Who Benefits from Retirement Saving Incentives in the U.S.? Evidence on Gaps in Retirement Wealth Accumulation by Race and Parental Income*

Taha Choukhmane Jorge Colmenares Cormac O'Dea
Jonathan Rothbaum Lawrence D.W. Schmidt

August 2024

Abstract

U.S. employers and the federal government devote over 1.5% of GDP annually toward promoting defined contribution (DC) retirement saving. Using a new employer–employee linked dataset covering millions of Americans, we show that this system of saving incentives benefits White workers and those with richer parents more than their similar-income coworkers who are Black or Hispanic or from lower-income families. Breaking the link between contribution choices and saving subsidies—through revenue-neutral reforms—can close the gaps in DC wealth between Black and White workers, between Hispanic and White workers, and between those with the richest and those with the poorest parents by close to a third.

*Choukhmane: MIT Sloan School of Management, tahac@mit.edu; Colmenares: Harvard University, jcolmenares@fas.harvard.edu; O'Dea: Yale University, cormac.odea@yale.edu; Rothbaum: US Census Bureau, jonathan.l.rothbaum@census.gov; Schmidt: MIT Sloan School of Management, ldws@mit.edu. We are very grateful to Jackson Howell and Cashua Wang who provided outstanding research assistance. We thank Franca Glenzer, Luigi Guiso, Pascal Noel, and Natasha Sarin for illuminating discussions of this paper. We also thank James Choi, Amir Kermani, Debbie Lucas, Olivia S. Mitchell, Jim Poterba, John Sabelhaus, Jason Seligman, Angelino Viceisza, and Stephen Zeldes for extremely helpful conversations and comments. We thank Cristina Tello Trillo for her assistance in building the bridge from the Form 5500 data to the individual-level data. We are grateful to numerous seminar and conference participants for their helpful comments and suggestions. This report is released to inform interested parties of ongoing research and to encourage discussion. Any views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau. The data in this paper have been cleared by the Census Bureau's Disclosure Review Board release authorization numbers CBDRB-FY22-SEHSD003-001, CBDRB-FY22-SEHSD003-017, CBDRB-FY22-SEHSD003-033, CBDRB-FY23-SEHSD003-043, CBDRB-FY23-0494, and CBDRB-FY24-0151. The research reported herein was performed pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement and Disability Research Consortium through the Michigan Retirement and Disability Research Center Award RDR18000002. The opinions and conclusions expressed are solely those of the author(s) and do not represent the opinions or policy of SSA or any agency of the Federal Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the contents of this report. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply endorsement, recommendation or favoring by the United States Government or any agency thereof. We are very grateful to the MIT Institute Community & Equity Office, the MIT Sloan School of Management, the Yale University Tobin Center for Economic Policy, and the Yale University Arthur M. Okun Research and Policy Fund for co-funding this work.

1 Introduction

Every year, the equivalent of 1.5% of U.S. GDP is devoted to encouraging contributions to retirement savings plans such as 401(k) and 403(b) accounts.[1] Around 100 million Americans have access to such plans through their employers, and these accounts offer an attractive vehicle for long-term saving. Contributions are taxed favorably, and over 80% of employers further subsidize savers by matching their employees' contributions (Arnoud et al., 2021). This institutional design, therefore, rewards those who can, and do, save more for retirement. Employees who do not contribute receive neither tax benefits nor employer-matching contributions.

In this paper, we ask how much retirement savings incentives contribute to racial and inter-generational wealth inequality in the United States.[2] To address this question, we create a unique dataset linking (1) the retirement contributions and withdrawals of millions of Americans from their federal tax filings, (2) demographic information from survey responses to 10 years of the American Community Survey (ACS), and (3) hand-collected data on the characteristics of over 6,000 DC retirement plans, covering approximately 40 million employees from employers' Form 5500 fillings. We use this new dataset to assess the distributional impacts of retirement saving incentives by race and parental income.

To do so, we first measure contributions to DC accounts (as groups that contribute more receive more tax and matching benefits) and measure withdrawals (as groups that take more early withdrawals forgo long-term tax benefits and face tax penalties). We then build a model to translate the observed saving differences into differences in lifetime wealth accumulation and to study the implications of budget-neutral reforms to tax and employer saving subsidies for wealth inequality. The model also allows us to study retirement saving incentives in the context of the broader tax and retirement system and, importantly, taking into account Social Security benefits.

We divide our analysis into three parts. First, we document significant differences in retirement contributions by race and parental income. As shown in Figure 1, Black and Hispanic workers with access to a 401(k) or a 403(b) plan contribute approximately 40% less than White workers: respectively, 1.8 p.p. and 1.6 p.p. of salary less (panel (a)). Workers with parents in the bottom and middle quintiles of the income distribution save 1.9 and 1.3 p.p. of salary less, respectively, than those with parents in the top quintile (panel (b)). These

[1] In 2021, federal tax expenditures on defined contribution (DC) retirement accounts amounted to \$119 billion (US Department of the Treasury, 2023). In 2021, private sector employers contributed more than \$212 billion into these accounts (Department of Labor, 2023)—mainly in the form of matching contributions.

[2] We focus on three racial and ethnic groups: non-Hispanic Whites, non-Hispanic Blacks, and Hispanics, who together make up 93% of the individuals in our sample. We will often use "race" to refer to both race and ethnicity, "White" to refer to non-Hispanic White, and "Black" to refer to non-Hispanic Black.

saving differences mean that median White earners receive more than double the matching benefits of their Black and Hispanic counterparts (panel (c)), while those with parents in the top income quintile receive three times the matching benefits of those with parents in the bottom quintile (panel (d)). These gaps are partially accounted for by income differences between groups; however, approximately half of the gap remains when comparing workers of the same age and income. Sizable contribution gaps persist even when accounting for other individual-level characteristics—such as education, occupation, tenure, and employer. Comparing workers with similar characteristics, and in particular similar ages and incomes, is important because these workers accrue similar Social Security entitlements. Therefore, residual differences in subsidies received are not directly mitigated by other aspects of the broader retirement system.

In the second part of the paper, we turn to early withdrawals. Among those with at least $1,000 in recent contributions, 12% of White savers, 15% of Hispanic savers, and 23% of Black savers take an early withdrawal from their retirement accounts each year. Similarly, 9% of savers with parents in the top income quintile take an early withdrawal, against 16% of those with parents in the bottom quintile. Most of the differences in early withdrawals across groups remain when comparing savers with similar ages and incomes. Savers who take an early withdrawal often forgo tax benefits and can face tax penalties; this, as we will show, is important in shaping the lifetime distributional impact of the DC federal tax expenditure.

In the third part of the paper, we develop a micro-simulation model that uses our data on flows of earnings, employee contributions, employer matches, and early withdrawals to compute measures of retirement wealth. We use this model for three purposes. First, we evaluate the lifetime distributional effect of savings subsidies. The value of these for any individual, especially the benefits from favorable taxation, depend on the entire trajectory of their incomes, contributions, and withdrawals, and so cannot be calculated directly from the data. We find, for instance, that the median White earner receives more than three times the simulated lifetime tax benefits of the median Black earner (Figure 1(c)). Second, the model allows us to incorporate taxation and Social Security benefit accrual explicitly and therefore evaluate the distributional effect of retirement incentives in the context of the broader U.S. retirement system. Finally, we use the model to assess the effect of revenue-neutral reforms that would redistribute employer-match dollars within each firm and federal tax expenditures across the population so that they are distributed 1) proportionally to earnings but 2) independently of workers' own contribution choices. We estimate that, for median earners, the mechanical impact of these reforms would close each of the gaps between Black and White workers, Hispanic and White workers, and those with the richest and poorest parents by close to a third. This estimated relative change in the gaps would be similar if workers reduced their contributions in response to the removal of match and tax incentives (assuming an elasticity of employee savings to incentives calibrated to match the upper end of empirical estimates).

The first branch of the literature to which we contribute is that on the design and impact of retirement savings incentives. A large literature has studied the effect of tax and employer matching incentives on private savings and finds small to insignificant behavioral responses to these incentives (see Choi (2015) and Friedman (2015) for reviews). Retirement incentives seem to mainly shift the location of existing savings rather than creating new savings (Chetty et al., 2014; Choukhmane and Palmer, 2023). In contrast to this large literature, the mechanical effect of these incentives has received less attention. While the distributional impact of the federal tax expenditure by income has been studied (Burman et al., 2004; Brown et al., 2022), quantitative evidence on the full effect of retirement saving subsidies (including employer subsidies) has been limited due to a lack of systematic data on retirement plan characteristics. The fact that we directly observe the match schedule means that we can precisely measure the contribution of employer subsidies to retirement wealth. In doing so, we are able to quantify the total contribution of these savings supports to retirement wealth. We find their contribution to be large: at the bottom of the lifetime earnings distribution, they account for 40% of DC wealth, while at the top, they account for 50% of (a much larger base amount of) DC wealth. In addition to improving the measurement of saving subsidies, we contribute to this literature by bringing in richer demographic characteristics. We show how the receipt of subsidies differs by race and parental income for individuals who have the same lifetime income and thus likely face similar tax incentives and expect similar Social Security benefits. The disparities in subsidy receipt among co-workers of similar ages, incomes, and geographic locations are less likely than disparities by income to be mitigated by other features of the broader retirement system and Social Security formula.

The second branch of the literature to which we contribute is that concerned with race, earnings, and wealth in the U.S. The gap between Black and White wealth is large (Oliver and Shapiro, 1989; Darity and Nicholson, 2005), stable since the 1980s (Derenoncourt et al., 2022), and cannot be fully accounted for by earnings differences (Blau and Graham, 1990; Barsky et al., 2002; Altonji and Doraszelski, 2005; Kuhn et al., 2020). There is also a large wealth gap between Hispanic and non-Hispanic White individuals, with the latter being approximately four times wealthier (Sabelhaus and Thompson, 2021). Our contribution to this literature is to study one channel that contributes to both earnings and wealth inequality by race. On the earnings side, our contribution is to measure an often-unmeasured component of earnings—the employer match—which gives a compensation premium to those who save more. In doing so, we contribute to a rich literature on racial earnings inequality.[3] On the wealth side, Derenoncourt et al. (2022) emphasize that differences in rates of return are the dominant factor shaping the lack of racial wealth convergence over the past 30 years—the period in which DC accounts have emerged as the main vehicle for private retirement savings. Our results shed light on an important mechanism generating such differences in rates of return across racial groups, even holding portfolio risk constant: differences in the take-up of employer match and tax incentives.

The differences by race in broad measures of wealth referenced above are also seen in retirement wealth (Hou and Sanzenbacher, 2021; Francis and Weller, 2021; Viceisza et al., 2022; Wolff, 2023). Closer to our paper, a number of studies have examined racial differences in savings rates. While studies using administrative retirement record-keeper data from one firm (Kuan et al., 2015) or a small number of firms (Ariel/AON Hewitt, 2009) find differences in saving rates by race, the literature using representative survey data has typically found that these patterns no longer hold after accounting for income differences.[4] However, surveys can suffer from significant measurement error, as shown by studies comparing self-reported measures of access and contributions to DC plans with the respondents' tax records.[5] Administrative tax data combines the strength of both data sources: they have better coverage and are more representative than record-keeper data while suffering from less measurement and misreporting error than surveys. Using such data yields a different result than that seen in prior studies: racial gaps in contributions persist even among co-workers and after controlling for individual and household incomes. Other recent papers also use tax data to study differences in participation in retirement plans (Yogo et al., 2023) and penalized early withdrawals (Coyne et al., 2022). We add to this literature, first by measuring contributions (in addition to participation and withdrawals), second by observing racial self-identification directly rather than imputing it, and third by bringing in detailed information on plan features—which are important determinants of differences in contribution rates.

Our paper contributes to a third, closely related literature that evaluates the disparate

[3]Altonji and Blank (1999) offer a comprehensive review of studies to that date, and Bayer and Charles (2018), Chetty et al. (2020), and Derenoncourt and Montialoux (2021) provide more recent evidence.

[4]A review of the evidence by Darity Jr et al. (2018) concludes that "the finding advanced in peer-reviewed articles in economic journals is clear: there is no evidence that black Americans have a lower savings rate than white Americans once household income is taken into account" We show a similar result using data from the Survey of Consumer Finances (SCF), albeit with wide confidence intervals.

[5]Dushi and Iams (2010) found that 24% of private-sector and 36% of public-sector respondents to the 2006 Survey of Income and Program Participation (SIPP) misreported making a tax-deferred DC contribution relative to their W2 records (with both false positives and negatives being common). Similarly, Dushi and Honig (2015) found that the average absolute difference between annual DC contributions reported in the Health and Retirement Study (HRS) and respondents' W2 records was approximately 1.5 times larger than the mean DC contribution in the W2s. Likewise, Bee and Mitchell (2017) show that survey respondents vastly underreport their DC plan withdrawals.

impact of policies by race. This includes research examining racial disparities in welfare programs (Darity and Myers, 1983, 1987), unemployment insurance (Kuka and Stuart, 2021; Skandalis et al., 2022), mortgage access (Myers Jr, 1995; Ross and Yinger, 2002; Bhutta and Hizmo, 2021), housing returns (Kermani and Wong, 2021), property tax assessments (Avenancio-León and Howard, 2022), and financial aid for college (Levine and Ritter, 2022). In the context of retirement plans, Brown (2021) argues that the design of retirement incentives favors activities that are more likely to be carried out by White Americans (retirement saving) and penalizes activities that are more likely to be carried out by Black Americans (early withdrawals). More broadly, Hamilton and Darity (2017) argue that "if the existing federal asset-promotion budget were allocated in a more progressive manner, federal policies would go a long way toward eliminating racial disparities and building an inclusive economy for all Americans." We quantify how much changing a major component of the U.S. asset-promotion budget, namely the design of retirement savings subsidies, could affect racial and intergenerational wealth inequality.

The fourth branch of literature that we contribute to is that on intergenerational persistence in wealth. The correlation in wealth across generations has been well documented (Charles and Hurst, 2003). Recent work emphasizes the importance of heterogeneity in rates of return for cross-sectional wealth inequality (Fagereng et al., 2020). Our paper draws a link between these two phenomena. While it has long been known that the rich save more (Dynan et al., 2004), we show that the *children* of the rich save more conditional on their own earnings. The saving in question here is, by virtue of matching, one with an extraordinary rate of return. This correlation between the resources of one generation and the rates of return of the next will directly contribute to intergenerational persistence in wealth. This channel also relates to a theme that has been emphasized in the literature on wealth gaps by race in the U.S. Chiteji and Hamilton (2002) and Charles and Hurst (2002) highlight the role of the family in savings decisions and the direction of intergenerational transfers: Black individuals are both more likely to provide financial support to their parents and less likely to receive support from their parents than White individuals. This is consistent with our finding that, even among those with similar *individual* characteristics, accounting for differences in parental and household resources reduces gaps in contributions by race.

The paper proceeds as follows. Section 2 discusses the institutional background. Section 3 introduces our new employer–employee linked data set. Section 4 gives our results on gaps in retirement saving rates by race and parental income. Section 5 turns to early withdrawals. Section 6 uses our data and a microsimulation model to study the distributional impact of the savings patterns we observe and the retirement saving subsidies we study on measures of wealth at retirement. Section 7 concludes.

2 Institutional Background

Defined contribution (DC) plans have become the dominant vehicle through which Americans save for retirement. Sixty percent of U.S. civilian workers now have access to an employer-sponsored DC plan (Myers and Topoleski, 2020). Participants in these plans can make pretax contributions to their accounts (up to an annual maximum employee contribution of $20,500 in 2022), thereby deferring income taxes to when they retire and when they will (likely) face lower tax rates. In addition to the advantages that this deferral brings, dividends and capital gains are untaxed provided that they remain in the account. Wealth held in DC plans is illiquid: participants generally face tax penalties on withdrawals made before the age of 59.5, though some plans permit borrowing against existing DC balances.

DC plans provide substantial flexibility and discretion to participants in deciding how much to save. This structure contrasts substantially with defined benefit (DB) plans, in which employees typically only choose whether to participate, and *employer* contributions do not depend on any choice that the *employee* makes. The shift away from DB toward DC plans in recent years moves considerable risk related to financing retirement income from employers to employees.[6] Whereas traditional pension plans insure against mortality risk and the lion's share of risks associated with fluctuations in investment returns, DC plans force households to self-insure against these risks.

In the vast majority of DC plans, employers match employee contributions at some rate up to a cap, meaning that the amount contributed by the employer depends on how much the employee chooses to save. Appendix Figure A.1 shows the full set of matching schedules in our data, the construction of which is described in the next section.

In contrast, the rules governing both employee and employer Social Security contributions are more rigid. Social Security payments are financed via non-discretionary FICA payroll tax contributions from both employers and employees on each dollar of labor earnings up to a taxable maximum. Social Security benefits to workers are then computed as a function of the worker's earnings history. These benefit amounts are progressive, implying that low-income workers generally receive larger benefit payments per dollar of payroll tax contributions than higher-income workers in the same cohorts. The progressivity of Social Security means that measures of wealth that include it display less inequality and narrower racial gaps than measures without (Catherine and Sarin (2023), Sabelhaus (2023))).

[6]Only a quarter of civilian workers now have access to a DB pension (Myers and Topoleski, 2020), a share that continues to fall. DC plans are becoming, alongside Social Security, one of the largest sources of income in retirement. Devlin-Foltz et al. (2016) show that over the past 30 years the dynamics of retirement wealth have had a moderating impact on overall wealth inequality. They also find, however, that DC wealth is more concentrated than DB wealth.

As DC plans become more dominant and DB coverage recedes, there is greater scope for individuals' decisions to affect retirement wealth, and employer plan design can amplify the implications of these decisions for wealth inequality. Endogenous DC participation also implies that the benefits paid to employees in the form of matching contributions will not be equally distributed across workers, even among those with identical earnings (and Social Security entitlements). To study the interplay between individual saving decisions and firm matches, we therefore need data that contain both the saving decisions made by individuals and the full match schedules offered by their employers.

3 Data

We build a new dataset linking administrative data on retirement saving and demographics of a large sample of U.S. employees with a newly constructed data set on employer-sponsored retirement plan characteristics. We discuss these data in sections 3.1 and 3.2, respectively. Section 3.3 defines the main outcomes we study, and section 3.4 defines our samples.

3.1 Employee data

The basis for our analysis is all individuals ever observed in the 2008 to 2017 waves of the American Community Survey (ACS).[7] We link ACS respondents to other administrative data using protected identification keys (PIKs).[8] With this method, 90%–94% of ACS respondents are successfully assigned a PIK in any given year (Ferrie et al., 2021).[9] Next, we link ACS respondents with their 1040, W-2, and 1099-R tax filings. The ACS provides individuals' race, year, age, education, gender, occupation, and location at the time of the survey. The 1040 and W-2 filings provide other socioeconomic and demographic indicators, including family structure, employer identification number (EIN), employment tenure, spousal income, and intergenerational linkages (for example, parental income). Appendices A.1 and A.2

[7]From 2005 to 2019, the ACS averaged over 3.2 million individuals surveyed, including a sample expansion from 2010 to 2012. Refer to `https://www.census.gov/acs/www/methodology/sample-size-and-data-quality/` for more information on the ACS sample and response rates over time; accessed 7/12/2024.

[8]PIKs are assigned by a probabilistic matching algorithm that compares the characteristics of records in Census, survey, and administrative data to those in a reference file constructed from the Social Security Administration Numerical Identification System and other federal administrative data. PIKs correspond one-to-one with SSNs and so allow us to link individuals over time and across data sources. For more information, see Wagner and Layne (2014).

[9]As noted in Bond et al. (2014), there is some selection into linkage, for example by age, race, and citizenship status. However, we do not believe that the magnitudes of these differences will bias our estimates substantially. For example, in 2010, the linkage rate for Black ACS respondents was 91.4%, compared to 93.5% for White respondents.

provide detailed overviews of our data build and variable construction, respectively.

3.2 Employer retirement plan data

All employers must submit an annual regulatory form (Form 5500) on their U.S. retirement plans to the federal government. Plans with over 100 participants provide narrative descriptions of plan characteristics, including match schedules, vesting schedules, and auto-features. We create a data set by extracting these descriptions from the original free-form text.[10] We do this for the largest 4,800 plans in the U.S. and a random sample of 1,000 smaller plans. These employers cover a substantial portion of the U.S. population; in 2017, 37 million employees were eligible for these large plans, constituting 55% of employees with access to private and nonprofit-sector DC retirement plans.[11] Appendix A.1.3 provides further details. These plan-level data, further detailed in Arnoud et al. (2021) and Choukhmane et al. (2023), include information on vesting schedules, auto-enrollment, and—crucially for our question—employer match schedules. These match schedules are typically concave functions of employee contribution rates, often linear up to a threshold (Appendix Figure A.1).

We match the retirement plan data with our employee data using numeric identifiers such as the EIN, telephone number, name, and address fields. We impose two restrictions on which employers we retain in the linked data (see Appendix A.3 for further details). First, so that a link to a firm is sufficient for a link to a particular match, the employer must use the same matching formula for all employees. Second, as a guardrail against incorrect matches or inaccurate formulas, we filter out plans for which there is too large of a discrepancy between the ratio of employer to total contributions from Form 5500 and an analogous measure computed from applying our matching rules to the linked W-2 data.

3.3 Retirement savings outcomes

Our four primary measures of saving and withdrawals are: i) **Employee contributions**: deferred compensation reported in Box 12 of the W-2 tax form. This amount generally corresponds to contributions to an employer-sponsored contribution plan (such as a 401(k)). We define the employee contribution *rate* as a percentage of salary using the ratio of the employee contribution reported in Box 12 to the sum of the taxable wage reported in Box 1 of the W-2 form and the Box 12 employee contribution. ii) **Employee plus employer matching contributions**: the sum of the employee contribution and the employer match

[10] Refer to `https://www.dol.gov/agencies/ebsa/about-ebsa/our-activities/public-disclosure/foia/form-5500-datasets`.

[11] Note that while employers include firms, hospitals, non-profits, and other non-firm employers, "firm" and "employer" are used interchangeably throughout this paper.

contribution we can calculate using match formulas hand-collected from Form 5500 filings, also expressed as a percentage of salary unless indicated otherwise. iii) **Participation**: a dummy equal to one if the individual makes a positive contribution to a retirement savings plan. iv) **Early withdrawal**: a dummy equal to 1 if an individual between the ages of 25 and 54[12] has a distribution from a retirement account reported in tax Form 1099-R. Appendix A.2.1 gives further details on variable construction.

3.4 Samples

We define two main samples: i) a "Form 5500 sample" of individuals for whom we have employer-level retirement plan data, and ii) a "Parent-Form 5500 sample" of individuals who can be linked to their parents. We restrict both samples to those between ages 24 and 59 and exclude those with very low labor income. Specifically, we impose that the sum of nominal W-2 Box 1 wages and deferred compensation exceeds $8,000, which corresponds to roughly 20 hours per week at the current Federal minimum wage. We also exclude those who have zero Box 1 wages as well as those for whom we have missing data on individual-level characteristics in the ACS.

The "Form 5500 sample" corresponds to the subset of ACS respondents who have access to an employer-sponsored DC plan and for whom we have collected retirement plan information from Form 5500 filings. This sample has approximately 1,722,000 unique individuals. This serves as the primary sample in most of our analysis. The "Parent-Form 5500 sample" corresponds to the subset of the "Form 5500 sample" for which we can link respondents to their parents and who are themselves in the 1978-1992 birth cohort. This sample is, on average, younger and has approximately 471,200 unique individuals. We weight estimates so as to be nationally representative of the population of US workers with employers offering retirement plans with at least 100 participants. Appendices A.3 and A.4 provide more details on the different samples, weighting, and representativeness of our estimates, respectively.

To assess whether selection into the Form 5500 linked employer-employee sample matters for our findings, we compare summary statistics and key results from our baseline sample to a broader sample of all ACS respondents who satisfy our sample restrictions (of which approximately 9,595,000 unique individuals work at an employer offering a DC plan). Appendix Table A.1 shows that the observable characteristics and contribution behavior of workers in our "Form 5500 sample" are broadly similar to those in the full ACS for respondents working at an employer offering a DC plan.

[12] After this age, withdrawal upon separation from an employer do not incur a tax penalty.

4 The Distribution of Retirement Contributions

Groups that contribute more to retirement accounts receive more from employer matching and tax benefits. In this section, we document differences by race and parental income in retirement contribution rates and the interplay with institutional saving incentives. First, we find that baseline employee contribution gaps are large and that they are amplified by employer matching. While differences in income and age account for about half of the gaps, sizable differences remain even among workers in the same firm with similar individual-level characteristics. Second, given the notable differences in parental income by race, we consider the interplay between contribution gaps by race and parental income. Finally, we conclude by discussing the role of other plausible correlates of contribution gaps.

4.1 Contribution gaps by race and parental income

Contribution gaps are large and amplified by employer matches. Figures 1(a) and 1(b), discussed in the introduction, plot average DC contribution rates by race and parental income groups, respectively. Among workers with access to a DC plan, White employees contribute an average of 4.2% of their salary to their DC plan, while Black (Hispanic) workers contribute 2.4% (2.6%). Those with parents in the top income quintile contribute 3.9%, while those with parents in the middle and bottom contribute 2.6% and 2.0%, respectively. These differences are substantial relative to the overall average in our sample (5.7%) and the average personal saving rate in the U.S. In Appendix Figure A.11, we show that these contribution gaps are also broad-based, focusing on race for brevity: contribution gaps persist (and get larger in terms of percentage points of earnings) along the income distribution, throughout the life cycle, and across all education groups.

Employer matching amplifies the effect of these differences in employee contribution rates. As shown in Figure 1(a), once employer matches are factored in, the Black-White DC contribution gap increases by 0.7 p.p., and the Hispanic-White increases by 0.6 p.p of salary. Figure 1(b) shows that the difference between the top and bottom quintiles of parental incomes increases by 0.8 p.p., while the top-middle grows by 0.5 p.p of salary. As a result, matching increases the gaps in contributions by roughly 40% across all these groups. Because those who earn more also contribute a larger share of their earnings, employer matching dollars are distributed more unequally than labor earnings. As shown in Figure 1, gaps in matching receipt by race (panel c) and by parental income (panel d) are significantly larger than gaps in earnings.[13]

Quantifying gaps conditional on characteristics. A natural question is the extent to which these sizeable gaps persist conditional on observable characteristics. To answer this, we estimate linear models for worker i of the form:

$$y_i = \alpha + \beta group_i + X_i^{'}\delta + \epsilon_i \tag{1}$$

in which we progressively add observable characteristics.

Equation (1) is fairly standard in the literature on wage gaps (see, e.g., Cahuc et al., 2014, Ch. 8). When studying racial gaps, $group_i$ is a race indicator for Black, Hispanic, Asian, Native American, Pacific Islander, or Two or More racial groups, and the omitted category is White workers. Our analysis restricts attention to the three largest racial groups in the U.S.; therefore the coefficients we report are the elements of β on the dummies for Black and Hispanic group membership. For the intergenerational analysis, $group_i$ is an indicator for the parental income quintile bins, and the top bin serves as the omitted category.

Figures 2(a) and 2(b) plot estimated values of $\hat{\beta}$—the difference in conditional means—which we also express as a percentage difference relative to the omitted category under each bar. The first set of bars (i.e., "Raw") corresponds to the univariate version of equation (1) without any X_i, and so represents the raw differences in the outcome. The subsequent sets of columns add, successively, additional co-variates. In each step of this "regression cascade," the addition of a new variable shrinks the gap only if both the addition incrementally predicts the outcome variable and there is a correlation between the variable and group membership. While our baseline approach, shown in Equation (1), relies on additive separability in the effect of the different mediating factors, we show in Appendix B.2 (see Figure A.6) that our results are quantitatively similar using a fully nonparametric approach.

Which mediating factors *should* we include? In an analysis such as ours, whether to partial out the influence of a given mediating factor depends on the question at hand.[14] Our objective is to study the distributional effect of retirement saving subsidies. These subsidies do not exist in isolation, and a challenge to the interpretation of unconditional differences across groups is that this mechanical impact of employer matching and retirement

[13] While our sample differs somewhat from that in other studies (e.g., we drop workers with very low earnings and focus on firms with DC plans) our results are largely consistent with findings in the most recent literature on the gaps in labor income across races (Bayer and Charles, 2018; Derenoncourt and Montialoux, 2021).

[14] See, for example, Neal and Johnson (1996), Lang and Manove (2011), and Carneiro et al. (2005) for discussions of test score and education controls in racial wage gap regressions.

tax benefits could be mitigated by other aspects of the broader U.S. retirement system. For instance, groups that receive less in employer matching and tax benefits might also receive a greater return (in terms of Social Security entitlement), for each dollar of payroll taxes paid, due to the program's progressivity.

Therefore, in addition to documenting raw saving differences, we emphasize two different sets of results with different mediating variables. The first set of results partials out the mediating effect of age and labor income since the accrual of Social Security benefits depends directly on one's age and earnings. Gaps in saving, holding earnings constant, are therefore less likely to be undone by differences in Social Security benefits. Our second set of results further includes location, education, occupation, gender, tenure, and firm fixed effects. These factors do not enter directly into the Social Security formula but could interact with other tax and benefit programs. It is also of independent interest to highlight the extent to which gaps (conditional on age and income) can be accounted for, in a statistical sense, by differences across groups along these other dimensions.

Age and earnings account for about half of the contribution gaps. The second set of bars in Figures 2(a) and 2(b) shows that contribution gaps are minimally affected by the inclusion of age and year in the regression. While older workers save more, there are only modest differences in the age distribution across racial groups and even smaller differences by parental income, so including these in X_i does little to the estimated gap.

Next, we include indicators for workers' own income deciles, constructed by sorting workers into deciles of labor income within each calendar year and age bin. We find that approximately half of the gap in contributions by race and slightly more than half of the gap by parental income can be accounted for by differences in labor income. This result reflects two forces. First, there are sizable and well-documented gaps in income across racial groups and between individuals with different levels of parental income. Second, high-income workers contribute more to DC accounts (Dynan et al., 2004). In Appendix Figure A.9, we visualize these two results more directly, focusing on gaps by race.

Differences in contributions remain quantitatively large even after accounting for differences in age and labor income: the Black-White gap is 1.1 p.p., the Hispanic-White is 0.96 pp, and the gap between the top and bottom parental income quintile is 1.1 p.p. of earnings. To put these numbers in context, these *residual* contribution gaps (conditional on age and earnings) are around half the *raw* difference between a high school and college graduate (2.3 pp). See Appendix Figure A.13 for further details.

Gaps remain even among co-workers with similar earnings, gender, occupation, tenure, education, and location. The final bars in Figures 2(a) and 2(b) report residual gaps after accounting for a number of additional differences in the individual characteristics of the workers in our sample. In particular, we construct, and include in the regression, indicators for gender, four different educational attainment levels, occupation codes, and four different employment tenure levels.[15] In addition, we absorb fixed effects for county of residence and the EIN of the employer. These EIN fixed effects are identified off of coworkers and absorb unobserved drivers of average contribution rates across firms, which includes plan characteristics such as the match schedule. We discuss potential rationales for each characteristic to impact savings rates further in Appendix B.1.

Accounting for the mediating role of this rich set of individual characteristics further reduces the Hispanic-White gap by 48%, the Black-White gap by 18%, and the gap between the top and bottom parental income quintile by around 28%. However, residual gaps displayed in the rightmost part of each cascade remain quantitatively large even when comparing co-workers who are similar on a broad range of characteristics (e.g., income, education, and occupation). The residual Black-White combined employee and employer contribution gap is close to 1 pp, and the Hispanic-White gap is close to 0.5 p.p. of earnings. The difference between those at the top of the parental income distribution and those at the bottom is approximately 0.75 p.p. of earnings. We show results under an alternative ordering of the cascade (including firm fixed effects before introducing income) in Appendix Figure A.4.

Intensive margin savings (not participation) differences account for most of the residual gaps. We next examine the roles of the extensive and intensive margins in driving contribution gaps. Panels (c) and (d) of Figure 2 show gaps in contributions among those who participate; panels (e) and (f) show gaps in the probability of participating. Differences along both margins contribute to the raw gaps, while differences in the gaps that remain after accounting for the mediating role of individual characteristics are largely driven by the intensive margin.

Taking stock and comparing our results with survey data. We document significant differences in DC retirement contributions even for workers with similar ages, earnings and other individual characteristics. In Appendix C, we reproduce our baseline analysis on contribution differences by race using data from the Survey of Consumer Finances, the gold standard source of survey information on wealth in the U.S. Consistent with findings

[15]In these graphs, we show results adding groups of regressor at a time; we detail the incremental impact on gaps from sequentially adding each individual characteristic in Appendix Figure A.3.

from previous research (see Darity Jr et al. (2018) for a review), we find estimates of gaps are imprecisely estimated. It is hard to make precise statements about the Black-White contribution difference once we account for differences in income. Confidence intervals are large and not only overlap with zero but also come close to and even overlap with some of our estimates using administrative data. This suggests that survey data that has been typically used to study this question might be under-powered to detect (even sizeable) differences in retirement contribution by race. For example, in the 2013 wave of the SCF, only 167 Black and 94 Hispanic respondents reported having access to a DC plan through their employer; in contrast, our sample includes an average of 18,100 Black and 19,400 Hispanic workers with access to a DC plan every year. This exercise highlights the gains from studying racial contribution gaps using large-scale administrative data. A second advantage of our administrative data is that they contain information on parental background and, therefore, allow us to also quantify contribution gaps by parental income, and study the interplay between race and family background in shaping savings behavior.

4.2 Household characteristics and the interplay between race and parental income

Our analysis thus far has analyzed gaps by race and parental income separately. In investigating mediators of the raw gaps, we focused solely on individual characteristics. We next investigate the potential mediating role of household factors in impacting contribution gaps and draw out connections between race and parental income.

Accounting for differences in family structure and spousal income shrinks racial contribution gaps. In Figure 3(a), we report racial gaps in saving for the sample for whom parental income is available (recall from Section 3 that this sample is younger than the sample we use to study racial gaps in isolation). The first bar reports residual gaps by race after including all individual-level variables from Figure 2(a). Gaps are similar in percentage terms, though the gaps in percentage points of earnings differ in this sample of younger people who have lower average saving rates.

The second bars in Figure 3(a) show racial gaps once we include variables capturing the structure of the household (number of adults and/or the presence of dependent children on form 1040) and spousal income (deciles if the spouse is working, plus an indicator for having a spouse with no income). Incorporating these household factors reduces the estimated Black-White (Hispanic-White) gap from 14% (8%) to 12% (7%) as a share of the average contribution rate of a White worker. These incremental reductions are obtained *after* accounting for the mediating effects of individual-level characteristics (including income, education, firm, and occupation). This effect arises as, relative to their White co-workers, Black and Hispanic workers are more likely to be single parents (who contribute less on average) and less likely to have a spouse with a high income (who contribute more on average).

The incentives to save can vary with the size and composition of a worker's household. For instance, the marginal utility from a given level of consumption is likely to be higher for workers who have larger families with more dependents. Further, dual-earner households may be better diversified and, therefore, more willing to invest in an illiquid retirement account. Consistent with this idea and our results, Oliver and Shapiro (1989) find marked gradients in wealth by family structure and argued that these patterns are relevant for understanding racial wealth inequality.

Accounting for differences in parental income reduces the residual Hispanic-White gap by 40% and the Black-White gap by around one quarter. Above, we document substantial differences in contribution rates for workers with different levels of parental income, even conditional on a rich set of individual characteristics; that is, the children of the rich save more. The strong association between parental income and own saving has relevance for differences in saving by race given that there are very large gaps in parental income by race. These gaps are illustrated in Figure 3(b), which shows the average income of a worker's parents at age 16 by race and decile of own income. These bins are computed in the population overall, not by race, so comparing differences within these income groups is analogous to controlling for individual income in our regressions. A White worker in the middle-income decile has parental income of around \$90,000 on average. For Black and Hispanic workers in the same income decile, average parental income is around \$50,000.

Taken together, these two facts—that children with high income parents save more, and that parents of White workers have much higher average incomes than those of Black and Hispanic workers—imply that parental income can play a mediating role in the racial savings gap. We quantify this effect in Figure 3(a). Including indicators for each decile of parental income at age 16 reduces the estimated Black-White gap from 0.57 p.p. to 0.44 p.p. and the Hispanic-White gap from 0.32 p.p. to 0.19 p.p. of earnings. Put differently, including parental income in the regression reduces the residual contribution gap (i.e., the estimated gap that remains after accounting for the part mediated by individual- and family-level characteristics) by 23% and 40% for Black and Hispanic workers, respectively, relative to their White counterparts.[16]

Taking Stock. These last results show that the two phenomena documented in Section 4.1—gaps in DC contributions by race and by parental income—are intertwined. The saving gaps by race can be partially attributed to differences in household structure and parental income. These broader household circumstances (which differ by race) can impact the cost of saving and the attractiveness of contributing to subsidized, but illiquid, DC accounts. Before turning to differences in early withdrawals by race and parental income, we report several results on other potential drivers of contribution gaps.

4.3 Other potential drivers of contribution differences

While the goal of this paper is not to explain the differences in contributions by race and parental income, we do explore other plausible correlates of contribution gaps.

The role of access differences. Our main analysis focuses on workers with access to an employer-sponsored retirement plan. To the extent that access to such plans varies by race and parental income, differences in saving subsidy receipts may be even larger than those we report. However, after accounting for individual characteristics, Black and Hispanic workers are, if anything, *more* likely to work at an employer that sponsors a DC retirement plan (see Appendix Figure A.5). In our micro-simulation model, introduced in section 6, we account for differences in access across employers in our measurement of lifetime employer matching and tax benefits.

The role of vesting. Not all employer contributions are immediately vested. When vesting is not immediate, differences in earnings risk could generate different incentives to contribute by race. To investigate this channel, Appendix Figure A.8 re-runs our baseline regressions for racial gaps on a sample where we include only those employees who are fully vested. The patterns in this sample (both the raw gaps and how they change as regressors are added) are very similar to our baseline results. From this, we conclude that vesting has a limited effect on contribution differences by race.

The role of auto-enrollment. Employer-sponsored savings plans are increasingly moving from an opt-in to an automatic enrollment regime. Earlier evidence has shown that, in

[16] We also evaluate the importance of including a dummy for parents having contributed to a DC account, a proxy for familiarity with and exposure to these accounts. This does not affect the size of the residual contribution gap conditional on the mediating factors included in Figure 3(a).

the short-run, the positive effect of automatic enrollment on participation rates is larger for Black and Hispanic workers relative to their White counterparts (Madrian and Shea, 2001). However, in the medium run, the savings gains from auto-enrollment are largely attenuated (Choukhmane, 2024). Comparing employees hired before and after the adoption of auto-enrollment, we find that automatic enrollment does not reduce residual contribution gaps. This is consistent with our finding that residual contribution gaps reflect differences in intensive-margin saving rather than differences in extensive-margin participation (the latter being more likely to be affected by auto-enrollment).

The role of life-expectancy. Those who have lower life expectancy have less of an incentive to save for retirement. Differences in life expectancy by race and parental income (see, e.g., Schwandt et al., 2021, for evidence on the Black-White gap) could therefore account for some of the contribution differences we document. We make two remarks on how this could relate to our findings. First, differences in life expectancy, conditional on income and other individual variables such as education and occupation, are likely to be smaller than unconditional differences. They are therefore less likely to fully account for the residual contribution gaps (after conditioning on individual characteristics) that we emphasize.[17]

Second, and most importantly, differences in life expectancy do not directly impact our study's primary focus, which is the distribution of employer-matching subsidies and tax benefits. Unlike in DB plans, where life expectancy determines the present value of benefits, the value of a DC account is not directly linked to survival. Savings can often be accessed (albeit subject to penalties) before age 59.5, after which the full balance is liquid. These assets can also be bequeathed.

5 Early Withdrawal Differences by Race and Parental Income

The previous section documented racial and intergenerational gaps in *in-flows* into DC accounts. The distributional incidence of the federal tax expenditure, as well as wealth at retirement, depends also on pre-retirement *out-flows* from DC accounts. In this section, we document gaps by race and parental income in the propensity to take early withdrawals. We further show that these gaps are at their widest when individuals face large income declines.

[17] Furthermore, to the extent that saving decisions are determined by subjective survival probabilities, there is evidence that Black Americans have higher subjective life-expectancy than Whites. (See Palloni and Novak (2016) for related results and a discussion of the literature.)

Early withdrawals are common. While employer-sponsored retirement savings plans are designed to be a vehicle to finance consumption in retirement, savers are allowed to access these resources early at a potential tax penalty. Unless the distribution qualifies for an exception, withdrawals before the age of 59.5 are subject to a 10% tax penalty.[18] Despite these restrictions, early withdrawals are common; Goodman et al. (2021) find that flows *out* of DC plans and Individual Retirement Accounts (IRAs) over a 12-year period amount to over 20% of the value of flows *in*. In our sample, we find that 13.5% of DC savers take an early withdrawal of more than $1,000 in a given year (Table 1).

We measure early withdrawal rates using data from the 1099-R tax forms of individuals older than 25 and strictly younger than 55.[19] Our sample is restricted to current workers, while early withdrawals are common during unemployment. To capture withdrawals during unemployment, we look at withdrawals taken in the next calendar year (i.e., in year t+1) during which the worker might have separated from their current employer. We define individuals as taking an early withdrawal if we observe a withdrawal of at least $1,000 (in 2017 dollars) in the next calendar year.[20] Due to data limitations, we cannot distinguish between penalized early withdrawals—those subject to the 10% tax penalty—and nonpenalized distributions. Because those who have never contributed cannot take early distributions we further restrict the sample to individuals who have made at least $1,000 of retirement contributions over the preceding 4 years.

Almost one-quarter of Black savers make an early withdrawal each year. On average, 12.3% of the White retirement savers in our sample take an early distribution each year. The rate at which Hispanic savers make early withdrawals is slightly larger, at 14.5%, while the rate at which Black savers withdraw early is almost twice as high, at 23.3%.

Savers with lower-income parents are significantly more likely to make an early withdrawal. Similar gaps exist between children of top-income parents and low-income parents. On average, 9% of savers with parents in the top income quintile take early withdrawals, whereas 12.3% and 16.2% of savers with parents in the middle and lowest income quintile make an early withdrawal, respectively.

[18]Before 2020, early withdrawals additionally triggered a minimum six-month suspension from contributing to the plan.

[19]Early withdrawals are not penalized for individuals who separate from their employer at or after age 55.

[20]Employers are allowed to implement an automatic cash-out for terminated employees with a balance smaller than $1,000. Therefore, early withdrawals smaller than $1,000 may not reflect the individual actively choosing to withdraw from their retirement account

Large gaps remain even among co-workers with similar earnings, gender, occupation, tenure, education, and location. In Figure 4, we plot a regression cascade similar to that in Figure 2 with the dependent variable instead being an indicator for taking an early withdrawal of more than $1,000. The Black-White gap in early withdrawals remains very large even after accounting for income differences (the "+Income" column) and after including our full set of individual characteristics in the regression (the "+ Indiv." column).[21]

Comparing those who are similar on our full set of mediators (including age, income, education, occupation, tenure, and firm), Black savers make early withdrawals at a rate that is 9.3 p.p. higher than that of White savers. Relative to the Black-White gap, the Hispanic-White gap and gaps by parental income are more substantially mediated by individual characteristics. However, sizeable gaps still remain in the fully saturated regression.

Early withdrawals are most common—and gaps in withdrawal probabilities by race and parental income widen—after large negative income shocks. Early withdrawals are known to be more common when income declines, which is often when liquidity demand is elevated (Coyne et al., 2022). To study how the relationship between early withdrawals and earnings changes varies by race and parental income, we sort workers into 20 ventile bins based on the growth rate of income between year t (the year in which the respondent fills out the ACS and meets a minimum earnings requirement) and year $t+1$. Figure 5, panel (a) reports differences by race; panel (b) reports the average propensity by parental income quintile. We make three observations. First, those who experience larger income declines are more likely to take an early withdrawal. Second, Black DC savers and those with the lowest-income parents are significantly more likely to take early withdrawals at almost all levels of earnings growth. Third, gaps by race and by parental income are largest among those with the largest income declines. Magnitudes are quite large; 52% (41%) of Black (Hispanic) workers in the bottom ventile of income growth take an early withdrawal of at least $1,000, compared to 35% of White workers. Analogously, 42% (36%) of workers with parents in the bottom (middle) income quintile who are in the bottom ventile of income growth take early withdrawals, compared to 25% of workers with parents in the top income quintile.

Accounting for household and parental characteristics reduces estimated racial withdrawal gaps. Next, we explore the interplay between race and both household and

[21] We show the incremental effect of sequentially adding each individual characteristic in Appendix Figure A.3 and an alternative ordering of the cascade (including firm fixed effects before introducing income) in Appendix Figure A.4.

parental backgrounds. For the subset of savers for whom we have data on parental incomes, we find that incorporating household characteristics shrinks the Hispanic-White residual gap by 28% and the Black-White gap by 13% (Figure 6). Likewise, accounting for racial differences in the distribution of parental income shrinks the Hispanic-White gap by an additional 21% and the Black-White gap by an additional 5%.

What channels potentially drive differential early withdrawal rates? Our evidence of large differences in early distributions is consistent with DC savers who are Black or with lower-income parents having stronger liquidity needs and less access to alternative sources of liquidity than other savers. Coyne et al. (2022) highlight that the propensity to tap into retirement accounts early—despite the potential tax penalties—can serve as a measure of differences in liquidity valuation. Consistent with this interpretation, Ganong et al. (2020) find that Black and Hispanic households cut their consumption substantially more than White households following a similarly sized income shock, interpreted there as suggestive of differences in liquidity constraints by race. The same mechanism can also potentially account for differences in early withdrawals by parental income and family structure. For instance, those with richer parents may take fewer early withdrawals because they benefit both from more access to liquidity through familial support and less need to provide financial support. Indeed, there is evidence that richer parents support their children financially and insure them against shocks (Andersen et al., 2020; Fagereng et al., 2023), while poorer parents rely more on their adult children for financial assistance (Chiteji and Hamilton, 2002; Francis and Weller, 2022). Also consistent with differences in liquidity contributing to differences in early withdrawals, we find that single-parent savers are more likely to take early withdrawals (as they likely have stronger liquidity needs relative to married couples, see Appendix Figure A.12).

Finally, the institutional design of DC loan options may also contribute to these patterns. While most DC plans allow active participants to take loans, these loans are typically required to be fully repaid at separation.[22] Any outstanding balance not repaid at separation is treated as an early distribution and is generally subject to a 10% tax penalty. In our data, large income declines often coincide with job separations (Appendix Figure A.10). In this regime, workers who want to avoid penalties might need to give up significant liquidity after a job loss, precisely when liquidity needs are often highest. This could, in turn, reduce workers' incentive to contribute in the first place (Mitchell et al., 2007; Briere et al., 2022).

[22]Among plans administered by Vanguard in 2022, 82% offered a loan option, and 61% required loans to be repaid at separation (Vanguard, 2023). There is no counter-party; loans are financed by liquidating assets from the participant's accounts, and interest payments are rebated into the DC account.

6 The Lifetime Impact of Current and Alternative Retirement Savings Policies

We have documented substantial heterogeneity in annual contributions to and withdrawals from DC accounts. In this section, we combine our data with a micro-simulation model to examine the distributional impact of retirement saving subsidies on wealth at retirement. The model, which is described in full in Appendix E, simulates data on wealth in retirement by bringing together i) our data on flows in and out of DC funds, ii) a specification of the federal tax code (from NBER TAXSIM), iii) Social Security rules, and iv) assumptions about portfolio composition, asset returns, and the draw-down of wealth in retirement.

We use this model for three purposes. First, we evaluate the lifetime distributional impact of both employer matching contributions and the federal tax expenditure. This analysis depends on the whole life-cycle trajectory of income, contributions, and withdrawals. Second, we use the model to place our study of the distributional impact in the context of the broader U.S. retirement system by comparing differences in the allocation of subsidies between individuals with similar Social Security benefits. Lastly, we study the distributional impact of revenue-neutral reforms that would redistribute employer-match dollars within each firm and federal tax expenditures across the population to be distributed 1) proportionally to earnings but 2) independently of workers' own contribution choices.

6.1 Micro-simulation model

6.1.1 The components of wealth at retirement

The full model is outlined in Appendix E. Here, we summarize key inputs and outputs. The model inputs are data on earnings over the life cycle, employee and employer matching contributions to employer-sponsored DC accounts, and withdrawals from those accounts over working life. We do not observe *full* life cycle paths (we have at most 13 years of data for any one individual), so we construct simulated full-life cycle paths using these data and a hot deck-based imputation procedure (described in Appendix E.2).

The key model outputs are:

- **DC wealth** (A_i^{DC}): We define this as the discounted value of after-tax withdrawals from the simulated DC account balance. We assume that savers employ a draw-down rule that keeps withdrawals constant in retirement. We divide DC wealth into three components:
  - **Lifetime tax expenditure**: A_i^T is the part of DC wealth arising from its favor-

able tax treatment. We define this as the difference between A_i^{DC} and the discounted value of withdrawals that worker i would have received if she had instead saved in a taxable account. A_i^T therefore represents the tax advantage obtained from deferring the taxation of contributions, having tax-free growth of assets, and being exempt from capital gains, net of tax penalties on early withdrawals from DC accounts.[23]

  - **Employee contributions**: A_i^{EE} is the portion of DC wealth, exclusive of tax benefits (i.e., $A_i^{DC} - A_i^T$), accruing from *employee* contributions.
  - **Employer contributions**: A_i^{ER} is the portion of DC wealth, exclusive of tax benefits (i.e., $A_i^{DC} - A_i^T$), accruing from *employer* contributions.

- **Social Security Wealth**: Social Security wealth (SS_i) is measured as the discounted stream of benefits an individual will receive through retirement. We calculate this benefit using the Social Security formula and their whole history of earnings.

- **Broad retirement wealth**: Broad retirement wealth (A_i^{BR}) is the sum of DC wealth (A_i^{DC}) and the discounted value of Social Security payments (SS_i).

Appendix Figures A.15 and A.16 illustrate model outputs (earnings, DC wealth, and Social Security) by race and parental income.[24] In these figures, and in our analysis in this section, we divide the population in to six groups based on their lifetime earnings: the bottom four income quintiles and the top two income deciles. We split the top income quintile in two as those in the top income decile are much more likely to be constrained by the annual federal contribution limit. Note that there is very little variation in average Social Security wealth SS_i by race and parental income within these lifetime earnings bins because Social Security payouts are calculated based on lifetime earnings.

[23]As a plausibility check of our model, we compare our estimates of aggregate tax expenditure to DC savings with official Treasury Department figures. In 2023, the Treasury estimated the net value of tax liability foregone because of DC treatment to be \$119 billion in 2021 (US Department of the Treasury, 2023). When we estimate a comparable figure for the US population using our simulations, we obtain \$117 billion. See Appendix E.8.1 for more details.

[24]Because parental income is only observed for a subset of observations (younger individuals), it cannot be used in the hot deck. The relationship between parental income and other lifetime outcomes will be captured through the correlation between parental income and child characteristics at age 25, including income and DC plan participation. To the extent that future income and plan participation are also positively correlated with parental income, our estimates of the relationship between parental income and lifetime outcomes of the children will be attenuated.

6.1.2 Measuring the lifetime effect of the federal tax expenditure

We start by discussing the distributional impact of retirement saving tax expenditures, before turning to a full decomposition of all components of DC wealth. While differences in matching benefits can be, to an extent, read from the data, how much an individual worker receives from the federal tax expenditure depends on trajectories of earnings, contributions, and withdrawals over the whole life cycle. Measuring these factors requires a model such as ours.

Figure 1 shows simulated differences in lifetime tax benefits between median White, Black, and Hispanic earners (panel (c)), as well as between median earners in each parental income group (panel (d)). For every dollar of tax benefits received by a White worker with median lifetime earnings, Black (Hispanic) workers in the middle of the (race-specific) lifetime earnings distribution receive 31 cents (62 cents). For Black (but not Hispanic) workers, these gaps are substantially larger than exist for matching contributions.[25] The reason for this difference is the substantially higher rates of early withdrawals made by Black savers than Hispanic (and White) savers. In addition to the potential tax penalty, those who take early withdrawals forgo the long-term benefits associated with keeping funds in a tax-favored account. Tax benefits are also unequally distributed across parental income groups. Median earners with parents in the bottom (middle) parental income decile receive 39 cents (63 cents) in lifetime tax benefits for each dollar received by those with parents in the top income quintile. Appendix Figure A.18 provides a more granular analysis by dividing each race and parental income decile into six groups based on their lifetime earnings. We observe differences across all lifetime earnings groups.

The differences in the amount of tax benefits across groups are due to a combination of earnings differences and differences in saving behavior conditional on earnings. In the next section, we use the model to examine the distributional effect of current and alternative retirement subsidies, focusing on differences between individuals with similar lifetime earnings.

6.2 Distributional impacts of current retirement savings subsidies

Employer and tax subsidies account for over 40% of DC balances at retirement. The model decomposes wealth balances at retirement into three components: those arising from employee saving elections, employer matches, and favorable tax treatment. Figure 7(a)

[25] This fact can be seen by comparing to the matching numbers in Figure 1, which shows that Black workers receive 45 cents for every dollar of matching received by White workers but only 31 cents for every dollar of tax benefits received by White workers.

reports the share of retirement wealth coming from each component across the six lifetime earnings groups.

Employee contributions account for the largest share of wealth at retirement–approximately three-fifths of DC wealth among those in the bottom quintile of the lifetime earnings distribution and one-half in the top decile. The remainder of DC wealth comes from a combination of employer matches and the tax expenditure. Matches account for approximately one-quarter of wealth across all groups, and the federal tax expenditure contributes between 15% (at the bottom) and 25% (at the top) of wealth.

While the *shares* of wealth arising from each source differ only modestly across the lifetime earnings bin, the *level* of each component scales, approximately, with the saving done by those in each group. Figure 7(b) gives the level of each component expressed as a share of lifetime earnings. Given that savings rates increase in income, the subsidies, as a share of earnings, also increase with earnings. In the bottom quintile of lifetime earnings, cumulative tax and employer subsidies are worth less than 70% of annual lifetime earnings. By comparison, these saving subsidies are worth more than 250% of annual lifetime earnings for individuals in the top decile of earnings. Differences are larger still in dollar terms (as shown in Table 2).

Next, we develop this distributional analysis beyond lifetime earnings to show how differences in saving conditional on earnings (as emphasized in Section 4) shape differences in receipt of saving subsidies.

Savings subsidies amplify racial and intergenerational retirement wealth inequality, even among those with similar earnings and Social Security. Figure 8 shows differences in the receipt of match and tax subsidies by race (panel (a)) and parental income (panel (b)). Earnings bins are defined in the population, so each cluster of bars compares those with the same income. We express the value of subsidies as a percentage of average annual lifetime earnings.

Differences between groups mirror the differences in saving rates documented in Section 4. In each of the bottom five income groups, White workers receive more than Hispanic workers with similar lifetime earnings, who in turn receive more than Black workers. Differences as a proportion of lifetime earnings are largest in the middle of the population earnings distribution; White workers in the middle quintile receive combined tax and matching subsidies worth 161% of their average annual earnings, compared to 119% for Black workers and 148% for Hispanic workers.[26]

[26] Appendix Figure A.17 shows a similar analysis with *race-specific* and parental income group-specific lifetime earnings. Differences here are even larger; for the middle quintile of White workers, the total subsidy is worth 171% of lifetime earnings, while it is 85% for Black workers and 134% for Hispanic workers.

Figure 8(b) shows the value of DC subsidies by parental income. Each successive set of five bars represents a group based on an individual's lifetime earnings. Within each group, the individual bars correspond to quintiles of parental income. Because the children of the rich save more, even conditional on their own earnings (as shown in Figure 2(b)), subsidies for savers advantage them. The differences across groups are substantial: among those in the middle population lifetime earnings quintile, subsidies range from 140% of average annual earnings for those with parents in poorest income decile to 168% for those with parents in the richest group. Gaps in the value of subsidies received by parental earnings are particularly large for higher-income individuals in the top two deciles of lifetime earnings.

6.3 Distributional impacts of alternative retirement savings policies

In this section, we use the micro-simulation model to evaluate a budget-neutral counterfactual exercise that would break the link between private saving and the amount of employer matching benefits and tax subsidies that individuals receive. The counterfactual we consider is an environment in which a) all employees in each firm receive a contribution that is the same percentage of their earnings and b) all workers in the economy get a share of the tax expenditure that is the same proportion of their lifetime earnings.

6.3.1 Description of the counterfactual policy exercise

The counterfactual exercise we discuss below changes the allocation of both employer and tax subsidies for retirement saving. In the Appendix we provide supplementary results for experiments where we separately, rather than simultaneously, redistribute matches and taxes. A full description of the exercise is given in Appendixes E.9-E.11; here we provide a summary. In both our reform of employer matches and the tax subsidy, it is important to note that, while we break the link between savings choice and subsidies, we impose that subsidies remain proportional to earnings (which limits the extent of redistribution along the income distribution).

Employer contributions. We first redistribute the employer matching contributions within each firm. That is, we calculate the aggregate employer matching contribution made by each employer, and we divide these contributions by aggregate compensation. This gives a counterfactual proportion of salary that, if given to all employees regardless of how much they elected to contribute, would cost the same as the status quo.[27]

Tax expenditures. Next, we calculate the aggregate tax expenditure on DC retirement savings. We redistribute this tax expenditure such that that every individual receives a direct government contribution to their retirement account calculated as a proportion of lifetime earnings. This proportion is uniform across individuals and keeps the aggregate tax expenditure constant.

Behavioral responses. In our baseline counterfactual exercise, we assume that individual saving rates (and therefore the the level of individual component of DC wealth) are unchanged across the different counterfactual exercises. Before showing results, we discuss the interpretation of our results under this assumption.

Our counterfactual exercise removes employer matching and tax incentives. As a result, individuals may choose to consume more during their working life and save less for retirement. Whether such behavioral responses change the conclusion of our distributional analysis depends on the policy's goal. On the one hand, if there is no concern about undersaving for retirement, abstracting from these behavioral responses may not change the conclusion of the distributional analysis: groups that receive more employer and tax resources in the counterfactual exercises are better off whether they decide to allocate these new resources toward consumption in working life or consumption in retirement. On the other hand, if undersaving for retirement is a concern, increasing consumption during working life and reducing saving could change the distributional impact of the counterfactual policies. The magnitude of this effect depends on the size of the behavioral response.

There is no consensus in the literature on how much private saving responds to employer matching and tax incentives. Engen et al. (1996) and Poterba et al. (1996) discuss the implications of the early literature on saving incentives. In a more recent contribution, Choi (2015) reviews the literature on matching and finds that it is associated with a small positive effect on participation and an ambiguous effect on average contribution rates.[28] Regarding tax incentives, a review by Friedman (2015) notes that "tax subsidies appear to primarily

[27] While employer-matching contribution formulas are chosen by employers, the government can encourage employers to adopt specific contribution formulas. Arnoud et al. (2021) estimate that a majority of employees are covered by plans with a safe-harbor matching formula. Our counterfactual can be thought of as a change in safe harbor rules that shifts all employers away from offering matching contributions and toward non-elective contributions.

[28] Engelhardt and Kumar (2007) use cross-sectional data to estimate that an increase in the match rate of 25 cents per dollar increases 401(k) participation rates by 5 pp, while Duflo et al. (2006), in a randomized controlled trial with a one-time saving subsidy, find that increasing the match rate from 0% to 50% increases take-up by 11 pp. However, the positive effect of matching on take-up and employee contributions may not translate into higher wealth accumulation if employees reduce their nonretirement saving or increase borrowing in response. Choukhmane and Palmer (2023) estimate that approximately two-thirds of increased employee pension contributions in the UK are financed through reduced nonretirement saving and increased credit card borrowing.

affect the allocation of savings across accounts, rather than the total amount of savings."[29]

Given the lack of consensus and overall small effects found in the empirical literature, our baseline assumption of no behavioral response in private saving will likely be a reasonable approximation. In an extension, we recalculate the results assuming that each dollar of employer matching or tax subsidies generates either 10 cents (which corresponds to the upper bound of the 95% confidence interval in Chetty et al. (2014)) or 30 cents of additional employee savings. The results are shown in Figure A.21, and indicate that, even with an elasticity of employee saving to financial incentives higher than empirical estimates, our counterfactual policy raises retirement wealth accumulation in the bottom half of the lifetime earnings distribution (and especially for those who are Black, Hispanic or have lower income parents). Furthermore, Figure A.22 shows that the relative change in DC wealth gaps by race and parental income is quantitatively very similar when assuming a 10% or 30% elasticity, reflecting the fact that the reduction in employees' saving is larger for groups that benefit more from saving incentives in the baseline.

6.3.2 The reform would reduce inequality in retirement wealth accumulation by own earnings, race, and parental income

Figure 9 summarizes the effect of our counterfactual exercise on DC wealth by race (top panels), and parental income (bottom panels). We express the effect of each reform in two ways. In the panels on the left-hand side ((a) and (c)), we give the change in DC wealth as a proportion of the average annual lifetime earnings of each group. In the panels on the right hand side ((b) and (d)), we express the change as a proportion of broad retirement wealth (DC wealth and Social Security payments). Tables 2 and 3 complement this analysis by expressing the outcomes in dollar values, both in levels and changes.

Results by own lifetime earnings. The take-up of matching and tax incentives is increasing in earnings; therefore, allocating tax expenditures and employer matching contributions in proportion to lifetime earnings (rather than in proportion to saving) would redistribute resources toward lower-income workers. We find that our revenue-neutral reform would significantly raise tax and employer transfers to the bottom 80% of the lifetime earnings distribution, resulting in more than $20,000 in additional DC wealth at retirement on average (Table 3). As shown in Figure 9, the relative gains are largest for the bottom 20% of lifetime earners (with gains of around 125% of annual earnings) and remain sizeable in the

[29]Ramnath (2013) finds no statistically significant effect of the U.S. saver's tax credit on the level of retirement contributions. Similarly, Chetty et al. (2014), using administrative data from Denmark, estimate an elasticity of net saving of less than 1 cent per Danish kroner (DKr) of tax expenditure on subsidies for retirement saving.

middle of the lifetime earnings distribution (with gains of around 60% of annual earnings).[30] These gains at the bottom and middle of the earnings distribution come at the expense of lower resources allocated to top earners (especially in the top decile of lifetime earnings). The loss in DC retirement wealth for the top decile is worth about 40% of annual average earnings. Because the losses are concentrated among those with higher lifetime income and the gains are concentrated among those with lower lifetime income, the relative gains (in percentage terms) from this counterfactual policy are much larger than the relative losses.[31]

Results by race. Among those with the same lifetime earnings, the counterfactual policy would redistribute more to those with lower saving rates (Black and Hispanic workers) and less to those with higher saving rates (White workers). The differences by race are largest in the middle of the lifetime earnings distribution. The counterfactual reform is worth close to a year's worth of average earnings for Black workers, and just over half a year's worth for White workers, with an intermediate effect for Hispanic workers. Gaps also exist in all other income groups except the very top.[32]

We can quantify how much such a policy would reduce racial gaps in DC wealth. Table 3 shows that the reform would increment wealth among Black, Hispanic, and White workers in the middle of the lifetime earnings distribution by about $37,000, $27,000, and $24,000, respectively. This would reduce the gap between the DC wealth of Black and White workers in this lifetime earnings group from 37% to 24%, and that between Hispanic and White workers from 16% to 11%.

Results by parents' income. The bottom two panels of Figure 9 show the effect of our reform by both own and parental income (with Appendix Figure A.20 giving a complementary analysis where lifetime earnings groups are defined within each parental income group). Across all income groups, those with lower-income parents benefit more from the reform than those with richer parents, and these differences by parental income (given own income) are larger for those with above-median lifetime earnings.

For instance, Table 3 shows that those in the bottom-income group with the lowest and highest-income parents have similar gains from the reform (around $19,000 in additional DC wealth). In contrast, in the fourth quintile of earnings, those with the lowest-income parents

[30] Appendix Figure A.23 gives selected results for experiments where we separately, rather than simultaneously, redistribute tax and matching subsidies.

[31] While these reforms are designed to be revenue neutral for the government and aggregate-compensation neutral for the firms, they lead to a net increase in wealth on retirement as matching resources are transferred from older workers to younger workers, who have more time to retirement to benefit from asset returns.

[32] The lifetime earnings groups in our baseline analysis are defined at the population level, and therefore a Black worker in the median earnings bin has higher earnings than the median-earning Black worker. Appendix Figure A.20 gives a complementary analysis where lifetime earnings categories are defined within each racial group.

gain an additional \$21,000 while those with the richest parents gain an additional \$8,000. Losses are concentrated among those with both high earnings and high parental income; in the top earnings decile, those with the poorest parents experience a \$28,000 drop in DC wealth, whereas those with parents in the top income quintile experience a \$99,000 drop in DC wealth. While these wealth losses are large in absolute terms, they represent only a 4% reduction in broad retirement wealth for top earners with the richest parents. Given that those with lower-income parents benefit more from the reform, we estimate that the reform can close the gap in DC wealth accumulation by parental income by approximately a third.

7 Conclusion

Since the introduction of the permanent income tax system in 1913, the U.S. has promoted retirement saving with tax subsidies and employer contributions. A long-standing concern is that these subsidies are regressive and largely favor higher-income individuals. This concern has sparked a long tradition of economics research studying the distributional effects and optimal design of the retirement system (Diamond, 1977; Kotlikoff et al., 1982; Geanakoplos et al., 1999; Moser and Olea de Souza e Silva, 2019). This concern is also reflected in the regulatory framework; since 1942, U.S. pension plans have been required to pass an annual nondiscrimination test to ensure that the benefits of the plan do not disproportionately accrue to highly compensated employees.[33] In addition to income-based nondiscrimination tests, the Social Security formula is progressive and offers higher replacement rates for individuals with lower lifetime earnings. One view is that these more progressive aspects of the U.S. retirement system balance the income-regressive nature of retirement saving subsidies.

In this paper, we challenge this view by studying the distributional properties of retirement saving subsidies among individuals who have similar incomes but differ along other dimensions (with a focus on differences by race and parental income). We find that the current system channels more tax and employer resources toward workers who are White and have richer parents than toward their similar-income coworkers who are Black or Hispanic and have lower-income parents. Analogous distributional comparisons could be made by other characteristics which are important for saving. While we do not emphasize them in our paper, analyses in Appendix D show that, conditional on income, those with more education save more than those with less education, while single parents save less than couples with children. The consequent effects on wealth accumulation are large and are not directly

[33]To pass the nondiscrimination test, the employer must show that differences between the average employee and employer contribution rates for highly compensated and non–highly compensated employees are sufficiently small. Employers can avoid these annual tests by adopting a set of plan features that qualify a plan as a safe harbor plan.

offset by other aspects of the retirement system. The Social Security formula does not vary by race, education, or parental background, and employer nondiscrimination tests only consider one's compensation. Our results thus suggest that future research on the distributional impact and optimal design of retirement systems would benefit from looking beyond income.

References

Altonji, J. G. and R. M. Blank (1999). Race and gender in the labor market. *Handbook of Labor Economics 3*, 3143–3259.

Altonji, J. G. and U. Doraszelski (2005). The role of permanent income and demographics in black/white differences in wealth. *Journal of Human Resources 40*(1), 1–30.

Andersen, A. L., N. Johannesen, and A. Sheridan (2020). Bailing out the kids: new evidence on informal insurance from one billion bank transfers.

Ariel/AON Hewitt (2009). 401(k) Plans in Living Color.

Arnoud, A., T. Choukhmane, J. Colmenares, C. O'Dea, and A. Parvathaneni (2021). The Evolution of U.S. Firms' Retirement Plan Offerings. Evidence from a New Panel Data Set. Technical report, NBER.

Avenancio-León, C. F. and T. Howard (2022). The assessment gap: Racial inequalities in property taxation. *The Quarterly Journal of Economics 137*(3), 1383–1434.

Barber, B. M. and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics 116*(1), 261–292.

Barsky, R., J. Bound, K. K. Charles, and J. P. Lupton (2002). Accounting for the black–white wealth gap: a nonparametric approach. *Journal of the American Statistical Association 97*(459), 663–673.

Bayer, P. and K. K. Charles (2018). Divergent paths: A new perspective on earnings differences between black and white men since 1940. *The Quarterly Journal of Economics 133*(3), 1459–1501.

Bee, A. and J. Mitchell (2017). Do Older Americans Have More Income Than We Think? *U.S. Census Bureau SEHSD Working Paper #2017-39*.

Bhutta, N. and A. Hizmo (2021). Do minorities pay more for mortgages? *The Review of Financial Studies 34*(2), 763–789.

Blau, F. D. and J. W. Graham (1990). Black-white differences in wealth and asset composition. *The Quarterly Journal of Economics 105*(2), 321–339.

Bond, B., J. D. Brown, A. Luque, A. O'Hara, et al. (2014). The nature of the bias when studying only linkable person records: Evidence from the American Community Survey. *Center for Administrative Records Research and Applications Working Paper 8*.

Briere, M., J. Poterba, and A. Szafarz (2022, May). Precautionary Liquidity and Retirement Saving. *AEA Papers and Proceedings 112*, 147–150.

Brown, D. A. (2021). *The Whiteness of Wealth: How the Tax System Impoverishes Black Americans–And How We Can Fix It.* Crown Publishing Group (NY).

Brown, J. R., J. M. Poterba, and D. P. Richardson (2022). Trends in Retirement and Retirement Income Choices by TIAA Participants: 2000–2018. Technical report, National Bureau of Economic Research.

Burman, L. E., W. G. Gale, M. Hall, and P. Orszag (2004). Distributional Effects of Defined Contribution Plans and Individual Retirement Accounts.

Cahuc, P., S. Carcillo, and A. Zylberberg (2014). *Labor economics.* MIT press.

Carneiro, P., J. J. Heckman, and D. V. Masterov (2005). Labor market discrimination and racial differences in premarket factors. *The Journal of Law and Economics 48*(1), 1–39.

Catherine, S. and N. Sarin (2023). Social Security and the Racial Wealth Gap. Technical report.

Charles, K. K. and E. Hurst (2002). The transition to home ownership and the black-white wealth gap. *Review of Economics and Statistics 84*(2), 281–297.

Charles, K. K. and E. Hurst (2003). The correlation of wealth across generations. *Journal of Political Economy 111*(6), 1155–1182.

Chen, A. and A. H. Munnell (2020, November). How Much Taxes Will Retirees Owe on Their Retirement Income? Technical Report Working Paper 2020-16, Center for Retirement Research at Boston College, Chestnut Hill, MA.

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. *The Quarterly Journal of Economics 129*(3), 1141–1219.

Chetty, R., N. Hendren, M. R. Jones, and S. R. Porter (2020). Race and economic opportunity in the United States: An intergenerational perspective. *The Quarterly Journal of Economics 135*(2), 711–783.

Chiteji, N. S. and D. Hamilton (2002). Family connections and the black-white wealth gap among middle-class families. *The Review of Black Political Economy 30*(1), 9–28.

Choi, J. J. (2015). Contributions to defined contribution pension plans. *Annual Review of Financial Economics 7*, 161–178.

Choukhmane, T. (2024). Default options and retirement saving dynamics. Technical report.

Choukhmane, T., L. Goodman, and C. O'Dea (2023). Efficiency in Household Decision Making: Evidence from the Retirement Savings of US Couples. Technical report, National Bureau of Economic Research.

Choukhmane, T. and C. Palmer (2023). How Do Consumers Finance Increased Retirement Savings? Technical report.

Congressional Budget Office (2021). The Distribution of Major Tax Expenditures in 2019. Technical report, Congressional Budget Office.

Coyne, D., I. Fadlon, and T. Porzio (2022). Measuring Valuation of Liquidity with Penalized Withdrawals. Technical report, National Bureau of Economic Research.

Darity, W. and S. L. Myers (1983). Changes in Black Family Structure: Implications for Welfare Dependency. *American Economic Review 73*(2), 59–64. Publisher: American Economic Association.

Darity, W. and M. J. Nicholson (2005). Racial wealth inequality and the Black family. *African American Family Life: Ecological and Cultural Diversity 78*, 85.

Darity, W. A. and S. L. Myers (1987). Do Transfer Payments Keep the Poor in Poverty? *American Economic Review 77*(2), 216–222. Publisher: American Economic Association.

Darity Jr, W., D. Hamilton, M. Paul, A. Aja, A. Price, A. Moore, and C. Chiopris (2018). What we get wrong about closing the racial wealth gap. *Samuel DuBois Cook Center on Social Equity and Insight Center for Community Economic Development 1*(1), 1–67.

Department of Labor (2023). Private Pension Plan Bulletin. Technical report, Department of Labor.

Derenoncourt, E., C. H. Kim, M. Kuhn, and M. Schularick (2022, June). Wealth of Two Nations: The U.S. Racial Wealth Gap, 1860-2020. Working Paper 30101, National Bureau of Economic Research.

Derenoncourt, E. and C. Montialoux (2021). Minimum wages and racial inequality. *The Quarterly Journal of Economics 136*(1), 169–228.

Devlin-Foltz, S., A. Henriques, and J. Sabelhaus (2016). Is the US retirement system contributing to rising wealth inequality? *RSF: The Russell Sage Foundation Journal of the Social Sciences 2*(6), 59–85.

Diamond, P. (1977). A framework for social security analysis. *Journal of Public Economics 8*(3), 275–298.

Duflo, E., W. Gale, J. Liebman, P. Orszag, and E. Saez (2006). Saving incentives for low-and middle-income families: Evidence from a field experiment with H&R Block. *The Quarterly Journal of Economics 121*(4), 1311–1346.

Dushi, I. and M. Honig (2015). How much do respondents in the health and retirement study know about their contributions to tax-deferred contribution plans? a cross-cohort comparison. *Journal of Pension Economics & Finance 14*(3), 203–239.

Dushi, I. and H. M. Iams (2010). The impact of response error on participation rates and contributions to defined contribution pension plans. *Soc. Sec. Bull. 70*, 45.

Dynan, K. E., J. Skinner, and S. P. Zeldes (2004). Do the rich save more? *Journal of Political Economy 112*(2), 397–444.

Engelhardt, G. V. and A. Kumar (2007). Employer matching and 401 (k) saving: Evidence from the health and retirement study. *Journal of Public Economics 91*(10), 1920–1943.

Engen, E. M., W. G. Gale, and J. K. Scholz (1996). The illusory effects of saving incentives on saving. *Journal of Economic Perspectives 10*(4), 113–138.

Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2020). Heterogeneity and persistence in returns to wealth. *Econometrica 88*(1), 115–170.

Fagereng, A., L. Guiso, L. Pistaferri, and M. A. Ring (2023). Insuring Labor Income Shocks: The Role of the Dynasty.

Farber, H. S. (1994). The Analysis of Interfirm Worker Mobility. *Journal of Labor Economics 12*(4), 554–593.

Feenberg, D. and E. Coutts (1993). An Introduction to the TAXSIM Model. *Journal of Policy Analysis and Management 12*(1), 189–194.

Ferrie, J., C. Massey, and J. Rothbaum (2021). Do Grandparents Matter? Multigenerational Mobility in the United States, 1940–2015. *Journal of Labor Economics 39*(3).

Fidelity (2023). A Lifetime Investment Approach in a Single Strategy.

Fidelity (2023). Fidelity Freedom Fund, Selected Years.

Francis, D. V. and C. E. Weller (2021). Retirement inequality by race and ethnicity. *Public Policy & Aging Report 31*(3), 83–88.

Francis, D. V. and C. E. Weller (2022). Interfamily Financial Dependence and Retirement Savings. *The Journal of Retirement 9*(4), 73–90.

Friedman, J. (2015). Tax Policy and Retirement Savings. *Paper prepared for the Economics of Tax Policy Conference at the Brookings Institution.*

Ganong, P., D. Jones, P. J. Noel, F. E. Greig, D. Farrell, and C. Wheat (2020). Wealth, race, and consumption smoothing of typical income shocks. Technical report, National Bureau of Economic Research.

Geanakoplos, J., O. S. Mitchell, and S. Zeldes (1999). Social Security Money's Worth. In O. S. Mitchell, R. Myers, and H. Young (Eds.), *Prospects for Social Security Reform*, pp. 79–151. Philadelphia, PA: Univ. of Pennsylvania Press.

Goldin, C. (2021). *Career and family: Women's century-long journey toward equity*. Princeton University Press.

Goodman, L., J. Mortenson, K. Mackie, and H. R. Schramm (2021). Leakage From Retirement Savings Accounts In The United States. *National Tax Journal 74*(3), 689–719.

Gourinchas, P.-O. and J. A. Parker (2002). Consumption over the life cycle. *Econometrica 70*(1), 47–89.

Guvenen, F., S. Ozkan, and J. Song (2014). The nature of countercyclical income risk. *Journal of Political Economy 122*(3), 621–660.

Hamilton, D. and W. A. Darity (2017). The Political Economy of Education, Financial Literacy, and the Racial Wealth Gap. *Federal Reserve Bank of St. Louis Review 99*(1), 59–76.

Hou, W. and G. T. Sanzenbacher (2021). Measuring racial/ethnic retirement wealth inequality. *The Journal of Retirement 8*(3), 12–28.

Internal Revenue Service (2023, September). SOI Tax Stats - Individual Information Return Form W-2 Statistics.

Joint Committee on Taxation (2019, July). Background on Revenue Sources for the Social Security Trust Funds. Technical Report JCX-41-19.

Jordà, O., K. Knoll, D. Kuvshinov, M. Schularick, and A. M. Taylor (2019, 04). The Rate of Return on Everything, 1870–2015. *The Quarterly Journal of Economics 134*(3), 1225–1298.

Kermani, A. and F. Wong (2021). Racial Disparities in Housing Returns. Technical report, National Bureau of Economic Research.

Kotlikoff, L. J., A. Spivak, and L. H. Summers (1982). The Adequacy of Savings. *American Economic Review 72*(5), 1056–1069.

Kuan, K. Y., M. R. Cullen, and S. Modrek (2015). Racial disparities in savings behavior for a continuously employed cohort. Technical report, National Bureau of Economic Research.

Kuhn, M., M. Schularick, and U. I. Steins (2020). Income and wealth inequality in America, 1949–2016. *Journal of Political Economy 128*(9), 3469–3519.

Kuka, E. and B. A. Stuart (2021). Racial Inequality in Unemployment Insurance Receipt and Take-Up. Technical report, National Bureau of Economic Research.

Lang, K. and M. Manove (2011). Education and labor market discrimination. *American Economic Review 101*(4), 1467–1496.

Levine, P. B. and D. Ritter (2022). The Racial Wealth Gap, Financial Aid, and College Access. `https://conference.nber.org/conf_papers/f162246.pdf`.

Lusardi, A., P.-C. Michaud, and O. S. Mitchell (2017). Optimal financial knowledge and wealth inequality. *Journal of Political Economy 125*(2), 431–477.

Lusardi, A. and O. S. Mitchell (2014). The economic importance of financial literacy: Theory and evidence. *American Economic Journal: 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. *The Quarterly Journal of Economics 116*(4), 1149–1187.

Mitchell, O. S., S. P. Utkus, and T. Yang (2007). Turning workers into savers? Incentives, liquidity, and choice in 401 (k) plan design. *National Tax Journal 60*(3), 469–489.

Moser, C. and P. Olea de Souza e Silva (2019). Optimal paternalistic savings policies. *Columbia Business School Research Paper* (17-51).

Myers, E. and J. Topoleski (2020). Worker Participation in Employer-Sponsored Pensions: Data in Brief. Technical report.

Myers Jr, S. L. (1995). Racial discrimination in housing markets: Accounting for credit risk. *Social Science Quarterly*, 543–561.

Neal, D. A. and W. R. Johnson (1996). The role of premarket factors in black-white wage differences. *Journal of Political Economy 104*(5), 869–895.

Oliver, M. L. and T. M. Shapiro (1989). Race and wealth. *The Review of Black Political Economy 17*(4), 5–25.

Palloni, A. and B. Novak (2016). Subjective survival expectations and observed survival: How consistent are they? *Vienna Yearbook of Population Research 14*, 187.

Poterba, J. M., S. F. Venti, and D. A. Wise (1996). How retirement saving programs increase saving. *Journal of Economic Perspectives 10*(4), 91–112.

Ramnath, S. (2013). Taxpayers' responses to tax-based incentives for retirement savings: Evidence from the Saver's Credit notch. *Journal of Public Economics 101*, 77–93.

Ross, S. L. and J. Yinger (2002). *The color of credit: Mortgage discrimination, research methodology, and fair-lending enforcement*. MIT press.

Sabelhaus, J. (2023). How Does Social Security Affect the Racial Wealth Gap? *Ann Arbor, MI. University of Michigan Retirement and Disability Research Center (MRDRC) Working Paper*.

Sabelhaus, J. E. and J. P. Thompson (2021). Racial Wealth Disparities: Reconsidering the Roles of Human Capital and Inheritance.

Schwandt, H., J. Currie, M. Bär, J. Banks, P. Bertoli, A. Bütikofer, S. Cattan, B. Z.-Y. Chao, C. Costa, L. González, et al. (2021). Inequality in mortality between Black and White Americans by age, place, and cause and in comparison to Europe, 1990 to 2018. *Proceedings of the National Academy of Sciences 118*(40), e2104684118.

Sialm, C. and H. Zhang (2020). Tax-Efficient Asset Management: Evidence from Equity Mutual Funds. *The Journal of Finance 75*(2), 735–777.

Skandalis, D., I. E. Marinescu, and M. Massenkoff (2022). Racial inequality in the us unemployment insurance system. *Available at SSRN 4097104*.

Social Security Administration (2023a, September). Benefit Formula Bend Points.

Social Security Administration (2023b, September). Maximum Taxable Earnings.

Topoleski, J. J. (2018). Worker participation in employer-sponsored pensions: A fact sheet.

US Department of the Treasury (2023). Tax Expenditures, FY2023. Technical report.

Vanguard (2023). How america saves 2023. Technical report.

Viceisza, A., A. Calhoun, and G. Lee (2022). Racial and Ethnic Disparities in Retirement Outcomes: Impacts of Outreach. *The Review of Black Political Economy*, 00346446231182343.

Wagner, D. and M. Layne (2014). The Person Identification Validation System: Applying the Center for Administrative Records and Research and Applications' Record Linkage Software. *Center for Administrative Records Research and Applications Working Paper*.

Watson, J. and M. McNaughton (2007). Gender differences in risk aversion and expected retirement benefits. *Financial Analysts Journal 63*(4), 52–62.

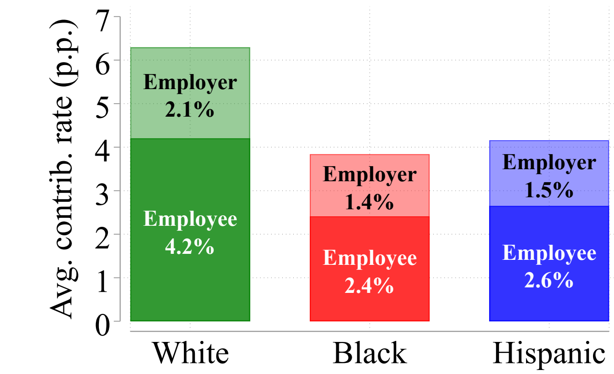
Wolff, E. N. (2023). Trends in the Retirement Preparedness of Black and Hispanic Households in the US. Technical report, National Bureau of Economic Research.

Yogo, M., N. Cox, and A. Whitten (2023). WI23-12: Financial Inclusion Across the United States.

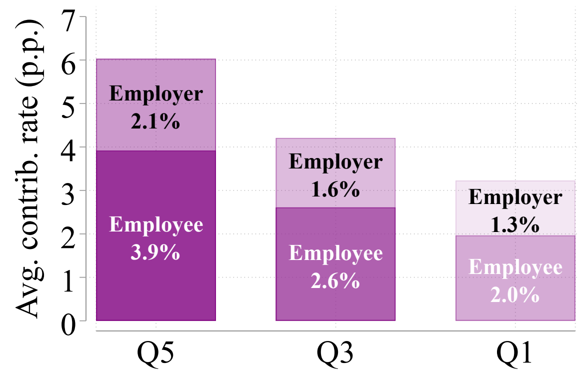
Figures and Tables

Figure 1: Differences by race and parental income in contributions, labor income, matching compensation, and DC tax benefits

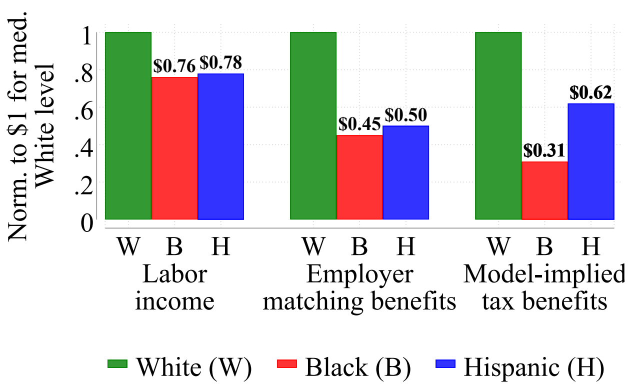
(a) Average Employee + Match DC Contrib. Rate, by race



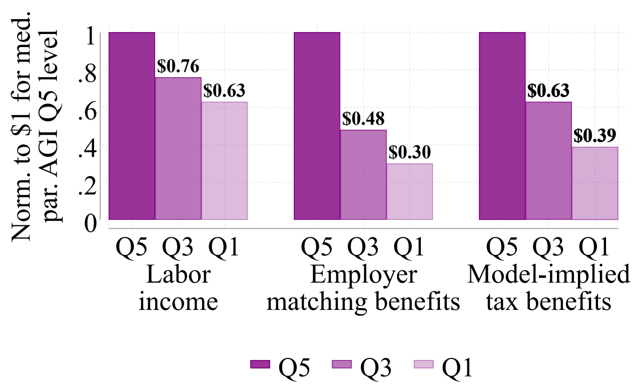
(b) Average Employee + Match DC Contrib. Rate, by parental income



(c) Median income, matching benefits, and model-implied tax benefits, by race

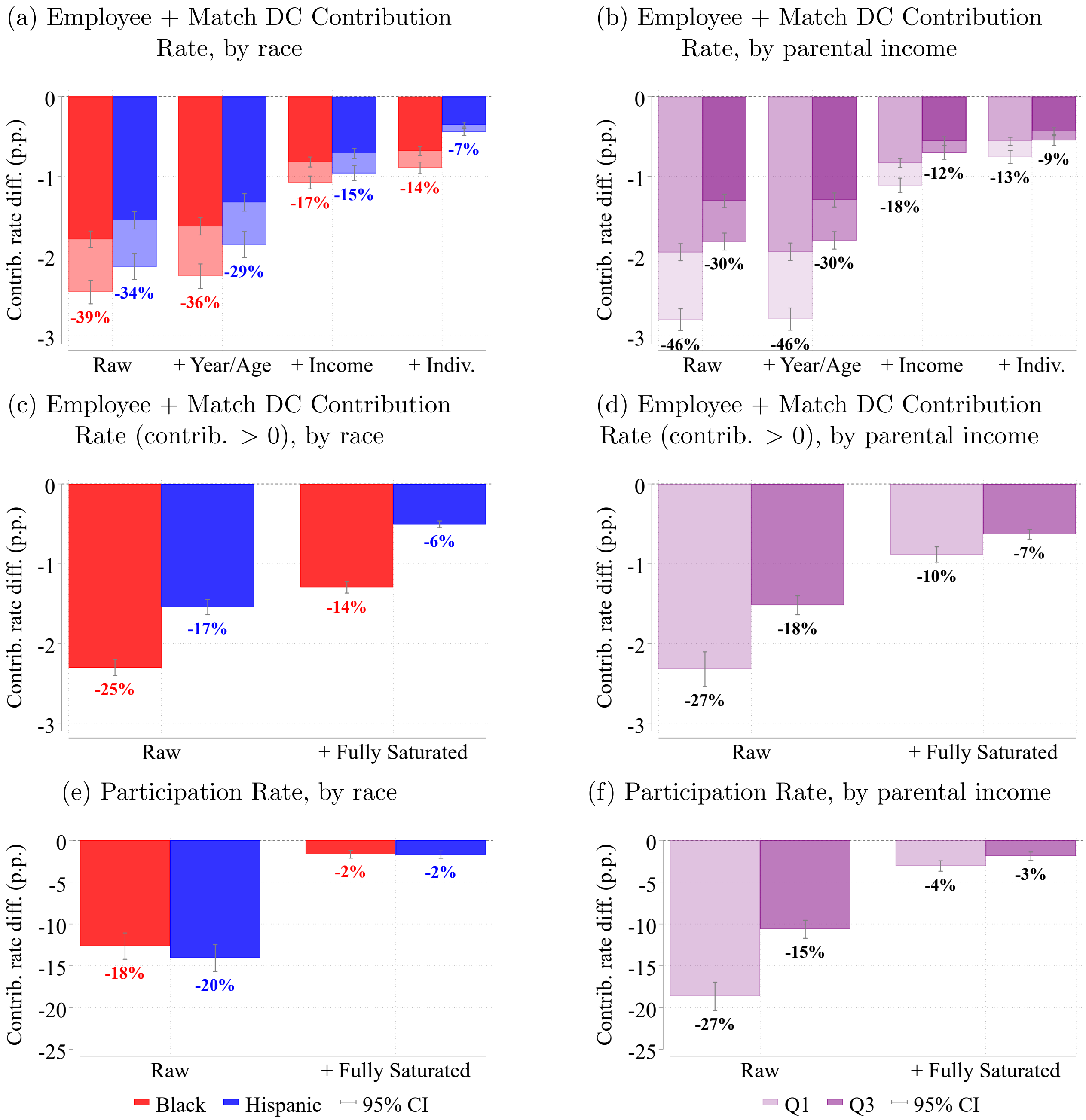


(d) Median income, matching benefits, and model-implied tax benefits, by par. income



Notes: Panels (a) and (b) show the average employee DC contribution rate (the darker shaded areas) and average employer matching rate (the lighter shaded) as a proportion of salary by race and parental income. The sample for panel (a) is restricted to workers with at least $8,000 in annual earnings and whose employer sponsors a DC plan. The sample for panel (b) has the same conditions but is further restricted to those whose parental income is observable (which limits the sample to those born after 1978). Appendix Section A.3.3 provides more information on these samples. Panels (c) and (d) document average gaps in labor income, matching contributions, and DC tax benefits for individuals around the median labor income of each group (with the White and Parental AGI Quintile 5 levels normalized to 1). The first (second) set of three bars shows mean labor income (employer matching contributions) for those between the 45th and 55th percentiles of the race- or parent-specific labor income distribution. The third set of three bars in panel (b) reports calculations from our life-cycle micro-simulation model in Section 6. It shows mean model-implied tax benefits for individuals in each group between the 40th and 60th percentiles of the race- or parent-specific lifetime earnings distribution. This statistic quantifies the present discounted value of the deferral of taxation and exemption of returns from taxation, net of tax penalties on early withdrawals.

Figure 2: Differences in contribution rates by race and parental income

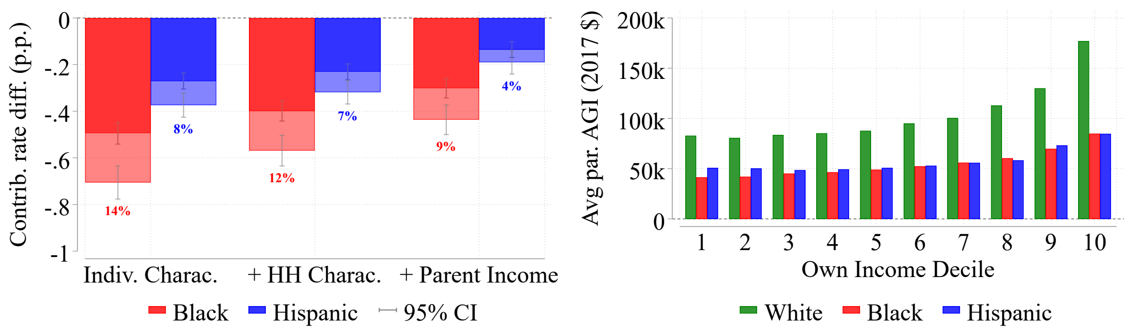


Notes: We measure the gaps between non-Hispanic White and Black (Hispanic) workers and between workers with parental incomes in quintile 5 and quintile 1 (quintile 3). To show the effects of potential mediating channels, we begin with Model (i), i.e., "Raw" which represents the univariate regression of the outcome variable on the categorical race or parental income variable: $y_{it} = \alpha + \beta_0 race_i + \epsilon_{it}$, where α and ϵ_{it} are the constant and error terms, and *race* (*parent_AGI_quintile*) identifies, among others, the non-Hispanic White, Black, and Hispanic groups (analogously, parent quintiles 1-5). In all models, White (Quintile 5) is absorbed as the omitted category, so the coefficient on the race (parent) term, β_0, plotted in the figures measures the average gap between White and Black or Hispanic (Quintile 5 and Quintile 3 or 1). We then sequentially include potential mediating channels: (ii) year and age, i.e., "Year/Age"; (iii) income; and (iv) other individual characteristics, i.e., "Indiv.", which includes gender, education, tenure, county, occupation, and employer. Please see Figure A.3 for a version of panel (a) with bars for each specification. In panels (a) and (b), the darker shaded regions represent the employee DC contribution rates, while the lighter region is the employer rate. In panels (c)-(f), "Fully Saturated" signifies the estimates that include all individual-level factors. Since our data set is a repeated cross-section, we calculate clustered standard errors by EIN. The percentages printed under the bars represent the percentage difference relative to the average level for the omitted category (White/parental quintile 5). For more information on how our variables are constructed, please see Appendix A.2.2. For an alternative cascade ordering, please see Figure A.4.

Figure 3: Understanding links between gaps by race and parental income

(a) Employee + Match DC Contribution Rate

(b) Average parental income by own income

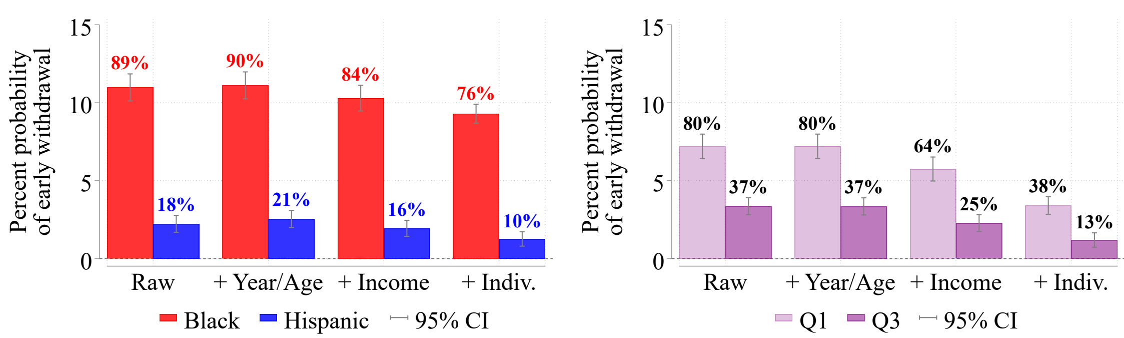


Notes: As in Figure 2, panel (a), here panel (a) shows the racial gaps in employee plus employer match contribution rates as we include different mediating channels. "Indiv. Charac." contains all the channels up to and included in "Indiv." as in panel (a) of the previous figure. "HH Charac." includes household-level family structure and spousal income. Lastly, we account for the worker's parental income decile. For more information on how our variables are constructed, please see Appendix A.2.2. It is important to note that the sample used for this analysis is a subset of the sample in Figure 2 (i.e., they are younger being from just the 1978-1992 birth cohort). This deviation accounts the differences in the estimates. For more detail about the different samples, please see Section 3 and Appendix A.3.3. The darker shaded regions represent the employee DC contribution rates, while the lighter region is the employer rate. The percentages printed under the bars represent the percentage difference relative to the average level for the omitted category (White workers). Panel (b) shows the average parental income by race in each income decile.

Figure 4: Differences in early withdrawal rates by race and parental income

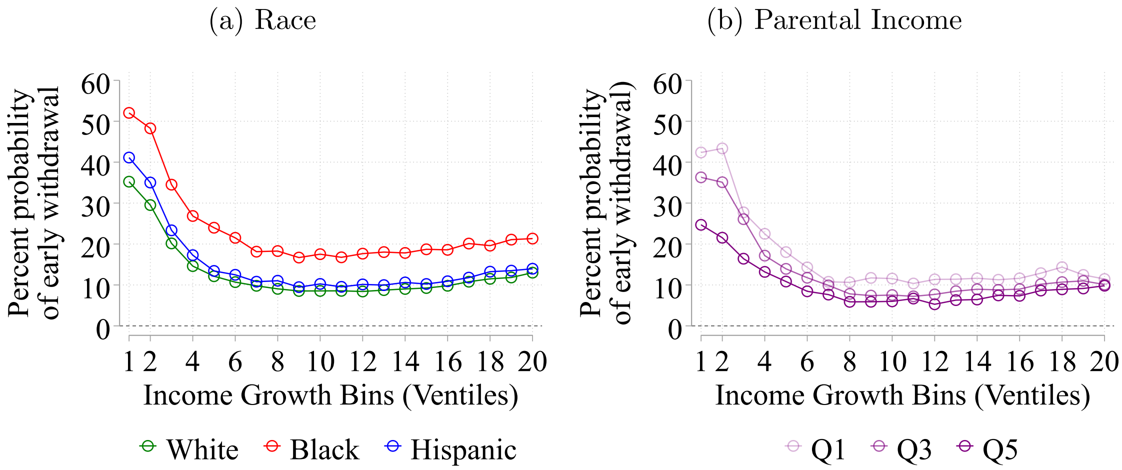
(a) Race

(b) Parental Income



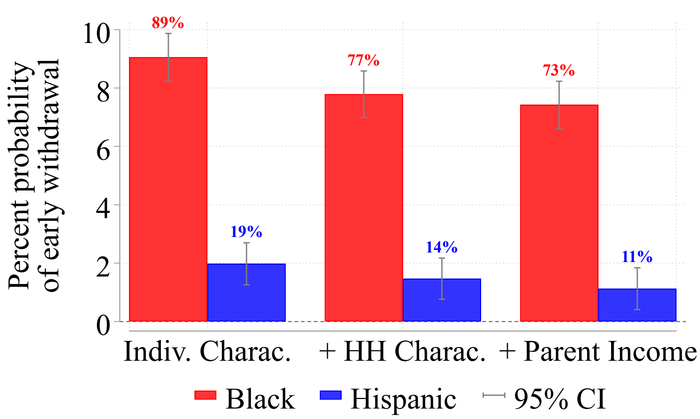
Notes: These figures show the probability of taking an early withdrawal of at least $1,000 by race and parental income quintiles among the Form 5500 and parent-Form 5500 samples. Both panels follow the same structure as Figure 2. Panel (a) shows the progression of the racial gaps relative to White workers; panel (b) shows the progression of parental income gaps relative to Q5 parental income as we add potential mediating channels. Early withdrawal dummies are equal to one for people who i) contributed greater than $1,000 in the prior 4 years, ii) withdraw at least $1,000 in the year following our survey year, and iii) were less than age 55 at the time of withdrawal, for more details please see Appendix A.2.1. All workers in our sample were employed in the survey year. The percentages printed above the bars represent the percentage difference relative to the average level for the omitted category (White workers). For an alternative cascade ordering, please see Figure A.4.

Figure 5: Early withdrawal probability by race/parental income and ventile of earnings growth



Notes: These figures show the breakdown of the probability of early withdrawals over $1,000 in year $t+1$ for race and parental income quintiles by income growth ventile from year $t+1$ to year t (where t is the year we observe individuals in the ACS). Please see Figure 4 notes for detailed information on early withdrawal dummies.

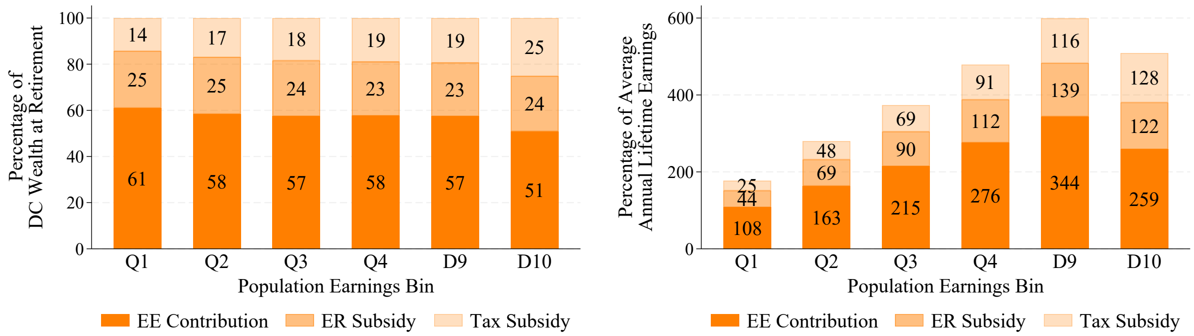
Figure 6: Understanding links between gaps by race and parental income for dissavings



Notes: Analogous to panel (a) in Figure 3, we bring racial and intergenerational analysis together here to understand their interplay with regards to early withdrawals. As in Figure 4, early withdrawal dummies are equal to one for people who i) contributed greater than $1,000 in the prior 4 years, ii) withdraw at least $1,000 in the year following our survey year, and iii) were less than age 55 at the time of withdrawal. All workers in our sample were employed in the survey year. "Indiv. Charac." and "HH. Charac." have the same definition as in Figure 3, and "Parent Income" signifies parental income deciles, likewise. For more information, please see the Figure notes of Figures 3 and 4.

Figure 7: Contributions of employee contributions and subsidies to retirement wealth

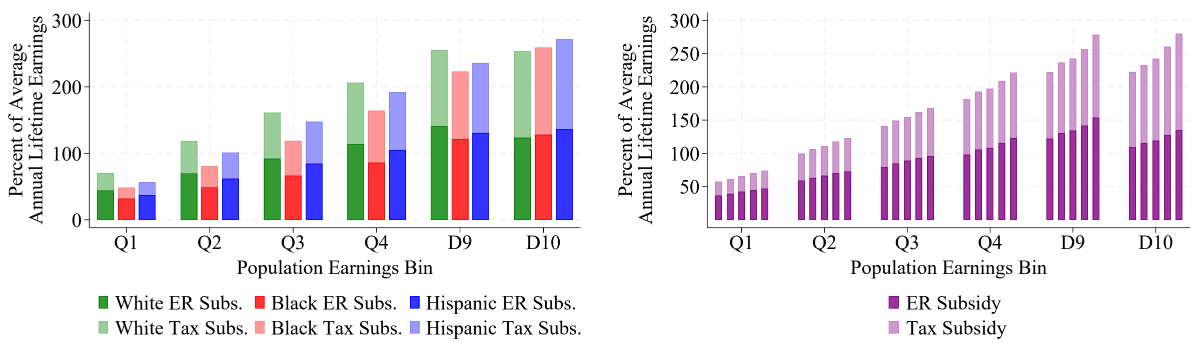
(a) Shares of wealth at retirement   (b) Values as a prop. of average annual earnings



Notes: These figures decompose DC wealth at retirement into components arising from employee contributions, employer matching contributions, and federal tax subsidies. Panel (a) shows shares of DC wealth arising from these three sources. Numbers may not add to 100 due to rounding. Panel (b) shows the value of each component as a proportion of the average annual lifetime earnings of each group. Both figures break down results by lifetime earnings. Lifetime earnings are grouped into the bottom four population quintiles and the top two population deciles.

Figure 8: Contributions of employer and tax subsidies to retirement wealth

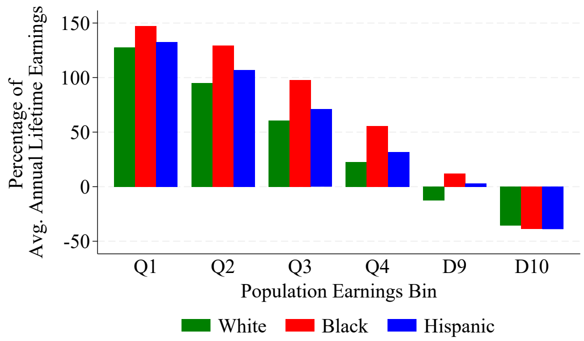
(a) By race   (b) By parental income



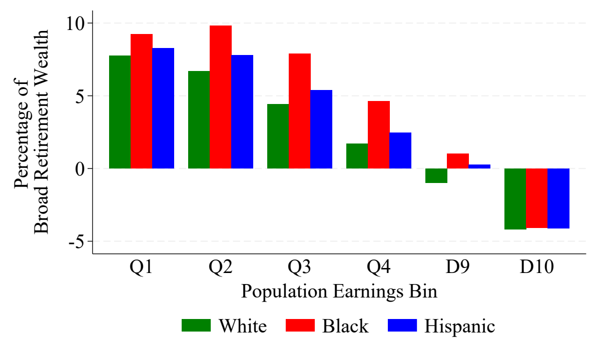
Notes: This figure shows lifetime employer and tax subsidies as a percentage of average annual lifetime earnings, by own earnings level and by either race or parental income. Panel (a) shows these subsidies by race, and panel (b) shows them by quintiles ("bins") of parental income. In both panels, the darker bars show average employer matching subsidies, and the lighter bars show average tax subsidies to retirement savings. Lifetime earnings groups are divided into six bins—the bottom four quintiles and the top two deciles. Earnings bins are calculated at the population level. Appendix Figure A.17 shows results where lifetime earnings bins are defined within race and parental income group.

Figure 9: Change in retirement wealth measures under counterfactual, by race and parental income

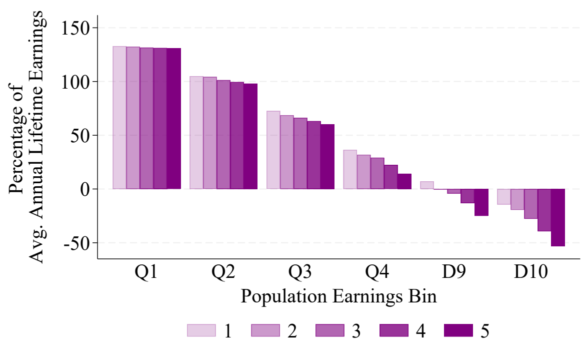
(a) Δ in ret. wealth by race



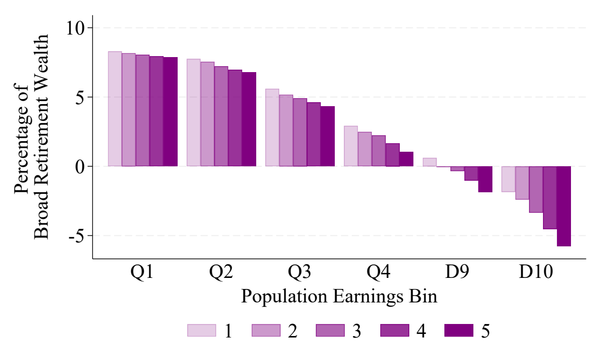
(b) Δ broad ret. wealth (DC + SS) by race



(c) Δ in DC wealth by par. inc.



(d) Δ broad ret. wealth (DC + SS) by par. inc.



Notes: This figure illustrates the impact of our baseline counterfactual exercise on measures of retirement wealth. This counterfactual exercise distributes the aggregate employer matches in each firm so that all workers in that firm receive the same proportion of their earnings. The counterfactual further distributes the aggregate federal tax expenditure so that all workers receive a contribution that is in proportion to their lifetime earnings. We show the effect on two outcomes: panels (a) and (c) show the change in DC wealth on retirement, with the change in wealth expressed as a proportion of average annual working life earnings. Panels (b) and (d) show proportionate change in broad retirement wealth, where broad retirement wealth is the sum of DC wealth and Social Security. Lifetime earnings groups are divided into six bins—the bottom four quintiles and the top two deciles. Earnings bins are calculated at the population level. Appendix Figure A.20 shows results where lifetime earnings bins are defined within race and parental income group.

Table 1: Worker-level summary statistics, by respondent race and parental income

	Aggregate		Race			Parent Income Quintiles		
Outcome	Form 5500	Parent-Form 5500	W	B	H	Q5	Q3	Q1
Average age	41.21	30	42.06	40.14	39.4	30.04	29.97	30.02
	(9.918)	(3.613)	(9.99)	(9.727)	(9.649)	(3.588)	(3.601)	(3.628)
Box 1 W-2 total compensation	$72,810	$50,050	$81,310	$46,250	$51,150	$70,490	$46,540	$38,280
	(211,200)	(116,100)	(247,200)	(79,220)	(121,400)	(232,700)	(40,790)	(32,180)
Participation dummy	65.2%	59.5%	68.9%	56.3%	54.8%	69.5%	58.9%	50.9%
	(47.6)	(49.1)	(46.3)	(49.6)	(49.8)	(46)	(49.2)	(50)
Avg employer match ($)	$1,707	$1,069	$1,974	$856.7	$992.8	$1,746	$953	$658.5
	(2,806)	(1,795)	(3,058)	(1,603)	(1,885)	(2,420)	(1,582)	(1,291)
Employee contribution rate	3.8%	2.8%	4.2%	2.4%	2.6%	3.9%	2.6%	2%
	(4.6)	(3.5)	(4.8)	(3.3)	(3.6)	(4.2)	(3.3)	(2.8)
Employer match rate	1.9%	1.6%	2.1%	1.4%	1.5%	2.1%	1.6%	1.3%
	(2)	(1.9)	(2)	(1.8)	(1.9)	(2)	(1.9)	(1.7)
1099r withdrawal > $1000 dummy	13.5%	12%	12.3%	23.3%	14.5%	9%	12.3%	16.2%
	(34.2)	(32.5)	(32.8)	(42.3)	(35.2)	(28.6)	(32.9)	(36.8)
Foregone matching (% of income)	1.7%	2%	1.6%	2.1%	2%	1.6%	2%	2.4%
	(1.9)	(1.9)	(1.8)	(1.9)	(1.9)	(1.8)	(1.9)	(1.9)
Emply. + match contrib. rate,	8.8%	7.4%	9.1%	6.8%	7.6%	8.7%	7.1%	6.3%
(contrib. > 0)	(5.5)	(4.6)	(5.6)	(4.5)	(4.8)	(4.9)	(4.4)	(4.2)
Avg. Parental AGI		$91,720	$110,900	$50,620	$56,440	$237,900	$66,860	$14,750
		(269,400)	(312,400)	(58,970)	(213,000)	(572,300)	(11,880)	(69,410)
Number of unique individuals	1,722,000	471,200	1,220,000	181,000	194,000	99,720	96,110	83,760

Notes: The table reports summary statistics for wage earnings data from the Form 5500 (merged employee and employer data), which covers the 2008–2017 period. For more information about the different samples, please see Section 3.4.

Table 2: Wealth and wealth components, by population lifetime earnings bins

(a) By race

Value	Group	Q1	Q2	Q3	Q4	D9	D10
Wealth from employee contributions ($'000)	White	16.4	46.0	87.2	162.8	306.2	512.1
	Black	8.6	24.9	49.2	101.2	229.8	416.3
	Hispanic	11.9	34.0	70.4	140.4	276.8	482
Wealth from employer contributions ($'000)	White	6.4	18.8	35.8	65.0	121.4	238.9
	Black	4.4	12.9	25.4	47.8	102.3	193.4
	Hispanic	5.5	16.3	32.3	59.0	112.0	225.3
Wealth from tax subsidies ($'000)	White	3.7	13.1	27.1	52.7	98.6	252
	Black	2.3	8.5	20.0	43.5	85.6	198.6
	Hispanic	2.9	10.4	24.2	49.1	90.2	224.0
Total DC Wealth ($'000)	White	26.5	77.9	150.1	280.5	526.2	1003.0
	Black	15.3	46.2	94.5	192.5	417.7	808.3
	Hispanic	20.2	60.7	126.9	248.5	479.0	931.3
Social Security Wealth ($'000)	White	210.3	304.9	383.0	479.4	569.8	645.0
	Black	205.1	301.2	378.0	474.4	568.1	631.7
	Hispanic	214.8	301.0	378.8	476.5	568.0	637.7

(b) By parental income

Value	Group	Q1	Q2	Q3	Q4	D9	D10
Wealth from employee contributions ($'000)	Bin 1	12.2	36.3	71.5	139.4	269.8	452.2
	Bin 3	14.5	41.5	81.0	151.6	290.2	489.9
	Bin 5	17.2	45.9	87.7	169.3	323.0	545.7
Wealth from employer contributions ($'000)	Bin 1	5.3	16.2	31.5	57.4	109.2	213.6
	Bin 3	6.1	17.6	34.4	61.5	116.7	230.3
	Bin 5	6.7	19.0	36.2	67.9	128.4	250.6
Wealth from tax subsidies ($'000)	Bin 1	2.9	11.2	24.8	48.8	89.4	220.4
	Bin 3	3.4	11.9	25.6	50.4	94.2	238.7
	Bin 5	3.9	13.0	27.3	54.7	103.8	268.7
Total DC Wealth ($'000)	Bin 1	20.4	63.8	127.7	245.6	468.4	886.2
	Bin 3	24.0	71.0	141.0	263.5	501.1	958.9
	Bin 5	27.7	77.9	151.2	291.9	555.2	1065.0
Social Security Wealth ($'000)	Bin 1	210.0	308.6	388.5	486.1	572.6	637.8
	Bin 3	209.8	302.4	381.4	478.7	569.9	643.1
	Bin 5	210.3	299.2	375.4	471.8	565.8	647.0

Notes: This table presents average DC wealth (total and decomposed into its three components) and Social Security wealth by race (panel (a)) and parental income (panel (b)). The first sub-panel of each table shows average values for each component of DC wealth. The middle sub-panel gives total DC wealth. The third sub-panel is the average value of Social Security. Please note in panel Bins 1, 3, and 5 correspond to the bottom, middle, and top parental income quintiles. Columns show results by own lifetime earnings. There are six lifetime earnings bins—the bottom four quintiles and the top two deciles. Earnings bins are defined at the population level. Appendix Table A.3 gives the same analysis with lifetime earnings bins defined within race/parental income group.

Table 3: Change in DC wealth at retirement under the counterfactual tax and employer contribution policy, population bins

(a) By race

Value	Group	Q1	Q2	Q3	Q4	D9	D10
Baseline Total DC	White	26.5	77.9	150.1	280.5	526.2	1003.0
Wealth ($'000)	Black	15.3	46.2	94.5	192.5	417.7	808.3
	Hispanic	20.2	60.7	126.9	248.5	479.0	931.3
Baseline DC	B-W Gap	42.4%	40.7%	37.0%	31.4%	20.6%	19.4%
Wealth Gap	H-W Gap	23.8%	22.1%	15.5%	11.4%	9.0%	7.1%
Absolute change in	White	+18.4	+25.6	+23.6	+12.9	-10.8	-69.1
DC Wealth ($'000)	Black	+20.3	+34.1	+37.3	+30.9	+10.1	-58.5
	Hispanic	+19.5	+28.2	+27.2	+17.8	+2.6	-64.3
Counterfactual DC	B-W Gap	20.7%	22.3%	24.1%	23.9%	17.0%	19.7%
Wealth Gap	H-W Gap	11.7%	14.1%	11.3%	9.2%	6.6%	7.2%
Relative change in the	B-W Gap	-51.2%	-45.1%	-34.9%	-24.0%	-17.6%	1.6%
racial DC wealth gap	H-W Gap	-50.9%	-36.1%	-27.0%	-19.0%	-26.9%	0.2%

(b) By parental income

Value	Group	Q1	Q2	Q3	Q4	D9	D10
Baseline Total DC	Bin 1	20.4	63.8	127.7	245.6	468.4	886.2
Wealth ($'000)	Bin 3	24.0	71.0	141.0	263.5	501.1	958.9
	Bin 5	27.7	77.9	151.2	291.9	555.2	1065.0
Baseline DC	1-5 Gap	26.5%	18.2%	15.5%	15.9%	15.6%	16.8%
Wealth Gap	3-5 Gap	13.6%	8.9%	6.7%	9.7%	9.7%	10.0%
Absolute change in	Bin 1	+19.1	+28.9	+28.9	+21.4	+6.4	-28.2
DC Wealth ($'000)	Bin 3	+18.8	+27.0	+25.7	+16.6	-3.7	-53.8
	Bin 5	+18.7	+25.6	+22.7	+7.9	-20.9	-98.8
Counterfactual DC	1-5 Gap	14.9%	10.5%	9.9%	10.9%	11.1%	11.2%
Wealth Gap	3-5 Gap	7.8%	5.4%	4.1%	6.6%	6.9%	6.3%
Relative change in the	1-5 Gap	-43.6%	-42.2%	-36.0%	-31.0%	-28.8%	-33.3%
parental income DC wealth gap	3-5 Gap	-42.3%	-39.7%	-38.6%	-32.5%	-29.1%	-36.5%

Notes: This table presents the effect on wealth of our counterfactual exercise. Panel (a) gives results by race, and panel (b) gives results by parental income quintiles (bins) with Bins 1, 3, and 5 shown. Value row 1 in each panel shows baseline wealth. Value row 2 gives the baseline gap as a percentage of the White level (panel (a)) and the average level for those with the richest parents (panel (b)). Value row 3 shows the absolute change in DC wealth under the counterfactual. Value row 4 gives the counterfactual gap as a percentage of the White level (panel (a)) and the average level for those with the richest parents (panel (b)). Value row 5 gives the relative change in the percentage gaps obtained in moving from the baseline (value row 2) to the counterfactual (value row 4). In both panels, each row is divided into six bins–the bottom four quintiles and the top two deciles. Earnings bins are defined at the population level. Appendix Table A.4 gives the same analysis with lifetime earnings bins defined within race/parental income group.

A Data

Appendix A.1 introduces our three data sources: the American Community Survey (ACS), the administrative tax data, and our codified Form 5500 filings. Appendix A.2 defines the variables used in our analysis. Appendix A.4 introduces our data construction and outlines which years of data we use, how we define our samples, and how we weight. Appendix A.4 discusses the representativeness of our data.

A.1 Data sources

A.1.1 American Community Survey (ACS)

Our individual-level build begins with all American Community Survey respondents from 2008-2017. The ACS provides data on respondent age, education, gender, occupation, and county. We supplement this with administrative tax records and Form 5500 regulatory filings, which we introduce in the next two subsections.

A.1.2 Administrative tax data

Our tax data comes from data from form W2s, form 1099Rs, and form 1040s.

We obtain data on earnings and deferred compensation from form W2. We use Form W-2 data to measure earnings and deferred compensation. The W-2 extracts available at the Census Bureau have information from Box 1 on taxable wages, tips, and other compensation. These W-2 extracts also have an aggregate measure of deferred compensation from Box 12 that primarily consists of employee contributions to DC retirement plans. We cannot distinguish between contributions to different plans, but aggregate IRS data indicates that 93% of contributions that are to 401(k), 403(b), or 457(b) plans, the dominant DC employer-sponsored plans offered by the employers in the private, non-profit, and public sectors, respectively.[34]

[34] Our aggregate measure of Box 12 aggregates elective deferrals to plans under Box 12 codes D: 401(k), E: 403(b), F: 408(k)(6), G: 457(b), and H: 501(c)(18)(D). The items in boxes E-F (403(b), 408(k), and 457(b) plans) are DC plans that primarily differ from 401(k)s in which employers can provide them (such as nonprofits and local, state, and federal governments). 501(c)(18)(D) contributions cover future payments under certain defined benefit (DB) plans. From 2008 to 2018, the average share of those dollars by Box 12 Code are D: 76 percent, E: 12 percent, F: 0.1 percent, G: 5.6 percent, and H: 0.02 percent. See IRS Statistics of Income Tax States for Individual Information Return Form W-2 Statistics, Table 7.A of Internal Revenue Service (2023), accessed 09/20/2023.

We obtain data on withdrawals from DC accounts using form 1099Rs. Form 1099-R filings ("Distributions From Pensions, Annuities, Retirement or Profit-Sharing Plans, IRAs, Insurance Contracts, etc.") contain information on withdrawals from DC plans and payments from DB pensions. On the 1099-R extracts available to us, we observe the sum of withdrawals and distributions in two categories: 1) gross distributions from employer-sponsored plans and 2) IRA withdrawals.[35]

We link individuals to their spouses and parents using form 1040. We link individuals to 1040 tax filings (from the years 1994, 1995, and 1998-2017), both contemporaneously (in the year we observe their earnings) and for a subset of younger workers (under age 42 in 2020), to the 1040 filings of their parents when they were claimed as dependents. We include non-filers who do not receive W-2s. From the contemporaneous 1040s of tax filers, we can observe marital status (from filing status) and link spouses through the PIK of the other filer on the tax return. We then link the spouses' W-2s to observe their earnings as well. Section A.2.2 provides more details about how we make this link.

To construct intergenerational linkages and observe parental resources, we use the dependent information on 1040 tax returns, which is available for 1994, 1995, and from 1998 onwards. We create a dependent claiming history that identifies any parent(s) that claimed each individual at all observed ages up to 18. Therefore, we can link individuals with their parents, conditional on the parents filing a 1040 in which they claim them as a dependent at some point during their childhood.

A.1.3 Retirement plan data

The data set that we construct in this paper uses the fact that that all retirement plans must submit an annual regulatory form (Form 5500) to the federal government. For plans with more than 100 participants, this form must include a narrative description of the retirement plan characteristics including details on the the match schedules, vesting schedules, and auto features. These descriptions have been made publicly available by the Bureau of Labor, but in their original form (free-form text) they are not amenable to empirical analysis.[36] The data set that we use (described further in Arnoud et al. (2021), and Choukhmane et al. (2023)) was constructed from these files for the largest 5,000 defined contribution plans and a random sample of 1,000 smaller plans. For completeness, we reproduce several details

[35]The IRS also excludes distributions, such as direct rollovers, Section 1035 exchanges, and Roth conversions from the 1099-R extract we use. For more information on the 1099-Rs, including separating DB and DC plans in the data, see Bee and Mitchell (2017).

[36]See `https://www.dol.gov/agencies/ebsa/about-ebsa/our-activities/public-disclosure/foia/form-5500-datasets`.

about the Arnoud et al. (2021) data construction here. Details for each plan were codified in a consistent fashion. The plan-level data contain details on the full matching schedule, the vesting schedule, and any automatic features (auto-enrollment or auto-escalation). These very large firms cover a large number of employees—in 2017, 37 million employees were eligible to contribute to one of these large plans, collectively accounting for 55% of the population of workers enrolled in private and non-profit sector DC retirement plans.

We link these plan-year level variables to the Census firm infrastructure via a multi-stage procedure which incorporates information on numeric identifiers such as EIN and telephone number as well as fuzzy matching on name and address fields. We are able to match around 5,000 plans and 35,500 plan-year combinations. We drop firms that have different match formulas for different employees, that change match formulas mid-year, or for which we cannot find match formulas. As a back-stop to our fuzzy linking, we further conduct internal consistency checks with our universe of W-2 filings, described further in section A.3.2 below, which leaves us with about 3,800 unique plans and 21,500 plan-year combinations.

A.2 Variable definitions

A.2.1 Outcome variables

All variables in dollar terms are deflated to base year 2017 using the Consumer Price Index provided by the Bureau of Labor Statistics.[37]

Employee contributions This is deferred compensation reported in Box 12 of the W-2 tax form. This amount generally corresponds to contributions to an employer-sponsored contribution plan (such as a 401(k) plan).

Employee contribution rate The employee contribution rate is the percentage of salary, using the ratio of the real employee contribution reported in Box 12 divided by the sum of the real taxable wage reported in Box 1 of the W-2 and the real employee contribution. The formula is $\frac{\text{employee deferred compensation}}{\text{employee deferred compensation} + \text{employee W-2 wages}}$. We additionally refer to this variable as "Own contrib. (% of inc.)" in the output above.

Participation rate A dummy equal to one if the individual makes a positive contribution to a retirement savings plan. This measures contributions on the extensive margin. We additionally refer to this variable as "Positive contribution dummy (%)".

[37]See `https://www.bls.gov/cpi/research-series/r-cpi-u-rs-home.htm`.

Employee contribution plus employer matching contributions This is the sum of real employee contributions and the imputed match contribution implied by the employer matching formula collected from the employer's Form 5500 filing. If an individual works more than one job, we match the employer matching formula to the highest-salary job. We apply the match formula to the three highest-earning jobs separately. We then aggregate the imputed contribution to generate the real employer match contribution. This is then added to the real employee contribution for the combined employee and employer matching contributions. The formula is $\frac{\text{employee deferred compensation + employer match}}{\text{employee deferred compensation + employee W-2 wages}}$. We additionally refer to this variable as "Employee plus employer matching contributions," "Own plus match contribution (% of income)," or "Employee contribution + employer match (% of income)".

Early withdrawals We observe DC-plan withdrawals (and payments from pension plans) in Form 1099-R filings, which we treat as potential early withdrawals from DC plans. We take early withdrawals from the year after individuals appear in the ACS survey. We apply three key restrictions: 1) individuals must contribute more than \$1,000 in deferred compensation in the four years prior to early withdrawal, 2) individuals must withdraw more than \$1,000 to be classified as an early withdrawal, and 3) individuals must be younger than 55 *at the time of* the early withdrawal. We apply the first and second restrictions as federal law allows employers to automatically disburse individuals with under \$1,000 in deferred compensation upon separation. The third restriction relates to the tax penalty for taking an early withdrawal-individuals 55 years and older are allowed to take early withdrawals without incurring the tax penalty. We additionally refer to this variable as "Positive withdrawal dummy (withdrawal >\$1,000)".

A.2.2 Additional Observable Mediating Variables

Year The ACS provides the survey year.

Age bin We generate age from the ACS birth years and the ACS survey year. We bin people into ages 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, and 55-59.5.

Income bin Income is defined as the sum of total real Box 1 wages and Box 12 deferred compensation on W-2 filings. We generate income deciles from the total compensation distribution per year and individual's age, incorporating ACS weights.

Education We generate four educational categories from the ACS education variable, corresponding to whether a respondent has completed less than a high school degree, is a high school graduate, has a college degree, or has a graduate degree. Those who have some college but did not graduate college are included in the high school graduate category.

Gender The ACS provides gender (categorized as either male or female) for the 2001-2019 surveys. We generate a dummy for female.

Occupation The ACS provides several hundred occupational categories. The IPUMS 2010 crosswalk provides occupation codes that are consistent over time. We match the ACS occupation codes to the consistent IPUMS 2010 codes, matching 12,260,000 out of 12,480,000 PIKs in our full ACS sample.

County The ACS provides the county of residence, so we construct categorical variables capturing the concatenation of a FIPS state code and the county code.

EIN W-2 forms are filed by employers who are required to report their Employer Identification Numbers (EINs). We take the EIN for the highest-earning job if an individual worked multiple jobs. We associate a worker with the retirement plan characteristics of the highest-earning EIN in the survey year.

Tenure Tenure is constructed by matching all ACS individuals with their employers from 2005-2020. Our W-2 filings report employers (by EIN) in order of most wages earned. We take the earliest known year for each individual-employer combination. We match the start year with the individual's first EIN (employer from whom the individual earned the highest earnings) during the ACS survey year. Since our universe of W-2s begins in 2005 and our build begins in 2008, to avoid censoring issues we classify tenure at the main employer into four main categories: 1) working less than one year, 2) between 1-2 years, 3) between 2-3 years, and 4) at least 3 years.

Family structure We construct family structure from 1040 filings. The five main groups are single filer, no kids; single filer, with kids; dual filer, no kids; dual filer, with kids; and non-filers. Non-filers are those individuals who may receive W-2s but either forget or choose not to file 1040s.

Spousal income Spousal income is linked using 1040 filings from the ACS observation year. Spousal income is the sum of total real Box 1 wages and Box 12 deferred compensation from W-2 filings. Spousal income bins are classified by year and age using ACS weights into 12 main indicators: i) 0 percentile (spouses who report $0 in earnings), ii) 10, 20, ..., 100 percentiles (spouses for whom we have nonzero earnings), and iii) missing (individuals who are either single, non-filers, or for whom we cannot match spousal income).

Auto-enrollment Auto-enrollment is taken from our universe of W-2 filings and matched firm data. Our Form 5500 filings report whether a 401(k) plan offers auto-enrollment in a given year. We classify Form 5500 filings that do not report an auto-enrollment start date after 2005 as not offering auto-enrollment. Individuals who start at their main firm after firms enact an auto-enrollment policy are classified as having auto-enrollment. Individuals who start at their main firm before an auto-enrollment policy begins or work at firms without auto-enrollment policies are classified as not having auto-enrollment. Due to censoring issues, individuals who are observed starting at a firm in 2005 and work at firms where auto-enrollment begins either before or during 2005 are classified as unknown.

Parental income Parental income is defined as real adjusted gross income for parents that we can link to ACS respondents in 1040 filings. They are linked closest to when a person is claimed at age 16. We generate parent income bins by year and child's birth year (as a proxy for child age) from W-2s. Note that we do not incorporate ACS weights in our calculation of parent income bins.

DC Access We construct an indicator for working at an employer offering a DC account from the universe of W-2 filings. We define a firm as offering a DC plan in a given year if at least 5% of its employees have positive deferred compensation.

A.3 Data construction

A.3.1 Years

While some of our data cover a broader time range, we restrict our analysis to individuals observed in survey years from 2008-2017 due to censoring issues. Given that our W-2 filings begin in 2005, two key variables, job tenure and early withdrawals, depend on having a panel of at least four years. Job tenure is categorized into < 1 year, 1 year, 2 years, and 3+ years. Early withdrawals condition on more than $1,000 in nominal deferred compensation over the four years prior to the early withdrawal. Both require W-2 Box 1 and EIN information from the three years prior to appearing in the ACS survey. Including pre-2008 individuals would select for higher income employees who can contribute more in a given year and would attenuate their tenure. We cap our observation years at 2017 due to our retirement plan-EIN crosswalk ending in 2017. Since the early withdrawal probability is computed in the year after the ACS year, our estimates for the probability of taking an early withdrawal are computed using data from 2009-2018.

A.3.2 Sample restrictions imposed throughout the paper

We apply three restrictions to compensation and control variables before running regressions. First, we restrict attention to survey respondents who are between ages 25 and 59.5 in the ACS survey year. Second, for compensation, we require the nominal sum of Box 1 wages and deferred compensation to be greater than $8000 and require Box 1 wages to be strictly greater than $0. This eliminates people who have zero wages but have deferred compensation (likely people with high wealth who are exploiting employer matches or instances in which box 1 wages were incorrectly parsed into the Census database). Finally, we require all potential mediating variables to be nonmissing to ensure consistency across all regressions.

For analyses which use the linked information about retirement plans, we also impose a restriction which checks for internal consistency of the administratively reported level of employer contributions coming from Form 5500 with a comparable measure that we calculate internally by applying the matching formulas to the population of deferred compensation levels coming from linked W2 forms. To reduce the potential impact of linking/measurement errors, we restrict analysis to plans for which difference in the calculated ratios of employee contributions to total contributions obtained from the two sources is smaller than 15 percentage points. Figure A.2 illustrates the strong concordance between two firm-level measures in our analysis sample by reporting a binned scatter plot comparing our imputations of employer contributions on the vertical axis with the actual reported Form 5500 measures on the horizontal axis. The two measures are very highly correlated, though there is a modest discrepancy for firms which have very low employee shares of contributions, likely due to the presence of additional, nonelective (non-matching) contributions which are excluded from our measurement.

A.3.3 Samples

We have two primary samples:

1. **Form 5500 sample** This is our main sample for our analysis of gaps in saving by race. It contains all individuals in the ACS for whom we match Form 5500 filings that meet our match formula and internal consistency restrictions. Analysis using this sample uses combined ACS and firm-level analytic weights, discussed below. The total number of unique individuals (after dropping missing individual variables required for our analyses) is approximately 1,722,000.

2. **Parent-Form 5500 sample** This is our main sample for our analysis of gaps in saving by parental income. It contains all individuals in the Form 5500 sample who are born after 1978 and to whom we can match a non-missing level of parental income. Analysis using this sample uses combined ACS and firm-level analytic weights, discussed below. The total number of unique individuals (after dropping those missing individual variables required for our analyses) is approximately 447,500.

To assess whether selection into the Form 5500 linked employer-employee sample matters for our findings, we compare summary statistics and key results from our baseline sample to two broader samples. These are our 'Full ACS sample', which contains all individuals in the ACS, and our 'DC Access sample' which contains all respondents at firms where at least 5% of employees report deferred compensation (and for which we therefore assume offer an employer-sponsored DC account).

A.3.4 Weighting

Individual weights. The ACS microdata include person-level analytic weights which enable researchers to produce estimates which are representative of the US population. For regressions and other summary statistics which do not use any linked retirement plan information, we use these person-level weights to construct estimates which are nationally representative.

Note that several of the individual-level income variables are converted into deciles or quintiles. To construct these categories, we first apply earnings and age restrictions (as explained in Appendix A.3.2), then compute weighted percentiles to use as breakpoints by year and age using the ACS sample weights. These decile assignments are therefore computed to be representative of all people in the ACS who match our sample requirements.

Retirement plan weights. Our employer data combines two different samples: a certainty sample of the five thousand largest firms and a random sample of two thousand plans from all remaining firms. For the random sample, we sampled firms with probabilities proportional to the number of participants. Since many of large firms above the certainty threshold were also selected in the random sample, the random sampling procedure yields an additional one thousand firms. To ensure that our estimates are representative of the full population of firms filing the long version of form 5500, we calculate firm-level weights which are equal to the inverse of the probability of being selected into our sample.

Combining individual and plan weights. In our analyses which link the ACS and plan-level data (e.g., the Form 5500 sample) we combine person and firm-level weights to compute a combined measure to use at the individual level. Each individual's probability of appearing in our matched build is the joint probability of being in a sampled firm and a sampled employee. Since the two samples are drawn independently of one another, the matched individual's probability weight is the product of the ACS probability weight with the plan probability weight.

A.4 Data representativeness

This section presents some additional information which speak to the representativeness of our results, most of which are computed for a sample of ACS workers who are linked with firms whose retirement plans are included in our sample. Here, we characterize some differences between our analysis samples and broader populations of US workers.

Table A.1 provides information on a number of summary statistics which are computed for various samples. We begin with the set of workers who are in the ACS and satisfy the basic income and age restrictions we impose throughout (see Appendix A.3.2). As we move from left to right in the table, we see how sample means as we impose additional restrictions which are required to perform our analysis. Moving from column 1 to 2 shows the impact of requiring that all individual observable characteristics are available, which is only associated with very modest reductions in the sample size and changes in sample means, respectively.

Moving from column 2 to column 3 imposes a more substantive restriction, namely that the worker receives income from an EIN for which at least 5% of its employees report positive levels of deferred compensation, which is our administrative proxy for having DC access. Imposing this restriction is associated with higher income and higher average savings rates, driven by higher participation on the extensive margin.

Column 4 imposes that we can successfully link ACS respondents to a plan in our form 5500 sample. Relative to column 3, this sample excludes workers in small firms because employers with less than 100 employees are not required to submit a detailed description of their plan alongside their Form 5500 filing. Our sample weights are intended to make our estimates representative of the set of workers who are employed at the set of firms that offer DC plans and have more than 100 employees.

Overall, we see some modest changes in sample means between the sample of all ACS respondents with DC access and our Form 5500 sample (for which we observe retirement plan details). In the latter sample of larger employers, workers earn about $5,000 more in labor income and save at slightly higher rates (mostly driven by a higher probability of having positive contributions), and we also see that the propensity to take early withdrawals is slightly higher in this sample. The similarities between the two samples suggest that our estimates from the Form 5500 are quite representative of the broader population covered by the nationally representative ACS. Appendix Figure A.7 further shows that racial gaps in contribution and withdrawal behavior are also very similar across our full sample of ACS respondents with DC access and our form 5500 sample.

Finally, column 5 additionally restricts to the subset of individuals with parental income available. Given that we can only match younger cohorts to their parents, this sample is unsurprisingly younger, has lower earnings, and saves at lower rates.

B Discussion of Mediating Variables in Regressions

B.1 Rationale for and impact of mediating variables in regression

In Section 4 of the main text, we report estimates of gaps in contributions by race and parental income which include a number of observable characteristics, including dummies for age, year, deciles of labor income, gender, educational group, occupation, county, employer identification number (EIN), and tenure bin. In this section, we discuss potential economic rationales for why each of the individual-level characteristics that we include may impact DC savings rates, as well as the relationship between these variables and average savings rates in our data.

Year: Recent years have seen a substantial evolution in the DC landscape (e.g., the growth of auto-enrollment). To account for these, as well as savings differences over the business cycle, we include year fixed effects. However, we do not expect (and do not find) that the inclusion of year fixed effects affects our gaps as the composition of race and parental income groups is quite stable over our sample period.

Age: Age is an important driver of retirement saving; financing consumption in retirement is likely to be a central financial objective for older workers, whereas younger workers face a number of other competing savings objectives. Black and Hispanic workers are, on average, younger than White workers, and so understanding the extent to which age differences account for the gaps we observe is important. We find, as expected, that savings rates are increasing in age (see, e.g., Gourinchas and Parker, 2002). Panel (a) of Appendix Figure A.9 helps visualize why accounting for age impacts differences in contribution gaps by race. It shows the age distribution among Black and Hispanic workers in our sample tends to skew younger (as shown in the bars, read off right axis), and younger workers save less on average (as shown in dots, read off left axis).[38]. These two facts together contribute to the inclusion of age having an attenuating effect on the race gaps.

Income: Income has been a traditional focus of the regulatory system and of the literature on the distributional analysis of the U.S. retirement system. It is well established that the rich save more (Dynan et al., 2004), and there are many reasons why this would be the case. Replacement rates from Social Security decline in income, the tax benefits are higher for those

[38]The former give the ratio of the share of each race group in each age bin to the the share of the population in that age group. The latter are point estimates from our fully saturated regression with combined employee and employer saving as the dependent variable

facing higher marginal tax rates (Congressional Budget Office, 2021), income risk tends to decline with income over most of the distribution outside of the top decile (Guvenen et al., 2014), and financial literacy is typically increasing in income (Lusardi and Mitchell, 2014; Lusardi et al., 2017). Furthermore, there are well-established differences in the distribution of income across races and by parental income (see Figure 3(b), and so we find (as expected) that including income in the regression attenuates the gaps. Panel (b) of Appendix Figure A.9 helps to visualize why accounting for income impacts differences in contribution gaps by race. It shows that Black and Hispanic workers are overrepresented at the bottom of the income distribution (as shown in the bars, read off the right axis), and that poorer workers save less on average (as shown in dots, read off the left axis). Taken together, these two facts illustrate why accounting for income differences reduces the estimated race gaps.

Education: Educational attainment could affect saving through channels beyond its correlation with income levels: life-cycle trajectories in expected income levels and income risk vary with education, and financial literacy increases in education. We consider the role of the highest degree attained, which we capture via four dummies for less than high school, a high school degree, a college degree, and a graduate degree. We find a strong relationship between educational attainment and savings. Conditional on other worker-level characteristics, those without a high school diploma or equivalent contribute 0.19 p.p. less to a DC account than those with a high school diploma, 0.83 p.p. less than those with a college degree, and 1.2 p.p. less than those with a graduate degree. Panel (c) of Figure A.9 shows these coefficients, along with an illustration of the education distribution by race. Those without a high school degree are over-represented in the Hispanic population, which means that education plays a quantitatively important mediating role in the Hispanic-White savings gap.

Gender: Men and women may save different amounts for a variety of reasons such as differences in life-cycle earnings profiles (Goldin, 2021), risk preferences, life expectancy, and/or expected retirement benefits (Barber and Odean, 2001; Watson and McNaughton, 2007). We find that, conditional on the full set of individual and household characteristics we include, female workers are 4.3 p.p. more likely to participate in and contribute 0.55 p.p. of salary more to DC accounts than men. Given that gender ratios are similar for workers across the racial and parental income groups we consider, gender has little impact on our estimated contribution gaps.

Occupation and County: Occupation may be relevant for savings as it can correlate with expected future earnings, income risk, and potential differences in risk or time preferences.

Racial and parental income distributions differ across space, which may correlate with various factors such as the cost of living in retirement, so we additionally absorb county fixed effects.

Employer (EIN): Our data allow us to absorb EIN fixed effects, which enables us to identify racial contribution gaps among coworkers within the same employer. In addition to a number of economic characteristics that may differ across firms (for example, expected income trajectories and employment stability), a natural possibility is that workers sort into firms that differ in terms of the quality of the retirement benefits that they offer. For example, there is substantial heterogeneity across firms in the generosity of matching incentives, the nature of vesting schedules, and auto-enrollment and other default policies. Absorbing EIN fixed effects allows us to hold many of these features constant.

Tenure: The final economic characteristic that we consider is job tenure, which we split into bins for less than 1, 1, 2, and 3+ years. Tenure may relate to saving through its correlation with employment risk (e.g., Farber, 1994, shows that the probability of job separation decreases for workers with higher tenure), the probability that a worker's contributions will vest, and workers' awareness of plan benefits, among other channels. Conditional on other individual characteristics, employees with one year of tenure contribute 0.47 p.p. of salary more to a DC account compared to employees with less than a year of tenure, while employees with at least three years of tenure save 1.8 p.p. more.

B.2 Relaxing the assumption of additive separability

It is important to recognize that in Sections 4 and 5 the gaps with the inclusion of individual characteristics are accounted for using an additive specification. To evaluate whether this additivity conceals consequential interactions between characteristics, we rerun the analysis by reweighting the cells based on observables, so the Black and Hispanic worker distributions match the White worker distribution. To ensure that cells have full overlap across groups, we present estimates for the first five regression controls, from raw gaps through gender. Figure A.6(a) shows a cascade with the reweighting that is similar to that in Figure 2(a) for all the individual characteristics up to gender, and the qualitative lessons are unchanged. We also conduct an analogous exercise for early withdrawals: the estimates from Figure A.6(c) are largely the same as the ones from Figure 4(a), which assumes additive separability.

C Comparison with Survey of Consumer Finances

We reproduce our baseline analysis using data from the Survey of Consumer Finances (SCF), the gold-standard source of survey information on wealth in the U.S. Given that the SCF does not contain information on parental background, we focus on differences in retirement contributions by race.

Sample. We use data from the 2010, 2013, and 2016 waves of the SCF, which cover a similar period to our administrative data. We impose the same restrictions as in our baseline analysis using administrative data: we focus on respondents aged 25 to 60 who make at least \$8,000 in wage income. For the regression analysis we further restrict the sample to those who report having access to a DC plan through their employer. This restricted sample contains 4,097 respondents across the three SCF waves, of whom 512 are Black and 338 are Hispanic.

Descriptive statistics. Table A.2 compares summary statistics across the SCF sample and our sample of ACS respondents linked with administrative tax records. Demographics are broadly similar across the two samples, although the SCF sample has slightly higher labor earnings. Access to, and participation, in DC plans are significantly lower in the SCF (respectively 49.6% and 35.3%) relative to our ACS sample (78.4% and 45.6%). This is consistent with Dushi and Iams (2010) finding that survey responses underestimate access and participation in DC plans. They find that access to and participation in a DC plan are measured to be, respectively, 17 p.p. and 11 p.p. higher when complementing responses to the 2006 Survey of Income and Program Participation (SIPP) with respondents' W2 records. The National Compensation Survey (NCS), which is based on responses from employers rather than employees, also reports higher levels of participation and access than the SCF (Topoleski (2018)). Among full-time civilian workers in the 2017 NCS—who are more comparable to our sample of workers earning more than \$8,000—68% have access to, and 48% participate in a DC plan.

Regression results. Figure A.14 compares the racial gaps in employee contributions estimated in the SCF to those estimated in the administrative data. Our specification using SCF data is the same as in the first three steps of the baseline regression cascade (reported in Figure 2(a)), with standard errors adjusted for both imputation and sample variability errors.[39] Contribution gaps are qualitatively similar across the two datasets, although confidence intervals are much wider for the estimates using the SCF. In particular, it is hard to make precise statements about gaps conditional on age and income using the SCF: confidence intervals are large and not only overlap with zero (in the case of the Black-

[39]Income bins in the SCF are constructed by year and 5-year age bins.

White gap) but also come close to and even overlap with our estimates using administrative data (in the case of the Hispanic-White gap). This suggests that survey data that has been typically used to study this question might be underpowered to detect (even sizeable) differences in retirement contribution rates by race.

D Gaps in retirement wealth accumulation by characteristics other than race and parental income

Our primary focus in this paper is to measure and better understand the distributional impact of savings gaps by race and parental income. We documented large gaps in saving along these characteristics within income groups. These differences in saving rates generate differences in remuneration across workers and the incidence of tax subsidies.

In Section 4.1, we discuss the impact of accounting for the potential mediating role of different observable characteristics to our estimates of residual contribution gaps by race and parental income. In this short appendix, we discuss the extent of the independent association of several of those mediators with saving, and therefore the extent of their association with saving incentives. The aim is to illustrate that the distributional point we make is qualitatively relevant for characteristics beyond race and parental income: just as the matching and tax subsidies associated with the current system will disproportionately accrue to White workers relative to their Black and Hispanic coworkers, these subsidies will also disproportionately accrue to other groups with higher saving rates.

Figure A.13 shows the coefficients obtained from estimating a version of equation (1) where we quantify contribution (left panels) and early withdrawal gaps (right panels) for different groups. Analogous to the bottom two panels of Figure 2, we report raw differences in means as well as coefficients from a "fully saturated" specification which includes additional individual and household-level variables.

We begin by reporting in panels (a) and (b) gaps by decile of parental income, which provides additional resolution about gaps we reported earlier using coarser bins and for a subset of parental income quintiles in Figure 2. Whereas raw gaps in both contributions and early withdrawals are close to linear in parental income ranks, residual contribution gaps by parental income from the saturated regression are more strongly increasing in parental income ranks for the higher deciles of the parental income distribution. By contrast, early withdrawal gaps increase fairly consistently across the parental income deciles. In our saturated specification which, there remains a gap of around 4 p.p. in the probability of taking an early withdrawal between individuals with parents in the bottom decile relative to the top decile of parental income, around 30% of the unconditional mean.

Panels (c) and (d) show a similar analysis where we sort workers into bins based on four categories of educational attainment. The raw gap in employer + employee contributions between high school graduates and those with a graduate degrees is almost 4 pp, which shrinks to around 1.5 p.p. in our saturated specification. Likewise, we also see evidence for a stronger demand for liquidity among those with lower levels of education.

Next, panels (e) and (f) analyze gaps by our measure of family structure, expressed relative to households with two adults but no children. Since non-filers and single parents tend to save less, they will tend to participate less and therefore enjoy fewer matching and tax subsidies relative to their coworkers. Likewise, we find that single parents and non-filers take early withdrawals at substantially elevated rates. These gaps are fairly similar in both raw and saturated specifications, suggesting that the association between liquidity demand and family structure is fairly distinct from other mediating channels.

These analyses further emphasize that the current system of retirement incentives redistributes in systematic ways, overall and conditional on income, across many economic and demographic characteristics which are correlated with savings.

E Micro-simulation Model

E.1 Overview

To understand the implications of differential saving and match patterns over the whole life cycle, we need full life cycles of data on retirement plan access and DC plan withdrawals in the population. However, we have a maximum of 13 years of observations per individual. We use these partial life cycles and a simple hot deck imputation strategy to construct panels of synthetic life cycles, described in Section E.2.

With this data, we develop a micro-simulation model, described in Section E.4, which has three objectives. The first is to use the data on observed flows (earnings, contributions to DC accounts, and withdrawals from DC accounts) and a model of the economic and policy environment to generate simulated data for objects that we do not directly observe: the stock of resources for retirement, Social Security entitlements in retirement, and the trajectory of withdrawals from retirement accounts.

The second objective is to evaluate the counterfactual differences in wealth at retirement in a world where the individual saved in a taxable brokerage account rather than the tax-advantaged DC account. This allows us to build a measure of the value of tax expenditure at the individual level and its distributional incidence.

The third is to evaluate the distributional impact of changes to retirement savings institutions in the U.S. We consider three counterfactual policies. In the first, we break the link between saving and remuneration by calculating each firm's counterfactual employer contribution that, if paid to every employee in proportion to their earnings, would cost the same to the employer as their current matching contributions. We evaluate the distributional impact of moving from the status quo to a system where all employees received that same proportional contribution. The second counterfactual setting breaks the link between government contributions to retirement accounts and savings choices by redistributing the tax expenditure so that it is proportional to lifetime income, once again regardless of the taxpayer's retirement savings choices. The third counterfactual combines both reforms. In the interest of brevity, in the main paper, we focus on the combined counterfactual, but show selected results in the appendix for the individual match and tax counterfactual.

E.2 Modeled lifetime paths of earnings, retirement plans, and withdrawals

To estimate our micro-simulation model and evaluate the distribution of tax and wealth impacts of Defined Contribution (DC) retirement plans, we need to capture the distribution of paths of earnings, retirement plan access, and DC plan withdrawals in the population. However, our data are limited in several respects. First, for many workers who are now close to retirement, DC plans were not in wide use at the onset of their working career. Furthermore, Form W-2s, our data source for individual wage and salary earnings and contributions to DC plans, are only available starting in 2005. Our information on plan characteristics from the Form 5500 is only available through 2017. This leaves us with at most 13 years (2005 to 2017) to simultaneously observe earnings and DC contributions from W-2s, plan characteristics and matching from the Form 5500s, and retirement account withdrawals on Form 1099-Rs. Our aim is to convert these shorter windows of information into plausible lifetime trajectories spanning a working life cycle from age 25 to 65.

To construct the plausible lifetime trajectories, we use a simple hot deck imputation strategy. We partition ages starting at age 25 into overlapping bins of 4 years (25-28, 27-30, 29-32..., 63-66). For a given age bin b, we observe their ages at t, $t+1$, $t+2$, and $t+3$. For individuals in bin $b+1$, we observe their ages in $t+2$, $t+3$, $t+4$, and $t+5$. We use the information from individuals in bin $b+1$ to impute earnings, DC plan access, contributions, characteristics, and withdrawals to individuals in bin b. We do so by matching individuals in bin b to similar individuals in $b+1$ using the information observed at the overlapping ages ($t+2$ and $t+3$) and appending the information from the later non-overlapping age ($t+4$

and $t + 5$) to bin b individuals.

As an example, suppose Person A had annual of \$25,000 at age 25, increasing \$1,000 each year to \$28,000 at age 28. Their employer did not offer a 401(k) plan and thus the person made no contributions to or withdrawals from a plan. Now suppose Person B earned \$26,500 at age 27, with annual increases of \$1,500 so that their salary was \$31,500 at age 30. Person B likewise had no access to a DC plan. Persons A and B had similar earnings and plan access during their observed overlapping ages, such that $y_{A,27} = \$27,000$ and $y_{A,28} = \$28,000$ compared to $y_{B,27} = \$26,500$ and $y_{B,28} = \$28,000$. As these workers had similar observable characteristics during their overlapping years, we impute to Person A the salary and contributions information from Person B for ages 29 and 30. This allows us to lengthen the number of years of "observed" earnings for Person A from four (covering ages 25-28) to six (covering ages 25-30). We can then repeat this process by imputing earnings for Person A at ages 31 and 32 using individuals in the next age bin covering ages 29 to 32. For a visual representation of how this works in practice, see Figure A.24. By repeating this process, we construct synthetic lifetime "histories" of earnings, DC plan access, and employee and employer plan contributions.

For early retirement withdrawals by working-age individuals, we conduct an additional imputation step. Here, we impute withdrawals relative to contributions in the prior years to better align withdrawal amounts to contributions. This helps reduce the number of cases in the model where withdrawals substantially exceed recent contributions. However, because we do not observe returns or contributions in the distant past, there will be many cases in the data where withdrawals exceed recent contributions even with contributions observed over a longer time horizon than we use in the imputation.

E.2.1 Imputing DC plan access and matching rules for all firms

The hot deck model described in Section E.2 requires information on firm matching rules and DC plan availability for all firms. However, our data set of firm matching schedules from publicly available Form 5500 filings covers only a subset of firms, including the largest approximately 5,000 firms and a random sample of 1,000 smaller firms. We use this data to impute DC access and plan matching rules for the remaining firms. Because we are interested in simulating lifetime trajectories for workers under the current system, we restrict to the plan characteristics in the most recent year for each firm linked to the Form 5500. For all firms, we summarize the distribution of deferred contributions across their workers. As an example, suppose that in a given firm 90% of workers have 0 deferred compensation and 10% contribute exactly 3 percent of their earnings to a DC plan. We summarize the share of workers in each firm that contribute between 0% and 10% of their earnings to DC plans with separate bins for 0 contribution and > 10 percent (i.e., bins of 0, (0-1) percent, [1-2) percent, [2-3) percent, and so on). We use kmeans clustering to separate firms into 10 distinct groups based on the distribution of worker deferred contributions in these bins. Finally, we impute DC plan access and firm match schedules to those firms without available Form 5500 data using a hot deck matching on the worker DC contribution clusters, firm size, and average earnings for workers. This means that if two firms, A and B, have a mass of contributions at around 3 percent of earnings, they are likely to be in the same worker contribution cluster. Suppose Firm A has plan details available from Form 5500, with matching contributions of 100 percent up to 3 percent of earnings and 0 percent thereafter. Suppose further that Firm B, on the other hand, does not have available Form 5500-based plan information. Firm A would then be a likely "donor" of its match schedule to Firm B.

E.3 Summary and Output

The result of this procedure is a simulated data set for individuals i age $t \in \{25, \ldots, 90\}$, where 90 is assumed to be the last age of life and in which mortality is deterministic.

Variables that we observe (with the associated notation given for objects that will feature in the treatment below) are:

- Demographic measures: age (t), race, and parental income
- Compensation measures: earnings (e) and contributions the employee elects to make to their employer-sponsored defined contribution account (dc^{ee}),
- Whether the individual works in a firm offering a DC plan and, if so, the match schedule ($dc_f(.)$), and
- Withdrawals from DC accounts before retirement (w).

E.4 Model Description

E.4.1 Savings Vehicles

Central to the exercise is to compare outcomes under the status quo (in which the deferred compensation is paid into a tax-deferred defined contribution account) with a counterfactual setting (in which tax-favored DC accounts are not available, and those same contributions are instead paid into a (taxable) brokerage account). We evaluate each individual's savings trajectory under two systems of taxation, indexed by $j \in \{DC, \ BK\}$. The superscript $j = DC$ indicates that the individual is saving in a tax-deferred 401(k) account, and $j = BK$

indicates that they are saving in a brokerage account. Savings in the tax-deferred (DC) account benefit from the fact that income tax is deferred until the funds are withdrawn and that investment returns accumulate free from income and capital gains taxes. Savings in the brokerage account come from taxed income, have returns that are subject to tax, and have income tax-free withdrawals.

Below we refer to the 'DC saver' and the 'brokerage saver' as shorthand for the saver in a setting where DC accounts are available and not, respectively.

E.4.2 Observable: Earnings, contributions, and withdrawals

Employees receive compensation that can be divided into earnings $e_{i,t}$ and deferred compensation $dc^{ee}_{i,t}$. Employees may also receive an employer match, which is a firm-varying function indexed by f: $dc^{er}_f(ee_{i,t})$. For ease of notation, we suppress the dependence of the employer contribution on the employee contribution and denote the employer contribution made on behalf of individual i at age t as $dc^{er}_{i,t}$.

Withdrawals from retirement accounts are denoted $w^j_{i,t}$, with j indexing the nature of the account (DC or brokerage). We observe withdrawals made by our agents up to age 65. These observed withdrawals in the data are from the DC account and recorded before the deduction of income tax.

E.4.3 Wealth

Wealth balance at the beginning of the period is given by $B^j_{i,t}$ and is initialized to zero at age 25. Net flows into the wealth vehicle are denoted by $f^j_{i,t}$:

$$f^j_{i,t} = dc^{ee}_{i,t} + dc^{er}_{i,t} - \tau^{c,j}_{i,t} - w^j_{i,t}, \tag{2}$$

where dc^{ee} and dc^{er} are, respectively, deferred compensation by the employee and the employer-match contributions. There are two deductions from these gross flows. The first ($\tau^{c,j}$) are taxes on these contributions. This object will be defined in detail below, but, in brief, note that dc^{ee} and dc^{er} are measured as gross-of-tax. For the DC saver, no income tax is owed on these flows and so $\tau^{c,DC}_{i,t} = 0$. For the brokerage saver, income tax must be paid before contributions are made. The second deduction, $w^j_{i,t}$, are withdrawals from the account. These are observed before the age of 65; in Section E.4.6, we propose a model of withdrawals which fills these in for after the age of 65.

The law of motion for wealth balance is given by:

$$B^j_{i,t+1} = (B^j_{i,t} + f^j_{i,t})(1 + \rho_t) - \tau^{r,j}_t, \tag{3}$$

where ρ_t is a rate of return that depends on age (with time dependence due to the changing mix of assets in the portfolio), and $\tau_t^{r,j}$ represents the taxes paid on that return in that period. This will be zero for the DC saver, and we will describe it for the brokerage saver in the next subsection.

E.4.4 Investment returns

Two comments are needed on the investment returns. First, they vary with age. Each age t is associated with a portfolio composition between equities, bonds, and bills, with shares given by s_t^k, s_t^b, and s_t^m. During working years, these shares are interpolated from Fidelity target date funds.[40] In retirement, we assume exclusive investment in bonds. The age profile of investment composition is shown in Figure A.25a, and the associated age profile of real rate of return is shown in Figure A.25b. Real rates of return for these asset types (ρ^k, ρ^b, and ρ^m, respectively) are taken from Jordà et al. (2019). The combination of these assumptions yields age-specific rates of return ρ_t:

$$\rho_t = \rho^k \cdot s_t^k + \rho^b \cdot s_t^b + \rho^m \cdot s_t^m. \tag{4}$$

The second comment on returns is the division of returns into unrealized capital gains, distributions taxed as long-term capital gains, and returns taxed as income (e.g., ordinary dividend income).[41] Distinguishing between the nature of the return will be important in our treatment of the brokerage saver's taxable returns. The share of returns represented by each of these is given by χ^g, χ^k, and χ^i, respectively, which sum to 1. The dollar flows associated with each of these three types of return are given below:

$$r_{i,t}^{g,j} = \left(B_{i,t}^j + f_{i,t}^j\right) \cdot \chi^g \cdot \rho_t, \tag{5}$$

$$r_{i,t}^{k,j} = \left(B_{i,t}^j + f_{i,t}^j\right) \cdot \chi^k \cdot \rho_t, \tag{6}$$

$$r_{i,t}^{i,j} = \left(B_{i,t}^j + f_{i,t}^j\right) \cdot \chi^i \cdot \rho_t. \tag{7}$$

[40] We use asset allocations of the Fidelity Freedom Funds ranging from retirement years 2005 to 2065 between equities, bonds, and short-term debt as of year-end 2022. Distance to retirement is thus the target date minus 2023. A one-dimensional Akima interpolator was used to calculate shares between observed age distances to retirement. Our shares may be compared to Fidelity's own description of their glide path (Fidelity, 2023).

[41] The second component–distributions taxed as long-term capital gains–does not represent returns which are realized for a withdrawal. Rather, they are the gains realized as mutual fund managers trade assets and passed on to investors. See Fidelity's description of these distribution types at `https://www.fidelity.com/learning-center/investment-products/mutual-funds/taxes`.

Accumulation and withdrawal of untaxed capital gains When individuals withdraw funds from their accounts, they realize some (previously unrealized) capital gains. This has tax implications for the brokerage saver, making it necessary for us to keep track of that part of the account balance formed of unrealized capital gains. We divide the account balance $B^{j}_{i,t}$ into principal $B^{p,j}_{i,t}$ and (thus far untaxed) capital gains $B^{g,j}_{i,t}$. We define the latter recursively as:

$$B^{g,j}_{i,t+1} = B^{g,j}_{i,t} + r^{g,j}_{i,t} - w^{k,j}_{i,t}, \tag{8}$$

where $B^{g,j}$ is the cash value of the stock of unrealized capital gains in the account balance, $r^{g,j}_{i,t}$ are additional untaxed gains attained in year t, and $w^{k,j}_{i,t}$ are gains actually realized when a withdrawal is made.

Whenever a withdrawal $w^{j}_{i,t}$ is made, we assume that the withdrawal comprises untaxed capital gains $w^{k,j}_{i,t}$ and principal $w^{p,j}_{i,t}$ in proportions that equal their share of the stock of wealth. That is, the share of any withdrawal by the brokerage saver that is subject to capital gains tax is equal to the share of unrealized capital gains in wealth:

$$\frac{w^{k,j}_{i,t}}{w^{j}_{i,t}} = \frac{B^{g,j}_{i,t}}{B^{j}_{i,t}}. \tag{9}$$

E.4.5 Social Security Income

We assume all individuals stop earning when they turn 66 and begin claiming Social Security benefits. Central to the determination of Social Security benefits is 'Average Indexed Monthly Earnings' (*aime*), calculated as the average of the best 35 years of total compensation.[42] Consistent with Social Security rules, the measure of earnings that enters this calculation is capped at a value e^{max}:

$$aime_i = \frac{1}{35} \sum_{k \in \text{best } 35} \left\{ \frac{\min(e + dc^{ee}_{i,t},\ e^{max})}{12} \right\}. \tag{10}$$

Monthly Social Security benefits are equal to 90% of *aime* up to the first 'bend point' ($895 in 2018), 32% of any *aime* above the first bend point and below the second point ($5,397 in 2018), and 15% of any *aime* above the second bend point.

[42] All variables are expressed in real terms, and we assume a stationary earnings distribution. As a result, there is no indexation of earnings in equation (10).

E.4.6 Withdrawals

We distinguish between 'early withdrawals' and 'retirement withdrawals.' The former are those taken before age 65, which we observe in our data. The latter are taken after age 65. These are not observed and so must be modeled.

Early withdrawals We define early withdrawals as all withdrawals before age 65.[43] The measure that we observe in our data (denoted $w_{i,t}^{DC}$) is that before income taxation, which must be paid on all withdrawals from DC accounts. For the equivalent withdrawal applied to the brokerage saver (denoted by $w_{i,t}^{BK}$), we calculate the after-tax quantity retained by the DC saver.

One complication arises when the early withdrawal that we see would lead to the brokerage saver having a negative balance. This occurs in only a small share of cases (14.2%). In these cases, we adjust the measure we see in our data to be the largest number that avoids the brokerage saver going negative. This adjustment reduces the withdrawal by approximately 17.6% for that share of savers.

Retirement withdrawals Individuals retire at the beginning of age 66 with balance in their account of $B_{i,66}^{j}$. They employ a consumption rule each year to determine how much to withdraw each period t. We set this rule such that consumption for the DC saver is constant each period.

In particular, the withdrawal each period is equal to:

$$w_t^j = \frac{1-\alpha}{1-\alpha^{90-t+1}} B_{i,t}^j, \tag{11}$$

where $\alpha = \frac{1}{(1+\rho^b)}$ is defined using the return on bonds ρ^b.[44] This rule, which we illustrate in Figure A.26, keeps pre-tax withdrawals constant. We assume that individuals consume their withdrawal, net of taxes:

$$c_t^j = w_t^j - \tau_{i,t}^{w,j}, \tag{12}$$

where $\tau_{i,t}^{w,j}$ are taxes incurred by withdrawing money from account j and defined in the next section. Constant (pre-tax) withdrawals keep post-tax consumption constant for the DC saver (as income does not change in retirement) and close to constant for the brokerage saver (for whom small changes in average tax rates will occur as wealth is decumulated).

[43] Not all of these will be subject to an early withdrawal penalty. We return to this when we discuss the taxation of withdrawals in Section E.6.1.

[44] This consumption rule is that obtained from a cake-eating problem in which life-span is deterministic and the discount rate is set equal to the interest rate.

E.5 Summary

The data that we construct, together with the features outlined above, yield two parallel data sets: one representing the earnings, savings, account balance, and withdrawals of the DC saver, and one representing the same objects for the brokerage saver. We represent these by the following:

$$\left\{ \{e_{i,t}, dc_{i,t}^{ee}, dc_{i,t}^{er}, B_{i,t}^{DC}, w_{i,t}^{DC}\}_{t=25}^{90}; \{c_{i,t}^{DC}\}_{t=66}^{90} \right\} \quad \left\{ \{e_{i,t}, dc_{i,t}^{ee}, dc_{i,t}^{er}, B_{i,t}^{BK}, w_{i,t}^{BK}\}_{t=25}^{90}; \{c_{i,t}^{BK}\}_{t=66}^{90} \right\},$$

where the first three objects are common across the two tuples, but the balances, withdrawals, and consumption profiles differ due exclusively to the different forms of taxation faced by the two savers.

E.6 Taxation

The previous section concludes by noting our data and micro-simulation model yield, for each individual in our data, two trajectories of wealth accumulation and decumulation– one if they save in a DC account and one if they saved the same quantities in a taxable brokerage account. Due to their access to preferential taxation, the DC saver will have higher consumption in retirement. This section shows how we measure these differences in tax treatment across the life cycle.

At the most general level, we take the flow of income, saving, and returns and use TAXSIM to evaluate the taxes. This allows us to construct our summary measure of wealth at retirement: the present discounted value of consumption facilitated by accumulated wealth at retirement. This section provides the interested reader full details on how we measure that.

E.6.1 Decomposing the overall tax burden into components

We denote our modelled tax function, which distinguishes between the three forms of income that agents in our model earn, as $T(N, K, S)$. N denotes inflows taxed according to the income tax schedule (e.g., wage income during working life and 401(k) distributions in retirement); K denotes income taxed as long-term capital gains; and S denotes Social Security benefits.[45]

[45] Note that effective tax rates in retirement are usually very low (Chen and Munnell, 2020) due in part to the favorable tax treatment of Social Security benefits, on which many households pay no tax at all (Joint Committee on Taxation, 2019).

We decompose the total tax burden (denoted T) into shares that can be ascribed to earnings (τ^e), contributions to retirement accounts (τ^c), investment returns (τ^r), taxes owed on Social Security benefits (τ^s), and withdrawals from retirement accounts (τ^w). Earnings, contributions, returns and withdrawals, of course, interact in a non-linear (and quite complex) manner to generate overall tax liability. This means that there is no unique decomposition such that the total tax burden T can written as the sum of these components. This section explains how we obtain one such decomposition.

We use rules for tax year 2018 according to NBER's TAXSIM 32 tool to calculate federal income tax owed by each simulated individual.[46]

Taxation of Earnings We first define taxes on earnings ($\tau_{i,t}^{e,j}$) as follows:

$$\tau_{i,t}^{e,j} = \begin{cases} T\big(e_{i,t},\ 0,\ 0\big) & \text{if } t < 66 \text{ for } j = DC, BK; \\ 0 & \text{if } t \geq 66 \text{ for } j = DC, BK. \end{cases} \tag{13}$$

This does not differ by the type of saver, and the second equality follows from our assumption of no earnings from the age of 66.

Taxation of Social Security We define the tax on Social Security as the tax that would be paid if an agent had their Social Security income and no other income:

$$\tau_{i,t}^{ss,j} = T\big(0,\ 0,\ ss_{i,t}\big) \quad \text{if} \ \ t \geq 66 \quad \text{and } j = DC, BK, \tag{14}$$

which also does not differ by type of saver.[47]

Taxation of Contributions Our definition of taxable earnings excludes that part of earnings which was saved for retirement: an employee's choice of deferred compensation and any associated employer match $dc_{i,t}^{ee} + dc_{i,t}^{er}$. For the DC saver, income contributed to the account is untaxed, so $\tau_{i,t}^{c,DC} = 0$. For the brokerage saver, the tax we ascribe to contributions is equal to the additional income tax the saver would have paid by taking compensation as

[46]The N, K, and S income sources are fed into the *pwages*, *ltcg*, and *gssi* fields in TAXSIM, respectively. We assume that all individuals take the standard deduction and do not claim any other credits or deductions. See Feenberg and Coutts (1993) for a description of the TAXSIM model.

[47]As it happens, $\tau_{i,t}^{ss,j}$ will be zero for everyone in our sample–an individual with maximum Social Security income and no other income will not face any income tax. We retain the variable for completeness and because its exclusion may obscure some features of the exposition.

earnings. This is given by the second line in:

$$\tau_{i,t}^{c,j} = \begin{cases} 0 & \text{for } j = DC, \\ T\left(e_{i,t} + dc_{i,t}^{ee} + dc_{i,t}^{er},\ 0,\ 0\right) - \tau_{i,t}^{e,BK} & \text{for } j = BK, \end{cases} \tag{15}$$

where the positive term in the second line gives the income tax owed from earnings that include deferred compensation, and the negative term nets off that tax already ascribed to earnings, defined in equation (13).

As we assume that there are neither earnings nor contributions after retirement, for both savers we obtain $\tau_{i,t}^{c,j} = 0$ for all $t \geq 66$.

Taxation of withdrawals The taxation of withdrawals depends on whether they are 'early withdrawals' (those made up to the age of 65) or 'retirement withdrawals' (from the age of 65). In the former case, the DC saver must pay income tax and may face a tax penalty. This penalty is incurred at a rate p_t, which is equal to 10% for non-exempt withdrawals before age 59.5 and 0 for withdrawals after age 65. The first line of equation (16) gives this quantity. The positive terms are the regular income tax on earnings and DC withdrawals and the tax penalty; the negative term subtracts taxes already ascribed to earnings.

The brokerage saver need not pay income tax on withdrawals but must pay capital gains taxes on gains realized to withdraw their funds ($w_{i,t}^{k,BK}$). This quantity is defined in the second line in equation (16), where the first term gives the tax liability from earnings, contributions, and capital gains and the negative term subtracts taxes already ascribed to earnings and contributions:

$$\tau_{i,t}^{w,j} = \begin{cases} T\left(e_{i,t} + w_{i,t}^{DC},\ 0,\ 0\right) + p_t w_{i,t}^{DC} \mathbb{1}(t < 60) - \tau_{i,t}^{e,TD} & \text{if } j = DC \text{ and } t < 66, \\ T\left(e_{i,t} + dc_{i,t}^{ee} + dc_{i,t}^{er},\ w_{i,t}^{k,BK},\ 0\right) - \left(\tau_{i,t}^{e,BK} + \tau_{i,t}^{c,BK}\right) & \text{if } j = BK \text{ and } t < 66. \end{cases} \tag{16}$$

In retirement, the DC saver pays regular income taxes on withdrawals (see the first line of equation (17)), while the brokerage saver pays capital gains taxes on the share of withdrawals that represent previously unrealized gains ($w_{i,t}^{k,BK}$). Both savers are also claiming their Social Security payments, which enter as the third argument of the tax function:

$$\tau_{i,t}^{w,j} = \begin{cases} T\left(w_{i,t}^{DC},\ 0,\ ss_{i,t}\right) - \tau_{i,t}^{ss,DC} & \text{if } t \geq 66 \text{ and } j = DC, \\ T\left(0,\ w_{i,t}^{k,BK},\ ss_{i,t}\right) - \tau_{i,t}^{ss,BK} & \text{if } t \geq 66 \text{ and } j = BK. \end{cases} \tag{17}$$

Taxes on investment returns All returns on funds in DC accounts are untaxed. That is, there is no taxation of unrealized gains ($r_{i,t}^{g,j}$), no income tax on dividend income ($r_{i,t}^{i,j}$), and no capital gains tax for distributions ($r^{k,j}_{i,t}$). The taxes paid by the DC saver on returns are therefore zero.

For the brokerage saver, while the unrealized capital gains ($r^{g,j}_{i,t}$) incur no immediate tax liability, income tax is paid on dividend income ($r^{i,j}_{i,t}$), and capital gains tax is paid on realized gains. As described in Section E.4.4, the latter come in two parts–that part of the return which is distributed even in the absence of a withdrawal ($r^{k,j}_{i,t}$) and that part of the return which is realized when a withdrawal is made ($w^{k,j}_{i,t}$).

Taxes on portfolio returns for the brokerage saver are given in (18). In both lines (representing taxes before and after retirement, respectively), the first term gives all taxes due in a particular period (on earnings, contributions, withdrawals, and returns), and the second term nets off those taxes already ascribed to earnings, contributions, and withdrawals:

$$\tau^{r,BK}_{i,t} = \begin{cases} T\left(e_{i,t} + dc^{ee}_{i,t} + dc^{er}_{i,t} + r^{i,BK}_{i,t},\ r^{k,BK}_{i,t} + w^{k,BK}_{i,t},\ 0\right) \\ \quad - \left(\tau^{e,BK}_{i,t} + \tau^{c,BK}_{i,t} + \tau^{w,BK}_{i,t}\right) & \text{if } t < 66, \\ T\left(r^{i,BK}_{i,t}, r^{k,BK}_{i,t} + w^{k,BK}_{i,t},\ ss_{i,t}\right) - (\tau^{ss,BK}_{i,t} + \tau^{w,BK}_{i,t}) & \text{if } t \geq 66. \end{cases} \tag{18}$$

E.7 Lifetime Measures

E.7.1 Implied post-tax interest rate

Our model contains multiple interest rates that could be used to evaluate the present value of future flows. To do this, we define an interest rate $\hat{r}_{i,t}$ as the post-tax rate of return that the brokerage saver would pay if their deferred gains each period were realized as long-term capital gains.[48] We first define the hypothetical taxes on portfolio returns in this case as:

$$\widehat{\tau^{r,BK}_{i,t}} = \begin{cases} T\left(e_{i,t} + dc^{ee}_{i,t} + dc^{er}_{i,t} + r^{i,BK}_{i,t},\ r^{g,BK}_{i,t} + r^{k,BK}_{i,t} + w^{k,BK}_{i,t},\ 0\right) \\ \quad - \left(\tau^{e,BK}_{i,t} + \tau^{c,BK}_{i,t} + \tau^{w,BK}_{i,t}\right) & \text{if } t < 66, \\ T\left(r^{i,BK}_{i,t}, r^{g,BK}_{i,t} + r^{k,BK}_{i,t} + w^{k,BK}_{i,t},\ ss_{i,t}\right) - (\tau^{ss,BK}_{i,t} + \tau^{w,BK}_{i,t}) & \text{if } t \geq 66, \end{cases} \tag{19}$$

where this expression is the same as that in equation (18) except for the inclusion of $r^{g,BK}_{i,t}$ each period in the second argument. The implied, post-tax interest rate is then

$$\hat{r}_{i,t} = r_t - \frac{\widehat{\tau^{r,BK}_{i,t}}}{B^{BK}_{i,t} + f^{BK}_{i,t}}. \tag{20}$$

[48]This assumption ensures that interest rate we choose for discounting does not depend on patterns of withdrawals that we observe in our data.

This rate is used across all counterfactuals.

E.7.2 Wealth

We have two measures of resources in retirement: a) 'DC wealth' (the value in DC accounts) and b) 'Broad Retirement Wealth,' which also includes Social Security wealth.

Wealth Our measure of wealth is the present discounted value of after-tax withdrawals facilitated by the account balance. We express this as recursively, backwards from age 90 with $A^j_{i,90} = 0$:

$$A^j_{i,t} = \begin{cases} \frac{A^{DC}_{i,t+1}}{1+\hat{r}_{i,t+1}} + \left(w^{DC}_{i,t+1} - \tau^{w,DC}_{i,t+1}\right) - \left(dc^{ee}_{i,t+1} + dc^{er}_{i,t+1} - \tau^{c,BK}_{i,t+1}\right) & \text{for } j = DC, \\ \frac{A^{BK}_{i,t+1}}{1+\hat{r}_{i,t+1}} + \left(w^{BK}_{i,t+1} - \widehat{\tau^{r,BK}_{i,t+1}}\right) - \left(dc^{ee}_{i,t+1} + dc^{er}_{i,t+1} - \tau^{c,BK}_{i,t+1}\right) & \text{for } j = BK, \end{cases} \quad (21)$$

as the present value of future post-tax withdrawals less future post-tax contributions.

This is private retirement wealth and does not include wealth held in the form of Social Security benefits. We define Social Security wealth as:

$$SS_{i,t} = \frac{SS_{i,t+1}}{1+\hat{r}_{i,t+1}} + \left(ss_{i,t+1} - \tau^{ss}_{i,t+1}\right). \quad (22)$$

Our broad measure of wealth takes into account both wealth in private accounts and Social Security wealth:

$$A^{BR}_{i,t} = A^{DC}_{i,t} + SS_{i,t}. \quad (23)$$

E.8 Decomposing retirement wealth

In this subsection we define how we decompose retirement wealth into three components: wealth that flows from employee contributions, wealth that can be ascribed to employer contributions, and wealth due to the favorable tax treatment of DC accounts.

E.8.1 Value of DC tax treatment

The total tax benefit to an individual i is defined as the difference between the retirement wealth of the DC saver and that of the brokerage saver:

$$A^T_i = A^{DC}_{i,65} - A^{BK}_{i,65}. \quad (24)$$

To find the retirement wealth concept attributable to the employee alone, we need to find the proportion of contributions that are from the employee for each individual in our data. The value at retirement of the respective contributions made by the employee and the employer are:

$$DC^{ee} = \sum_{t=25}^{65} dc_{i,t}^{ee}\left(\prod_{\tau=t}^{65}(1+\hat{r}_\tau)\right) \quad DC^{er} = \sum_{t=25}^{65} dc_{i,t}^{er}\left(\prod_{\tau=t}^{65}(1+\hat{r}_\tau)\right). \tag{25}$$

These can then be used to calculate the respective proportions of retirement wealth for the brokerage saver (i.e., after tax benefits have been removed) coming from employee and employer contributions:

$$A_i^{EE} = \frac{DC^{ee}}{(DC^{ee}+DC^{er})} \cdot A_{i,65}^{BK} \quad A_i^{ER} = \frac{DC^{er}}{(DC^{ee}+DC^{er})} \cdot A_{i,65}^{BK}. \tag{26}$$

The Treasury Department estimates the aggregate tax benefit given to DC savers in 2021 was \$119 billion(US Department of the Treasury, 2023). As a check on our model, we compare our estimate of the tax benefit to the official estimate. Using an annuitization factor based on our model interest rate, we transform the mean lifetime tax benefit $A_i^T = \$52,936$ to an annual measure by dividing it by a factor of approximately 50. This results in a mean *annual* tax benefit of about \$1,054 for the population represented by our simulated data, where population DC coverage is estimated to be of those currently in their 20s.[49] To convert our number to one which can be considered reflective of the current US population (who are the basis for the Treasury's numbers), we multiply our average annual tax benefit by the ratio of the population DC savings rate to the hot deck sample DC savings rate (approximately 1.5). This yields a comparable mean annual tax benefit of \$694 per worker. Finally, we multiply this by an estimate of the civilian population engaged in work at any time in 2018 from the public CPS-ASEC, around 168 million people. Our model estimate of aggregate annual tax benefit to DC savers is then \$117 billion.

E.9 Tax Counterfactual

The tax counterfactual considers how distributing the aggregate tax expenditure proportionally to lifetime earnings would affect wealth. This would break the link between saving decisions and a worker's share of this tax expenditure but would not otherwise increase redistribution across lifetime income groups. Every individual would receive a government contribution to her DC account proportional to her lifetime earnings. This uniform propor-

[49] Our hot deck imputation model matches younger people to older people based in part on DC access. The fact that younger people are more likely to have access to DC plans makes DC access more prevalent in our sample than in the population.

tion is chosen such that the total cost of these government contributions matches the total cost incurred under the existing tax-favored system.

Let the value of lifetime total earnings be:

$$LE_i = \sum_{t=25}^{65} (comp_{i,t}) \left(\prod_{\tau=t}^{65} (1 + \hat{r}_\tau) \right), \tag{27}$$

where $comp_{i,t} = e_{i,t} + dc_{i,t}^{ee}$ is the sum of earnings and deferred compensation. We define a redistributed tax advantage that allocates the total tax benefit in the economy proportionally to lifetime income:

$$A_i'^T = \frac{LE_i}{\sum_n LE_n} \cdot \sum_n A_n^T \tag{28}$$

where the first term is an individual's share of aggregate lifetime earnings and the second term is the aggregate tax expenditure. We assume that these redistributed tax benefits are fully illiquid prior to retirement and cannot be withdrawn early. In our baseline, we assume no behavioral response to the change in the tax treatment of retirement contributions so that employee and employer contributions are unchanged. We indicate aggregates under this counterfactual with a $'$ superscript. DC wealth and broad retirement wealth in this counterfactual are therefore:

$$A_i'^{DC} = A_i^{EE} + A_i^{ER} + A_i'^T \qquad A_i^{BR'} = SS_i + A_i'^{DC}.$$

E.10 Match counterfactual

In the presence of an employer match for retirement contributions, those who save more receive higher total compensation from their employer. Our employer match counterfactual breaks this link and considers the effect of a noncontingent employer contribution that is proportional to employee earnings. Every worker receives an employer contribution to her DC account proportional to her current earnings, regardless of her contributions. This percentage would be the same for all workers at the same employer but varies across employers.

For our employer-match counterfactual, we calculate the proportional contribution that, if given to all employees in the firm, would cost the same as their actual matching contributions. That is, for each time period t we calculate the ratio of total matching contributions to total income for each firm and multiply that by individual income. Denoting an employee i

working in firm f with an employer match of $dc^{er}_{i,t}$,[50] instead of receiving $dc^{er}_{i,t}$ in period t, the employee receives:

$$dc^{*er}_{i,t} = \frac{comp_{i,t}}{\sum_{n\in f} comp_{i,t}} \cdot \sum_{n\in f} dc^{er}_{n,t}, \tag{29}$$

where the first term is individual i's share of compensation in their firm in period t and the second is the aggregate matching contribution made by their employer in period t. We then calculate all modeled objects as described above assuming that employees receive the counterfactual match $dc^{*,er}_{i,t}$. Accounting for this and for the fact that taxation trajectories will be different will yield different levels of wealth at retirement. All stocks in this model are denoted as in the baseline model but with the addition of a $*$ superscript. We denote the employers' counterfactual contributions and due to the tax expenditure as ($A^{*,ER}_i$ and A^{*T}_i), respectively, so that the new levels of wealth and broad wealth in retirement equal:

$$A^{*DC}_i = A^{EE}_i + A^{*ER}_i + A^{*T}_i \qquad A^{BR*}_i = SS_i + A^{*DC}_i. \tag{30}$$

E.11 Combined Counterfactual

Our combined counterfactual equalizes the employer match contribution and the tax subsidy. To do this, we first obtain the brokerage saver's wealth under the employer match counterfactual $C^{\dagger BK}_{i,t}$. We add the redistributive tax subsidy calculated in tax counterfactual ($A^{\dagger ER}_i$). Denoting all aggregates under the combined counterfactual with an † superscript (though note that $A^{\dagger ER}_i = A^{*ER}_i$), we obtain:

$$A^{\dagger DC}_i = A^{EE}_i + A^{\dagger ER}_i + A^{\dagger T}_i \qquad A^{BR\dagger}_i = SS_i + A^{\dagger DC}_i. \tag{31}$$

E.12 Parameterization

E.12.1 Rates of return

Total investment return is given by an age-varying interest rate r_t. Each age t is associated with a portfolio composition between equities, bonds, and bills, parameterized by σ^k_t, σ^b_t, and σ^m_t. During working years, these shares are interpolated from Fidelity target date funds (see, for example, the 2040 Target Date Fund in Fidelity (2023)). In retirement, we assume exclusive investment in bonds. The age-profile of investment composition is shown in Figure A.25a, and the associated age-profile of the real rate of return is shown in Figure A.25b.

[50] This will be linked to the employee's contribution ($dc^{ee}_{i,t}$) by a function that gives the employer match: $dc^{er}_{i,t} = m_f(dc^{ee}_{i,t})$.

Real rates of return for these asset types (ρ^k, ρ^b, and ρ^m) come from Jordà et al. (2019). The combination of these assumptions yields age-specific rates of return r_t according to:

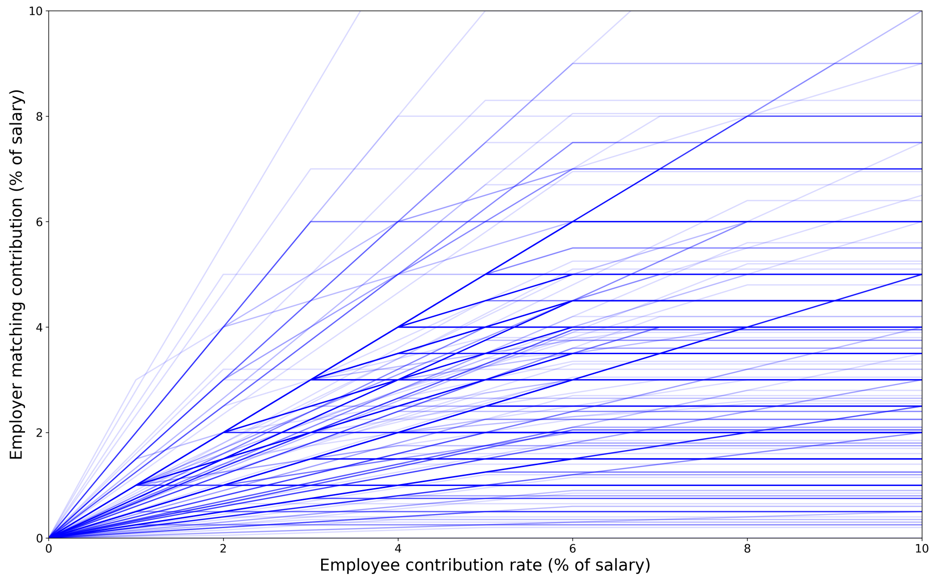
$$r_t = \rho^k \cdot \sigma_t^k + \rho^b \cdot \sigma_t^b + \rho^m \cdot \sigma_t^m. \tag{32}$$

Note that in retirement $r_t = \rho^b$. We derive the decomposition of returns into these shares by studying the historical price trends and distributions of the Fidelity Freedom Funds Fidelity (2023).[51]

[51] Our breakdown of 50% price change, 40% distribution taxed as long-term capital gains, and 10% taxed as income is very similar to the 48/43/9 breakdown found by Sialm and Zhang (2020) under the assumption that 95% of dividends are non-qualified.

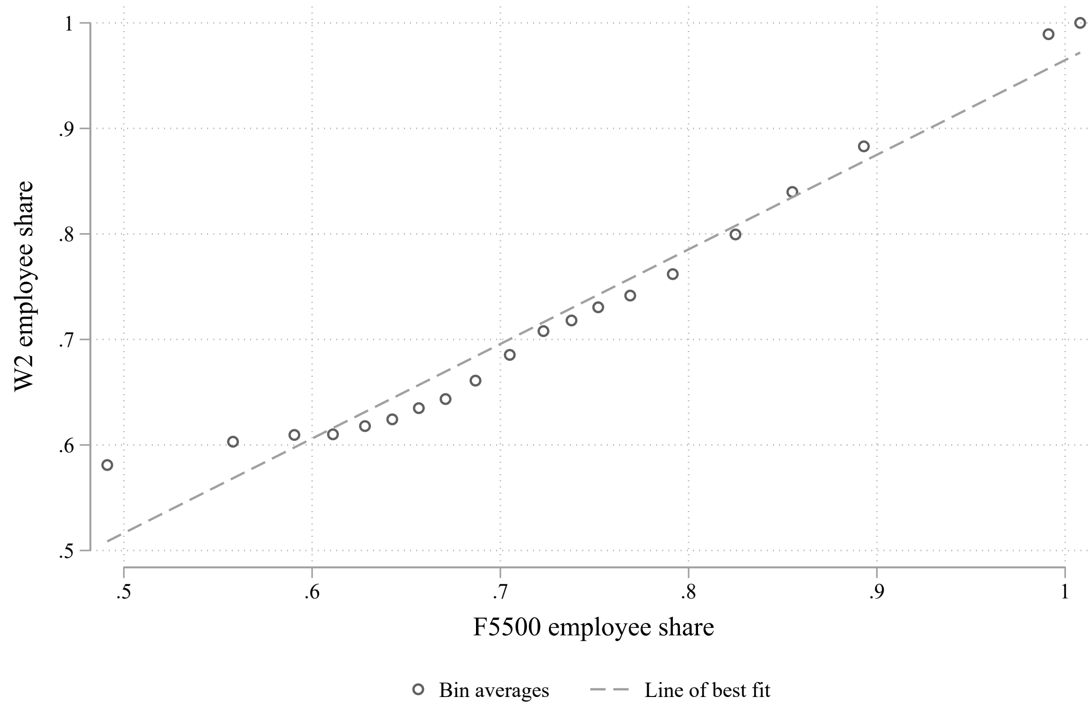
F Supplemental Figures and Tables

Figure A.1: Matching schedules



Notes: The sample is all employer match schedules for plans in a particular year. Each line represents a match schedule, and the depth of shade represents the frequency of the match schedule.

Figure A.2: Bin scatter of W2-imputed vs. Form 5500-reported employee contribution share

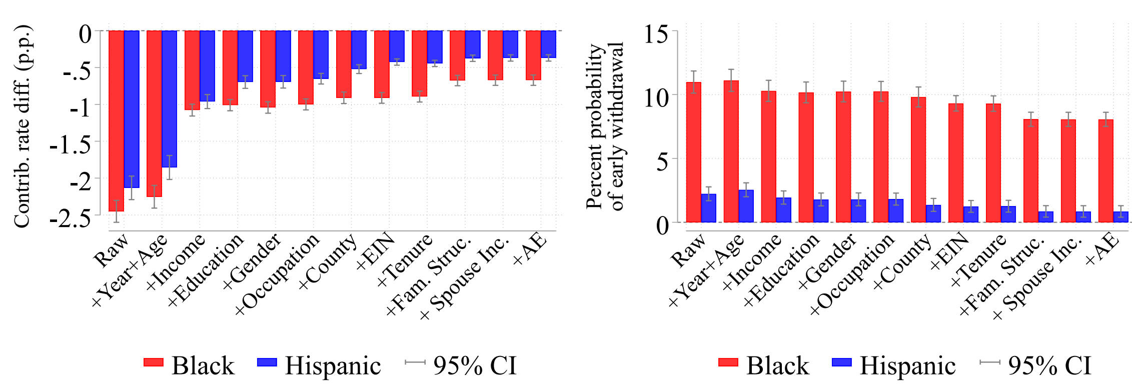


This graph presents a binned scatter plot comparing the W2-imputed firm-level employee share of contributions ($\frac{\text{total employee deferred compensation}}{\text{total employee deferred compensation + total employer match}}$) against the publicly-filed Form 5500 average employee share of contributions, which is used as the running variable to compute 20 ventiles on the horizontal axis. The dashed line is a line of best fit through the individual points reported on the scatter plot, and coefficients of this line are reported on the graph.

Figure A.3: Savings and Early Withdrawal Gaps with all variables, by race

(a) Employee+Match DC Contribution Rate

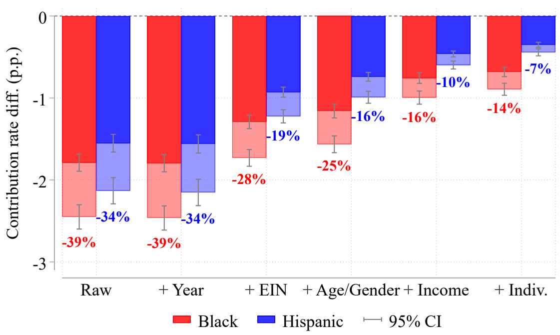
(b) Early withdrawal gaps



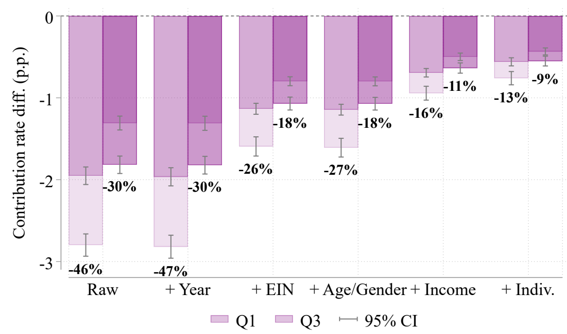
Notes: Please see Figure Notes for Figures 2 and 4. Here, we show each layer of the cascade with all the individual specifications (education, gender, occupation, county, EIN, and tenure) and all household specifications (family structure and spousal income). We also include a control for auto-enrollment.

Figure A.4: Savings Gaps under Alternative Ordering

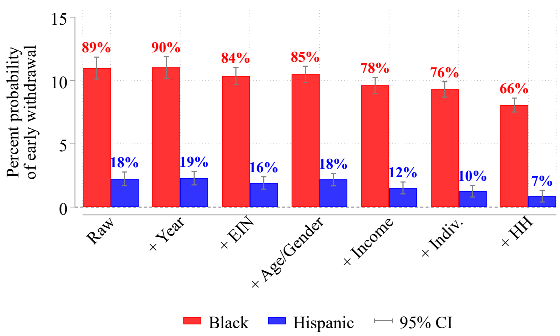
(a) Employee + Match DC Contribution Rate, by race



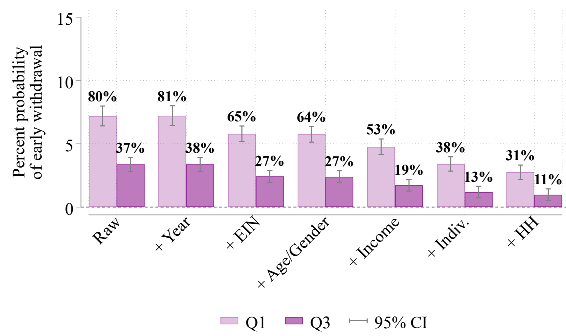
(b) Employee + Match DC Contribution Rate, by parental income



(c) Early withdrawal gaps, by race

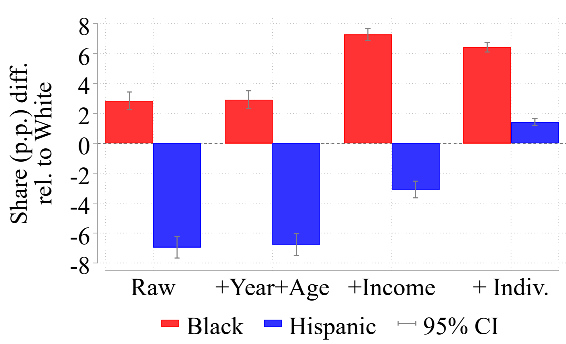


(d) Early withdrawal gaps, by par. income



Notes: Please see Figure Notes for Figures 2 and 4. Here, the EIN FE comes earlier in the analysis.

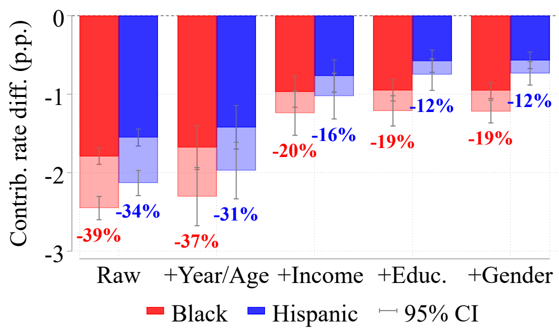
Figure A.5: Decomposing DC plan accessibility and savings by race



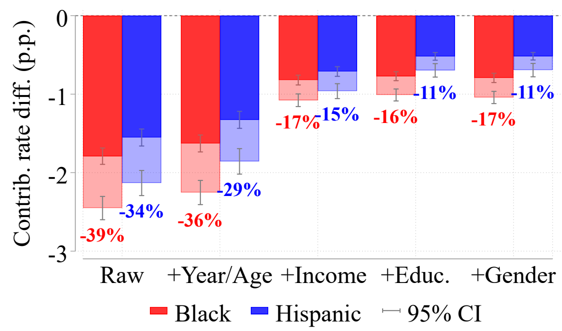
Notes: The dependent variable is an indicator for DC plan access. We use the specification defined in Equation 1, omitting EIN due to perfect collinearity.

Figure A.6: Racial gap estimates re-weighted using the characteristic shares of White employees

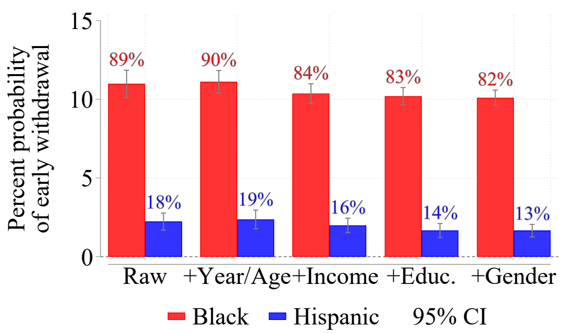
(a) Employee + Match DC Contribution Rate, Reweighted



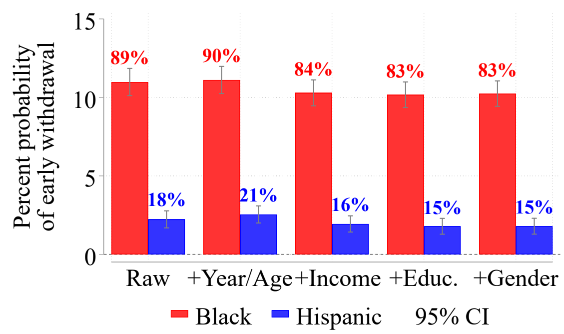
(b) Employee + Match DC Contribution Rate, Baseline



(c) Early withdrawal gaps, Reweighted

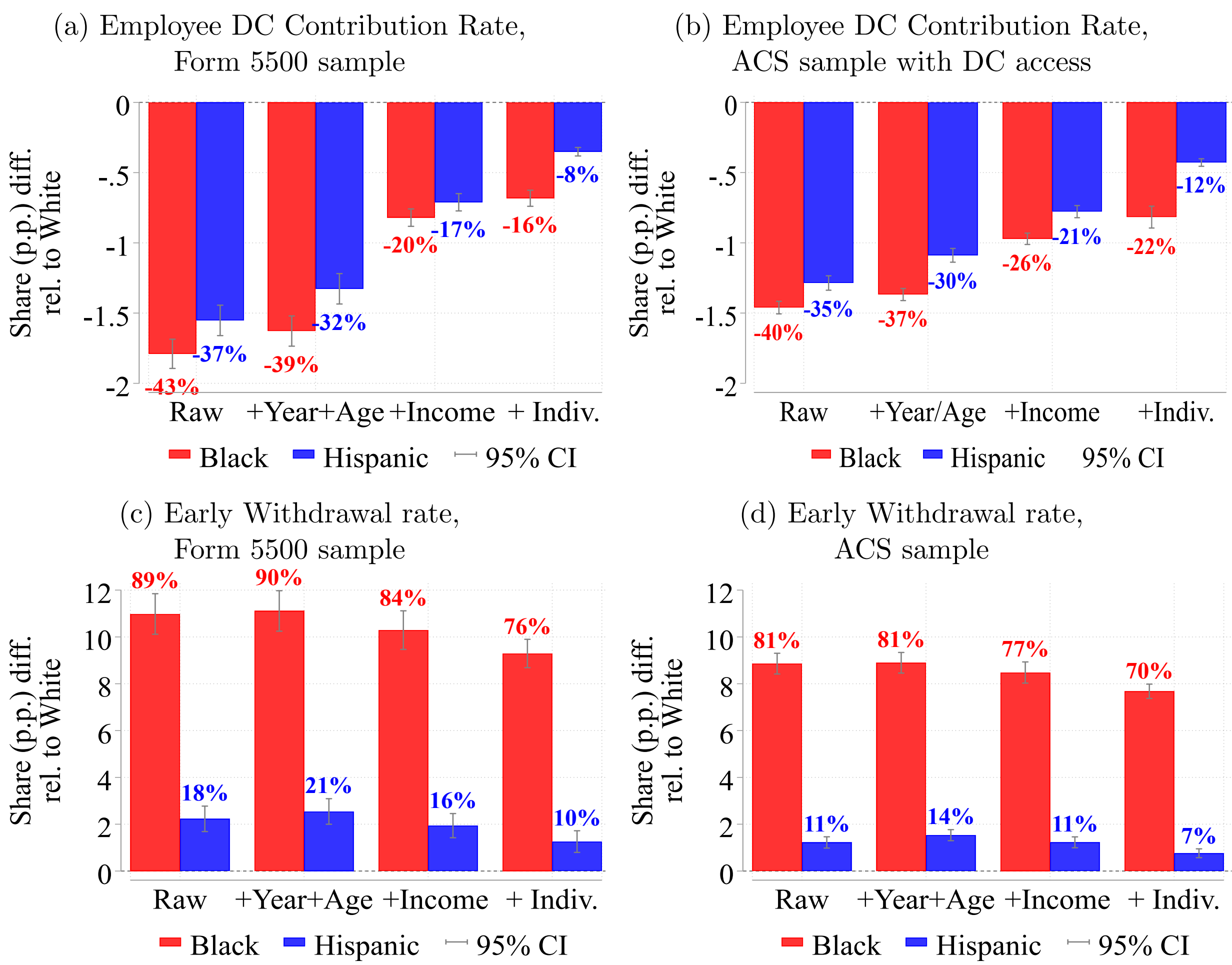


(d) Early withdrawal gaps, Baseline



Notes: This figure presents a robustness check of our main results. We re-weight the Black and Hispanic worker distributions according to the White worker distribution. The left column of figures presents gaps for our reweighted Form 5500 sample gaps; the right column of figures presents gaps for our Form 5500 sample results, as presented in Figures 2 and 4. We use the progressive specification defined in Equation 1. Due to cell size constraints in U.S. Census disclosure requirements, we present estimates for the first five regression controls from raw gaps through education. Panels (a) and (b) show the employee contribution in opaque bars and overlays the employer match in transparent bars. Panels (c) and (d) shows the early withdrawal rates. For further discussion, please see Appendix B.2.

Figure A.7: Racial gap estimates in Form 5500 vs. ACS samples

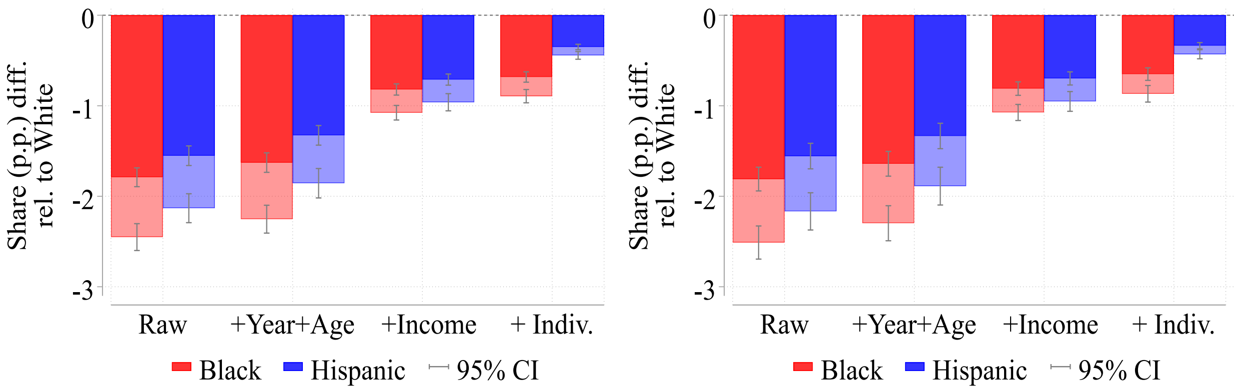


Notes: This figure presents robustness checks for our main results (Figure 2). We use the progressive specification (Equation 1) for the Form 5500 sample and full ACS sample of employees (conditional on non-missing control variables so the sample is consistent across columns) and the subset of those with DC Access. In panels (a) and (c), we report the same figures as in Figures 2(a) (note, here with just the employee DC contribution rate) and 4(a). In panel (b), we show the same estimates as in panel (a) but on a larger sample, which is the ACS sample with DC access. In panel (d), we similarly show the same estimates as in panel (c) but on the entire ACS sample. We report the ACS sample in panel (d) because our sample inclusion restriction for analysis of early withdrawals conditions on having recently made contributions and therefore implies having had DC access (please see Appendix A.2.1 for more information). As in previous figures, the numbers in bold represent the percentage difference relative to the average level for the omitted category (i.e., White workers). For more information on this methodology, please refer to the notes of Figures 2 and 4, and for more information on the differences between the samples, please see Appendix A.3.3.

Figure A.8: Racial gap estimates in Form 5500 vs. Fully-Vested sample

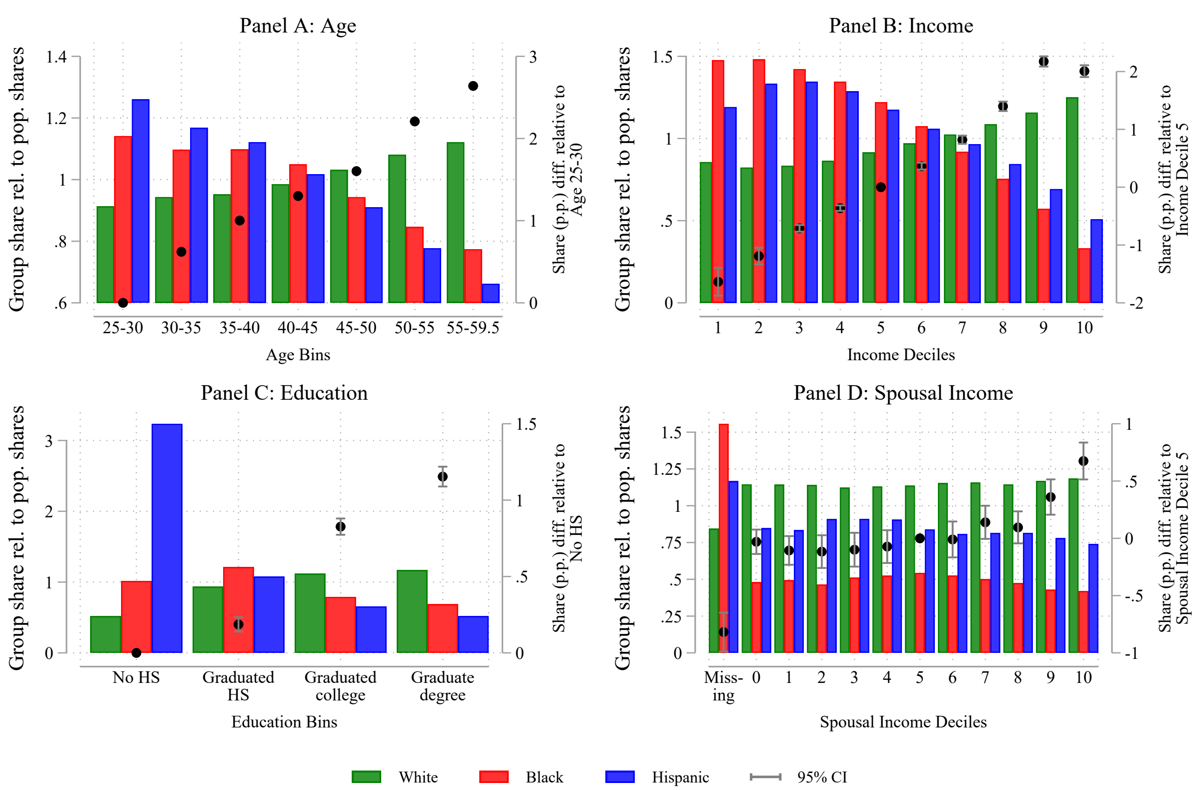
(a) Employee + Employer DC Contribution Rate, Form 5500 sample

(b) Employee + Employer DC Contribution Rate, Fully-vested sample



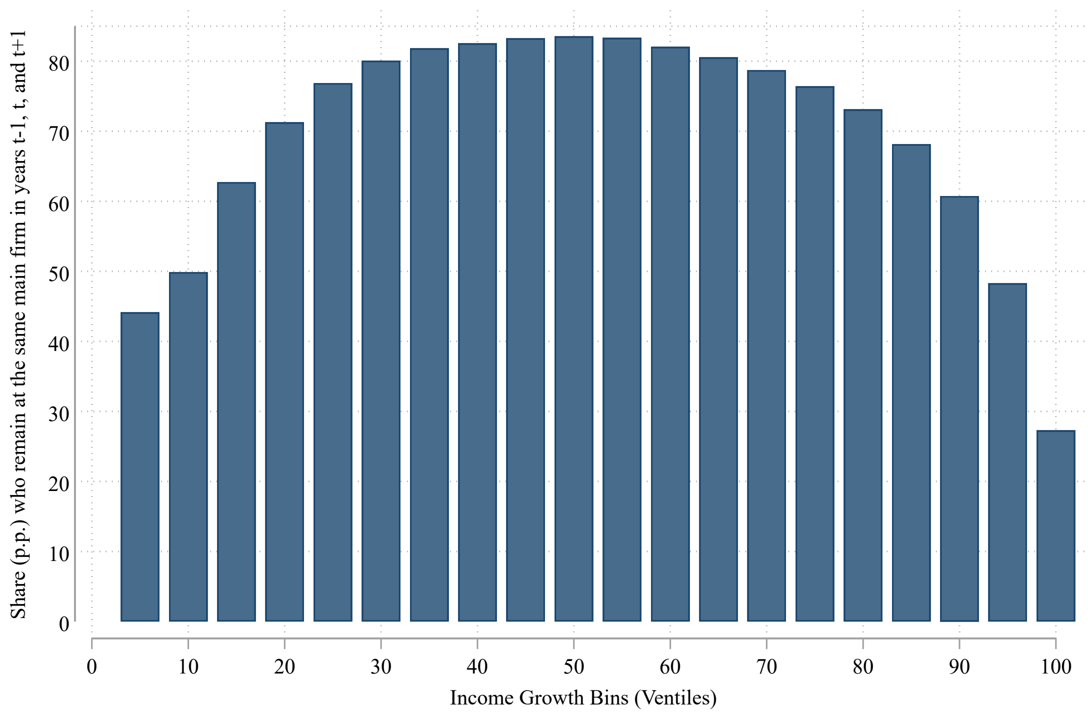
Notes: This figure is analogous to Figure A.7 comparing the Form 5500 results with the sub-sample that are fully-vested. For more information, please refer to the notes of Figure A.7.

Figure A.9: Regression coefficients and groups shares



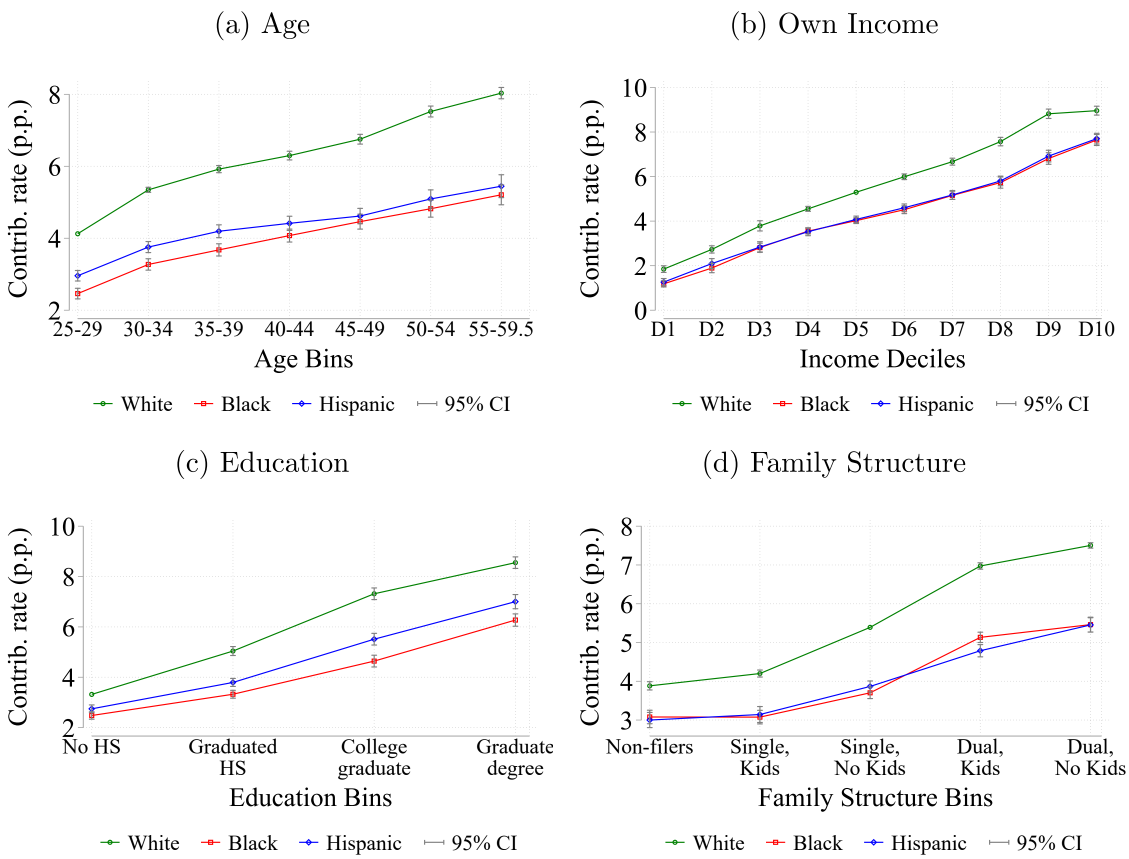
Notes: This figure presents i) the racial composition (bars, right axes) and ii) the regression coefficients (dots, left axes), from our fully saturated model (defined in Equation 1) for four important mediating channels: age (panel (a)), income (panel (b)), education (panel (c)), and spousal income (panel (d)). The regression outcome is employee contribution plus employer matching rate (% of income). Appendix A.2.2 provides definitions for the outcome and mediating channels.

Figure A.10: Share of workers who stay at their firm across the income growth distribution



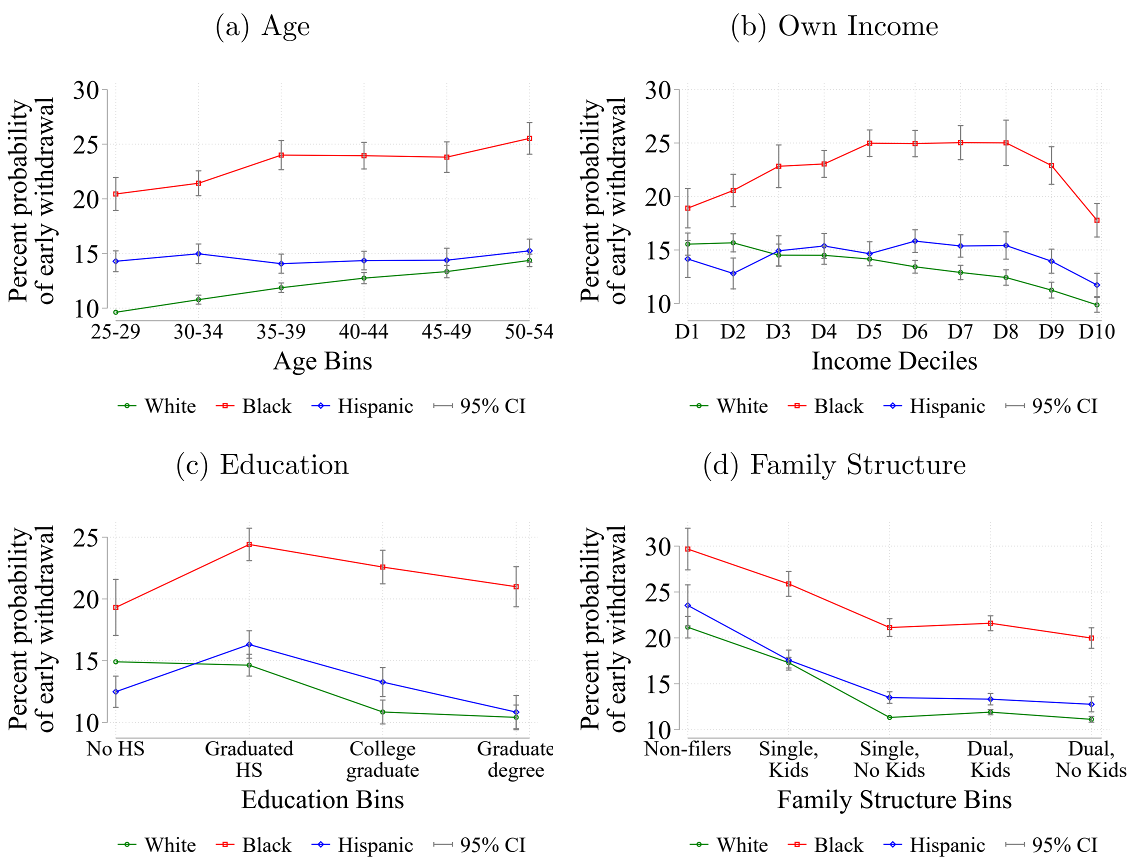
Notes: Of the ACS respondents we observe in year t and who satisfy our requirements to be in the early withdrawal sample, we plot the share of workers who remain at the same main firm (i.e., the firm who pays them the most) in years $t-1$, t, and $t+1$ over income growth ventiles.

Figure A.11: Racial contribution rates across different specifications



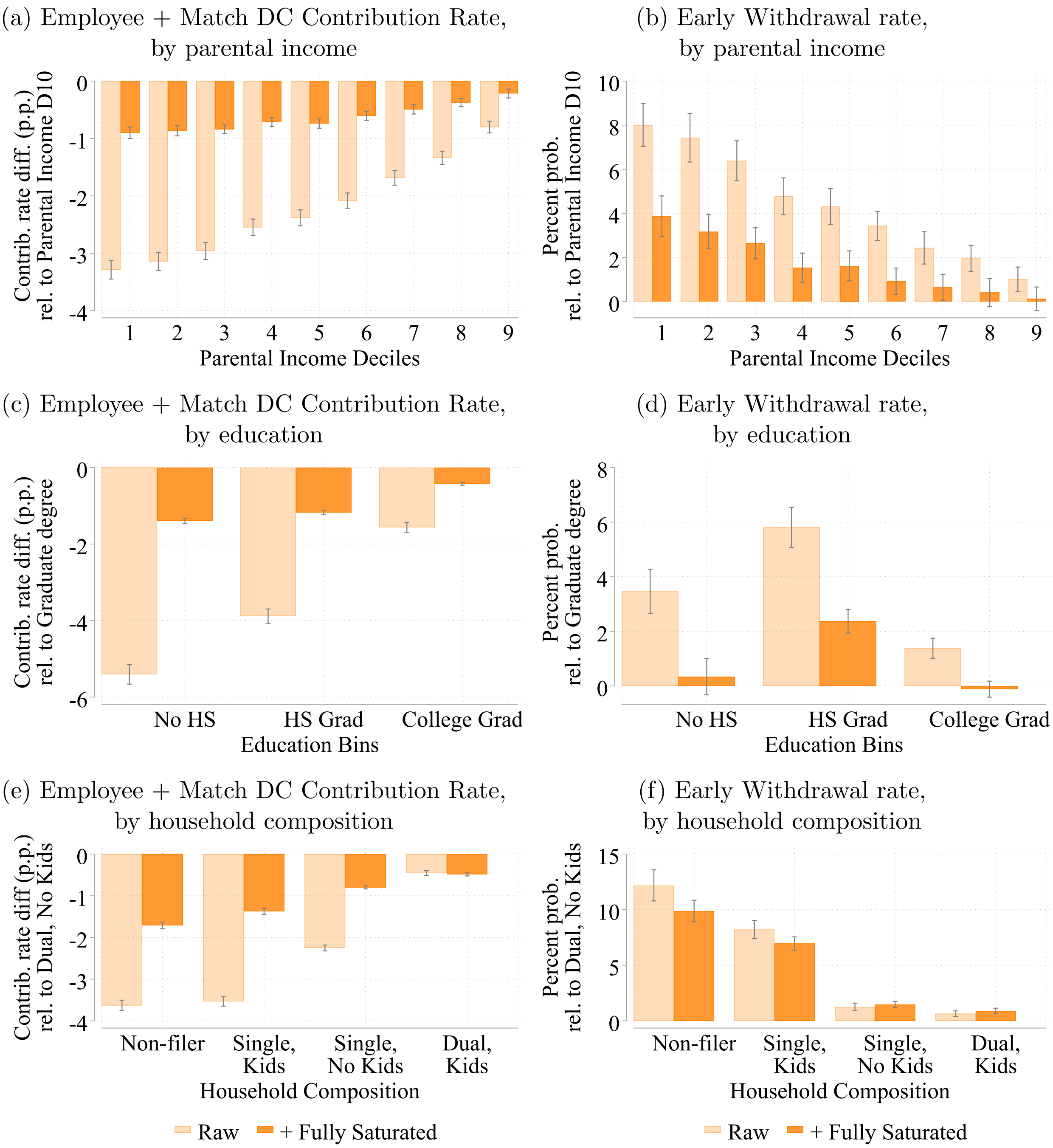
Notes: This figure presents the average racial contribution rates (employee plus employer contribution rates) across different age bins (panel (a)), own income deciles (panel (b)), education bins (panel (c)), and household composition groups (panel (d)). The estimates come from the raw specification without other mediating channels. The model is $y_{it} = \alpha + \beta_1 group\_i + \zeta(group\_i \cdot race_i) + \epsilon_{it}$. We add the intercept to put these numbers into levels rather than differences. 95% confidence intervals are included; standard errors are clustered by EIN. Please note that the confidence intervals reflect differences relative to the base category, which varies depending on the specification. For age, it is White workers of ages 25-29. For own income, it is White employees in the 5th income decile. For education, it is White workers without a high school degree. For family structure, it is White employees who are single and have no children.

Figure A.12: Racial early withdrawal rates across different specifications



Notes: This figure, analogous to Figure A.11, presents the average racial rates in leakage (probability of an early withdrawal > $,1000) across different age bins (panel (a)), own income deciles (panel (b)), education bins (panel (c)), and household composition groups (panel (d)). Please see figure notes of Figure A.11 for more details.

Figure A.13: Gaps by parental income, education, and household composition

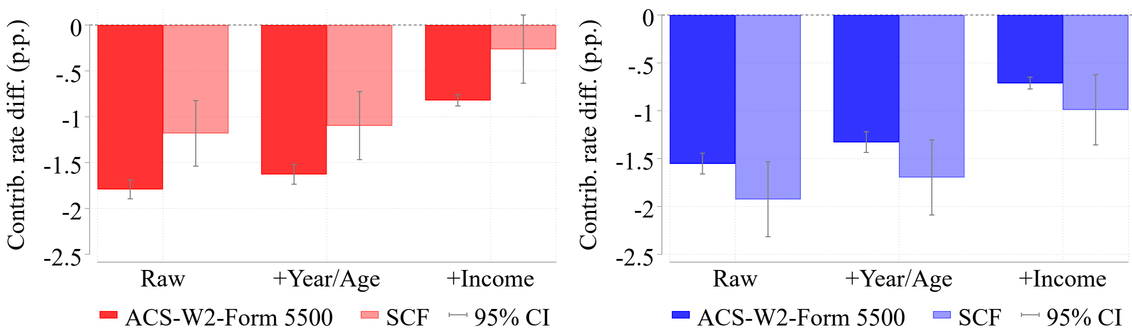


Notes: In this figure, we explore gaps in employee plus employer contribution and early withdrawal rates by various observable variables. Using the framework from eq. (1), we modify *group* here to be indicators for parental income deciles, education bins, and family structure groups, respectively. In each case, the omitted category is explicitly stated on the y-axis. For example, for parental income, the omitted category are workers with parents in the top income decile. With the bars, we graph the estimated coefficients for the *group* indicators from the univariate regression (i.e., "Raw") and the regression with the full set of individual-level characteristics (i.e., "Fully Saturated") which is consistent with the definition in Figure 2.

Figure A.14: Comparison of our results with the Survey of Consumer Finances

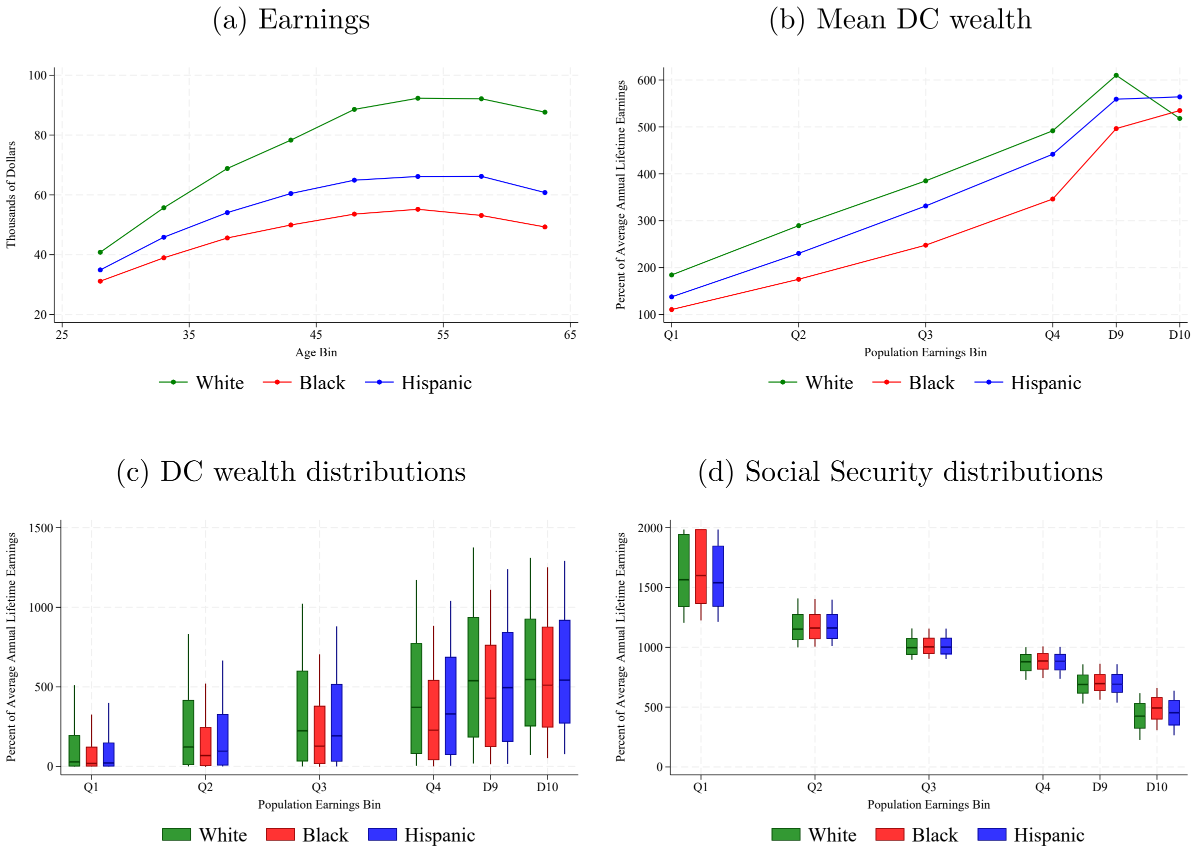
(a) Employee DC contribution rate, Black-White gap

(b) Employee DC contribution rate, Hispanic-White gap



Notes: This figure compares our baseline results from the Form 5500 sample with those from the SCF. Panel (a) shows the estimates of the Black-White gap, while panel (b) is for the Hispanic-White gap. Standard errors are calculated by bootstrapping with 500 replications, using the scfcombo command. The darker bars are the same estimates for the employee DC contribution rate as in Figure 2(a), while the lighter bars represent the estimates from the SCF. For more information, please see the figure notes for Figure 2.

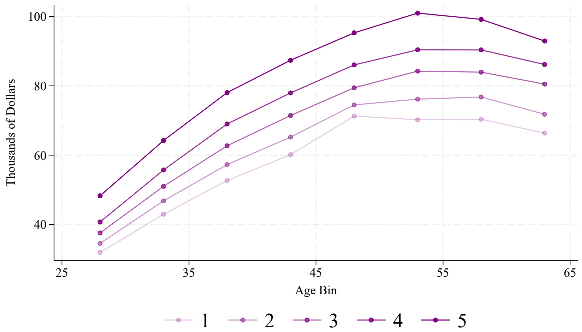
Figure A.15: Microsimulation model: Key outputs by Race



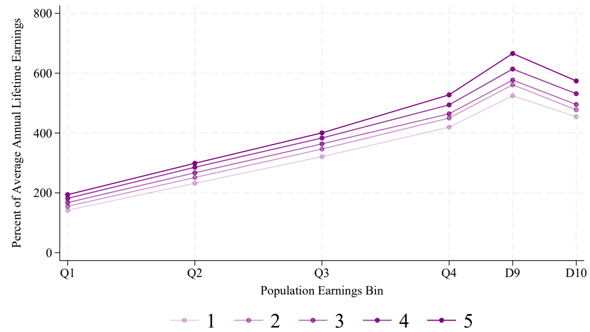
Notes: This figure illustrates the features of main outputs from our micro-simulation model. Panel (a) shows mean values by race and age bins 25-29, 30-34, . . . , and 60-65. Note that the last age bin contains six ages. In panel (a), earnings are the sum of wage income and deferred compensation. Panel (b) shows DC wealth at retirement divided by the simple average of earnings during working years for each race and population earnings bin group. Lifetime earnings groups are divided into six bins–the bottom four quintiles and the top two deciles. Panel (c) illustrates the heterogeneity in DC wealth at retirement within each race and lifetime earnings group. Percentiles shown are p10, p25, p50, p75, and p90. The measure of wealth shown is the same as in panel (b), DC wealth at retirement divided by average lifetime earnings. Panel (d) shows the same percentiles for the present value of all Social Security distributions over average lifetime earnings.

Figure A.16: Microsimulation model: Key outputs by Parental Income

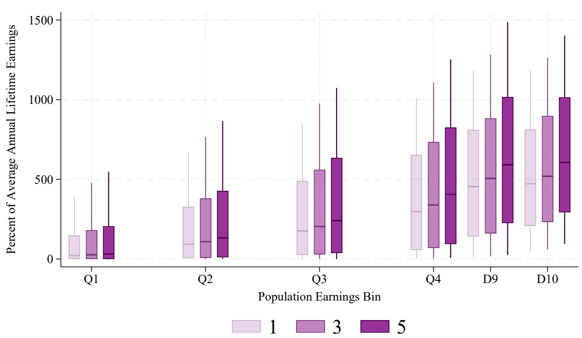
(a) Earnings



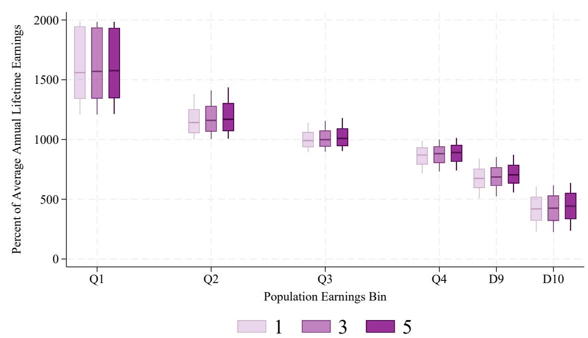
(b) Mean DC wealth



(c) DC wealth distributions



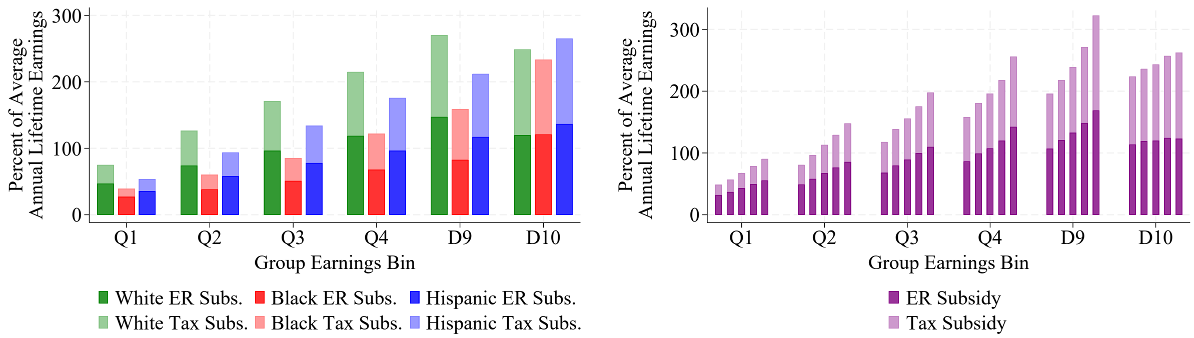
(d) Social Security distributions



Notes: This figure illustrates the features of main outputs from our micro-simulation model. Panel (a) shows mean values by parental income bin and age bins 25-29, 30-34, . . . , and 60-65. Note that the last age bin contains six ages. In panel (a), earnings are the sum of wage income and deferred compensation. Panel (b) shows DC wealth at retirement divided by the simple average of earnings during working years for each parental income bin and population earnings bin group. Lifetime earnings groups are divided into six bins–the bottom four quintiles and the top two deciles. Panel (c) illustrates the heterogeneity in DC wealth at retirement within each parental income bin and lifetime earnings group. Percentiles shown are p10, p25, p50, p75, and p90. The measure of wealth shown is the same as in panel (b), DC wealth at retirement divided by average lifetime earnings. Panel (d) shows the same percentiles for the present value of all Social Security distributions over average lifetime earnings.

Figure A.17: Contributions of employer and tax subsidies to retirement wealth, group-specific bins

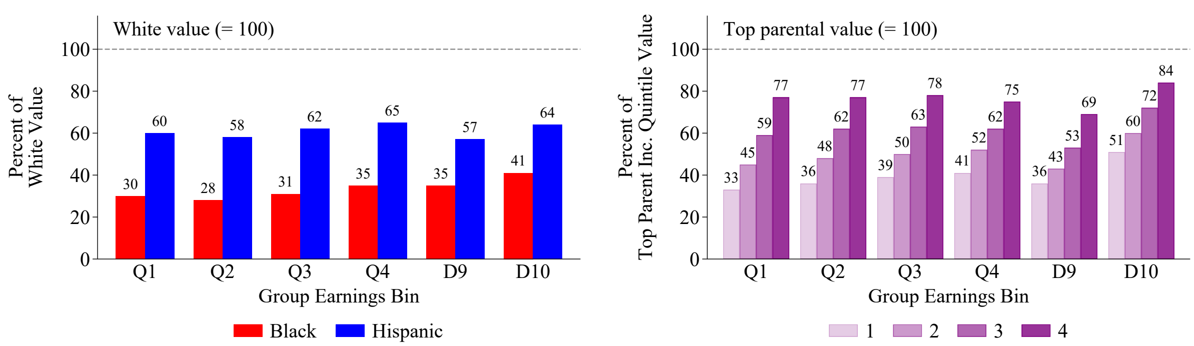
(a) By race (b) By parental income



Notes: These figures show lifetime employer and tax subsidies as a percentage of average annual lifetime earnings by lifetime earnings group and by either race or parental income. Panel (a) shows these subsidies by race, and panel (b) shows them by quintiles ('bins') of parental income. In both panels, the darker bars show average employer matching subsidies and the lighter bars show average tax subsidies to retirement savings. Lifetime earnings groups are divided into six bins–the bottom four quintiles and the top two deciles. Earnings bins are calculated within race and parental income groups. This analysis differs from Figure 8 which shows the analysis for the case where lifetime earnings bins are defined at the population level.

Figure A.18: Total tax subsidy relative to top value by earnings, own group earnings bins

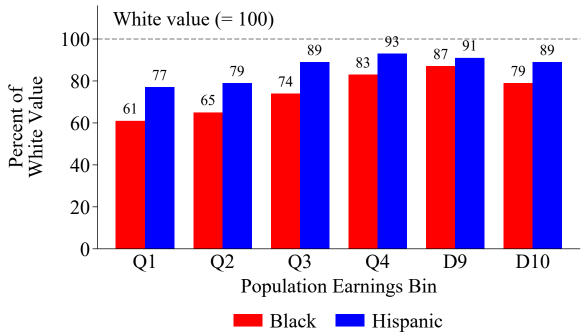
(a) By race and own-race quintiles (b) By parental income and own-group quintiles



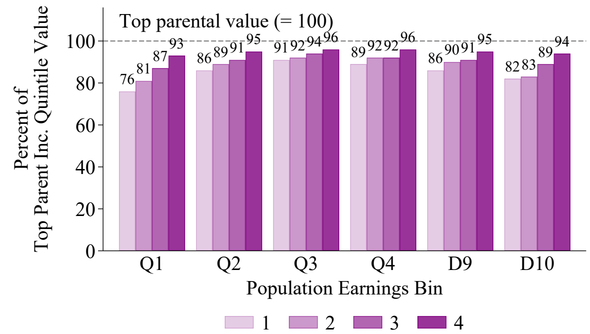
Notes: Panel (a) shows the the subsidies earned by Black and Hispanic workers relative to White workers in the same earnings bin. Panel (b) shows the subsidies earned by workers with parents in different income quintiles relative to workers with parents in the top parental income quintile. Lifetime earnings groups are divided into six bins–the bottom four quintiles and the top two deciles. Earnings bins are calculated within each race and parental income group. Figure A.19 shows results where lifetime earnings bins are at the population level.

Figure A.19: Total tax subsidy relative to top value by earnings, population earnings bins

(a) By race and population quintiles



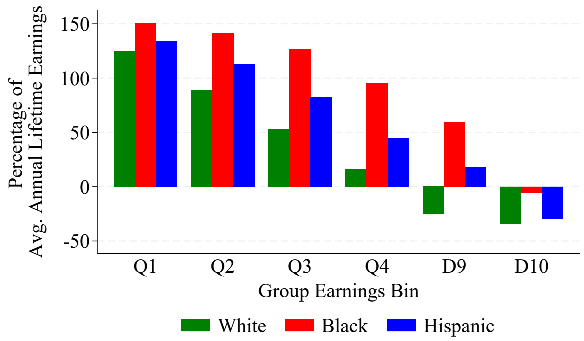
(b) By parental income and population quintiles



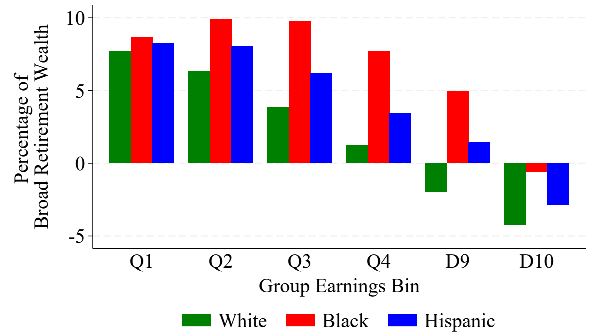
Notes: Panel (a) shows the the subsidies earned by Black and Hispanic workers relative to White workers in the same earnings bin. Panel (b) shows the the subsidies earned by workers with parents in different income quintiles relative to workers with parents in the top parental income quintile. Lifetime earnings groups are divided into six bins–the bottom four quintiles and the top two deciles. Earnings bins are calculated at the population level. Figure A.18 shows results where lifetime earnings bins are defined within race and parental income group.

Figure A.20: Change in retirement wealth measures, by race and parental income, group-specific bins

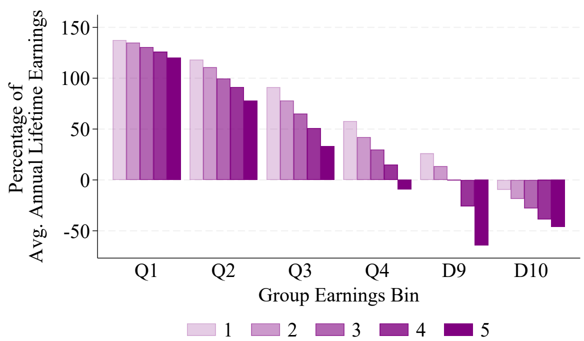
(a) Change in ret. wealth by race



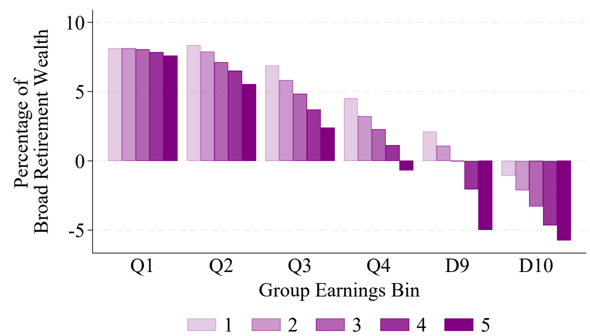
(b) Δ broad ret. wealth (DC + SS) by race



(c) Change in ret. wealth, by par. inc.

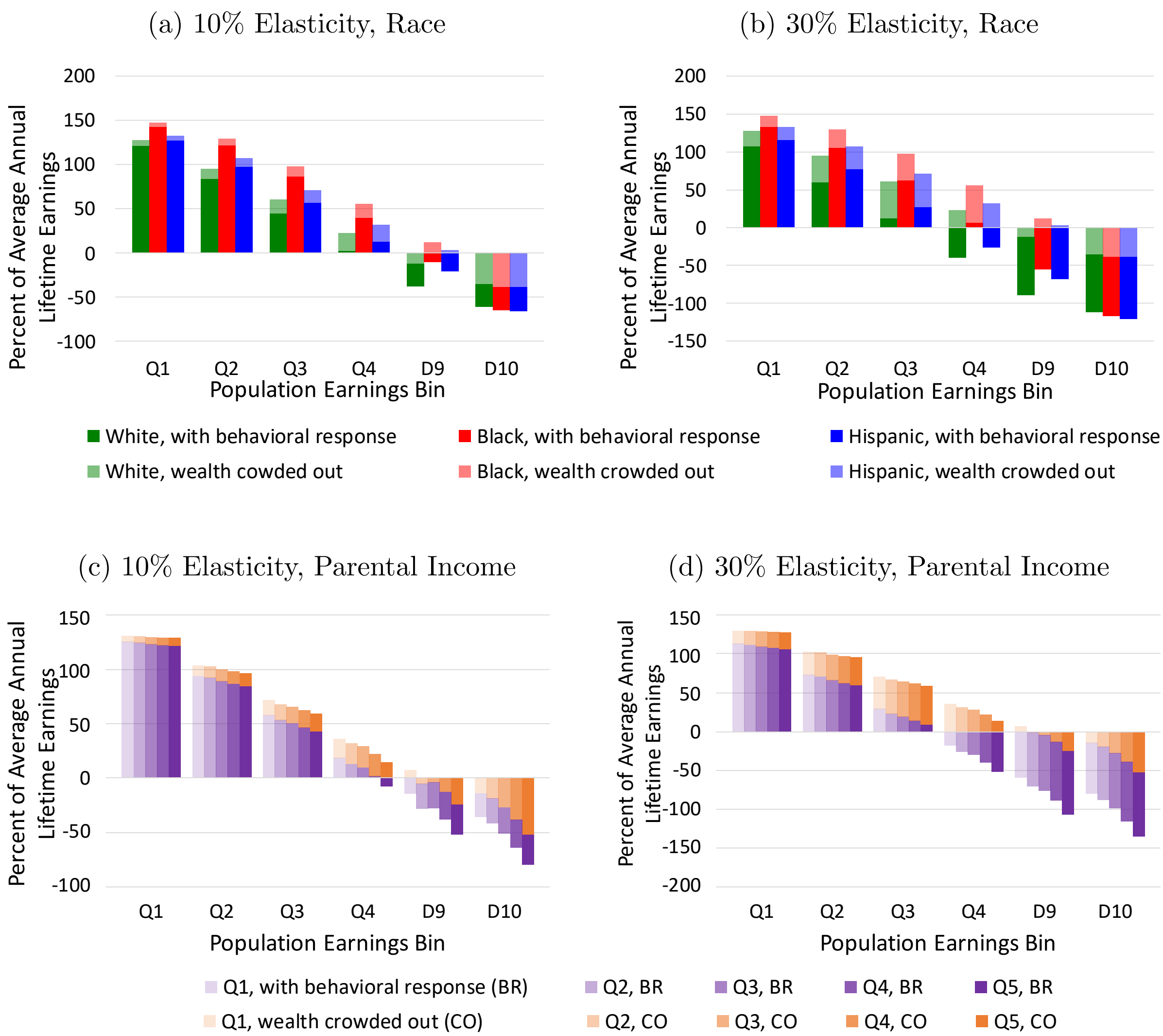


(d) Δ broad ret. wealth (DC + SS), by par. inc.



Notes: This figure illustrates the impact of our baseline counterfactual exercise on measures of retirement wealth. This counterfactual exercise distributes each firm's aggregate employer matches such that all workers in that firm receive the same contribution as a proportion of their earnings. It further distributes the aggregate federal tax expenditure so that all workers receive the same contribution in proportion to their lifetime earnings. We show the effect on two outcomes: panels (a) and (c) show the change in DC wealth on retirement, with wealth expressed as a proportion of average annual working life earnings. Panels (b) and (d) show proportionate change in broad retirement wealth (where broad retirement wealth is the sum of DC wealth and Social Security). Lifetime earnings groups are divided into six bins–the bottom four quintiles and the top two deciles. Earnings bins are calculated within race and parental income group. Figure 9 shows results where lifetime earnings bins are defined at the population level.

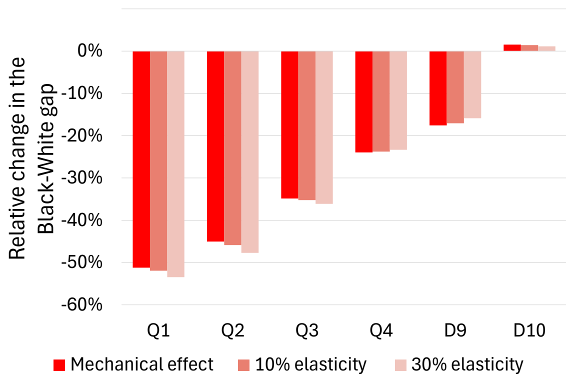
Figure A.21: Change in retirement wealth measures under alternative assumptions about the elasticity of employees' savings to incentives



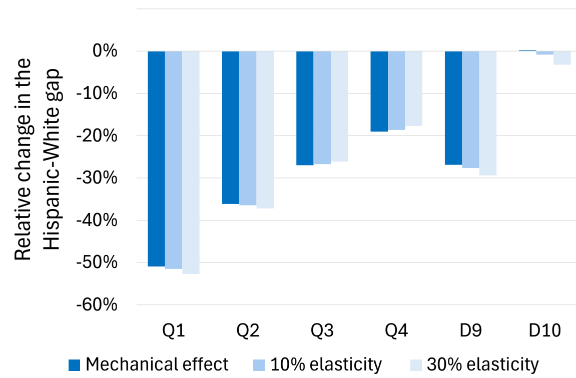
Notes: This figure shows alternate versions of Figure 9, panels (a) and (c), under different assumptions about behavioral responses to a reduction in savings incentives. We assume that each additional dollar of tax or employer matching incentives creates an additional \$0.1 (left panels) or \$0.3 (right panels) of additional employee savings. Moving from subsidies to flat contributions, therefore, crowds out some employee contributions (i.e., equal to either 10% or 30% of the baseline level of tax and matching subsidies). The change in retirement wealth under the counterfactual policy is expressed as a percentage of average annual lifetime earnings by population lifetime earnings quintiles, with the top quintile split into two deciles. In all panels, the height of bars in solid colors represents the change in wealth, taking into account the behavioral responses, while the transparent portion represents the wealth crowded out due to these behavioral responses.

Figure A.22: Relative change in DC wealth gaps under alternative assumptions about the elasticity of employees' savings to incentives

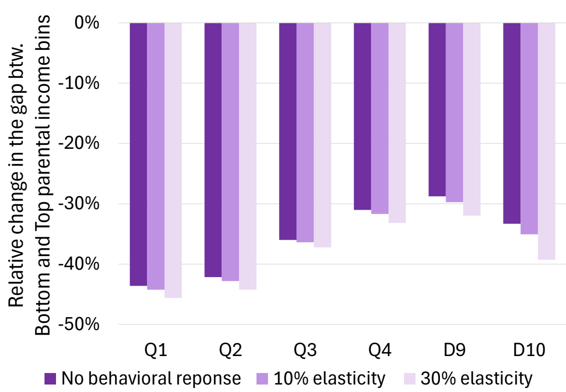
(a) Black-White gap



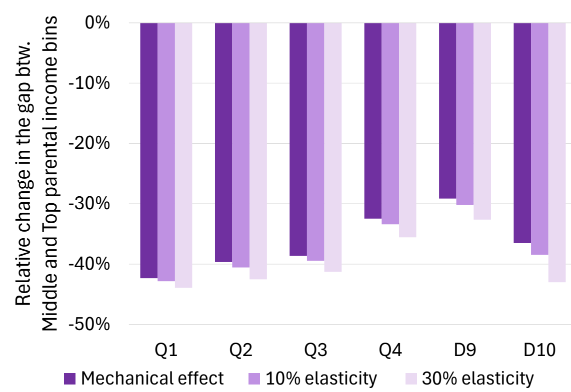
(b) Hispanic-White gap



(c) Bottom-Top parental inc. bin gap



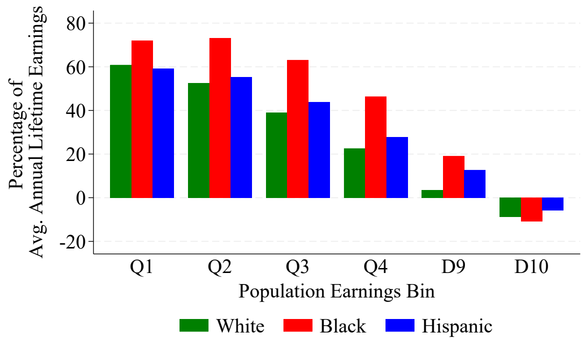
(d) Middle-Top parental inc. bin gap



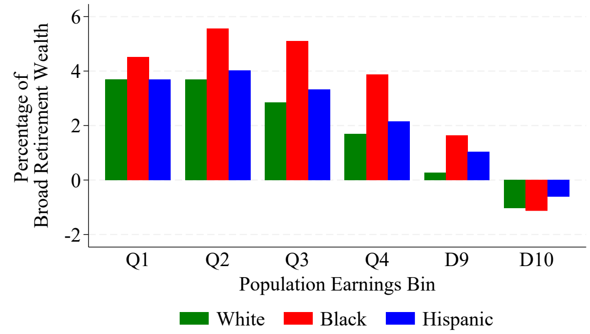
Notes: This figure shows the relative change in the DC wealth gaps under different assumptions about the behavioral response to a reduction in savings incentives. We assume that each additional dollar of tax or employer matching incentives creates an additional \$0, \$0.10 or \$0.30 of additional employee savings. The first of these returns the mechanical effect of removing and redistributing the subsidies, and the change in the gaps under this assumption are those in our baseline (see the last rows in both panels of Table 3). The second two assumptions imply that moving from subsidies to flat contributions crowd out some employee contributions (i.e., equal to either 10% or 30% of the baseline level of tax and matching subsidies). The change in DC wealth gaps under the counterfactual policy is expressed as a percentage of the racial and parental income gap for each lifetime earnings quintiles, with the top quintile split into two deciles. The lifetime earnings quintiles and deciles are defined at the population level.

Figure A.23: Change in retirement wealth measures under different counterfactuals by race

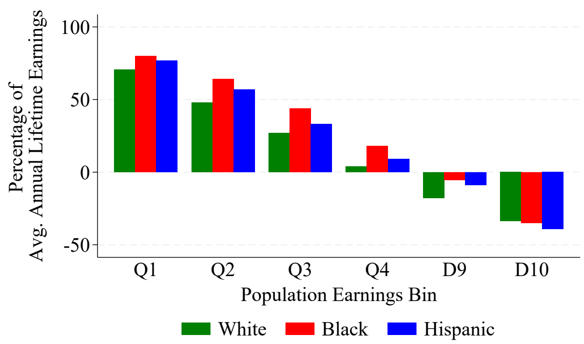
(a) Δ in ret. wealth, match counterfactual



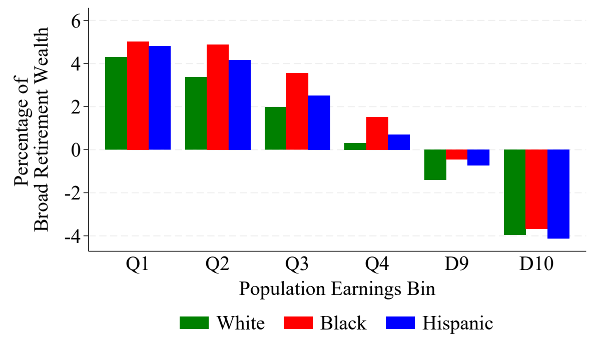
(b) Δ broad ret. wealth (DC + SS), match counterfactual



(c) Δ in ret. wealth, tax counterfactual

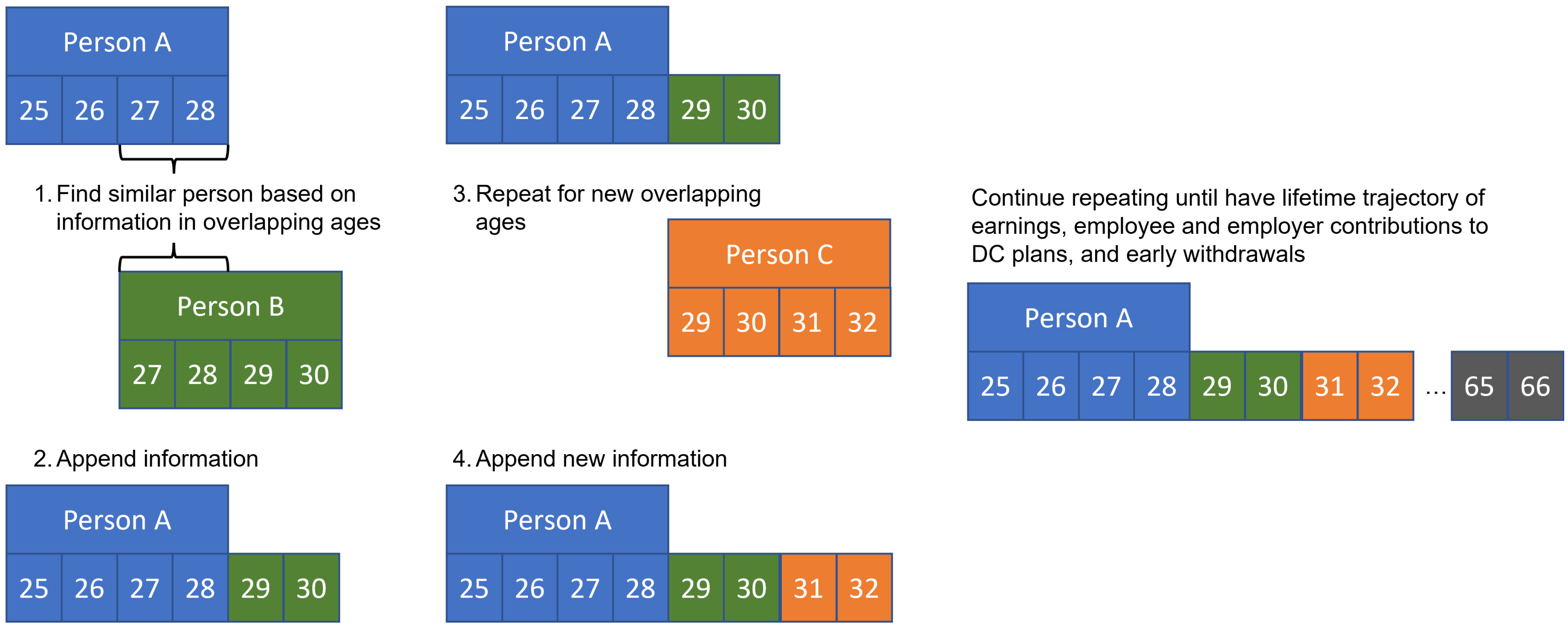


(d) Δ broad ret. wealth (DC + SS), tax counterfactual



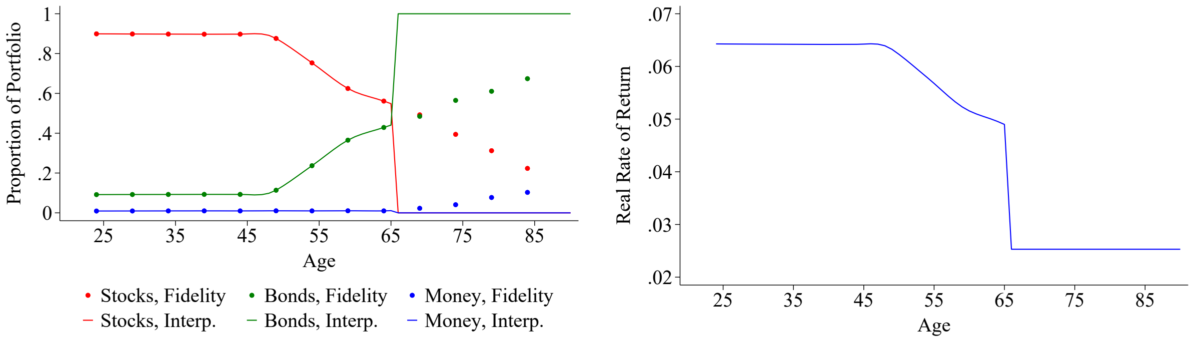
Notes: This figure illustrates the impact of two supplementary counterfactual exercises on measures of retirement wealth. The 'match counterfactual' (panels (as) and (b) exercise distributes the aggregate employer matches in each firm so that all workers in that firm receive the same proportion of their earnings. The 'tax' counterfactual (panels (c) and (d) distributes the aggregate federal tax expenditure so that all workers receive a contribution that is in proportion to their lifetime earnings. We show the effect on two outcomes: panels (a) and (c) show the change in DC wealth on retirement, with the change in wealth expressed as a proportion of average annual working life earnings. Panels (b) and (d) show proportionate change in broad retirement wealth, where broad retirement wealth is the sum of DC wealth and Social Security. Lifetime earnings groups are divided into six bins—the bottom four quintiles and the top two deciles. Earnings bins are calculated at the population level.

Figure A.24: Simulating lifetime trajectories from shorter panels



Notes: This figure shows a schematic of the imputation model used to simulate lifetime trajectories of earnings, deferred compensation, and DC plan withdrawals for workers aged 25 to 65 from the shorter panels available to us for individual workers. We construct full lifetime trajectories by repeatedly matching individuals across overlapping age bins. For example, in 1. Persons A and B have similar earnings and job characteristics in the overlapping ages (27 and 28), so we append Person B's information at 29 and 30 to Person A and therefore add two additional years of data to the trajectory of Person A. We repeat this process at increasing ages (31-32, 33-34, ..., 65-66) to create a full lifetime path of earnings, employee and employer contributions to DC plans, and early withdrawals from ages 25 to 65.

Figure A.25: Portfolio shares and rate of return

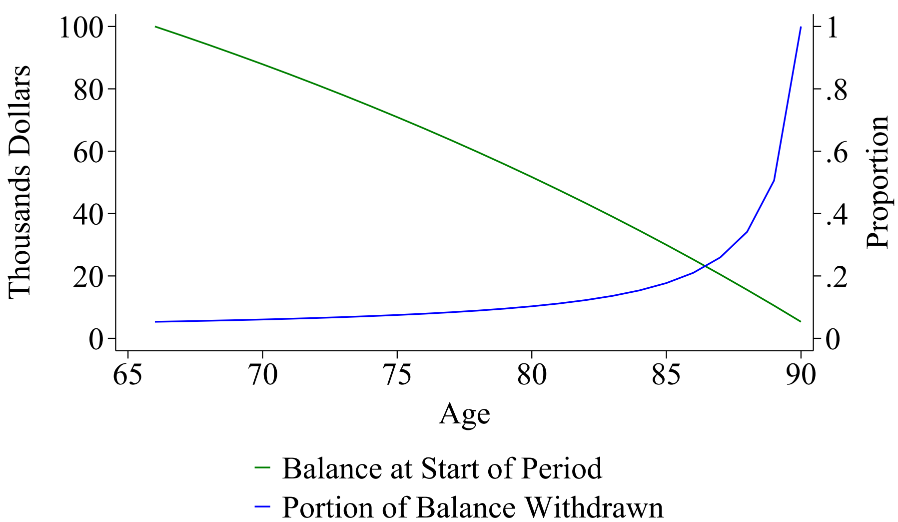


(a) Interpolated Portfolio Shares

(b) Real Rate of Return over the Lifecycle

Notes: This figure shows underlying parameterizations for portfolio composition and returns in the micro-simulation model described in Appendix E. The points of panel (a) show actual portfolio shares for Fidelity Freedom Funds at each age. We interpolate shares between these observations for each integer age, given by the lines in panel (a), although we assume exclusive investment in bonds at retirement. Panel (b) shows the real rate of return by age, which is determined by the portfolio composition and the associated returns of each component.

Figure A.26: Withdrawal Path



Notes: This figure shows the process for estimating withdrawals in retirement in the micro-simulation model described in Appendix E. For the purposes of illustration, we suppose an individual retires with wealth balance of $100,000, which they draw down until their last year of life at age 90. The left axis corresponds to the green line, showing the wealth balance at the start of each period. The right axis corresponds to the blue line, showing the proportion of remaining wealth balance that is withdrawn each period. This process ensures constant withdrawals each period and a smooth draw-down of wealth in retirement.

Table A.1: Summary statistics by sample

Variable	(1) ACS	(2) ACS, restricted	(3) DC Access	(4) Form 5500	(5) Parent-Form 5500
Average age	41.63	41.66	41.81	41.21	30
Employee contribution ($)	$2,213	$2,248	$2,855	$3,351	$1,882
Box 1 W-2 total compensation	$61,140	$61,880	$67,780	$72,810	$50,050
Spousal Box 1 W-2 total compensation	$9,915	$9,914	$10,000	$9,842	$9,741
Own contrib. (% of inc.)	2.7%	2.7%	3.4%	3.8%	2.8%
Own contrib. (% of inc., contrib. >0)	5.9%	5.9%	6%	5.8%	4.7%
Positive contrib. dummy (%)	45%	45.6%	57.3%	65.2%	59.5%
Match contrib. (% of inc.)				1.9%	1.6%
Own + match contrib. (% of inc.)				5.7%	4.4%
Max match - own contrib. (% of inc.)				1.7%	2%
Positive withdr. dummy	14.3%	14.2%	14.2%	16.5%	15.4%
Positive withdr. dummy (withdr. >$1000)	11.8%	11.8%	11.8%	13.5%	12%
Number of unique individuals	12,480,000	12,140,000	9,595,000	1,722,000	471,200

From left to right, this table presents our (1) entire sample of ACS respondents linked to administrative tax records, (2) ACS sample restricted on non-missing individual characteristics, (3) full DC access sample, (4) Form 5500 sample, and (5) parent-Form 5500 sample. For more information about the different samples, please see Section 3.4. Spousal income includes spouses claimed on Form 1040 who made $0 in earnings (41% of our Form 5500 sample).

Table A.2: Administrative Data and SCF Comparison

Outcome	Statistic	Worker Samples		DC Access Samples	
		ACS-W2	SCF Worker	ACS-W2 + Form 5500	SCF DC Access
Average age	Mean	41.7	42.6	41.2	42.5
Wage Compensation	Mean	\$61,880	\$70,797	\$72,810	\$86,813
DC Access	Mean	78.4%	49.6%	100%	100%
DC Participation	Mean	45.6%	35.3%	65.2%	71.3%
Employee contribution rate	Mean	2.7%	2.3%	3.8%	4.8%
	Count	12,140,000	8,430	1,722,000	4,097

Notes: This table presents means and observation counts for our administrative and SCF samples. The worker samples compare ACS and SCF workers, independent of DC plan access. The DC access samples compare our Form 5500 sample with SCF workers who have DC access.

Table A.3: Distribution of DC wealth at retirement by source and by group - baseline model, group bins

(a) By race

Value	Group	Q1	Q2	Q3	Q4	D9	D10
Wealth from employee contributions ($'000)	White	18.4	52.3	98.9	185.8	349.7	529.2
	Black	5.5	14.9	27.5	51.1	95.0	262.3
	Hispanic	10.9	29.8	57.4	107.8	201.8	413.6
Wealth from employer contributions ($'000)	White	7.2	21.5	40.5	73.7	140.1	251.6
	Black	2.9	7.5	14.3	26.4	44.8	119.5
	Hispanic	5.0	14.2	26.9	46.9	81.9	184.9
Wealth from tax subsidies ($'000)	White	4.3	15.2	31.3	59.7	116.9	271.2
	Black	1.3	4.3	9.6	21.0	41.2	111.2
	Hispanic	2.6	8.8	19.4	38.7	66.3	174.1
Total DC Wealth ($'000)	White	30.0	89.0	170.7	319.2	606.7	1052.0
	Black	9.8	26.8	51.4	98.6	181.0	493.0
	Hispanic	18.5	52.8	103.7	193.4	350.0	772.6
Social Security Wealth ($'000)	White	218.1	318.8	401.7	499.7	586.3	652.0
	Black	177.8	255.6	312.1	382.6	468.9	575.0
	Hispanic	209.9	288.4	354.9	439.8	528.6	615.4

(b) By parental income

Value	Group	Q1	Q2	Q3	Q4	D9	D10
Wealth from employee contributions ($'000)	Bin 1	8.9	25.1	49.7	93.2	179.4	379.6
	Bin 3	15.2	43.2	82.2	149.0	277.1	481.6
	Bin 5	23.9	66.7	127.2	254.9	451.7	588.9
Wealth from employer contributions ($'000)	Bin 1	3.9	11.1	22.1	40.1	72.7	170.5
	Bin 3	6.3	18.3	35.0	60.5	111.5	224.2
	Bin 5	9.6	27.6	51.8	101.0	185.5	284.9
Wealth from tax subsidies ($'000)	Bin 1	2.0	7.2	16.2	33.2	60.4	165.0
	Bin 3	3.6	12.4	26.0	49.8	89.3	230.6
	Bin 5	6.0	20.1	41.5	80.7	168.8	322.2
Total DC Wealth ($'000)	Bin 1	14.8	43.3	88.0	166.5	312.5	715.1
	Bin 3	25.1	73.9	143.2	259.3	477.9	936.4
	Bin 5	39.5	114.4	221.0	436.6	806.0	1196.0
Social Security Wealth ($'000)	Bin 1	190.8	275.0	341.9	426.9	521.5	612.9
	Bin 3	212.7	306.2	383.8	476.7	564.1	640.6
	Bin 5	233.3	340.1	430.8	531.0	612.0	665.0

Notes: This table gives average DC wealth, its components, and Social Security wealth by race (panel (a)) and parental income (panel (b)). Value rows 1 - 3 of each panel show average values for each component of DC wealth. Value row 4 gives total DC wealth. Value row 5 is the average value of Social Security. Panel (a) shows results separately for White, Black and Hispanics workers. Panel (b) shows results separately by parental income quintiles (bins), with Bins 1, 3, and 5 shown. In both tables, columns show results by own lifetime earnings. There are six lifetime earnings bins—the bottom four quintiles and the top two deciles. Earnings bins are defined within each race and parental income group. Table 2 gives the same analysis with lifetime earnings bins defined at the population level.

Table A.4: Change in DC wealth at retirement under the counterfactual tax and employer contribution policy, group-specific bins

(a) By race

Value	Group	Q1	Q2	Q3	Q4	D9	D10
Baseline Total DC	White	30.0	89.0	170.7	319.2	606.7	1052.0
Wealth ($'000)	Black	9.8	26.8	51.4	98.6	181.0	493.0
	Hispanic	18.5	52.8	103.7	193.4	350.0	772.6
Baseline DC	B-W Gap	67.5%	69.9%	69.9%	69.1%	70.2%	53.1%
Wealth Gap	H-W Gap	38.4%	40.7%	39.3%	39.4%	42.3%	26.6%
Absolute change in	White	+19.1	+25.9	+22.1	+10.1	-23.7	-72.2
DC Wealth ($'000)	Black	+16.3	+27.9	+35.5	+37.0	+32.1	-5.9
	Hispanic	+18.9	+27.5	+28.5	+21.9	+12.5	-39.8
Counterfactual DC	B-W Gap	47.0%	52.4%	55.0%	58.8%	63.4%	50.3%
Wealth Gap	H-W Gap	24.0%	30.0%	31.4%	34.6%	37.8%	25.2%
Relative change in the	B-W Gap	-30.3%	-25.1%	-21.4%	-14.9%	-9.6%	-5.4%
racial DC wealth gap	H-W Gap	-37.6%	-26.1%	-19.9%	-12.2%	-10.6%	-5.1%

(b) By parental income

Value	Group	Q1	Q2	Q3	Q4	D9	D10
Baseline Total DC	Bin 1	14.8	43.3	88.0	166.5	312.5	715.1
Wealth ($'000)	Bin 3	25.1	73.9	143.2	259.3	477.9	936.4
	Bin 5	39.5	114.4	221.0	436.6	806.0	1196.0
Baseline DC	1-5 Gap	62.6%	62.2%	60.2%	61.9%	61.2%	40.2%
Wealth Gap	3-5 Gap	36.6%	35.4%	35.2%	40.6%	40.7%	21.7%
Absolute change in	Bin 1	+16.7	+26.6	+29.6	+26.8	+17.7	-14.4
DC Wealth ($'000)	Bin 3	+19.1	+27.1	+25.5	+16.8	-0.5	-52.1
	Bin 5	+20.7	+25.1	+15.7	-6.7	-70.8	-107.0
Counterfactual DC	1-5 Gap	47.7%	49.9%	50.3%	55.0%	55.1%	35.7%
Wealth Gap	3-5 Gap	26.6%	27.6%	28.7%	35.8%	35.1%	18.8%
Relative change in the	1-5 Gap	-23.8%	-19.7%	-16.4%	-11.0%	-10.0%	-11.3%
parental income DC wealth gap	3-5 Gap	-27.4%	-22.0%	-18.4%	-11.9%	-13.9%	-13.4%

Notes: This table summarizes the effect on wealth of our counterfactual exercise. Panel (a) gives results by race, and panel (b) gives results by parental income quintiles (bins) with results for Bins 1, 3, and 5 shown. Value row 1 in each table shows baseline wealth. Value row 2 gives the baseline gap as a percentage of the White level (panel (a)), and the average level for those with the richest parents (panel (b)). Value row 3 shows the absolute change in DC wealth under the counterfactual. Value row 4 gives the counterfactual gap as a percentage of the White level (panel (a)), and the average level for those with the richest parents (panel (b)). Value row 5 gives the relative change in the percentage gaps obtained in moving from the baseline (value row 2) to the counterfactual (value row 4). In both panels, each row is divided into six bins—the bottom four quintiles and the top two deciles. Earnings bins are defined within each race and parental income group. Table 3 gives the same analysis with lifetime earnings bins defined at the population level.

Table A.5: Parameter and variable definitions

Earnings, Wealth, Social Security		**State Variables**	
$e_{i,t}$	Earnings	i	Individual
α	Discount rate in retirement	$t \in \{25, \ldots, 90\}$	Age
c_t^j	Consumption in retirement	$j \in \{DC, BK\}$	Type of savings vehicle
$aime_i$	Average indexed monthly earnings		
e^{max}	Social Security taxable maximum	**Taxes**	
δ_1	First PIA bend point	$T(\cdot\,,\,\cdot\,,\,\cdot)$	Federal income tax function
δ_2	Second PIA bend point	$\tau_{i,t}^{e,j}$	Taxes owed on earnings
$ss_{i,t}$	Annual Social Security benefits	$\tau_{i,t}^{ss,j}$	Taxes owed on Social Security Benefits
		$\tau_{i,t}^{c,j}$	Taxes owed on savings
Wealth Flows		$\tau_{i,t}^{r,j}$	Taxes owed on returns
$dc_{i,t}^{ee}$	Employee savings	$\tau_{i,t}^{w,j}$	Taxes owed on withdrawals
$dc_{i,t}^{er}$	Employer savings	$\hat{\tau}_{i,t}^{r,j}$	Hypothetical taxes owed on returns
$w_{i,t}^j$	Savings account withdrawals		
$f_{i,t}^j$	Flow into retirement account	**Lifetime Measures**	
$B_{i,t}^j$	Wealth balance	$A_{i,t}^{DC}$	DC Wealth
$B_{i,t}^{p,j}$	Principal part of wealth balance	$SS_{i,t}$	Social Security Wealth
$B_{i,t}^{g,j}$	LTCG part of wealth balance	$A_{i,t}^{BR}$	Broad Retirement Wealth
$w_{i,t}^{k,j}$	LTCG portion of withdrawal	$A_{i,t}^{BR,BK+SS}$	Broad Retirement Wealth brokerage WC
		A_i^T	DC tax subsidy
Rate of Return		DC_i^{EE}	Value of employee contributions
ρ_t	Rate of return at age t	DC_i^{ER}	Value of employer contributions
$r_{i,t}^{g,j}$	Return from unrealized capital gain	A_i^{EE}	Wealth attributable to employee
$r_{i,t}^{k,j}$	Return from LTCG distributions	A_i^{ER}	Employer subsidy
$r_{i,t}^{i,j}$	Return from interest distributions	LE_i	Value of lifetime income
s_t^k	Portion of assets invested in stocks		
s_t^b	Portion of assets invested in bonds	**Counterfactuals**	
s_t^m	Portion of assets invested in money	$C_i'^T$	Counterfactual tax subsidy
ρ^k	Real rate of return on stocks	$A_i'^{DC}$	DC Wealth under tax CF
ρ^b	Real rate of return on bonds	$A_i^{BR\prime DC}$	Broad Retirement Wealth under tax CF
ρ^m	Real rate of return on money	dc^*	Counterfactual employer match
χ^g	Share from unrealized capital gain	$A_{i,t}^{*DC}$	DC wealth under ER CF
χ^k	Share from LTCG distributions	$A_{i,t}^{BR*DC+SS}$	Broad Retirement Wealth under ER CF
χ^i	Share from interest distributions	$A_{i,t}^{\dagger DC}$	DC Wealth under combined (CB) CF
$\hat{r}_{i,t}$	Implied post-tax rate of return	$A_{i,t}^{BR\dagger}$	Broad Retirement Wealth under CB CF

Notes: This table gives the variables that enter the micro-simulation model, outlined in Appendix E.

Table A.6: Parameter values and sources

Parameter	Value	Source
e^{max}	\$128,400	Social Security Administration (2023b)
δ_1	\$895	Social Security Administration (2023a)
δ_2	\$5,397	Social Security Administration (2023a)
ρ^k	0.0688	Jordà et al. (2019)
ρ^b	0.0253	Jordà et al. (2019)
ρ^m	0.0103	Jordà et al. (2019)
σ^k_t, σ^b_t, σ^m_t	Figure A.25a	Fidelity (2023)
χ^g	0.5	Yahoo Finance, Sialm and Zhang (2020)
χ^k	0.4	Yahoo Finance, Sialm and Zhang (2020)
χ^i	0.1	Yahoo Finance, Sialm and Zhang (2020)

Notes: This table gives the values for parameters used in the micro-simulation model, outlined in Appendix E.