

The Growth Process of Individual Retirement Accounts: Evidence from Administrative Data

Asher Dvir-Djerassi

University of Michigan's Ford School of Public Policy

January 26, 2026

The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and do not necessarily reflect the views or the official positions of the U.S. Department of the Treasury or the Internal Revenue Service. All results have been reviewed to ensure that no confidential information is disclosed.

Abstract

Over the past two decades, Individual Retirement Accounts (IRAs) have grown faster than any other component of the retirement system—now standing at over one-third of U.S. retirement assets. Yet the sources of this growth remain difficult to assess with existing data. This paper contributes to filling this gap by leveraging IRS administrative microdata to decompose IRA asset growth. Using a representative 1% sample of the 2019 U.S. adult resident population and leveraging the direct reporting of IRA fair-market values to the IRS (Form 5498, Box 5)—a feature not shared by other major asset classes—together with information on direct contributions and rollovers from Forms 1099-R and 5498, the paper decomposes IRA asset growth from 2000 to 2019 into three components: direct contributions, rollovers from employer-sponsored plans, and compounded returns on pre-2000 balances. The results show that rollovers from employer-sponsored plans—primarily defined contribution (DC) plans such as 401(k)s—account for approximately 50% of asset growth over the period, compared to 20% from direct contributions and 30% from pre-2000 legacy balances. These findings imply that rollover contributions from DC plans, rather than direct contributions, are the core engine of IRA growth. A central implication of this finding is that treating DC plans and IRAs as distinct systems may warrant reconsideration. Leading measures of household wealth—such as the U.S. Financial Accounts—and federal tax-expenditure estimates—such as those produced by the Office of Management and Budget (OMB) and the Joint Committee on Taxation (JCT)—treat the DC and IRA systems as separate. By treating these systems as separate, estimates of the combined fiscal cost of the retirement system, for instance, are likely lower than they would be if DC plans and IRAs were treated as a single, integrated system. In sum, this paper highlights the central role of employer-plan rollovers in shaping IRA asset growth and clarifies the close integration of the DC and IRA systems, with important implications for tax administration.

1 Introduction

Individual Retirement Accounts (IRAs) are the fastest-growing component of the U.S. retirement system, now comprising more than one-third of all retirement assets. Despite their growing importance, the underlying drivers of IRA asset accumulation remain poorly understood. In particular, the relative roles of direct contributions and rollovers from employer-sponsored retirement plans—such as 401(k)s and 403(b)s—have not yet been systematically quantified using administrative data at scale.

This paper addresses two core questions: first, what share of IRA asset accumulation is attributable to direct contributions versus rollovers from employer-sponsored plans; and second, how does the relative importance of these sources vary across the population? By answering these questions using linked administrative data, the analysis also speaks to a broader set of measurement issues, which have broad implications for tax administration. Specifically, the results inform how IRAs and defined contribution (DC) plans are treated in the Financial Accounts of the United States, household wealth measures, and tax expenditure estimates, where assumptions about the independence of the DC and IRA systems shape characterizations of household wealth and the fiscal cost of retirement policies.

To answer these questions, the paper leverages a 1% random sample of the 2019 adult U.S. resident population, constructed from full-population IRS administrative data. Using linked records from 1999 to 2019, the analysis reconstructs the evolution of IRA balances over time for this cohort. It decomposes asset growth into three components: (1) direct contributions, (2) rollovers from DC and defined benefit (DB) plans, and (3) the compounded value of initial balances held in 1999. The analysis isolates the long-run impact of each component by estimating counterfactual scenarios in which contributions or rollovers are removed.

Assessing the importance of rollovers and direct contributions to IRA accumulation cannot be inferred from aggregate flow volumes alone. The contribution of rollovers to long-term IRA growth depends not only on the scale of incoming transfers, but also on whether those assets are retained, earn investment returns, and their pattern of decumulation. If rollovers are disproportionately concentrated among older individuals, rapidly withdrawn, annuitized, or allocated conservatively, their role in sustained asset accumulation may be greater than their aggregate volume would suggest. Conversely, if rollover assets are broadly distributed across the population and exhibit retention and return

profiles similar to other IRA balances, their central role in asset growth would be warranted. Because these behavioral and demographic patterns are not observable from aggregate volume data alone, this paper employs counterfactual simulations applied to micro-data to quantify the role of rollovers versus direct contributions to IRA growth.

The results show that approximately 50% of IRA asset growth over the 2000–2019 period is attributable to rollovers, less than 20% to direct contributions, and slightly more than 30% to legacy balances—i.e., IRA assets held prior to the study period. This pattern is broadly consistent across economic subgroups. However, among individuals with lower net worth and smaller IRA balances—likely early-career individuals—direct contributions account for a relatively larger share of growth. Taken together, these findings indicate that the employer-based DC system and the IRA system operate as an integrated pipeline of asset accumulation, rather than as separate retirement vehicles.

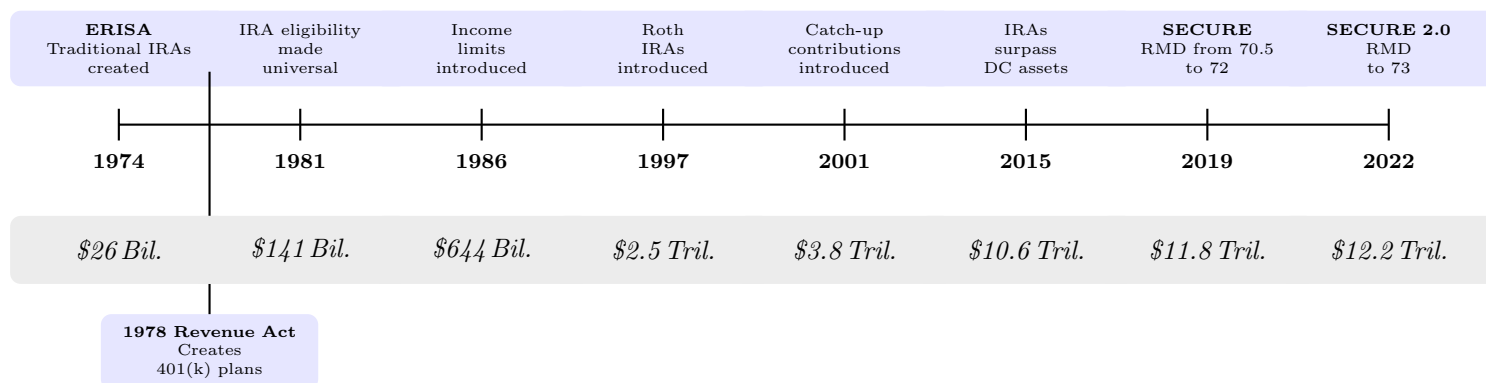
By providing the first population-wide accounting of the processes underlying IRA accumulation, this paper offers new empirical clarity on how IRA assets grow over time. The results and methods developed here have direct relevance for revenue forecasting, national accounting, and the estimation of retirement-related tax expenditures by clarifying the extent to which employer-sponsored plans and IRAs operate as a single, integrated system. In doing so, the paper highlights the limits of frameworks that treat the DC and IRA systems as institutionally or economically independent.

2 Background & Literature

IRAs were established under the Employee Retirement Income Security Act of 1974 (ERISA) to give workers without access to employer-sponsored pension plans a means to save for retirement with comparable tax advantages. Since their creation, IRAs have been repeatedly expanded and modified through successive pieces of legislation and administrative rulings, substantially broadening both eligibility and the forms of tax-advantaged saving available through these accounts.

The Economic Recovery Tax Act of 1981 (ERTA) expanded IRA eligibility to all workers with earned income, including those already covered by workplace pensions. The Tax Reform Act of 1986 introduced income-based limits on the tax deductibility of contributions, leading to the creation of nondeductible IRAs, in which contributions are made with after-tax dollars but investment

Figure 1: IRA Policy Evolution and Asset Growth, 1974–2022



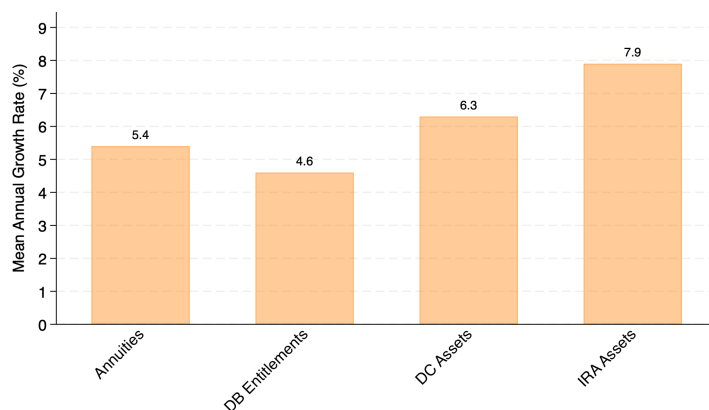
Notes: Asset values are expressed in constant 2019 U.S. dollars using the CPI-U.
Source: United States Financial Accounts.

gains accumulate outside of the tax system; the reform also introduced partially or fully deductible IRAs based on income. The Taxpayer Relief Act of 1997 introduced Roth IRAs—another type of nondeductible IRA—allowing tax-free withdrawals in retirement, provided certain conditions are met, and permitted nonworking spouses to contribute to IRAs within specified income thresholds.

In addition to direct contributions, rollover contributions from employer-sponsored retirement plans became increasingly common over the 1980s and 1990s. Prior to the early 1980s, DC plans were rare in the United States. This changed with the emergence of the 401(k), authorized under a 1978 amendment to the Internal Revenue Code and clarified by IRS guidance in 1981, which permitted the widespread use of elective pre-tax salary deferrals from regular wages into employer-sponsored retirement plans. As access to 401(k) and 403(b) plans expanded across the workforce, so too did the volume of rollovers from these plans into IRAs. Rollovers typically occur when individuals change jobs or retire and choose to move their retirement savings out of an employer-sponsored plan and into an IRA, which offers greater individual control and investment flexibility.

Over the period from 1999 to 2020, IRA assets experienced the fastest growth among all major components of retirement assets. As shown in Figure 2, the mean annual growth rate of IRA assets was 7.9%, exceeding that of DC assets (6.3%), annuities (5.4%), and DB entitlements (4.6%). This rapid growth has transformed the composition of private retirement assets. In 2000, the total value of IRA assets was well below that of DC assets—such as 401(k) and

Figure 2: Mean annual growth rate of retirement assets, 1999–2020.



Source: Author’s analysis of the Financial Accounts of the United States, Table L.117.

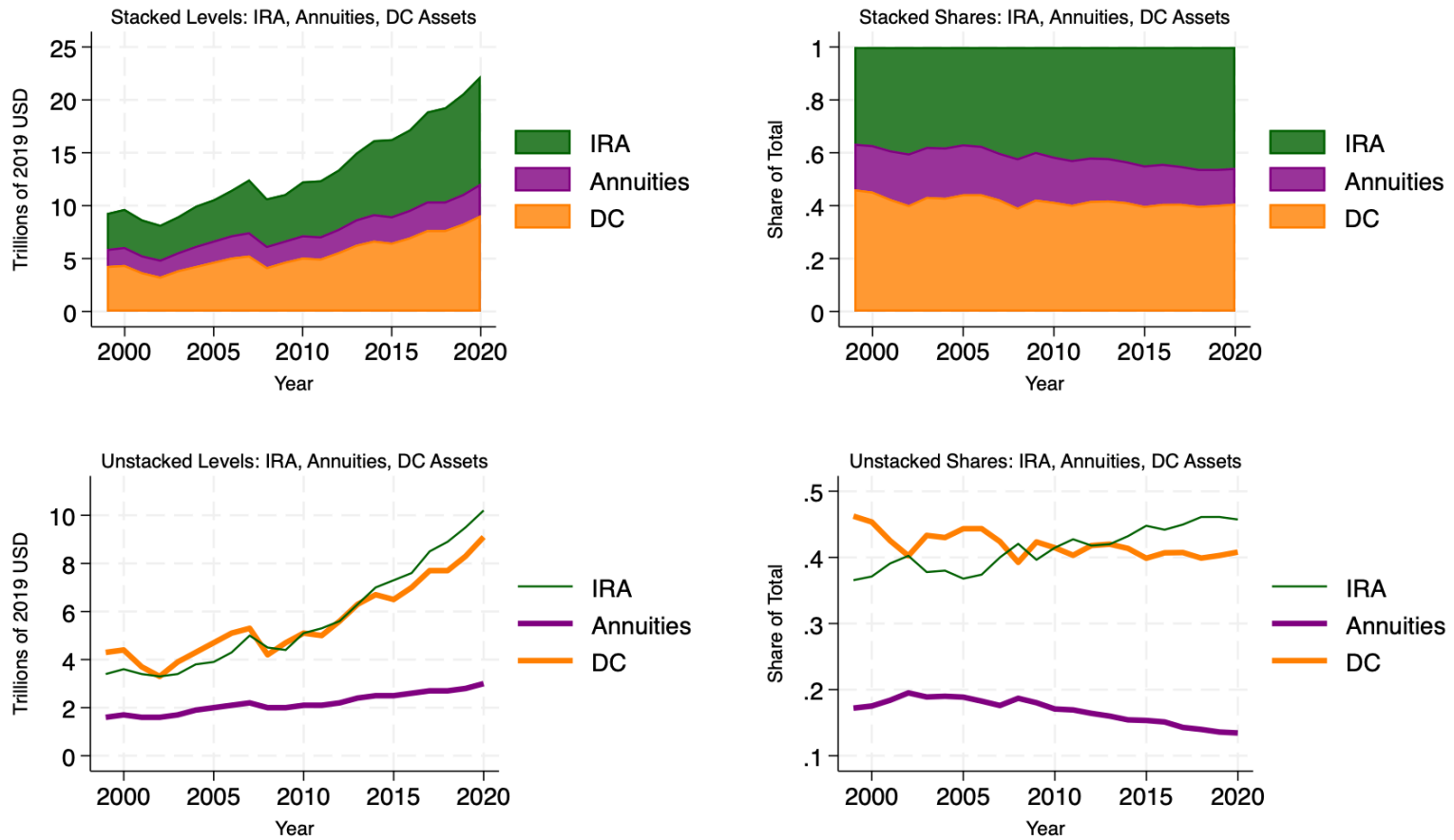
403(b) plans—but IRA balances grew steadily, overtaking DC assets around 2015. Likewise, IRA assets were roughly twice the size of annuity holdings in 2000 and are now nearly four times as large. These shifts are illustrated in Figure 3, which shows the relative trajectories of the main retirement asset categories over time. IRAs have thus outpaced the growth of other retirement assets and have become the largest component of the private retirement system outside of public DB entitlements.

2.1 IRA Varieties and their Features

There are two primary types of IRAs available to individuals:

1. **Traditional IRAs:** Anyone with earned income, along with a nonworking spouse under specific conditions, can contribute to a Traditional IRA. Whether contributions are tax-deductible depends on the individual’s income and their participation in a workplace retirement plan. Investment earnings grow tax-deferred, and withdrawals taken after age $59\frac{1}{2}$ are taxed as regular income. Withdrawals must begin no later than April 1 following the year in which the account holder turns $70\frac{1}{2}$.
2. **Roth IRAs:** These accounts allow for tax-free retirement savings. Contributions are made with after-tax income, meaning they are not deductible, but investment earnings and withdrawals after age $59\frac{1}{2}$ are tax-free if the account has been maintained for at least five years. Unlike Traditional IRAs, Roth IRAs do not require minimum withdrawals at any age.

Figure 3: Retirement Assets (Excluding DB Assets): Levels and Shares Over Time



Source: Author's analysis of the Financial Accounts of the United States, Table L.117.

Table 1: Number and Share of Households that Own IRAs by Type in 2019

IRA Type	Year Created	Households	
		Number	Percent
Trad. IRA	1974 (ERISA)	36.1 million	28.1%
SEP IRA	1978 (Revenue Act)	7.8 million	6.1%
SAR-SEP IRA	1986 (Tax Reform Act)		
SIMPLE IRA	1996 (Small Business Job Protection Act)		
Roth IRA	1997 (Taxpayer Relief Act)	24.9 million	19.4%
Any IRA	–	46.4 million	36.1%

Notes: Households may own more than one type of IRA. SEP IRAs, SAR-SEP IRAs, and SIMPLE IRAs are employer-sponsored IRAs.

Sources: Holden and Schrass (2019).

There are annual limits on how much individuals can contribute to IRAs, with a higher limit for those age 50 and older, allowing them to make “catch-up” contributions. A tax unit’s adjusted gross income may restrict the maximum contribution amount for Roth IRAs and affect whether Traditional IRA contributions are deductible.

Then there are employer-sponsored IRAs:

1. **Simplified Employee Pension (SEP) Plans:** These plans allow employers to make tax-deferred contributions on behalf of their employees. Self-employed individuals can also use SEP IRAs to save for retirement.
2. **Savings Incentive Match Plans for Employees (SIMPLE) IRAs:** These allow both employer contributions and employee salary deferrals to be made on a tax-deferred basis. Employers are required to either match employee contributions or make fixed contributions to employee accounts.

2.2 IRA Decumulation

A substantial empirical literature examines changes in IRA assets through the lens of withdrawals and decumulation, drawing primarily on high-quality administrative data from the IRS Statistics of Income (SOI) division and industry sources such as the Employee Benefit Research Institute (EBRI) IRA Database. This work has made significant progress in documenting how tax

rules and institutional design shape the timing and magnitude of IRA withdrawals.

Withdrawal behavior varies sharply across account types and age groups, with tax rules governing required minimum distributions (RMDs) playing a central role. RMDs impose mandatory withdrawals from tax-deferred retirement accounts beginning at a specified age, a threshold that has shifted over time—from $70\frac{1}{2}$ prior to 2020, to 72 under the SECURE Act of 2019, and to 73 following SECURE 2.0. Argento, Bryant, and Sabelhaus (2015) show that many IRA balances remain untouched until RMDs take effect. Subsequent work demonstrates that households respond strategically to these requirements, adjusting withdrawal behavior around statutory thresholds (Brown, Poterba, and Richardson 2017; Mortenson, Schramm, and Whitten 2019). Lifecycle models developed by Horneff, Maurer, and Mitchell (2023) further illustrate how RMD incentives interact with broader retirement planning, shaping asset drawdowns over the retirement period.

A related strand of the literature focuses on *leakage*, defined as withdrawals from retirement accounts prior to retirement age. Leakage typically reflects responses to income shocks or liquidity needs and entails both tax penalties and the loss of tax-advantaged growth (Argento et al. 2015; Goodman et al. 2021). Munnell and Webb (2015) estimate that roughly 1.5 percent of assets leak annually from the combined 401(k)/IRA system, implying that aggregate retirement balances are at least 20 percent lower than they would be in the absence of current withdrawal rules. This work highlights the cumulative impact of small, repeated withdrawals and underscores the importance of distinguishing between voluntary early withdrawals and mandatory RMD-driven decumulation (Sabelhaus 2000; Mortenson et al. 2019).

2.3 IRA Accumulation

On the accumulation side, evidence from SOI annual reports and the EBRI IRA Database consistently shows that direct IRA contributions are relatively rare. In any given year, only a minority of IRA holders make contributions, and fewer than half of contributors reach the statutory maximum (Statistics of Income 2018). Although traditional IRAs hold the majority of aggregate IRA assets, Roth IRAs receive more frequent new contributions. Contribution activity is concentrated among younger individuals, whereas older account holders primarily accumulate IRA balances through rollovers from employer-sponsored plans (Copeland 2015). Related work highlights how tax incentives, such as the Saver’s Credit, influence participation and contribution patterns

Table 2: Aggregate IRA Flows and Assets (2019 Adult Resident Population)

Year	Contributions	Rollovers	Withdrawals	Assets
2019	73.2B	509.4B	383.1B	10.76T
2018	67.0B	461.8B	352.9B	8.88T
2017	65.9B	417.7B	305.1B	9.04T
2016	62.8B	378.6B	269.7B	7.66T
2015	61.6B	400.2B	263.7B	6.99T
2014	60.2B	365.4B	245.0B	6.79T
2013	57.2B	328.8B	231.1B	6.21T
2012	51.6B	275.6B	213.6B	5.19T
2011	50.2B	236.5B	193.3B	4.57T
2010	51.4B	231.5B	231.9B	4.87T
2009	49.6B	193.3B	140.8B	3.84T
2008	53.8B	218.7B	154.2B	3.12T
2007	56.1B	240.9B	135.1B	3.99T
2006	54.4B	196.8B	113.9B	3.44T
2005	51.3B	159.1B	95.9B	2.74T
2004	43.5B	143.6B	84.5B	2.56T
2003	39.1B	113.5B	73.3B	2.11T
2002	37.6B	123.6B	71.9B	2.65T
2001	31.8B	112.7B	67.4B	2.09T
2000	31.8B	138.7B	70.8B	1.89T

Notes: All values are in nominal dollars. B denotes billions and T denotes trillions.

Source: Author's analysis of IRS administrative data.

across groups (Bryant 2020).

While this literature provides a detailed accounting of contribution behavior, comparatively less research has examined the mechanisms through which IRA assets have accumulated over time, particularly in light of the rapid growth of IRAs relative to employer-sponsored DC plans. Existing empirical work nevertheless suggests that rollovers play a central role in shaping aggregate IRA balances. As shown in Table 2, rollover flows consistently exceed direct contributions by a wide margin: in 2019, households transferred over \$500 billion from employer-sponsored plans into IRAs through rollovers, compared with less than \$75 billion in direct contributions. Aggregate volumes suggest that rollovers account for a substantial share of new IRA funding and thus accumulation. This perspective is reflected in industry analyses noting that rollovers from employer-sponsored retirement plans have been a major contributor to IRA growth (Holden and Schrass 2019).

Further evidence underscores the prominence of rollovers in IRA accumulation. In 2019, 59 percent of households with traditional IRAs had experienced at least one rollover, and nearly half of newly opened traditional IRAs in 2020

Table 3: Traditional IRA Investors with Rollovers by Investor Age, 2020

Age	Aggregate (Millions \$)	Share (%)	Median (\$)	Mean (\$)
18 to 24	901	0.2	2,090	4,250
25 to 29	3,975	1.0	2,780	7,820
30 to 34	9,209	2.2	4,110	16,900
35 to 39	14,568	3.5	6,220	30,480
40 to 44	19,054	4.6	9,130	47,090
45 to 49	25,100	6.1	13,770	66,220
50 to 54	36,032	8.7	18,240	87,710
55 to 59	59,994	14.5	29,000	127,730
60 to 64	109,522	26.5	58,300	189,270
65 to 69	94,043	22.8	68,130	200,720
70 to 74	40,896	9.9	32,380	152,510
All	413,295	100.0	11,270	87,520

Source: The IRA Investor Database™

Notes: This group consists of traditional IRA investors aged 18 to 74 who had rollovers into their traditional IRAs in tax year 2020. Components may not add to totals because of rounding.

were initiated solely with rollover funds (Holden and Schrass 2019). By contrast, direct contributions are infrequent: between 2007 and 2019, only about 12 percent of IRA holders made a contribution in the survey year.

Table 3 documents the age distribution of rollover activity. Although rollovers occur across the age spectrum, they are concentrated among individuals aged 55 to 74, who account for more than 70 percent of aggregate rollover volume. Mean rollover amounts peak near retirement age, exceeding \$200,000 among those aged 65 to 69. Together, these patterns suggest that most traditional IRAs originate through discrete rollover events rather than sustained contribution behavior.

2.4 Existing Research Based on Administrative Data

Existing empirical work on IRA accumulation and decumulation is often grounded in high-quality administrative data produced by the IRS. These data enable careful measurement of IRA balances, contributions, and withdrawals and are widely used in both academic research and policy analysis.¹

At the same time, the structure of these data shapes the types of questions that are addressed. The workhorse survey used in much research based on

¹IRS-based estimates are also used by the Federal Reserve in constructing the Financial Accounts of the United States, including the estimates of IRA assets reported in Table L.117 (Board of Governors of the Federal Reserve System, 2024).

IRS data is the INSOLE (INdividual and SOLE proprietor), a repeated cross-sectional, stratified random sample of individual income tax returns rather than a longitudinal panel. Analyses based on these data typically emphasize aggregates or cross-sectional patterns, rather than longitudinal dynamics, such as tracking the evolution of IRA balances for the same individuals over time. In addition, because the data are drawn from tax filers, IRA holdings among non-filers are not directly observed. These features do not detract from the value of existing work, but they do point to constraints—rather than immutable limitations—on the extent to which existing data structures can be used to decompose the growth of IRA assets into its underlying components.

In particular, while aggregate evidence points to an important role for rollovers, existing research has not directly assessed how rollover flows translate into sustained asset accumulation at the individual level, or how their contribution compares with other sources of growth over the life cycle. One reason for this gap is data limitations. The data used in this study are designed to complement this literature by enabling longitudinal analysis of IRA balances and transactions and by incorporating information on non-filers.

3 Data

This paper estimates the processes through which IRA assets accumulated for the 2019 U.S. resident population.² To do so, I construct a retrospective balanced panel that enables the study of long-term IRA asset accumulation and the relative role of different inflow mechanisms. This retrospective panel follows the 2019 U.S. resident adult population across a twenty-year window, allowing for detailed accounting of the inflows (contributions and rollovers), outflows (withdrawals), and investment returns that shaped their IRA balances over time.

By reconstructing the histories of IRA-related flows for this fixed population, the analysis sheds light on the role of rollovers and direct contributions in IRA accumulation. Since rollovers often occur as large, one-time events and contributions tend to be smaller and recurring (when they occur at all), a

²Only U.S. residents are included, although individuals who were not residents in 2019 could also have held IRAs. However, this definition aligns with the population concept used by the Federal Reserve’s Financial Accounts. While this choice excludes a subset of individuals, it ensures consistency with national accounts. Moreover, this exclusion would only affect the analysis if non-resident IRA holders were both systematically different from residents and held sizable IRA assets; neither condition is likely.

dynamic, multi-year perspective is essential for assessing their long-run impact.

The construction of this panel rests on administrative tax data that report IRA balances and transactions at the individual level. The next subsections describe the data sources, the method used to define the 2019 population, the procedures for linking individuals and accounts over time to form a consistent and analyzable panel, and the estimation of household wealth used for analysis of IRA accumulation processes by economic subgroups.

3.1 Determining the 2019 Population

This paper uses a 1% sample of the U.S. adult resident population in 2019. Constructing such a sample is nontrivial in the IRS data environment. While creating a 1% sample of tax filers is relatively straightforward, developing a representative 1% sample of the entire U.S. adult resident population—including both filers and non-filers—requires a more complex approach that triangulates multiple IRS and administrative data sources.

The target population includes all individuals who were at least 20 years old in 2019 (or younger if married) and who resided in the United States for the majority of that year. The core challenge lies in accurately identifying both residency and survival to 2019, particularly for individuals who did not file taxes that year.

The starting point for constructing the population is the Social Security Administration’s Death Master File, which contains year of birth and year of death information. Individuals are initially selected based on age and recorded survival to 2019. However, the absence of a recorded year of death is not sufficient to conclude that someone was alive in 2019.

To more rigorously establish evidence of life in that year, individuals are considered present in 2019 if they either have a recorded year of death after 2019 or appear between 2019 and 2023 in at least one qualifying data source. These include tax returns (filed as primary, secondary, or dependent), any of 45 different informational returns (such as W-2s and 1099s), passport applications, records of Economic Impact Payments in 2020 or 2021, or information returns associated with the Affordable Care Act. For individuals aged 85 or older, only tax returns and ACA forms are considered acceptable evidence of life. This process involves searching across billions of records to determine presence in the country during the reference year.

An additional challenge involves identifying marital status and linking spouses. This is particularly important for constructing household-level variables (e.g., household wealth). The standard approach in literature using tax data is to infer marital status from the filing status reported on tax returns, and, in the case of joint or separately filed married returns, to use the return to identify the spouse (e.g., Saez and Zucman 2016). However, this approach is imperfect. In other cases, spousal identifiers are missing from returns filed as married. There are also instances in which the identified spouse is deceased, and some individuals appear married to multiple people or are classified as both single and married across different returns.

For non-filers, the difficulty is greater. In many studies using tax data—including Chetty et al. (2014)—non-filers are simply assumed to be single. While this simplifying assumption eases sample construction, it is not appealing.

To overcome these limitations, a sequential algorithm is utilized that draws on a multiplicity of data sources to identify marital status and establish consistent spousal linkages. This procedure results in a demographically plausible and internally consistent marital structure for the 2019 population.

To align more closely with a widely used population concept in existing research, a further adjustment is made following the procedures used by Piketty, Saez, and Zucman (2018). For the years 1999 to 2019, they construct their population using the SOI’s INSOLE and supplement it with “synthetic” non-filers to ensure full population coverage—*see* Saez (2016) for further details. Importantly, for their synthetic non-filers they exclude those who died during the reference year. In line with this logic, the sample used for this paper eliminates all non-filers and their spouses (when those spouses are not directly in the 1% sample) if they died in 2019. Death is determined using multiple indicators: individual and spousal death dates, as well as a separate verification process to confirm survival at least through the first day of 2020.³

This sample construction procedure identifies 248.1 million individuals as part of the U.S. adult resident population in 2019. For comparison, the American Community Survey (ACS) yields a slightly smaller population under

³This adjustment is not required for analyses focused exclusively on IRAs, but is adopted for wealth-based comparisons across economic subgroups. Handling within-year deaths poses a measurement problem for wealth estimates based on capitalized income, since many income flows are observed only for the portion of the year an individual is alive. Mechanically capitalizing such income can severely understate underlying asset holdings for those who die within the year. There is no uniquely correct solution to this issue. Excluding non-filers who die during the reference year prioritizes a consistent population concept and avoids mechanically distorted wealth estimates driven by the timing of death, reflecting a pragmatic response to a limitation of annual administrative data.

the same age and residency conditions: 246.7 million, or about 0.5% lower than the IRS-based estimate. The modest discrepancy between the two figures may reflect definitional differences between data sources. For instance, the ACS applies a two-month residence rule, which excludes individuals temporarily abroad—such as those living overseas for ten weeks during the survey year—whereas such individuals may still appear in IRS administrative records. In general—and particularly for the purposes of this paper—such a small discrepancy is negligible and does not materially affect the analysis or its alignment with national population benchmarks.

3.2 Construction of IRA Flows and Balances

A longitudinal panel is constructed that consists of each member of the 2019 population and tracks key features of individual IRA activity over the period 1999–2019. The panel includes four core elements for each individual and year: (1) rollovers into IRAs, (2) contributions to IRAs, (3) year-end IRA asset balances, and (4) withdrawals from IRAs. These measures are derived from IRS information returns—specifically Forms 5498 and 1099-R—and provide the basis for evaluating the accumulation of IRA assets over time.

For each individual in the 2019 population, all relevant information returns filed between tax years 1999 and 2019 are collected. To be included in the 2019 population, individuals must be both living and U.S. residents in 2019. However, the residency requirement does not apply to earlier years. Nonetheless, individuals who were not US residents for any year from 1999 to 2018, but residents in 2019, remains relevant for understanding patterns of asset accumulation.

All information returns used in this analysis are filtered to remove duplicate forms and originals replaced by amended form. After this data cleaning process, each core element—rollovers, contributions, withdrawals, and year-end balances—is aggregated at the individual-level.

Rollovers from Employer-Sponsored Plans into IRAs

Rollovers from employer-sponsored retirement plans into IRAs are identified using a joint procedure based on Form 1099-R and Form 5498. Form 1099-R reports distributions from retirement accounts, including employer-sponsored DC plans and, less commonly, certain DB plans, while Form 5498 reports

rollover contributions received by IRAs (Box 2).

Rollover-eligible distributions from DC and DB plans are identified on Form 1099-R (Box 1) using the distribution codes G and H reported in Box 7. The analysis is further restricted to forms for which the *IRA/SEP/SIMPLE* checkbox is not marked, thereby excluding distributions originating from IRAs and limiting attention to rollovers plausibly sourced from employer-sponsored plans. In parallel, rollover contributions reported on Form 5498 (Box 2) are aggregated to the individual level. These contributions reflect amounts received by IRAs that are designated as rollovers but do not, on their own, identify the source account.

For each individual and year, employer-plan rollovers into IRAs are defined as the minimum of the rollover-eligible distribution amount reported on Form 1099-R and the rollover contribution amount reported on Form 5498. This minimum-based construction treats the overlap between distributions and contributions as the portion that can be reliably attributed to completed rollovers from employer-sponsored plans into IRAs. Any excess on either side—such as distributions not redeposited into IRAs, rollover contributions sourced from IRAs or other accounts, or timing and reporting discrepancies—is intentionally excluded.

By design, this approach eliminates intra-IRA rollovers and other transactions that cannot be confidently classified as DC/DB-to-IRA transfers using information returns alone. Bounding rollover activity by both reported distributions and reported contributions produces a conservative measure that prioritizes classification certainty.

Applying this procedure yields aggregate employer-sponsored plan rollovers into IRAs of approximately \$509 billion in 2019. This total is lower than sum reported in SOI aggregates (\$554 billion) and the sum obtained if exclusively reliant on Form 1099-R information (\$598 billion). The differences partially reflect the exclusion of transactions that cannot be unambiguously attributed to employer-to-IRA rollovers.

Lastly, indirect rollovers cannot be cleanly identified using administrative tax records alone. Neither Form 1099-R, which reports distributions from retirement accounts, nor Form 5498, which reports rollover contributions to IRAs, indicates whether a rollover was executed as an indirect transfer, nor do the forms reliably trace the movement of funds across accounts. Brady and Bass (2020) develop a substantially more complex methodology that combines tax return information with account-level flows to approximate indirect rollovers. Their estimates indicate that such transactions are quantitatively small: in

Table 4: IRA Contribution Limits (1999–2019)

Year	Trad/Roth IRA	Trad/Roth Catch-Up	SIMPLE IRA	SIMPLE Catch-Up	SEP IRA
2019	\$6,000	\$1,000	\$13,000	\$3,000	\$56,000
2018	\$5,500	\$1,000	\$12,500	\$3,000	\$55,000
2017	\$5,500	\$1,000	\$12,500	\$3,000	\$54,000
2016	\$5,500	\$1,000	\$12,500	\$3,000	\$53,000
2015	\$5,500	\$1,000	\$12,500	\$3,000	\$53,000
2014	\$5,500	\$1,000	\$12,000	\$2,500	\$52,000
2013	\$5,500	\$1,000	\$12,000	\$2,500	\$51,000
2012	\$5,000	\$1,000	\$11,500	\$2,500	\$50,000
2011	\$5,000	\$1,000	\$11,500	\$2,500	\$49,000
2010	\$5,000	\$1,000	\$11,500	\$2,500	\$49,000
2009	\$5,000	\$1,000	\$11,500	\$2,500	\$49,000
2008	\$5,000	\$1,000	\$10,500	\$2,500	\$46,000
2007	\$4,000	\$1,000	\$10,500	\$2,500	\$45,000
2006	\$4,000	\$1,000	\$10,000	\$2,500	\$44,000
2005	\$4,000	\$500	\$10,000	\$2,000	\$42,000
2004	\$3,000	\$500	\$9,000	\$1,500	\$41,000
2003	\$3,000	\$500	\$8,000	\$1,000	\$40,000
2002	\$3,000	\$500	\$7,000	\$500	\$40,000
2001	\$2,000	N/A	\$6,500	N/A	\$35,500
2000	\$2,000	N/A	\$6,000	N/A	\$30,000
1999	\$2,000	N/A	\$6,000	N/A	\$30,000

Source: IRS Publication 560 and Publication 590-A, various years.

Notes: (1) Catch-up contributions apply to those 50 and older. They do not apply to SEP IRAs but do apply to SIMPLE IRAs, Traditional IRAs, and Roth IRAs. (2) SEP contributions reported on Form 5498 (box 8) include only employer contributions. Employees cannot contribute directly to a SEP IRA. (3) SIMPLE IRA contributions reported on Form 5498 (box 9) include both employee salary deferrals and employer contributions. (4) Traditional & Roth IRA limits are combined: one can contribute a total of \$6,000 (or \$7,000 if age 50+), but not \$6,000 to each. (5) An individual cannot contribute to both a SEP IRA and a SIMPLE IRA within the same employer. Multiple employers may allow contributing to both.

tax year 2010, only \$0.7 billion in indirect rollovers from DC and DB plans to IRAs were completed, compared with \$267.7 billion in direct rollovers, and were associated with approximately 55,000 filers. Given their limited aggregate magnitude, greater susceptibility to failed redeposits and classification inconsistencies, and declining prevalence as direct rollovers have become increasingly automated, no separate procedure is implemented to estimate indirect rollovers explicitly—and could not be implemented for non-filers, for whom validation opportunities are limited.

Direct Contributions to IRAs

Contributions are drawn from Form 5498, which records total contributions to traditional IRAs (Box 1), Roth IRAs (Box 10), SEP IRAs (Box 8), and SIMPLE IRAs (Box 9). Contributions are first aggregated across plan types within individuals. They are then subjected to year-specific statutory limits

based on age and plan rules. The contribution caps vary by account type and change over time, as shown in Table 4. For those aged 50 or older, catch-up provisions allow for higher limits.

To reflect legal contribution caps, each individual’s total contributions are compared to the allowable limit in a given year, and excess contributions are capped accordingly. This procedure addresses a small number of observations with implausibly large reported contributions, which are likely driven by rare instances of reporting error, miscoding, or other forms of data contamination rather than genuine violations of statutory limits. The impact of this adjustment is limited: applying contribution caps reduces aggregate IRA contributions by approximately 3 percent.

The allowable limit is defined as:

$$\text{Cap}_{i,t} = \begin{cases} C_t^{\text{base}} + C_t^{\text{catchup}}, & \text{if } \text{age}_{i,t} \geq 50 \\ C_t^{\text{base}}, & \text{otherwise} \end{cases}$$

where C_t^{base} and C_t^{catchup} are the base and catch-up limits for year t . These limits vary by plan type and are applied accordingly. An additional rule ensures that SIMPLE IRAs and SEP IRAs are not double-counted when plan participation overlaps.

Fair Market Value of IRA Assets

End-of-year market values of IRA accounts are reported on Form 5498 (Box 5). These values reflect the total IRA holdings of each individual, including assets accumulated through contributions, rollovers, and investment returns. Reported balances represent the market values of IRA investments as of December 31 for each tax year.

However, several caveats apply to the interpretation of these reported values. In rare instances, particularly in the case of decedents, the market values reported in Box 5 of Form 5498 may reflect the value of the account on the date of death rather than at year-end. Additionally, custodians may occasionally report a missing or zero market value when they are unable to assign a market value to the assets in the account. This typically arises when IRAs hold non-publicly traded or illiquid assets—such as closely held business interests, real estate, or limited partnerships—for which no readily ascertainable market price exists. While these cases are uncommon, they introduce potential measurement challenges in assessing IRA balances in a small subset of the pop-

ulation; this issue is also addressed in greater detail below in the discussion of implicitly missing market values.

Non-Rollover IRA Withdrawals

Withdrawals are also captured using Form 1099-R, but under a separate classification from rollovers. To isolate true non-rollover distributions taken by the account holder, records are restricted to those in which the *IRA/SEP/SIMPLE* checkbox is marked—indicating that the distribution originated from an IRA—and where Box 7 does not include codes associated with rollovers, corrections, or exceptional cases. Specifically, the analysis excludes forms with Box 7 distribution codes 4, 5, 6, C, G, H, N, and R, which reflect death distributions, prohibited transactions, rollovers, recharacterizations, and return of contributions, among others. Withdrawals are then aggregated at the individual-by-year level.

3.3 Construction of Economic Subgroups

To support additional analyses, individuals in the 2019 population are classified according to the value of their IRA holdings and their total household net worth under the equal split assumption. These classifications are created by sorting individuals based on the reported value of IRA accounts and the estimated value of total assets minus liabilities, respectively. Each classification is fixed relative to the 2019 population and remains unchanged over the retrospective panel. From this, estimates are produced by group in order to determine whether IRA growth processes are uniform across the population.

The categorization of individuals by IRA holdings is based on observed account balances. Using these values, individuals are grouped based on ordered levels of account size. The classification of individuals based on household wealth is derived from an estimation framework that builds upon that used by Saez and Zucman (2020), but which introduces a set of methodological modifications. The scope of asset and liability components is broadened to include both funded and unfunded DB pension entitlements, which is consistent with the wealth concept used in the Financial Accounts. Additionally, IRA assets are directly assigned from Form 5498 rather than estimated via capitalization of Form 1040 IRA distributions and imputation.

Unlike prior efforts to estimate household net worth, which rely solely on

Table 5: Information Returns and Fields Used to Produce Non-Filer Estimates

Form	Field	Form	Field
1	Box 4a Net long-term capital gain/loss	25	Box 2: Net rental real estate income (loss)
2	Box 3: Net short-term capital gain/loss	26	Box 5: Fishing boat proceeds
3	Box 1: Interest income	27	Box 6: Medical and health care payments
4	Box 8: Net short-term capital gain/loss	28	Box 7: Nonemployee compensation
5	Box 5: Interest income	29	Box 14: Gross proceeds paid to an attorney
6	Box 3: Other income	30	Box 1a: Total ordinary dividends
7	Box 11: Other income (loss)	31	Box 3: Early withdrawal penalty
8	Box 9a Net long-term capital gain/loss	32	Box 10: Purchase price above face value
9	Box 1: Ordinary business income (loss)	33	Box 3: Interest on US Savings Bonds
10	Box 3: Other net rental income (loss)	34	Box 11: Tax-exempt interest
11	Box 2a: Ordinary dividends	35	Box 11: Exempt-interest dividends
12	Box 4: Guaranteed payments	36	Box 1: OID interest
13	Box 1f: Accrued Market Discount	37	Box 2: Other periodic interest
14	Box 1d: Proceeds	38	Box 2a: Taxable amount
15	Box 11: Aggregate profits on contracts	39	Box 1: Gross distribution
16	Box 13: Bartering	40	Box 7: Distribution codes
17	Box 1e: Cost Basis	42	Box 8a Net long-term capital gain/loss
18	Box 2a: Total capital gain distribution	43	Box 7: Net short-term capital gain/loss
19	Box 1: Rents	44	Box 1: Ordinary business income (loss)
20	Box 5a: Ordinary dividends	45	Box 3: Other net rental income (loss)
21	Box 1b: Qualified dividends	46	Box 6a: Ordinary dividends
22	Box 2: Early withdrawal penalty	47	Box 4: Interest income
23	Box 8: OID on Treas	48	W-2 Box 1: Wages
24	Box 1: Interest income		

simulated or imputed data for non-filers (e.g., Saez and Zucman 2016, 2020; Smith, Zidar, and Zwick 2023), this approach uses information returns to estimate non-filers’ income flows as they would appear on a Form 1040 had they filed. These income flows are then used to estimate non-filers’ net assets using the same method applied to filers. A detailed account of the fields and forms used in this reconstruction is provided in Table 5.

To validate the non-filer estimation procedure, single tax filers in 2019 are processed twice: once using full Form 1040 data and once using only the information returns that would be available if they had not filed; the resulting aggregates are then compared to assess accuracy. As shown in Table 6, the information return approach reproduces the vast majority of reported income with a high degree of accuracy. Wages—the largest component—are captured almost exactly (99.9% match). Aggregate totals for income relevant to asset and liability assignment differ by just 3.4%. These results suggest that the non-filer estimation procedure reliably approximates key financial characteristics, lending confidence to its application across the broader population.

To estimate net worth for all individuals, the framework begins by adopting a comprehensive accounting of asset and liability categories as used in the Financial Accounts of the United States Table B.101.h—excluding consumer durables, as is standard in comparable approaches—such that aggregate net asset levels produced through this procedure mechanically match those reported by the Federal Reserve.⁴

⁴See Board of Governors of the Federal Reserve System, *Financial Accounts of the United States*, Table B.101.h: “Balance Sheet of Households and Nonprofit Organizations.” Documentation available at <https://www.federalreserve.gov/apps/FOF/Guide/B101h.pdf>.

Table 6: Validation of Information Return Based Estimates Using Single Filers

Income Variable	F1040 Measure / Information Return Measure	F1040 Measure (Billions USD)	Information Return Measure (Billions USD)
Wages	100%	\$2,215	\$2,212
Taxable Retirement Distributions	129%	\$348	\$271
Capital Gains	121%	\$152	\$126
Profits from Pass-Throughs	93%	\$114	\$123
Income from Sole-Proprietorships	95%	\$101	\$107
Rental Payments on Tenant Occupied Real Estate	104%	\$88	\$85
Qualified Dividends	100%	\$62	\$62
Interest Income	81%	\$54	\$66
Property Taxes on Owner-Occupied Real Estate	n/a	\$30	n/a
Income for Fixed-Income Assets held in Mutual Funds	88%	\$23	\$26
Mortgage Payments on Tenant Occupied Real Estate	94%	\$12	\$13
Income Relevant to Asset and Liability Assignment	104%	\$3,199	\$3,089
F1040 Total Income net of OASDI Income and UC	99%	\$2,907	\$2,939

Source: Author’s analysis of IRS administrative data.

Notes: (1) Estimates use 2019 data; all dollar figures are in nominal 2019 dollars. (2) Mortgage payments on tenant-occupied real estate cannot be directly observed and are estimated using filer-level average ratios of rental income from tenant-occupied properties to associated mortgage payments. (3) Interest income combines taxable and tax-exempt interest; for non-filer estimation, these components are disaggregated. (4) Pass-through profits combine S-corporation and partnership income; because passive and non-passive profits cannot be directly observed in information returns, filer-level averages are used to disaggregate these flows. (5) Property taxes paid are observed only for itemizers and are therefore unavailable in the information-return-based approach.

Administrative records are then used to approximate individual asset or liability levels. As stated, IRA balances are directly assigned using information from Form 5498. Where feasible, income flows are linked to underlying assets and flows are capitalized to recover underlying stocks. In cases where this link is noisy or unavailable—such as for many liabilities or non-income-generating assets—imputation is performed using the Survey of Consumer Finances (SCF), following the approach in Saez and Zucman (2020). Specifically, individuals are sorted into cells defined by age, marital status, and income grouping. Within each cell, the probability of owning a given asset or liability and the average holding size are drawn from the Survey of Consumer Finances. Each individual is probabilistically assigned asset or liability ownership based on these probabilities. If ownership is assigned, the mean value for the cell is applied.

4 Methods

This section outlines the method used for systematically quantifying how account balances evolve over time based on observable flows—contributions, rollovers, and withdrawals. At its core, the method relies on a transition framework in which the IRA balance for individual i at the end of year t , denoted $IRA_{i,t}$, is modeled as the result of four components: the prior year’s

ending balance $IRA_{i,t-1}$, inflows (contributions and rollovers), outflows (withdrawals), and investment returns. While balances and flows are directly observed in administrative data, individual-year-specific returns are not. The unknown rate of return, $r_{i,t}$, is treated as a residual term that captures the change in balances not explained by observed flows.

To operationalize this framework, consider the following general formulation:

$$IRA_{i,t} = f(IRA_{i,t-1}, C_{i,t}, Roll_{i,t}, W_{i,t}, r_{i,t})$$

where $C_{i,t}$ denotes contributions, $Roll_{i,t}$ rollovers into IRAs, and $W_{i,t}$ withdrawals. The function $f(\cdot)$ summarizes how investment returns apply to both prior balances and net flows over the year and may depend on assumptions about the timing of flows (e.g., mid-year versus end-of-year).

4.1 Estimating Year-on-Year Rates of Return

Within this framework, year-on-year rates of return, $r_{i,t}$, are estimated for each individual from 2000 to 2019. The preferred return specification follows what is referred to here as the *mid-year weighted* assumption. This approach builds on standard methods for estimating returns with annual data (Fagereng et al. 2020), which assume that contributions and withdrawals occur evenly throughout the year, but adapts them to the specific timing patterns of IRA flows. Rather than assuming that all net contributions are evenly distributed, the specification treats half of direct IRA contributions as occurring at year-end and the other half as spread evenly throughout the year. This adjustment reflects administrative data from the Investment Company Institute showing that 40–50% of direct IRA contributions are made between January and April of the *following* calendar year—after the tax year to which they apply—and thus do not earn returns in the year they are attributed to.⁵

This mid-year weighted specification—the preferred specification used throughout this paper—treats rollovers and withdrawals as occurring evenly throughout the year, while direct contributions are treated differently: half of direct contributions are assumed to arrive at the end of the year and earn no return, while the remaining half are assumed to be distributed evenly throughout the year. As a result, only half of contributions receive return exposure, and those exposed contributions earn only half of the annual return, yielding an effective return weight of 0.25.

⁵IRA contributions made before tax day in April can still be designated for the prior tax year.

Table 7: Rate of Return Variants and Formulas

RoR Variant	Adjustment and Formula
Mid-Year Weighted	Assumes rollovers and withdrawals are spread evenly throughout the year. Half of direct contributions ($C_{i,t}$) are treated as arriving at year-end and therefore do not earn returns; the other half are assumed to earn returns for half the year. $r_{i,t}^{\text{mid-wgt}} = \frac{IRA_{i,t} - IRA_{i,t-1} - C_{i,t} - Roll_{i,t} + W_{i,t}}{IRA_{i,t-1} + 0.25 \cdot C_{i,t} + 0.5 \cdot Roll_{i,t} - 0.5 \cdot W_{i,t}}$
Beginning-of-Year	Assumes all flows occur at the start of the year. $r_{i,t}^{\text{beg}} = \frac{IRA_{i,t} - IRA_{i,t-1} - C_{i,t} - Roll_{i,t} + W_{i,t}}{IRA_{i,t-1} + C_{i,t} + Roll_{i,t} - W_{i,t}}$
Mid-Year	Assumes flows are evenly distributed throughout the year. $r_{i,t}^{\text{mid}} = \frac{IRA_{i,t} - IRA_{i,t-1} - C_{i,t} - Roll_{i,t} + W_{i,t}}{IRA_{i,t-1} + 0.5 \cdot (C_{i,t} + Roll_{i,t} - W_{i,t})}$
End-of-Year	Assumes no intra-year flows. $r_{i,t}^{\text{end}} = \frac{IRA_{i,t} - IRA_{i,t-1} - C_{i,t} - Roll_{i,t} + W_{i,t}}{IRA_{i,t-1}}$

Under these assumptions, the transition equation takes the following form:

$$IRA_{i,t} = IRA_{i,t-1} + C_{i,t} + Roll_{i,t} - W_{i,t} + r_{i,t} \cdot (IRA_{i,t-1} + 0.25 \cdot C_{i,t} + 0.5 \cdot (Roll_{i,t} - W_{i,t})).$$

Solving for $r_{i,t}$ gives the mid-year weighted rate of return formula:

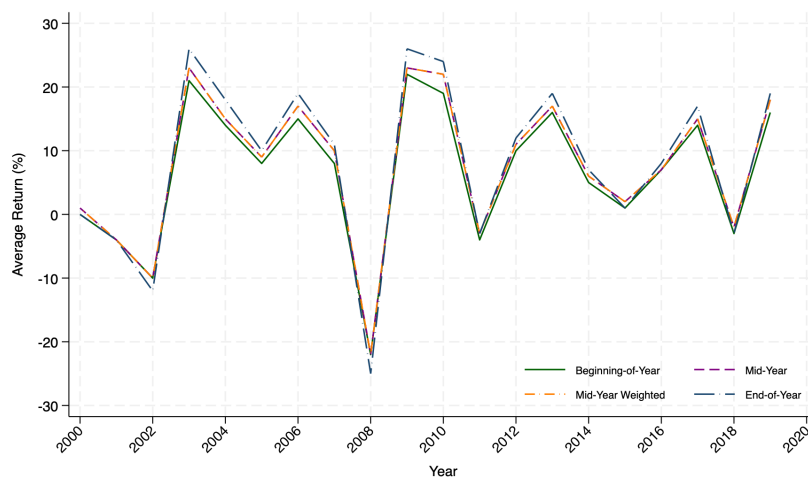
$$r_{i,t}^{\text{mid-wgt}} = \frac{IRA_{i,t} - IRA_{i,t-1} - C_{i,t} - Roll_{i,t} + W_{i,t}}{IRA_{i,t-1} + 0.25 \cdot C_{i,t} + 0.5 \cdot (Roll_{i,t} - W_{i,t})}$$

In addition to this preferred specification, three alternative return variants are estimated based on different assumptions about the timing of flows (*see* Table 7). Each variant is implemented using the same administrative data on IRA balances, contributions, rollovers, and withdrawals. While these return formulations differ in how they treat the timing of flows, the resulting estimates are nearly identical in practice. Figure 4 shows the average annual return on IRAs assets by year and return formula specification. Variation across the specifications is minimal. For consistency, reasons of parsimony, and the justifications provided above, the mid-year weighted return formulation is solely used throughout the subsequent analysis.

Capping Extreme Returns

To limit the influence of extreme outliers, annual rates of return are capped at $\pm 1000\%$, with estimates exceeding this threshold in absolute value set to

Figure 4: Average Individual IRA Return by Year and Return Formula



Source: Author’s analysis IRS administrative data.

Note: $\pm 1000\%$ cap has been imposed on all returns. The figure shows the mean individual-level rate of return, calculated as the average of each individual’s annual return.

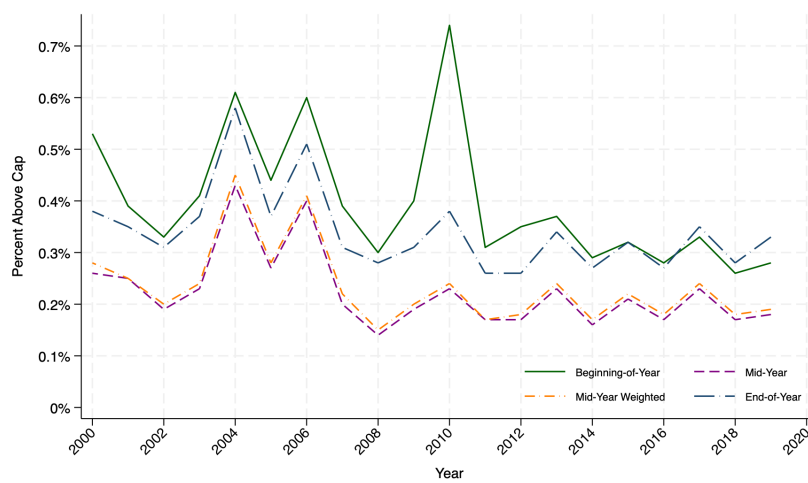
the cap. In addition, an alternative set of return estimates is generated in which returns are capped at $\pm 100\%$. These alternative estimates, reported in the appendix, serve as a robustness check and are nearly identical to those obtained using the preferred and less restrictive cap.

Extreme estimates are exceedingly rare. When they do occur, they typically reflect data limitations or mechanical artifacts—often arising from indirect rollovers not observed in the data or from missing IRA values for which market valuation is difficult—rather than genuinely extraordinary investment performance. Figure 5 shows the frequency of these extreme estimates by year and return specification. Under the preferred mid-year weighted specification, only about 0.2% of observations in any year have returns greater than +1000% or less than -1000%.

Validating Estimates Against an External Benchmark

Figure 6 compares the estimated annual mean returns on IRAs for the 2019 population with the aggregate annual mean return on DC accounts as reported in Form 5500 data from 2000 to 2019. Form 5500 is submitted annually by ERISA covered employer-sponsored retirement plans to the U.S. Department of Labor to comply with ERISA reporting requirements. The dataset includes information from all DC plans sponsored by private employers with 100 or

Figure 5: Percent of RoR Estimates Subject to the $\pm 1000\%$ Cap



Source: Author’s analysis of IRS administrative data.

more participants, making it a high-quality external benchmark that captures the vast majority of assets held in the DC system.

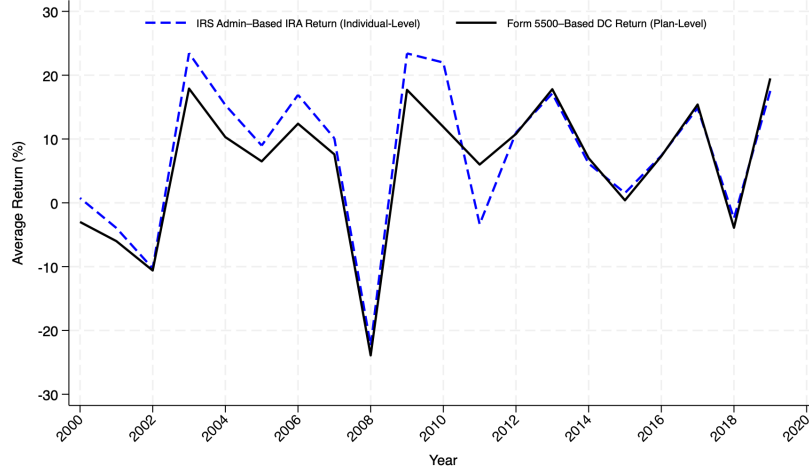
As shown in Figure 6, the estimated IRA return series closely tracks the Form 5500-based DC return benchmark throughout the entire period. The alignment is strong in all years and becomes nearly identical in the years closest to 2019. While some divergence is expected—since the Form 5500 returns reflect the full U.S. DC participant population, whereas the IRA estimates pertain to a subset of individuals alive and residing in the U.S. in 2019—the close correspondence provides additional confidence in the reliability of the estimated mean returns, the underlying data, and the method.

Addressing Implicitly Missing Prior-Year Balances

In a small number of cases (0.72% in 2019), individuals exhibit positive IRA balances in year t , but no recorded IRA assets in year $t - 1$ and neither direct contributions or rollovers in year t . If used directly in the return formulae, this combination would yield undefined values due to division by zero in the denominator or implausible estimates when only withdrawals are present.

These cases imply either an unobserved rollover and/or contribution in year (t) , or a missing IRA balance in year $(t - 1)$. In most instances, they likely reflect unobserved indirect rollovers, as evidenced by the declining incidence

Figure 6: Annual Mean IRA Returns for the 2019 Cohort Compared to Aggregate DC Plan Returns (Form 5500), 2000–2019



Source: Author’s analysis of IRS administrative data and U.S. Department of Labor’s analysis of Form 5500 information in annual Private Pension Plan Bulletin.

of such observations over time—consistent with the declining frequency of indirect rollovers—or, alternatively, the absence of reliable market valuation for IRA balances (e.g., when accounts contain hard-to-value assets).

While this represents a relatively small share of the total sample—typically fewer than 1% of cases—it still warrants a principled correction. To address these cases, a retrospective imputation procedure is adopted. The idea is to estimate a plausible prior-year IRA balance using the inverse of the annual IRA transition equation, under the assumption that the individual experienced the average *mid-year weighted rate of return* in year t . The imputation formula is:

$$IRA_{i,t-1} = \frac{1}{1 + \bar{r}_t^{\text{mid-wgt}}} \cdot \left(IRA_{i,t} - C_{i,t} - Roll_{i,t} + W_{i,t} - \bar{r}_t^{\text{mid-wgt}} \cdot (0.25C_{i,t} + 0.5(Roll_{i,t} - W_{i,t})) \right)$$

This procedure for recovering missing IRA asset information is applied first to 2019 and then proceeds iteratively backward to recover information for earlier years, through 1999. This sequencing is necessary because the current-year information is used to estimate the prior-year balance. Beginning with the most recent year ensures that the imputation process maximizes the recovery of earlier-year values by drawing on the fullest available data.

The aggregate magnitude of this correction is modest (0.43% of observed IRA assets in 2019) but important for internal consistency. From 1999 to 2019, the share of observations requiring imputation increases as one moves backward

in time, with the highest concentration in the early 2000s. This pattern reflects both the greater prevalence of indirect rollovers in earlier years and the fact that imputation is more feasible when information from a future year is available. Additional details are provided in the appendix.

4.2 Baseline Asset Accumulation Micro-Simulation

The evolution of IRA balances over time is simulated by applying a transition equation that maps asset values from one period to the next. This transition rule implies a forward accumulation process that expresses IRA balances in any future year as a function of the initial balance, net flows, and investment returns.

Recall the period-to-period transition equation under the mid-year weighted return assumption, defined for individual, i , between years $t - 1$ and t :

$$IRA_{i,t} = IRA_{i,t-1} + C_{i,t} + Roll_{i,t} - W_{i,t} + r_{i,t} \cdot (IRA_{i,t-1} + 0.25 \cdot C_{i,t} + 0.5 \cdot (Roll_{i,t} - W_{i,t})).$$

Iterating this transition equation yields an accumulation formula for IRA assets of each individual, i . Starting from the base-year balance in 1999, the implied IRA balance in 2019 can be written as:

$$IRA_{i,2019} = IRA_{i,1999} \prod_{t=2000}^{2019} (1 + r_{i,t}) + \sum_{t=2000}^{2019} \left[(0.25 \cdot C_{i,t} + 0.5 \cdot (Roll_{i,t} - W_{i,t})) \prod_{s=t}^{2019} (1 + r_{i,s}) + (0.75 \cdot C_{i,t} + 0.5 \cdot (Roll_{i,t} - W_{i,t})) \prod_{s=t+1}^{2019} (1 + r_{i,s}) \right]$$

This expression represents the 2019 IRA balance of each individual, i , as the sum of three components: (1) the compounded growth of the initial 1999 balance, (2) the return-weighted accumulation of within-year inflows and outflows according to their assumed timing, and (3) the non-compounding portion of net flows.

In implementation, the simulation applies additional observational anchoring and boundedness constraints in rare edge cases to preserve numerical stability, including a zero lower bound on simulated balances and an upper bound in extreme divergence cases. These adjustments introduce mild nonlinearities but do not alter the structure of the underlying accumulation equation characterized above. These further adjustments are discussed below in detail.

Table 8: Simulated Aggregate IRA Assets for the 2019 Population, 2000–2019

Year	(A) Observed IRA (T)	(B) Simulated IRA (T)	(C) Diff. (T) Col(B) – Col(A)	(D) % Diff. Col(C)/Col(A)
2019	10.81	10.38	-0.43	-4.01%
2018	8.94	8.55	-0.39	-4.36%
2017	9.14	8.72	-0.42	-4.63%
2016	7.79	7.44	-0.35	-4.47%
2015	7.11	6.79	-0.32	-4.50%
2014	6.94	6.60	-0.34	-4.83%
2013	6.36	6.06	-0.30	-4.71%
2012	5.34	5.09	-0.25	-4.62%
2011	4.71	4.49	-0.22	-4.67%
2010	5.04	4.80	-0.23	-4.62%
2009	3.99	3.82	-0.17	-4.24%
2008	3.25	3.08	-0.16	-5.02%
2007	4.17	3.91	-0.25	-6.02%
2006	3.67	3.45	-0.22	-5.92%
2005	2.99	2.88	-0.11	-3.52%
2004	2.73	2.63	-0.09	-3.34%
2003	2.32	2.27	-0.06	-2.38%
2002	2.77	1.85	-0.92	-33.19%
2001	2.25	1.98	-0.27	-12.07%
2000	2.05	2.04	-0.01	-0.71%

Source: Author’s analysis of IRS administrative data.

Notes: All monetary values are expressed in trillions (T) of nominal USD. Simulated values in Column (B) are generated using the iterative method outlined in this section. Column (C) reports the difference between simulated and observed IRA assets. Column (D) reports the percent difference relative to observed assets.

By parameterizing this accumulation equation with observed rollovers ($Roll_{i,t}$), contributions ($C_{i,t}$), withdrawals ($W_{i,t}$), estimated annual rates of return on IRA assets ($r_{i,t}$), and the base-year account value ($IRA_{i,1999}$), the model can be used to assess how closely the accumulation process approximates observed IRA balances in each year. Table 8 compares the simulated aggregate market value of IRA assets for each year from 2000 to 2019—constructed using the iterative summation formula for the analytic sample—with the corresponding observed aggregates derived from administrative records. The simulated values closely track the observed totals, with differences generally on the order of 4%.⁶

⁶Larger discrepancies appear in 2001 and 2002. These differences reflect irregularities in the underlying administrative data in the early years of IRA reporting, including incomplete identification of duplicate or amended records in 2000–2002. Importantly, the micro-simulation incorporates both observational anchoring and explicit boundedness constraints, ensuring that such early-period measurement error does not propagate mechanically through the accumulation process. As an additional safeguard, the annual aggregate percent difference between simulated and observed IRA assets is used to rescale counterfactual simulated values in subsequent analyses, preventing baseline structural error in the accumulation model from influencing estimated counterfactual results.

4.3 IRA Growth: Rollovers vs. Direct Contributions

To evaluate the contribution of specific account flows to the growth of IRA assets, the simulation model allows for counterfactual manipulation of key parameters: rollovers ($Roll_{i,t}$) and contributions ($C_{i,t}$). The analysis focuses on two central counterfactuals—see the appendix for supplemental simulations. The primary objective is to estimate upper-bound effects of rollovers and contributions on long-term IRA asset accumulation.

The two core simulations proceed as follows:

- In the *No Rollovers* simulation, all rollovers between 2000 and 2019 are set to zero, while all other components remain unchanged. This isolates the cumulative impact of rollover flows on asset accumulation and, when compared to observed 2019 balances, allows for an evaluation of their long-run effect.
- In the *No Contributions* simulation, all contributions over the same period are set to zero, while other components are unchanged. This isolates the role of direct contributions in shaping account balances and enables assessment of their long-run impact relative to observed outcomes.

To ensure internal consistency and to prevent mechanically implausible outcomes, several limited adjustments are applied during the iterative simulation procedure.

First, simulated IRA balances are constrained to be non-negative. In rare cases—most commonly under counterfactual scenarios in which rollovers and/or contributions are removed while withdrawals remain—simulated values can become negative as a mechanical consequence of applying observed outflows in the absence of prior inflows. In these cases, balances are reset to zero.

Second, the simulation handles a small set of observations that are inconsistent with the transition equation given the observed data. Specifically, when an individual has no observed prior-year balance ($IRA_{i,t-1} = 0$), no observed flows in year t ($C_{i,t} = 0$, $Roll_{i,t} = 0$, $W_{i,t} = 0$), but a positive observed balance in year t ($IRA_{i,t} > 0$), the simulated value is reset to the observed current-year balance. These cases are consistent with unobserved indirect rollovers and/or missing market valuation of IRA balances that were not handled through the imputation procedure. Replacing the simulated value with the observed balance prevents artificial attrition of IRA assets arising from missing data in the transition equation, rather than from genuine economic processes. Without this correction, mechanically zeroing such accounts would cause them to drop out of subsequent periods and artificially lowering aggregate IRA assets.

Third, simulated values are monitored for divergence from observed balances. Extreme divergence cases are constrained: when simulated assets exceed observed balances by a factor of 100 or more (conditional on a positive observed balance), the simulated value is capped at the observed level. This threshold-based correction is rare—affecting roughly 0.01% of observations—and curbs implausible compounding driven by edge cases (e.g., small initial balances combined with timing assumptions). The cap is applied independently each year to avoid compounding distortions.

Together, these adjustments reflect a conservative modeling strategy: corrections are introduced only to prevent mechanically unstable or empirically implausible outcomes, while preserving the underlying accumulation dynamics implied by observed flows and the return specification.

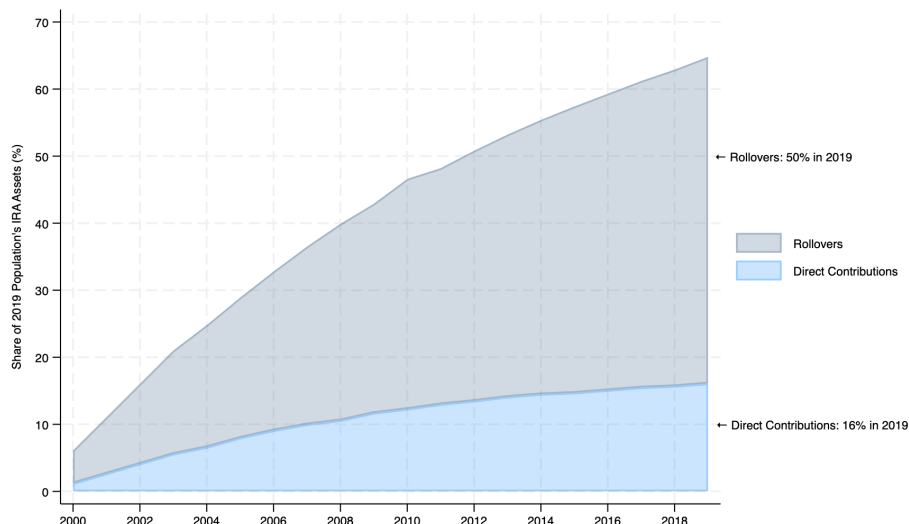
5 Results

Two simulations are used to assess the sources of IRA asset growth between 2000 and 2019, each introduces counterfactual changes to isolate the contribution of specific inflow mechanisms. In the first counterfactual, all direct IRA contributions are set to zero, while rollovers, withdrawals, and returns remain unchanged. In the second, all rollovers from employer-sponsored retirement plans—predominantly defined contribution plans, and in rare cases DB plans—are set to zero, with all other components held fixed. Together, these counterfactuals allow the overall growth of IRA assets over the 2000–2019 period to be decomposed into three components: (i) direct contributions, (ii) rollover contributions from employer plans, and (iii) legacy balances held prior to 2000, which themselves reflect earlier contributions or rollovers. The results from these simulations are then examined both in the aggregate and across economic subgroups.

5.1 The Cumulative Role of Rollovers and Contributions

The aggregate consequences of removing rollovers differ substantially from those of removing direct contributions. As shown in Table 9, when all rollovers from employer-sponsored retirement plans into IRAs are removed, simulated aggregate IRA assets in 2019 fall to \$5.55 trillion—roughly half of the observed \$10.81 trillion. This divergence emerges early in the simulation period and

Figure 7: Decomposition of IRA Asset Growth by Source



Source: Author’s analysis of IRS administrative data.

Notes: The figure decomposes aggregate IRA asset growth into direct contributions, rollovers from employer-sponsored retirement plans, and the compounded value of pre-existing balances. The decomposition is based on linked IRS administrative records and follows a 1% representative sample of the U.S. adult resident population observed in 2019 backward over the 2000–2019 period.

widens steadily over time. By the mid-2000s, simulated balances are already more than 25 percent below observed levels, and by the late 2010s the shortfall approaches 50 percent.

In contrast, removing direct IRA contributions produces a considerably smaller reduction in aggregate balances. Table 10 shows that under the no-contributions counterfactual, simulated IRA assets in 2019 decline to \$9.07 trillion, corresponding to a 16 percent reduction relative to the observed total. The contribution-driven gap is negligible in the early 2000s and grows gradually over time, reflecting the incremental nature of annual contributions and their more limited capacity to generate large balances. While contributions clearly matter for IRA accumulation, their aggregate impact is modest relative to that of rollovers.

Taken together, the two counterfactuals account for roughly two-thirds of observed IRA assets in 2019, with the remaining one-third attributable to balances already held prior to 2000. Figure 7 summarizes this decomposition visually.

Legacy balances cannot be separately decomposed into contributions and rollovers given data limitations. If pre-2000 balances were assigned using the same

Table 9: Counterfactual IRA Assets Without Rollovers

Year	(A) Observed IRA Assets (T)	(B) No Rollovers IRA Assets (T)	(C) Difference (T) Col(B) – Col(A)	(D) % Difference Col(C)/Col(A)
2019	10.81	5.55	-5.26	-48.7%
2018	8.94	4.72	-4.22	-47.2%
2017	9.14	4.97	-4.18	-45.7%
2016	7.79	4.35	-3.44	-44.2%
2015	7.11	4.08	-3.04	-42.7%
2014	6.94	4.10	-2.84	-40.9%
2013	6.36	3.87	-2.49	-39.1%
2012	5.34	3.35	-1.99	-37.3%
2011	4.71	3.05	-1.66	-35.2%
2010	5.04	3.31	-1.73	-34.3%
2009	3.99	2.75	-1.24	-31.2%
2008	3.25	2.30	-0.95	-29.3%
2007	4.17	3.06	-1.10	-26.5%
2006	3.67	2.80	-0.87	-23.7%
2005	2.99	2.36	-0.63	-20.9%
2004	2.73	2.23	-0.50	-18.2%
2003	2.32	1.96	-0.36	-15.4%
2002	2.77	2.42	-0.35	-12.7%
2001	2.25	2.05	-0.19	-8.6%
2000	2.05	1.95	-0.10	-4.9%

Source: Author’s analysis of IRS administrative data.

Notes: All monetary values are expressed in trillions (T) of nominal USD. Column (A) reports observed IRA asset totals for the 2019 population. Column (B) reports counterfactual totals under the removal of all DC/DB-to-IRA rollovers, scaled to maintain aggregate consistency. Column (C) reports the absolute difference between counterfactual and observed assets, and Column (D) reports the percentage difference relative to observed assets.

rollover-to-contribution mix observed over the 2000–2019 period, rollovers would account for approximately three-quarters of legacy assets. Under this assumption, rollovers would explain roughly 76 percent of total IRA assets in 2019, with the remaining 24 percent attributable to direct contributions.

5.2 Results by Economic Subgroups

Table 11 reports the role of rollovers and direct contributions in shaping IRA accumulation over the 2000–2019 period across economic subgroups, with outcomes measured as final (2019) balances. Results are shown for the full population (Panel A), IRA asset quantiles (Panel B), and net worth quantiles (Panel C). Across nearly all groups, rollovers account for a substantially larger share of accumulated IRA assets, while the contribution channel plays a more limited role, especially as IRA assets and net worth increase.

Panel B shows that the removal of rollovers substantially reduces IRA balances throughout the IRA asset distribution. For individuals in the bottom 80 percent of IRA balances, eliminating rollovers reduces simulated balances

Table 10: Counterfactual IRA Assets Without Direct Contributions

Year	(A) Observed IRA Assets (T)	(B) No Contributions IRA Assets (T)	(C) Difference (T) Col(B) – Col(A)	(D) % Difference Col(C)/Col(A)
2019	10.81	9.07	-1.74	-16.1%
2018	8.94	7.54	-1.40	-15.7%
2017	9.14	7.73	-1.42	-15.5%
2016	7.79	6.62	-1.18	-15.1%
2015	7.11	6.06	-1.05	-14.7%
2014	6.94	5.93	-1.00	-14.5%
2013	6.36	5.46	-0.90	-14.1%
2012	5.34	4.62	-0.72	-13.5%
2011	4.71	4.10	-0.61	-13.0%
2010	5.04	4.42	-0.62	-12.3%
2009	3.99	3.52	-0.47	-11.7%
2008	3.25	2.90	-0.35	-10.6%
2007	4.17	3.75	-0.42	-10.0%
2006	3.67	3.34	-0.33	-9.1%
2005	2.99	2.75	-0.24	-8.0%
2004	2.73	2.54	-0.18	-6.6%
2003	2.32	2.19	-0.13	-5.6%
2002	2.77	2.66	-0.11	-4.0%
2001	2.25	2.19	-0.06	-2.5%
2000	2.05	2.03	-0.02	-1.2%

Source: Author's analysis of IRS administrative data.

Notes: All monetary values are expressed in trillions (T) of nominal USD. Column (A) reports observed IRA asset totals for the 2019 population. Column (B) reports counterfactual totals under the removal of all direct IRA contributions, scaled to maintain aggregate consistency. Column (C) reports the absolute difference between counterfactual and observed assets, and Column (D) reports the percentage difference relative to observed assets.

by roughly 28 percent relative to observed values. The magnitude of the effect increases sharply moving up the distribution: balances fall by 46 percent in the 80–90th percentile, nearly 58 percent in the 90–99th percentile, and more than 60 percent among the top 0.1 percent. These patterns indicate that rollovers are especially central to the accumulation of large IRA balances, but remain quantitatively important even for the broad majority of IRA holders.

Panel C shows a closely related pattern when groups are defined by net worth rather than IRA holdings. Removing rollovers reduces IRA balances by roughly 37 percent for households in the bottom 80 percent of the wealth distribution, and by more than 50 percent for households in the top decile.

In contrast, the effects of removing direct contributions are smaller and vary more across groups. As shown in Panels B and C, eliminating contributions reduces IRA balances most sharply for lower- and middle-ranked groups, with declines of roughly 20–30 percent among the bottom 80 percent of the IRA and wealth distributions. The effect attenuates steadily moving up the distribution, falling below 15 percent for the top IRA and wealth groups and below 5 percent for the very top.

Table 11: Counterfactual 2019 IRA Balances by Simulation Type & Group

Panel A: Full Population Total					
Simulation Type	Group	Observed IRA (T)	Simulated IRA (T)	Diff. (T)	% Diff.
No Rollovers	All	10.81	5.55	-5.26	-48.7%
No Contributions	All	10.81	9.07	-1.74	-16.1%

Panel B: By IRA Asset Quantile Bins					
		Observed IRA (T)	Simulated IRA (T)	Diff. (T)	% Diff.
No Rollovers	0-80%	2.487	1.786	-0.701	-28.2%
	80-90%	1.986	1.081	-0.905	-45.6%
	90-99%	4.501	1.893	-2.607	-57.9%
	99-99.9%	1.387	0.526	-0.861	-62.1%
	99.9+%	0.404	0.163	-0.241	-59.6%
No Contributions	0-80%	2.487	1.714	-0.773	-31.1%
	80-90%	1.986	1.626	-0.360	-18.1%
	90-99%	4.501	4.003	-0.498	-11.1%
	99-99.9%	1.387	1.318	-0.069	-5.0%
	99.9+%	0.404	0.390	-0.014	-3.4%

Panel C: By Net Worth Quantile Bins					
		Observed IRA (T)	Simulated IRA (T)	Diff. (T)	% Diff.
No Rollovers	0-80%	1.689	1.070	-0.619	-36.6%
	80-90%	2.224	1.168	-1.057	-47.5%
	90-99%	5.516	2.615	-2.902	-52.6%
	99-99.9%	1.204	0.583	-0.622	-51.6%
	99.9+%	0.177	0.099	-0.078	-44.1%
No Contributions	0-80%	1.689	1.326	-0.363	-21.5%
	80-90%	2.224	1.832	-0.392	-17.6%
	90-99%	5.516	4.693	-0.823	-14.9%
	99-99.9%	1.204	1.067	-0.137	-11.4%
	99.9+%	0.177	0.159	-0.018	-10.4%

Source: Author's analysis of IRS administrative data.

Notes: All dollar values are expressed in trillions (T) of nominal USD. Groups are defined using 2019 values. IRA asset quantiles are defined among individuals with positive IRA balances; wealth quantiles are defined over the full adult population. The 0-80 percent group aggregates the bottom four quintiles.

Taken together, the subgroup results reinforce the aggregate findings in Panel A: both rollovers and contributions matter for IRA accumulation, but rollovers largely dominate across the distribution. Their influence is large, persistent, and broadly shared, while the contribution channel plays a secondary and more heterogeneous role.

6 Conclusion

This paper provides the first population-level accounting, using administrative microdata, of how IRA assets accumulated between 2000 and 2019. By decomposing IRA balances into direct contributions, rollovers from employer-sponsored plans, and the compounded value of pre-2000 balances, the analysis clarifies the mechanisms underlying the rapid growth of IRAs over the past two decades. The results show that rollovers—primarily from DC plans—are the dominant driver of IRA asset growth, accounting for roughly half of accumulated balances, compared to a substantially smaller role for direct contributions.

6.1 DC plans and IRAs are an integrated system

A central implication of these findings is that DC plans and IRAs should not be understood as standalone or parallel savings vehicles. In practice, they operate as tightly linked stages of a single accumulation process, with assets routinely flowing from DC plans into IRAs over the life course. IRA balances therefore reflect not only saving behavior within the IRA system, but also the cumulative outcomes of participation in employer-sponsored retirement plans.

Treating DC plans and IRAs as separate systems obscures their close integration and can lead to misleading inferences about both household wealth dynamics and the fiscal footprint of retirement saving. Increases in IRA assets should therefore not be interpreted mechanically as evidence of expanded IRA contributions or greater reliance on IRAs as a primary savings vehicle. The results instead indicate that a substantial share of observed IRA asset growth reflects the reclassification of assets initially accumulated in DC plans and later transferred through rollovers.

Recognizing the integrated nature of the DC–IRA system has important consequences for how retirement wealth is measured, how tax expenditures are

evaluated, and how retirement policy is assessed over time.

6.2 Implications for Household Wealth Measurement

Leading measures of household wealth—including the Financial Accounts of the United States—do not treat IRAs symmetrically with other forms of retirement wealth. In the main household balance sheet (Table B101.h), DB entitlements and DC assets are reported explicitly as retirement wealth, whereas IRA assets are not. Instead, the underlying assets held in IRAs are allocated across standard asset categories such as corporate equities, mutual fund shares, and bonds. This convention carries over the Distributional Financial Accounts (DFA), which inherits the same accounting structure as the Financial Accounts.⁷

As a result, DC plans and IRAs—despite forming a continuous accumulation pipeline in practice—are treated as analytically distinct in the core infrastructure used to measure household wealth. The findings in this paper highlight the tension between this accounting convention and the underlying economic reality. This separation can lead to misleading interpretations of household wealth dynamics. Given this accounting structure in the Financial Accounts, growth in IRA balances may be read as a shift toward non-retirement saving, even though the assets remain fully embedded within the retirement system.

Taken together, these results suggest that DC plans and IRAs should be conceptualized as components of a unified retirement wealth pipeline rather than as separate asset classes. Incorporating this perspective into household wealth frameworks would improve consistency across accounts, clarify the sources of observed asset growth, and provide a more accurate picture of how retirement wealth accumulates over the life course.

6.3 Implications for Tax Expenditure Accounting

Conceptualizing DC plans and IRAs as an integrated accumulation pipeline has important implications for how retirement-related tax expenditures are measured. In both systems, tax expenditures operate through two core mechanisms: the deductibility of contributions (for non-Roth accounts) and the exclusion of investment returns from current taxation. Contributions reduce taxable income at the time they are made, while returns on accumulated as-

⁷See Federal Reserve Board, Distributional Financial Accounts, and Batty et al. (2019).

sets—interest, dividends, and realized capital gains—accrue on a tax-free basis rather than being taxed annually, as they would be in a taxable brokerage account.

Current tax-expenditure estimates produced by the Office of Management and Budget (OMB) and the Joint Committee on Taxation (JCT) evaluate the revenue effects of DC plans and IRAs separately (U.S. Department of the Treasury 2025; Joint Committee on Taxation 2024). This practice follows statutory distinctions across account types. It does not, however, reflect how tax-deferred assets are accumulated and insulated from taxation across the retirement system as a whole, particularly when assets move between DC plans and IRAs through rollovers.

Evaluating DC and IRA tax expenditures in isolation can understate their combined fiscal cost, especially under a progressive income tax schedule. Contribution limits apply separately to DC plans and IRAs, allowing tax units to deduct income across both systems. When contributions are made to both accounts in the same year, the combined deductions can shift a larger share of income out of higher marginal tax brackets than would be implied by either contribution alone. As the OMB itself notes, when tax provisions interact, the total revenue effect of the system can differ from the sum of its individual components.

In addition to contribution deductions, the insulation of investment returns also operates jointly across account types. Consider a tax unit whose retirement assets are split between DC plans and IRAs. Returns that would otherwise generate taxable interest, dividends, or realized capital gains instead accrue on a tax-free basis across both accounts. Under a progressive tax schedule, the combined tax benefit associated with insulating returns across both systems is larger than what would be obtained by estimating DC and IRA return exclusions separately and adding them together. Put differently, the effective tax subsidy associated with jointly holding assets in DC plans and IRAs exceeds the subsidy that would be inferred by evaluating each system independently.

Taken together, these dynamics imply that the combined tax advantage of DC plans and IRAs exceeds the sum of their separately estimated tax expenditures whenever a tax unit contributes to, or holds assets in, both systems. Given, as this paper has shown, that most IRA assets originate in DC plans through rollovers, and that DC plans and IRAs operate in practice as a connected accumulation system, treating them jointly in tax-expenditure accounting would therefore better reflect observed patterns of asset accumulation, and would imply a larger total fiscal cost than estimates that evaluate the two systems

separately.

As IRAs continue to grow as a share of household wealth, incorporating accumulation-based and system-level perspectives into tax-expenditure analysis will become increasingly important. The framework developed here provides a foundation for such analyses and can inform future work on retirement saving, tax policy, and the governance of long-term wealth accumulation in the United States.

References

- [1] Argento, Robert, Victoria L. Bryant, and John Sabelhaus. 2015. “Early Withdrawals from Retirement Accounts During the Great Recession.” *Contemporary Economic Policy* 33, no. 1: 1–16.
- [2] Batty, Michael, Jesse Bricker, Joseph Briggs, Sarah Friedman, Danielle Nemschoff, Eric Nielsen, Kamila Sommer, and Alice Henriques Volz. 2020. “The Distributional Financial Accounts of the United States.” In *Measuring and Understanding the Distribution and Intra/Inter-Generational Mobility of Income and Wealth*. Chicago: University of Chicago Press.
- [3] Brady, Peter, and Steven Bass. 2020. *Decoding Retirement: A Detailed Look at Retirement Account Transactions*. Washington, DC: Internal Revenue Service, Statistics of Income Division. <https://www.irs.gov/pub/irs-soi/20rpdecodingretirement.pdf>.
- [4] Brown, Jeffrey R., James Poterba, and David 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.
- [5] Bryant, Victoria L. 2020. *The Outbreak of a Tax Break: Essays on the Participation and Impact of the Saver’s Credit Across Time and Distance*. PhD dissertation, George Mason University.
- [6] Bryant, Victoria L., Sarah Holden, and John Sabelhaus. 2010. “Qualified Retirement Plans: Analysis of Distribution and Rollover Activity.” In *Proceedings of the Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, vol. 103, 81–90. Washington, DC: National Tax Association.
- [7] Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. “Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States.” *Quarterly Journal of Economics* 129, no. 4: 1553–1623.
- [8] Congressional Budget Office. 2025. *The Budget and Economic Outlook: 2025 to 2035*. Washington, DC: Congressional Budget Office. <https://www.cbo.gov/publication/60870>.
- [9] Copeland, Craig. 2015. “Individual Retirement Account Balances, Contributions, and Rollovers, 2013; With Longitudinal Results 2010–2013: The

- EBRI IRA Database.” *EBRI Issue Brief* No. 414. Washington, DC: Employee Benefit Research Institute.
- [10] Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. 2020. “Heterogeneity and Persistence in Returns to Wealth.” *Econometrica* 88, no. 1: 115–170.
- [11] Goodman, Lucas, Jacob Mortenson, Kathleen Mackie, and Heidi R. Schramm. 2021. “Leakage from Retirement Savings Accounts in the United States.” *National Tax Journal* 74, no. 3: 689–719.
- [12] Holden, Sarah, and Daniel Schrass. 2019. *The Role of IRAs in U.S. Households’ Saving for Retirement, 2019*. ICI Research Perspective 25, no. 10. Washington, DC: Investment Company Institute. <https://www.ici.org/doc-server/pdf:per25-10.pdf>.
- [13] Horneff, Vanya, Raimond Maurer, and Olivia S. Mitchell. 2018. “How Persistent Low Expected Returns Alter Optimal Life-Cycle Saving, Investment, and Retirement Behavior.” NBER Working Paper No. 24311.
- [14] Horneff, Vanya, Raimond Maurer, and Olivia S. Mitchell. 2023. “Do Required Minimum Distribution 401(k) Rules Matter, and for Whom? Insights from a Lifecycle Model.” *Journal of Banking & Finance* 154: 106941.
- [15] Joint Committee on Taxation. 2024. *Estimates of Federal Tax Expenditures for Fiscal Years 2024–2028* (JCX-48-24). Washington, DC: U.S. Congress, Joint Committee on Taxation. <https://www.jct.gov/getattachment/765709fb-9a4b-430a-8f9e-4d342ec97f7e/x-48-24.pdf>.
- [16] Mortenson, Jacob A., Heidi R. Schramm, and Andrew Whitten. 2019. “The Effects of Required Minimum Distribution Rules on Withdrawals from Traditional IRAs.” *National Tax Journal* 72, no. 3: 507–542.
- [17] Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. 2018. “Distributional National Accounts: Methods and Estimates for the United States.” *Quarterly Journal of Economics* 133, no. 2: 553–609.
- [18] Sabelhaus, John. 2000. “Modeling IRA Accumulation and Withdrawals.” *National Tax Journal* 53, no. 4: 865–875.
- [19] Saez, Emmanuel. 2016. *Statistics of Income Tabulations: High Incomes, Gender, Age, Earnings Split, and Non-Filers*. SOI Working Paper. Washington, DC: Internal Revenue Service. <https://www.irs.gov/pub/irs-soi/16rpsaeztabulations.pdf>.

- [20] Saez, Emmanuel, and Gabriel Zucman. 2016. “Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data.” *Quarterly Journal of Economics* 131, no. 2: 519–578.
- [21] Saez, Emmanuel, and Gabriel Zucman. 2020. *Trends in U.S. Income and Wealth Inequality: Revising after the Revisionists*. NBER Working Paper No. 27921.
- [22] Smith, Matthew, Owen M. Zidar, and Eric Zwick. 2023. “Top Wealth in America: New Estimates under Heterogeneous Returns.” *Quarterly Journal of Economics* 138, no. 1: 515–573.
- [23] Statistics of Income. 2018. *Individual Retirement Arrangements, Tax Year 2018*. Washington, DC: Internal Revenue Service. <https://www.irs.gov/pub/irs-pdf/p5498.pdf>.
- [24] Stuart, Ellen, and Victoria L. Bryant. 2024. “The Impact of Withdrawal Penalties on Retirement Savings.” *Journal of Public Economics* 232: 105083.
- [25] U.S. Department of Labor, Employee Benefits Security Administration. 2022. *Private Pension Plan Bulletin: Abstract of 2019 Form 5500 Annual Reports*. Washington, DC: U.S. Department of Labor. <https://www.dol.gov/agencies/ebsa/researchers/statistics/retirement-bulletins/private-pension-plan-bulletins-abstract-2019>.
- [26] U.S. Department of the Treasury. 2023. *Managing the Nation’s Finances: The Unsustainable Fiscal Path*. Washington, DC: Bureau of the Fiscal Service. <https://www.fiscal.treasury.gov/reports-statements/financial-report/mda-unsustainable-fiscal-path.html>.
- [27] U.S. Department of the Treasury, Office of Tax Analysis. 2025. *Tax Expenditures: Fiscal Year 2027*. Washington, DC: U.S. Department of the Treasury. <https://home.treasury.gov/system/files/131/Tax-Expenditures-FY2027.pdf>.