (conditional on moving). The left two panels use the $1,000 threshold; the right two panels use the $5,000 threshold. See text for details of the calculation and sample restrictions. Data Source: U.S. tax and information records.

Figure B.4: Estimated abandonment of default compliers: robustness to bandwidth



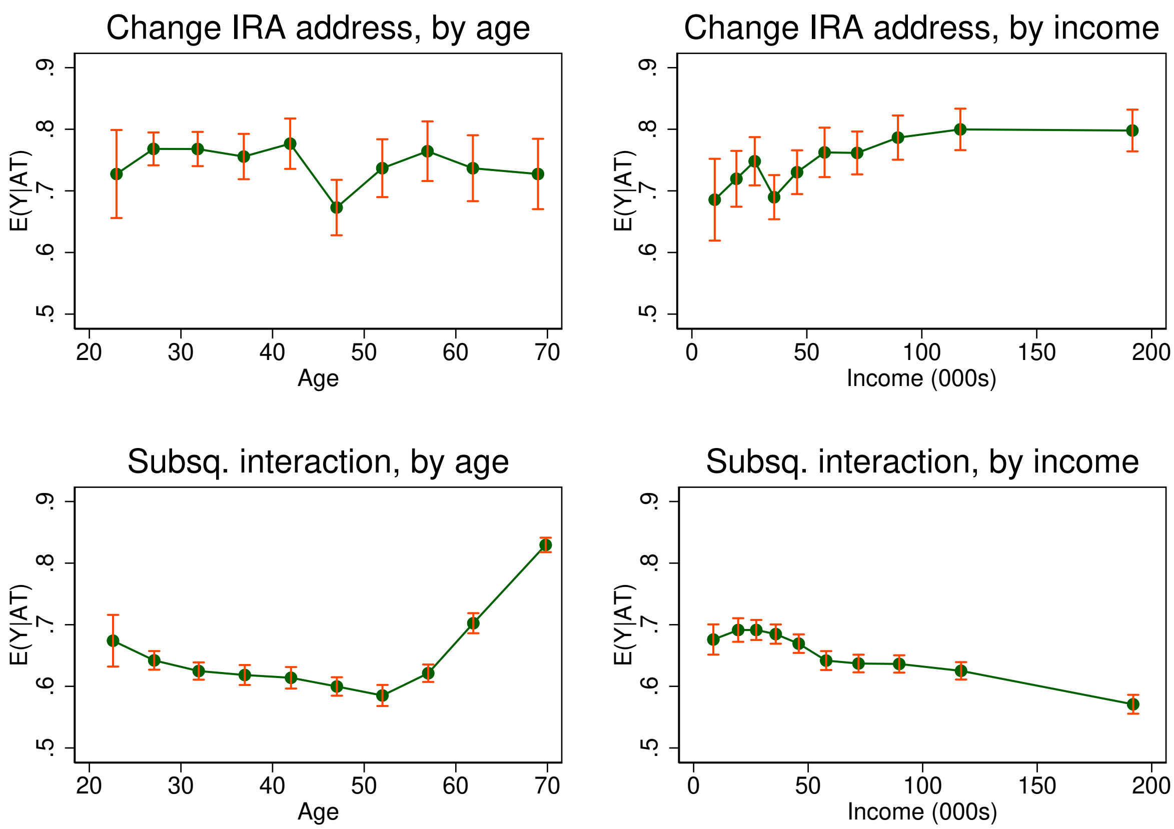
Notes: This figure plots the probability of abandonment for forced transfer compliers, as estimated using the method of Section 5, varying the bandwidth. The default donut holes ($250 at $1,000 and $400 at $5,000) are used. The top two panels use having any interaction with the account over the next five years as the key outcome, while the bottom two panels use changing the address with the IRA custodian as the key outcome (conditional on moving). The left two panels use the $1,000 threshold; the right two panels use the $5,000 threshold. See text for details of the calculation and sample restrictions. Data Source: U.S. tax and information records.

Figure B.5: Variation in always-taker means as a function of age and income: $1,000 threshold



Notes: This figure plots the mean outcomes for always-takers at the $1,000 threshold as a function of age or income. These objects are computed as the constant of separate regressions of the outcome on distribution amount, using observations only to the left of $1,000 within a certain bin of age or income, weighted by a triangular kernel. The outcome in the top panel is a dummy for changing the IRA address within 5 years; this analysis is restricted to those who move, based on address reported on Form 1040 and/or Form 1099-SSA. The outcome in the bottom two panels is a dummy for having any further interaction (distribution or contribution) with the account within 5 years of the initial distribution. Data Source: U.S. tax and information records.

Figure B.6: Variation in always-taker means as a function of age and income: $5,000 threshold



Notes: This figure plots the mean outcomes for always-takers at the $5,000 threshold as a function of age or income. These objects are computed as the constant of separate regressions of the outcome on distribution amount, using observations only to the right of $5,000 within a certain bin of age or income, weighted by a triangular kernel. The outcome in the top panel is a dummy for changing the IRA address within 5 years; this analysis is restricted to those who move, based on address reported on Form 1040 and/or Form 1099-SSA. The outcome in the bottom two panels is a dummy for having any further interaction (distribution or contribution) with the account within 5 years of the initial distribution. Data Source: U.S. tax and information records.

Table B.1: NAUPA codes categorized as retirement accounts

Code	Description
CK11	PENSION CHECKS
IR01	TRADITIONAL IRA - CASH
IR02	TRAD IRA - MUTUAL FUNDS
IR03	TRAD IRA - SECURITIES
IR04	RESERVED FOR TRADITIONAL IRA
IR05	ROTH IRA - CASH
IR06	ROTH IRA - MUTUAL FUNDS
IR07	ROTH IRA - SECURITIES
IR08	RESERVED FOR ROTH IRA
IR09	IRA OTHER - RESERVED 1
IR10	IRA OTHER - RESERVED 2
MS14	PENSION & PROFIT SHARING PLANS
05	IRA's-Securities
55	Annuities
71	IRAs
78	Pensions, retirement funds

Notes: NAUPA (National Association of Unclaimed Property Administrators) Codes used to categorize unclaimed property. Codes starting with "IR" were introduced in 2010 and gradually adopted by states. Pension checks are uncashed checks sent by plans to encourage required minimum distributions. Pension and profit-sharing plans, despite their name, consist of defined contribution plans (GAO, 2019). Codes without any alphabetic characters are exclusive to California.

Table B.2: Characteristics of escheated account owners

	Unclaimed retirement account owners	Overall 74 year old population
Name analysis		
Percent female	0.45	0.54
Percent Hispanic	0.24	0.12
County-level analysis		
Average county population	2,116,785	158,137
Percent age 65+	0.13	0.14
Percent white	0.69	0.72
Percent married	0.48	0.48
Percent bachelor degree	0.31	0.30
Percent born in state	0.57	0.57
Percent moved across state in past year	0.02	0.02
Percent own home	0.63	0.63
Median household income ($)	60,167	58,881
Percent of families below poverty line	0.11	0.11
Unemployment rate	8.60	7.60

Notes: Table reports average characteristics of individuals with unclaimed retirement assets (column 1) and average characteristics of 74 year olds (column 2). Data Sources: State unclaimed property records, American Community Survey county-level data, names databases derived from 2000 Census data and Social Security Administration data, and the U.S. Census Bureau Estimated State Population by Characteristics for 2016.

Table B.3: Escheated account summary and extrapolation, 2016

Sample	# accts	Total funds ($)	Mean amount ($)	Pop. age 74	Accts per 74	% US 74 pop
Data	36,529	18,347,524	633	898,227	.041	.43
National (extrap.)	69,507	38,012,684	547	2,094,035	.033	1.00

Notes: Table reports aggregate retirement-related unclaimed properties by our sample of 13 states and the extrapolated national sample. Data Source: State unclaimed property records, American Community Survey county-level data, and the U.S. Census Bureau Estimated State Population by Characteristics for 2016.

Table B.4: Impact of account value on claiming an escheated account (MA)

	Claimed?						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(Account Value)	0.0106	0.0159	0.0205	0.0235	0.0257	0.0272	0.0282
	(0.0004)	(0.0005)	(0.0005)	(0.0006)	(0.0006)	(0.0007)	(0.0007)
R-squared	0.172	0.154	0.145	0.139	0.136	0.138	0.133
Observations	50,063	46,524	43,204	41,436	39,709	36,025	34,356
Years Unclaimed	1	2	3	4	5	6	7
	(8)	(9)	(10)	(11)	(12)	(13)	
ln(Account Value)	0.0295	0.0313	0.0319	0.0374	0.0388	0.0388	
	(0.0008)	(0.0009)	(0.0009)	(0.0010)	(0.0011)	(0.0012)	
R-squared	0.130	0.132	0.135	0.139	0.129	0.129	
Observations	32,760	30,777	28,063	26,362	25,709	24,376	
Years Unclaimed	8	9	10	11	12	13	

Notes: Table shows coefficients on ln(Account Value) from separate regressions of whether the account was claimed within 1, 2,..., or 13 years (denoted by the Years Unclaimed row), including property code and year reported unclaimed fixed effects. Since the columns subsequently increase the number of years of possible claiming from 1 to 13, the sample size decreases due to the window of relevant data. Data Source: Massachusetts Unclaimed and Claimed Property Data, 1998 to 2018.