NBER WORKING PAPER SERIES

FINANCIAL INCLUSION ACROSS THE UNITED STATES

Motohiro Yogo
Andrew Whitten
Natalie Cox

Working Paper 33256
http://www.nber.org/papers/w33256

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2024

This research was conducted while Whitten was an employee of the U.S. Department of the Treasury. Any confidential taxpayer data used in this research were kept in a secured IRS data repository, and all results have been reviewed to ensure that no confidential information is disclosed. This research was derived in part from research activities performed pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement and Disability Research Consortium. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and do not necessarily reflect the views or the official positions of the Treasury, SSA, any agency of the U.S. Government, or the National Bureau of Economic Research. Neither the U.S. 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 paper. 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 U.S. Government or any agency thereof. For comments and discussions, we thank Edith Brashares, Adam Cole, Alice Henriques Volz, Ben Meiselman, Brian Melzer, Neviana Petkova, John Sabelhaus, Irina Stefanescu, and Alexander Zentefis. We also thank seminar participants at CUNY Graduate Center, Dartmouth College, Federal Reserve Bank of Chicago, Federal Reserve Bank of New York, Federal Reserve Bank of Philadelphia, Harvard University, MIT, National University of Singapore, Peking University, Princeton University, Rutgers University, Temple University, UCLA, University of Michigan, University of Wisconsin, U.S. Department of the Treasury, Wharton, the 2021 Office of Tax Analysis Research Conference, the 2021 NBER Conference on Innovative Data in Household Finance: Opportunities and Challenges, the 2022 Q-Group Spring Seminar, the 2022 NBER Public Economics Program Meeting, the 2022 UCLA Fink Center Conference on Financial Markets, the 2022 Chicago Household Finance Conference, and the 2023 AEA Annual Meeting.

Financial Inclusion Across the United States
Motohiro Yogo, Andrew Whitten, and Natalie Cox
NBER Working Paper No. 33256
December 2024
JEL No. D14, G51

ABSTRACT

We study retirement and bank account participation for the universe of U.S. households with a member aged 50 to 59 in the administrative tax data. ZCTA-level average income, income inequality, and racial composition predict retirement account participation for low-income households, conditional on household income and regional price parities. Income inequality also predicts bank account participation for low-income households. We estimate the causal effect of access to an employer retirement plan on participation. Recent policy proposals for universal access with automatic enrollment could increase participation by 19 percentage points in the lowest income quintile over ten years.

Motohiro Yogo
Department of Economics
Princeton University
Julis Romo Rabinowitz Building
Princeton, NJ 08544
and NBER
myogo@princeton.edu

Natalie Cox
Princeton University
nbachas@princeton.edu

Andrew Whitten
Department of the Treasury
andrew.whitten@treasury.gov

1. Introduction

An inclusive society should strive for financial participation of all households, regardless of income or race. Survey evidence shows that retirement and bank account participation are much lower for low-income and nonwhite households. However, household surveys have key limitations including small samples, a limited panel dimension, and potential measurement error. Big administrative data without these limitations could help us better understand and hopefully improve financial participation for low-income and nonwhite households. Toward that effort, we study retirement and bank account participation for the universe of U.S. households with a member aged 50 to 59 in the 2015 to 2019 administrative tax data. These data contain virtually all tax returns and information returns and have the same population count as the census.

In the lowest income quintile in 2019, 21 percent of households had retirement accounts, and 38 percent of households had access to an employer retirement plan. We define retirement accounts comprehensively to include employer retirement plans (both defined benefit and defined contribution plans) and individual retirement arrangements (IRAs). In comparison, 80 percent of the same group of households had bank accounts, according to the Survey of Consumer Finances. Although the previous literature portrays financial inclusion as a banking issue in poor areas or developing countries and often focuses on young adults, we think of access to employer retirement plans among low-income households in their prime retirement saving years as a financial inclusion issue in the United States. The heterogeneity in financial participation conditional on income implies that low participation is not a simple matter of not having enough income to save. Why do some low-income households manage to participate? We find that geography and access to an employer retirement plan are important determinants of financial participation.

We use the large sample to tabulate financial participation at the level of ZIP Code Tabulation Areas (ZCTAs), which is impossible with the smaller sample size of household surveys. Our maps reveal significant geographic variation in financial participation among low-income households. We then study whether ZCTA-level characteristics such as average income, income inequality, racial composition, and economic connectedness (Chetty et al., 2022) explain the geographic variation in financial participation. In the lowest income quintile in 2019, retirement account participation is positively related to ZCTA-level average income, conditional on household income and regional price parities (i.e., geographic differences in the cost of living). Furthermore, we find that ZCTA-level income inequality as well as Hispanic and Black population shares are negatively related to retirement account participation. We find that economic connectedness, which measures friendships between different

socioeconomic groups, is positively related to retirement account participation. However, economic connectedness does not reduce the explanatory power of average income or income inequality, suggesting that these income measures contain additional information about peer effects. That is, low-income households benefit from living in neighborhoods with higher average income and less income inequality. We rule out the hypothesis that ZCTAs with higher minority population shares have lower retirement account participation because of worse access to employer retirement plans.

In the lowest income quintile, bank account participation is negatively related to ZCTA-level income inequality, conditional on household income and regional price parities. In fact, income inequality is the only ZCTA-level characteristic that reliably predicts bank account participation. We do not find geographic variation in bank account participation by ZCTA-level average income or economic connectedness. We do not find a negative relation between bank account participation and minority population shares. We find that bank branch presence at the ZCTA level has a small positive effect on bank account participation. This finding complements the importance of other supply factors such as banking fees (Dlugosz et al., 2021) or spatial discrimination at a more granular geographic level (Sakong and Zentefis, 2022).

We use the panel dimension to estimate the causal effect of access to an employer retirement plan on retirement account participation. Our empirical design improves upon previous studies, based on a single or a small sample of employers, by studying significantly more employers and workers. In addition, we are able to link spouses that work for different employers and track workers that switch employers. We start with the sample of households that lacked access to an employer retirement plan in 2010. The identifying assumption is that whether or when an employer starts a retirement plan after 2010 is an unexpected treatment, conditional on the employer characteristics in 2010. We also address an endogeneity problem that arises from workers switching employers to gain access to a retirement plan. We construct an intent-to-treat instrument, which is the counterfactual access to an employer retirement plan had the worker remained with the same employer from 2010 to 2019. On the extensive margin, access to an employer retirement plan increases retirement account participation by 32 percentage points in the lowest income quintile. Automatic enrollment further increases retirement account participation by 28 percentage points. On the intensive margin, each additional year (beyond the first) of access to an employer retirement plan without automatic enrollment increases retirement account participation by a percentage point.

Following the early evidence that automatic enrollment boosts retirement account participation (Madrian and Shea, 2001), more recent work provides a nuanced view of the ef-

fectiveness of nudges over the long run and at large scale. For example, workers without automatic enrollment catch up to those with automatic enrollment after three years, by contributing at a higher rate (Choukhmane, 2024). Many workers cash out on retirement savings at job separation, and the likelihood of cashing out increases in the employer matching contribution rate (Wang et al., 2022). Finally, nudges that work at the scale of a single employer may not work as well at a national scale (DellaVigna and Linos, 2022).

In light of these findings, it is important to consider more heavy-handed government intervention at a national scale, such as universal access to employer retirement plans. Starting with Oregon in 2017, ten states now have mandates requiring most employers to enroll all workers in a state-sponsored retirement savings program if they do not already offer a retirement plan. Because these mandates are relatively recent, we cannot yet observe their long-run impact on retirement account participation. However, we could estimate the counterfactual retirement account participation in 2019 if universal access to an employer retirement plan with automatic enrollment had already been in effect nationally since 2010. In this counterfactual, retirement account participation increases by 19 percentage points in the lowest income quintile and 16 percentage points in the second income quintile.

Policymakers encourage retirement savings through a variety of tax incentives for both employers and workers. Eligible employers can claim tax credits for the cost of starting a retirement plan. Workers can deduct retirement contributions from taxable income and earn tax-deferred returns. Tax incentives affect even low-income workers if they have sufficient tax liabilities to claim a Saver's Credit of up to 50 percent of retirement contributions. However, Ramnath (2013) finds that the Saver's Credit has a limited causal effect on retirement contributions. Our findings suggest that tax incentives for employers to offer retirement plans may be more effective than those for workers to save in retirement accounts.

Based on the administrative tax data, we can measure the extensive margin of whether a household has a retirement or bank account but not the intensive margin of the account balance. The extensive margin is important from the perspective of life-cycle saving. First, households are forgoing valuable tax savings by not participating in retirement accounts. For qualifying households with sufficient tax liabilities, the Saver's Credit is like a matching contribution of 100 percent (up to $2,000 for joint filers) by the federal government. Second, participation in a retirement account with automatic contributions or a bank account with electronic deposit of wages could get households with behavioral biases in the habit of saving (Mullainathan and Shafir, 2009). Consistent with this hypothesis, automatic contributions increase savings without adverse effects on debt or credit outcomes (Beshears et al., 2022), and bank account participation increases wealth accumulation and durable good purchases (Célerier and Matray, 2019; Stein and Yannelis, 2020). Third, households without retire-

ment, bank, or brokerage accounts do not own any risky financial assets. With smooth preferences and no fixed costs, portfolio theory predicts that some equity exposure is optimal. By focusing on the extensive margin, we do not resolve tensions in the nudges literature that arise on the intensive margin, such as catch-up contributions (Choukhmane, 2024) and partial cash outs at job separation (Wang et al., 2022).

Our study contributes to the literature on wealth inequality. Wealth inequality is greater than income inequality in the United States (Bricker et al., 2020). According to the 2019 Survey of Consumer Finances, the income share was 15 percent for the bottom half of households and 19 percent for the top 1 percent. In comparison, the wealth share was 2 percent for the bottom half of households and 33 percent for the top 1 percent. Therefore, the bottom half of households earns 19 percent of income but owns only 3 percent of wealth in the subpopulation that excludes the top 1 percent. Because retirement and bank accounts are the most important means of accumulating financial wealth, studying low participation at the bottom of the income distribution is important for a better understanding of wealth inequality.

The remainder of the paper is organized as follows. Section 2 describes how we construct the sample of households and measures of income and financial participation. Section 3 summarizes financial participation by income. In Section 4, we study geography as a potential determinant of financial participation. In Section 5, we estimate the causal effect of access to an employer retirement plan on retirement account participation. We also estimate the potential impact of universal access to employer retirement plans. Section 6 concludes.

2. Data construction

We describe how we construct the sample of households and measures of income and financial participation with the administrative tax data and the Survey of Consumer Finances.

2.1. Administrative tax data

We use the administrative tax data of the Internal Revenue Service, which contain tax returns (Form 1040) and information returns, for tax years 2006 to 2019. The relevant information returns are Forms W-2 (Wage and Tax Statement), 1099-INT (Interest Income), 1099-DIV (Dividends and Distributions), 1099-R (Distributions from Pensions, Annuities, Retirement or Profit-Sharing Plans, IRAs, Insurance Contracts, etc.), 1099-MISC (Miscellaneous Income), 1099-G (Certain Government Payments), SSA-1099 (Social Security Benefit Statement), 1099-B (Proceeds from Broker and Barter Exchange Transactions), 5498 (IRA Contribution Information), and 1095 (Health Insurance Marketplace Statement, Health Cov-

erage, or Employer-Provided Health Insurance Offer and Coverage).

2.1.1. Sample

For each year from 2015 to 2019, we sample all individuals aged 50 to 59, who have either a tax return or an information return with a ZIP Code within the U.S. states or Washington, DC. As we show in Appendix A, the sample includes over 43 million individuals each year, which is essentially the same population count as the resident population aged 50 to 59 in the census. The sample starts in 2015 to coincide with the start of Form 1095, which is necessary to achieve the same population count as the census (Lurie and Pearce, 2019).

Three considerations motivate our sample selection. First, we can only analyze a subsample of the U.S. population due to computational limitations. Second, the 50 to 59 age group is the most relevant part of the life cycle for retirement saving. Third, the most important life-cycle decisions (e.g., residence, employment, and family structure) are usually made before age 50, which gives us a stable and comprehensive sample to analyze income and financial participation.

We use a crosswalk file to map the ZIP Codes to ZIP Code Tabulation Areas (ZCTAs).[1] The Census Bureau constructed ZCTAs by assigning census blocks to approximately 32,000 geographic areas. In most cases, the ZCTA assigned to a census block is the same as its ZIP Code. However, they could be different if a census block contains multiple ZIP Codes.

For each sampled individual, we also obtain her tax data for the previous nine years. For example, for a sampled individual in 2015, we obtain her tax data for 2006 to 2015. We also sample the spouses of sampled individuals (regardless of age). We define a spouse as a current joint filer on a Form 1040. For those who do not currently file a Form 1040, we also define a spouse as a joint filer in the previous ten years, who currently has the same household identifier (i.e., the same address) or appears on the same Form 1095 as the sampled individual.[2] Thus, we continue to link individuals as households, even if they stop filing tax returns. We intentionally do not link individuals with the same household identifier if they have not filed tax returns together. Our goal is to measure joint access to financial accounts, and it is unclear to what extent non-spousal household members (e.g., parents or children living at the same address) share financial accounts.

[1] When the ZIP Code is not available on Form 1040, we use the ZIP Code from information returns prioritized in the order listed above. If the ZIP Code is still not available, we use the most commonly reported ZIP Code on all other information returns.

[2] Larrimore et al. (2021) constructed the household identifiers, based on a textual analysis of the addresses on tax returns and information returns.

2.1.2. Income

We construct pre-tax household income following Larrimore et al. (2021). For tax filers who have a Form 1040, we start with total income (e.g., line 7b on the 2019 form), which includes wages and salaries, pass-through business income (including self-employment income), taxable interest, dividends, realized capital gains, taxable private retirement income, taxable Social Security benefits, rents, royalties, unemployment compensation, and alimony. We adjust total income by adding tax-exempt interest, replacing taxable private retirement income with gross private retirement income (excluding rollovers) on Form 1099-R, and replacing taxable Social Security benefits with total Social Security benefits on Form SSA-1099 (including disability insurance). Finally, we truncate pre-tax household income at zero to limit the impact of business losses.

For nonfilers who do not have a Form 1040, we again follow Larrimore et al. (2021). Pre-tax individual income is the sum of wages and salaries on Form W-2, interest income on Form 1099-INT, dividends on Form 1099-DIV, realized capital gains on Form 1099-B, gross private retirement income (excluding rollovers) on Form 1099-R, total Social Security benefits on Form SSA-1099, unemployment benefits on From 1099-G, and 30 percent of income on Form 1099-MISC (assuming 70 percent for offsetting expenses). For nonfilers who form households through a common household identifer or Form 1095, pre-tax household income is the sum of the pre-tax individual incomes.

We adjust income to 2019 dollars, based on the consumer price index for all urban consumers. We then define usual household income as the moving average of inflation-adjusted household income over a five-year history (i.e., the current and the previous four years). Usual household income is meant to capture permanent income that smoothes out transitory shocks.

We construct five income groups, based on the national distribution of usual household income within each year. We refer to the income quintiles as the lowest quintile (0–20 percentiles), the second quintile (20–40 percentiles), the third quintile (40–60 percentiles), the fourth quintile (60–80 percentiles), and the highest quintile (80–100 percentiles).

2.1.3. Retirement account participation

We define retirement account participation comprehensively to include employer retirement plans and IRAs. We measure participation in an employer retirement plan if retirement plan (box 13) is checked on Form W-2. This covers all employer retirement plans including defined benefit and defined contribution plans. We measure participation in an IRA, based on the presence of Form 5498, which is annually filed with the Internal Revenue Service

even when no contributions are made. We supplement our measure of retirement account participation, based on a retirement distribution on Form 1099-R. For some of our analyses, we use the distribution codes on Form 1099-R to distinguish an employer retirement plan from an IRA.

The resulting variable for retirement account participation could have gaps in the panel dimension if an individual does not receive a Form W-2 or a Form 1099-R in a given year. Therefore, we use a ten-year history (i.e., the current and the previous nine years) to improve our measure of retirement account participation. We observe that a retirement account is closed when total distribution (box 2b) is checked on Form 1099-R, excluding cases of rollovers and recharacterized IRA contributions. Thus, the definition of retirement account participation in 2015 is having an account in the previous ten years that is still open as of 2015.

We measure access to an employer retirement plan if retirement plan (box 13) is checked on *any* Form W-2 issued by an individual's employer in a given year. We search through all Forms W-2 (not just sampled individuals) to construct this variable. Furthermore, we require that the income from the employer is greater than the federal minimum wage times 1,000 hours to infer that the individual is eligible for retirement benefits. We define access to an employer retirement plan as access through any employer in the previous ten years. As we describe in Appendix B, we observe whether an employer retirement plan has automatic enrollment by merging the administrative tax data with the Department of Labor Form 5500.

For married households, we take the maximum over the two indicator variables for retirement account participation, access to an employer retirement plan, and bank account participation. That is, a household has (access to) an account if either spouse has (access to) an account.

2.1.4. Bank account participation

We measure bank account participation, based on electronic funds transfer for payment of taxes or receipt of refunds on Form 1040. According to the instructions for Form 1040, about 80 percent of tax filers who receive refunds do so by direct deposit. Moreover, the name on the tax filing must match the name on the bank account, which rules out tax filers receiving refunds in a bank account that they do not own. We also measure bank account participation, based on taxable (box 1) or tax-exempt (box 8) interest on Form 1099-INT. Form 1099-INT has incomplete coverage because it is required only for accounts with at least $10 of annual interest. Thus, we measure bank account participation primarily through electronic funds transfer as part of a tax filing.

The resulting variable for bank account participation could have gaps in the panel dimension if an individual does not file tax returns or receive a Form 1099-INT in a given year. Therefore, we use a ten-year history to improve our measure of bank account participation. For example, we measure bank account participation in 2015 if the criteria for having a bank account are satisfied in any year between 2006 and 2015. Thus, the definition of bank account participation in 2015 is having an account in the previous ten years, even if that account is closed as of 2015. From an economic perspective, an individual who has ever had a bank account is different from one who has never had a bank account.

During the initial research design, we considered other measures of financial participation. We could define financial participation more broadly to include mutual funds and brokerage accounts, based on Forms 1099-DIV and 1099-R. However, we have verified that virtually all households that have these accounts already have a bank account. Participation in stocks and equity mutual funds is interesting from the perspective that all households should participate under smooth preferences and no fixed costs. However, the administrative tax data do not contain any information about stocks and equity mutual funds in retirement accounts. We could measure mortgage participation, based on Form 1098 (Mortgage Interest Statement). Furthermore, we could measure home ownership, based on Form 1098 and itemized deductions for property taxes. However, these measures of financial participation are more difficult to interpret because a household may prefer to rent. Similarly, we cannot tell whether a household does not have a mortgage because it does not need one or has been denied. In related work, Lurie and Pearce (2019) use the administrative tax data to study health insurance coverage.

2.2. Survey of Consumer Finances

We benchmark our summary statistics to the 2016 and 2019 Survey of Consumer Finances (Board of Governors of the Federal Reserve System, 2016, 2019). We restrict the sample to households with a respondent aged 50 to 59 in these years. Usual income in the Survey of Consumer Finances is a self-reported measure of permanent income that smoothes out transitory shocks. It is broader than our measure of usual income by including food stamps and other government support that are not observed in the administrative tax data. We define bank account participation as ownership of a checking account, savings account, or money market fund. We define retirement account participation as either ownership of or payments from an employer retirement account or an IRA.

3. Summary of financial participation

We summarize retirement and bank account participation by income in the administrative tax data and the Survey of Consumer Finances. For retirement accounts, the overall participation rates match in the two datasets, and the participation rates conditional on income align closely. For bank accounts, both overall participation rates and participation rates conditional on income align, except that we underestimate the participation rate in the lowest income quintile.

3.1. Retirement account participation

Table 1 reports retirement account participation for households with a member aged 50 to 59 in the administrative tax data and the Survey of Consumer Finances. The overall retirement account participation rates match in the two datasets. In 2016, 70 percent of households had retirement accounts in the administrative tax data, which is close to 72 percent of households in the Survey of Consumer Finances. In 2019, 69 percent of households had retirement accounts in both datasets.

Retirement account participation in the administrative tax data has a slightly steeper income gradient than that in the Survey of Consumer Finances. Thus, retirement account participation for low-income households is slightly lower in the administrative tax data. In the lowest income quintile in 2019, 21 percent of households had retirement accounts in the administrative tax data, compared with 24 percent of households in the Survey of Consumer Finances. In the second income quintile in 2019, 54 percent of households had retirement accounts in the administrative tax data, compared with 63 percent of households in the Survey of Consumer Finances.

The administrative tax data has the same population count as the census, and the Survey of Consumer Finances is a random sample of the census population. Thus, the fact that the overall participation rates match in the two datasets implies that our sample construction and measurement assumptions mimic the survey data. Several factors could explain why retirement account participation in the administrative tax data has a steeper income gradient. As we discussed in Section 2, the Survey of Consumer Finances uses a broader definition of income that includes food stamps and other government support, which could lead to a different correlation with retirement account participation. Alternatively, because the Survey of Consumer Finances is a survey, both retirement account participation and income are subject to measurement or imputation error when households misreport or refuse to answer survey questions. Random measurement or imputation error in income attenuates the true relation between retirement account participation and income, which could explain

the flatter income gradient in the Survey of Consumer Finances.

3.2. Bank account participation

Table 2 reports bank account participation for households with a member aged 50 to 59 in the administrative tax data and the Survey of Consumer Finances. The overall bank account participation is slightly lower in the administrative tax data. In 2016, 92 percent of households had bank accounts in the administrative tax data, compared with 93 percent of households in the Survey of Consumer Finances. In 2019, 92 percent of households had bank accounts in the administrative tax data, compared with 95 percent of households in the Survey of Consumer Finances.

Bank account participation in the administrative tax data has a steeper income gradient than that in the Survey of Consumer Finances. Thus, bank account participation for low-income households is lower in the administrative tax data than in the Survey of Consumer Finances. In the lowest income quintile in 2019, 70 percent of households had bank accounts in the administrative tax data, compared with 80 percent of households in the Survey of Consumer Finances.

We view the 70 percent estimate as a lower bound. As we discussed in Section 2, we measure bank account participation primarily through electronic funds transfer on Form 1040 in the previous ten years. Thus, we could underestimate bank account participation for households that did not file tax returns in the previous ten years, called "never filers" hereafter, that are more prevalent among low-income households. We offer a rough estimate of bank account participation for never filers in the lowest income quintile. Suppose that we take the 80 percent participation rate in the Survey of Consumer Finances at face value. We know that 25 percent of households in the lowest income quintile in 2019 were never filers. We can fully reconcile the gap between the Survey of Consumer Finances and the administrative tax data if $(80 - 70)/25 = 40$ percent of never filers in the lowest income quintile have bank accounts that we do not observe.[3]

3.3. Additional facts about retirement accounts

Table 3 breaks down retirement account participation into employer retirement plans versus IRAs for households with a member aged 50 to 59 in 2019. In the lowest income quintile, 11 percent of households have only an employer retirement plan, 5 percent have only an IRA, and 4 percent have both. Thus, low-income households have retirement accounts primarily

[3]Another possibility is that a tax filer with a bank account chooses to pay taxes or receive refunds through checks instead of electronic funds transfer. We could underestimate bank account participation for low-income households if they are more likely to choose checks.

through their employers. Higher-income households are more likely to have both an employer retirement plan and an IRA. In the highest income quintile, 23 percent of households have only an employer retirement plan, 5 percent have only an IRA, and 69 percent have both.[4]

Table 4 reports access to an employer retirement plan for households with a member aged 50 to 59 in the administrative tax data. In the lowest income quintile in 2019, only 38 percent of households had access to an employer retirement plan (at any point in the previous ten years). Access to an employer retirement plan increases to 79 percent of households in the second income quintile and 93 percent of households in the third income quintile. Although the time series are short, access to an employer retirement plan apparently declines in the lowest income quintile from 42 percent of households in 2015 to 38 percent of households in 2019.

4. Geography of financial participation

We use the large sample to tabulate financial participation at the ZCTA level, which is impossible with the smaller sample size of household surveys. Our maps reveal significant geographic variation in financial participation among low-income households, which implies that geography is an important determinant of financial participation. By regression analysis, we document the characteristics of ZCTAs with higher participation rates among low-income households.

4.1. Maps of financial participation

Figure 1 is a map of retirement account participation by ZCTA for households in the lowest income quintile with a member aged 50 to 59 in 2019. We focus on the lowest income quintile because it has the greatest geographic variation in financial participation and attracts the most policy interest. The colors range from yellow (40–100 percent participation) to red (0–10 percent participation). The shade depends on the population aged 50 to 59, where a darker shade represents a more populous ZCTA. For example, a dark shade of red represents a populous ZCTA with low retirement account participation.

Figure 2 is a similar map of bank account participation by ZCTA for households in the lowest income quintile with a member aged 50 to 59 in 2019. The colors range from yellow (90–100 percent participation) to red (0–60 percent participation).

Figures 1 and 2 show that some low-income households manage to participate, depending on where they live. However, this geographic variation in financial participation is not a

[4]Table 3 could understate the importance of employer retirement plans if some households have IRAs that were funded entirely by rollovers from employer retirement plans.

simple matter of the north versus the south or the coasts versus the heartland. Within geographic areas smaller than states, red areas of low participation are mixed with yellow areas of high participation.

4.2. Geographic characteristics related to financial participation

We study whether ZCTA-level characteristics such as average income, income inequality, racial composition, economic connectedness, and bank branch access explain the geographic variation in financial participation among low-income households. We provide details about the regressors and their data sources in Appendix B.

4.2.1. Retirement account participation

Table 5 reports regressions of retirement account participation on household and ZCTA-level characteristics for households in the lowest income quintile with a member aged 50 to 59 in 2019. In column 1, we start with a benchmark regression on household characteristics, which are log income and marital status, and regional price parities to control for geographic differences in the cost of living. A standard deviation increase in log income predicts a 6 percentage point increase in retirement account participation.

In column 2 of Table 5, we add average log income and income inequality (i.e., standard deviation of log income) at the ZCTA level, which are statistically significant predictors of retirement account participation. A standard deviation increase in average log income predicts a percentage point increase in retirement account participation. A standard deviation increase in income inequality predicts a 2 percentage point decrease in retirement account participation. Since the regression already controls for household income, the importance of average income and income inequality suggests the presence of peer effects, where low-income households benefit from living in neighborhoods with higher average income and less income inequality.

In column 3 of Table 5, we add racial composition at the ZCTA level. We find that all minority population shares are negatively related to retirement account participation. Relative to the omitted white category, a percentage point increase in the Hispanic population share predicts a 11 basis point decrease in retirement account participation. The corresponding estimates for the Black, Asian, and other nonwhite population shares are respectively 3, 6, and 3 basis points.

In column 4 of Table 5, we add economic connectedness, which measures friendships between different socioeconomic groups, to test for peer effects (Chetty et al., 2022). A standard deviation increase in economic connectedness predicts a percentage point increase

in retirement account participation. However, economic connectedness does not reduce the explanatory power of average income or income inequality. Although economic connectedness is a direct measure of peer effects, these income measures could contain additional information if peer effects are complex and difficult to measure.

A hypothesis for the importance of racial composition is that ZCTAs with higher minority population shares have worse employment opportunities and worse access to employer retirement plans. In column 5 of Table 5, we test this hypothesis by including a household-level measure of access to an employer retirement plan in the previous ten years. Access to an employer retirement plan predicts a 41 percentage point increase in retirement account participation. All minority population shares continue to be negatively related to retirement account participation, except for the Asian category. Therefore, we rule out the hypothesis that ZCTAs with higher minority population shares have lower retirement account participation because of worse access to employer retirement plans.

In column 6 of Table 5, we add bank account participation as an additional household characteristic. Households with a bank account are 4 percentage points more likely to have a retirement account, conditional on all other household and ZCTA-level characteristics. This finding suggests the presence of a common factor driving financial participation, such as a mistrust of financial institutions.

4.2.2. Bank account participation

Table 6 reports regressions of bank account participation on household and ZCTA-level characteristics for households in the lowest income quintile with a member aged 50 to 59 in 2019. We further restrict the sample to households that have filed tax returns in the previous ten years (i.e., a 75 percent subsample), for which we have the most reliable measures of bank account participation. In column 1, we start with a benchmark regression on household characteristics, which are log income and marital status, and regional price parities. A standard deviation increase in log income predicts a 6 percentage point increase in bank account participation.

In column 2 of Table 6, we add average log income and income inequality at the ZCTA level. Average log income has little effect on bank account participation. However, a standard deviation increase in income inequality predicts a 3 percentage point decrease in bank account participation.

In column 3 of Table 6, we add racial composition at the ZCTA level. We do not find a negative relation between bank account participation and minority population shares, except for the other nonwhite category (i.e., American Indian, Alaska Native, Native Hawaiian, other Pacific Islander, and multiple race). Relative to the omitted white category, a percent-

age point increase in the other nonwhite population share predicts a 10 basis point decrease in bank account participation.

In column 4 of Table 6, we add economic connectedness to test for peer effects. Economic connectedness does not predict bank account participation. Moreover, economic connectedness does not reduce the explanatory power of income inequality.

In column 5 of Table 6, we add two ZCTA-level characteristics that relate to bank branch access. On the extensive margin, the presence of a bank branch predicts a percentage point increase in bank account participation. On the intensive margin, bank branch density does not predict bank account participation.

This finding complements the importance of other supply factors beyond the mere presence of bank branches. Dlugosz et al. (2021) find that banking fees are an important supply factor that could explain bank account participation. In particular, a cap on overdraft fees could constrain the supply of overdraft credit and deposit accounts through a higher minimum deposit requirement. Using more granular data at the census block-group level, Sakong and Zentefis (2022) find lower bank branch access in Black neighborhoods but not in low-income neighborhoods. Finally, the small positive effects of bank branch access is a cross-sectional correlation and does not establish causality. Célerier and Matray (2019) find that deregulation between 1994 and 2005 led to an expansion of bank branches and increased bank account participation among low-income households.

4.3. Interpretation and open questions

The fact that income matters for financial participation is unsurprising. Households without enough income are unable to save. Low-income households may not have sufficient tax incentives to open a retirement account if they face a low marginal income tax rate, a zero capital gains tax rate, and do not have sufficient tax liabilities to qualify for the Saver's Credit. Fees on accounts with low balances may disincentivize low-income households from opening and keeping a bank account. In addition to these direct effects of income, lower-income households tend to have lower educational attainment and financial literacy, which could further hinder financial participation (Lusardi and Mitchell, 2011).

Mullainathan and Shafir (2009) hypothesize that low-income households are less likely to have bank accounts because of institutions that shape behavior. For example, low-income households may not have incentives to open a bank account if their employers do not use electronic deposits. If their income is too low to file tax returns, low-income households have no need for an electronic funds transfer to pay taxes or receive refunds. Under this hypothesis, a nudge such as an electronic deposit of wages as the default option could boost bank account participation.

The fact that average income, income inequality, and economic connectedness matter for retirement account participation suggests the presence of peer effects (Duflo and Saez, 2002, 2003; Ouimet and Tate, 2020). In neighborhoods with higher average income and less income inequality, lower-income households may be more likely to socialize with higher-income households and learn about retirement and bank accounts (Hong et al., 2004; Brown et al., 2008).

We do not have a complete explanation for why the minority population shares are negatively related to retirement account participation but unrelated to bank account participation. A hypothesis is that racial composition is proxying for household-level race, which we do not observe in the administrative tax data. This hypothesis is consistent with the known fact that retirement account participation and retirement contribution rates are lower for Hispanic and Black households, conditional on income (Bhutta et al., 2020; Choukhmane et al., 2024). However, it is inconsistent with the known fact that bank account participation is lower for Hispanic and Black households, conditional on income (Hayashi and Minhas, 2018; Federal Deposit Insurance Corporation, 2020). Because we exclude never filers from the analysis of bank account participation, a potential explanation is that the relation between bank account participation and race exists only for the subpopulation of never filers. Future research could separately identify the importance of race at the household versus the ZCTA level if the administrative tax data were to be merged with the census data on household-level race.[5]

Another hypothesis is racial discrimination or lingering effects of historical discrimination, such as a mistrust of financial institutions that could persist as a cultural norm of a geographic area (Brown et al., 2019). Such mistrust could explain why the minority population shares are negatively related to retirement account participation. However, it is not clear why such mistrust does not affect bank account participation.

5. Access to an employer retirement plan

As reported in Table 3, the vast majority of households have retirement accounts through employer retirement plans rather than through only IRAs. Opening an IRA requires more effort and financial literacy than enrolling in an employer retirement plan. Therefore, access to an employer retirement plan could be a primary determinant of retirement account participation. Moreover, universal access to an employer retirement plan could be an effective policy intervention that boosts retirement account participation for low- and middle-income

[5]This data merge is currently impossible. The U.S. Department of the Treasury does not have access to the census data on household-level race, and the census research data centers do not have access to the full set of information returns that are necessary for our study.

households, which have lower access to employer retirement plans according to Table 4.

5.1. Identifying assumptions

Retirement account participation and access to an employer retirement plan could be jointly endogenous. Workers who care about retirement savings may choose to work for an employer with a retirement plan, leading to a positive selection bias. Conversely, some workers may already have retirement security through an IRA, their spouse's retirement savings, or Social Security. Workers with retirement security may choose to work for an employer without a retirement plan (e.g., a small employer or self-employment), leading to a negative selection bias.

We use the panel dimension to estimate the causal effect of access to an employer retirement plan on retirement account participation. Our empirical design improves upon previous studies, based on a single or a small sample of employers, by studying significantly more employers and workers. In addition, we are able to link spouses that work for different employers and track workers that switch employers. We start with a sample of individuals aged 50 to 59 in 2019, who lacked access to an employer retirement plan in 2010. The identifying assumption is that whether or when an employer starts a retirement plan after 2010 is an unexpected treatment, conditional on the employer characteristics in 2010. For employer characteristics, we use log employees to capture firm size and indicator variables for the two-digit North American Industry Classification System (NAICS) code to capture industry.

There are two arguments in favor of identification. First, most workers in our sample joined their 2010 employer long before 2010. The argument is that, at the time that the worker joined the firm, she would not have been able to predict that her employer would eventually start a retirement plan after 2010. Second, larger firms or those in certain industries may be more likely to start a retirement plan. The argument is that, conditional on firm size and industry, the residual variation in whether or when an employer starts a retirement plan is as-if random from the worker's perspective.

We must also address a separate endogeneity problem that arises from workers switching employers to gain access to a retirement plan. For example, Panel A of Figure 3 illustrates the case of a worker who stays with employer A from 2010 to 2019 and gains access to an employer retirement plan in 2016. Panel B illustrates the case of a worker who switches from employer A to employer B and gains access to an employer retirement plan in 2013. The realized access to an employer retirement plan is exogenous in Panel A but potentially endogenous in Panel B. To address this problem, we construct an intent-to-treat instrument, which is the counterfactual access to an employer retirement plan had the worker remained

with the same employer from 2010 to 2019. In Figure 3, the realized access to an employer retirement plan is in black, and the counterfactual access is in red.

An employer's decision to start a retirement plan is not random and could be a consequence of fits outcomes. For example, employer A in Figure 3 could have started a retirement plan in 2016 as a consequence of fast growth between 2010 and 2016. Our identifying assumption is that this growth would not have been predictable at the time that the worker joined the firm (long before 2010), conditional on firm size and industry in 2010. The threat to identification is that the employer's propensity to start a retirement plan is correlated with worker characteristics. For example, more productive workers could cause the firm to grow faster, and these workers may have a stronger saving motive and be more likely to enroll in an employer retirement plan. We control for income to address this particular channel. However, any residual variation in worker productivity that is orthogonal to income remains a threat to identification.

We estimate the instrumental variables regression separately by income quintile to allow for heterogeneous treatment effects. The sample size is larger for lower income quintiles because more of these households lacked access to an employer retirement plan in 2010. The dependent variable is retirement account participation in 2019. The first endogenous regressor is an indicator variable for access to an employer retirement plan by either spouse from 2011 to 2019. The second endogenous regressor is additional years (beyond the first) that either spouse had access to an employer retirement plan without automatic enrollment. This specification captures a potential nonlinearity from the first year of access having a larger impact than each subsequent year of access to an employer retirement plan. The third endogenous regressor is an indicator variable for automatic enrollment in an employer retirement plan by either spouse from 2011 to 2019.

We construct three instruments corresponding to the three endogenous regressors, based on the counterfactual access to an employer retirement plan. For example, in Panel B of Figure 3, the endogenous regressors are 1 for the indicator variable for access to an employer retirement plan, 6 for the additional years of access, and either 0 or 1 for the indicator variable for automatic enrollment (depending on employer B's plan). The corresponding instruments are 1 for the indicator variable for access to an employer retirement plan, 3 for the additional years of access, and either 0 or 1 for the indicator variable for automatic enrollment (depending on employer A's plan).

5.2. Causal effects of access to an employer retirement plan

Table 7 reports the instrumental variables regression of retirement account participation on access to an employer retirement plan. Before we discuss the estimates, we check for weak

identification. In all income quintiles, the test static for rejecting the null of weak instruments is far greater than the critical value at the 5 percent significance level (Stock and Yogo, 2005).[6] Furthermore, we report the first-stage regressions in Appendix C.

The constant in the model is the baseline participation rate for married households with no access to an employer retirement plan. The baseline participation rate increases from 8 percent of households in the lowest income quintile to 63 percent of households in the highest income quintile. For single households, the baseline participation rate increases from 4 percent of households in the lowest quintile to 57 percent of households in the highest quintile. Higher-income households are more likely to already have retirement accounts in 2010, through previous employers or IRAs.

On the extensive margin, access to an employer retirement plan increases retirement account participation by 32 percentage points in the lowest income quintile. Automatic enrollment further increases retirement account participation by 28 percentage points in the lowest income quintile. On the intensive margin, each additional year (beyond the first) of access without automatic enrollment increases retirement account participation by a percentage point. Thus, the treatment effect of access to an employer retirement plan with automatic enrollment is 60 percentage points in the lowest income quintile. Our estimates confirm the effectiveness of automatic enrollment for retirement account participation (Madrian and Shea, 2001; Derby et al., 2022; Choukhmane, 2024).

5.3. State-sponsored retirement savings programs

The causal effects in Table 7 suggest that universal access to an employer retirement plan is a policy intervention that could boost retirement account participation. Table 8 lists ten states that have mandates requiring most employers to enroll all workers in a state-sponsored retirement savings program if they do not already offer a retirement plan. Because the mandates do not require employer contributions, employers incur only the administrative costs of compliance. The mandates apply to all employers with a minimum number of workers and minimum years in business.[7] Bloomfield et al. (2024) find that 27 percent of Oregon employers comply by offering their own retirement plan instead of joining OregonSaves. The state-sponsored retirement savings programs are legally IRAs and subject to the IRA contribution limits. The default contribution rate is 3 or 5 percent (depending on the state), but workers can adjust the contribution rate or entirely opt out. Chalmers et al. (2021) find

[6] Although Stock and Yogo (2005, Table 5.2) do not report the critical values for the case of three endogenous regressors, the critical value is less than 7.03 for the reported case of two endogenous regressors.

[7] In December 2022, the Senate released legislative text that would require retirement plans for all employers with at least ten workers and in business for at least three years. It is too early to tell whether universal access to an employer retirement plan could become federal law.

a 34 percent participation rate for OregonSaves, which implies that a majority of workers actually opt out.

Because these mandates are relatively recent, we cannot yet observe their long-run impact on retirement account participation. Moreover, these mandates have a limited scope, applying to a subset of employers in the ten participating states. However, we could estimate the counterfactual retirement account participation in 2019 if universal access to an employer retirement plan with automatic enrollment had already been in effect nationally since 2010. For each worker in each year, we define counterfactual access to an employer retirement plan as 1 if she worked (i.e., received a Form W-2) and 0 otherwise. For each household, we then compute the indicator variable for access to an employer retirement plan by either spouse from 2011 to 2019. Finally, we use the instrumental variables regression model in Table 7 to predict the counterfactual probability of participation under automatic enrollment for each household.

Table 9 reports the predicted change in retirement account participation by income quintile. Had universal access to an employer retirement plan with automatic enrollment already been in effect nationally since 2010, retirement account participation in 2019 would have increased by 19 percentage points in the lowest income quintile and 16 percentage points in the second income quintile. In the third to the highest income quintiles, the predicted change declines respectively to 7, 2, and 1 percentage points. Universal access has the largest impact on low- and middle-income households, which initially have lower access to an employer retirement plan. However, universal access only raises the participation rate to 40 percent of households in the lowest income quintile. Universal access does not achieve universal participation because not all households are working, and the takeup rate is far less than one.

The causal effects in Table 7 are based on employers that voluntarily start retirement plans. Therefore, these causal effects may not have external validity for a policy that mandates universal access through all employers. First, employers that voluntarily start retirement plans may be different from other employers, perhaps viewing retirement plans more positively and nudging workers into participation. Second, universal access to an employer retirement plan could have macroeconomic effects, such as a higher takeup rate from both implicit and explicit government endorsement and greater public awareness. For these reasons, it would be interesting to revisit this question ten years later when we have enough data to evaluate the state-sponsored retirement savings programs.

6. Conclusion

We study retirement and bank account participation for the universe of U.S. households with a member aged 50 to 59 in the administrative tax data. In the lowest income quintile in 2019, 21 percent of households had retirement accounts, and 80 percent of households had bank accounts. The heterogeneity in financial participation conditional on income implies that low participation is not a simple matter of not having enough income to save. An important policy question is how to increase financial participation for low-income households.

Access to an employer retirement plan is a primary determinant of retirement account participation. By instrumental variables, we estimate that access to an employer retirement plan increases retirement account participation by 32 percentage points in the lowest income quintile. Automatic enrollment further increases retirement account participation by 28 percentage points in the lowest income quintile. Therefore, universal access to an employer retirement plan could be an effective policy intervention that boosts retirement account participation for low- and middle-income households. Universal access to an employer retirement plan with automatic enrollment could increase retirement account participation by 19 percentage points in the lowest income quintile and 16 percentage points in the second income quintile over ten years.

In hope of improving financial participation, we have constructed interactive maps of financial participation in the lowest income quintile. The interactive maps for retirement accounts and bank accounts are available from the first author's webpage. Users can search for specific locations or zoom in and out to visualize heterogeneity in financial participation across the United States. As we describe in Appendix A, we also construct aggregate datasets by year, ZCTA, and income quintile. We hope that the maps and datasets are useful for researchers, policymakers, banks, and financial advisors to identify geographic areas with the greatest opportunity for improvement.

Our findings have important implications for the Social Security program. First, 79 percent of households in the lowest income quintile without retirement accounts primarily rely on Social Security benefits during retirement. The wellbeing of these households is especially sensitive to Social Security programs that provide a safety net for low-income workers, such as the special minimum benefit and Supplemental Security Income. Policymakers need to be especially careful when considering changes that could cut benefits in these programs. Second, the geographic disparities in retirement account participation imply different degrees of reliance on Social Security benefits across the United States. Policymakers could use our interactive maps to target advertisement and outreach to communities that are in most need of Social Security benefits, making sure that all households that are eligible apply for

the benefits. Third, universal access to an employer retirement plan could boost retirement account participation and reduce reliance on Social Security benefits. However, households that are unable to save enough or do not have sufficient work history will continue to rely on Supplemental Security Income as a safety net. Policymakers need to continue evaluating the relevance of these programs in the changing retirement landscape.

References

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., Skimmyhorn, W. L., 2022. Borrowing to save? The impact of automatic enrollment on debt. Journal of Finance 77 (1), 403–447.

Bhutta, N., Chang, A. C., Dettling, L. J., Hsu, J. W., 2020. Disparities in wealth by race and ethnicity in the 2019 Survey of Consumer Finances, FEDS Notes.

Bloomfield, A., Goodman, L., Rao, M., Slavov, S., 2024. Why do employers establish retirement savings plans? Evidence from state "auto-IRA" policies, NBER Working Paper 32817.

Board of Governors of the Federal Reserve System, 2016, 2019. Survey of Consumer Finances, https://www.federalreserve.gov/econres/scfindex.htm.

Bricker, J., Goodman, S., Moore, K. B., Volz, A. H., 2020. Wealth and income concentration in the SCF: 1989–2019, FEDS Notes.

Brown, J. R., Cookson, J. A., Heimer, R. Z., 2019. Growing up without finance. Journal of Financial Economics 134 (3), 591–616.

Brown, J. R., Ivković, Z., Smith, P. A., Weisbenner, S., 2008. Neighbors matter: Causal community effects and stock market participation. Journal of Finance 63 (3), 1509–1531.

Bureau of Economic Analysis, 2019. Regional price parities by state, https://www.bea.gov/data/prices-inflation/regional-price-parities-state-and-metro-area.

Célerier, C., Matray, A., 2019. Bank-branch supply, financial inclusion, and wealth accumulation. Review of Financial Studies 32 (12), 4767–4809.

Chalmers, J., Mitchell, O. S., Reuter, J., Zhong, M., 2021. Auto-enrollment retirement plans for the people: Choices and outcomes in OregonSaves, NBER Working Paper 28469.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenberg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Barberá, P., Bhole, M., Wernerfelt, N., 2022. Social capital I: Measurement and associations with economic mobility. Nature 608 (7921), 108–121.

Choukhmane, T., 2024. Default options and retirement saving dynamics, working paper, MIT.

Choukhmane, T., Colmenares, J., O'Dea, C., Rothbaum, J. L., Schmidt, L. D. W., 2024. Who benefits from retirement saving incentives in the U.S.? Evidence on gaps in retirement wealth accumulation by race and parental income, NBER Working Paper 32843.

DellaVigna, S., Linos, E., 2022. RCTs to scale: Comprehensive evidence from two nudge units. Econometrica 90 (1), 82–116.

Derby, E., Mackie, K., Mortenson, J., 2022. Worker and spousal responses to automatic enrollment, working paper, Joint Committee on Taxation.

Dlugosz, J., Melzer, B., Morgan, D. P., 2021. Who pays the price? Overdraft fee ceilings and the unbanked, working paper, Board of Governors of the Federal Reserve System.

Duflo, E., Saez, E., 2002. Participation and investment decisions in a retirement plan: The influence of colleagues' choices. Journal of Public Economics 85 (1), 121–148.

Duflo, E., Saez, E., 2003. The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. Quarterly Journal of Economics 118 (3), 815–842.

Federal Deposit Insurance Corporation, 2019. Annual survey of branch office deposits, https://banks.data.fdic.gov/docs/.

Federal Deposit Insurance Corporation, 2020. How America banks: Household use of banking and financial services, 2019 FDIC Survey.

Hayashi, F., Minhas, S., 2018. Who are the unbanked? Characteristics beyond income. Economic Review: Federal Reserve Bank of Kansas City 2, 55–70.

Hong, H., Kubik, J. D., Stein, J. C., 2004. Social interaction and stock-market participation. Journal of Finance 59 (1), 137–163.

Larrimore, J., Mortenson, J., Splinter, D., 2021. Household incomes in tax data using addresses to move from tax-unit to household income distributions. Journal of Human Resources 56 (2), 600–631.

Lurie, I. Z., Pearce, J., 2019. Health insurance coverage from administrative tax data, Office of Tax Analysis Working Paper 117.

Lusardi, A., Mitchell, O. S., 2011. Financial literacy and retirement planning in the United States. Journal of Pension Economics and Finance 10 (4), 509–525.

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

Mullainathan, S., Shafir, E., 2009. Savings policy and decisionmaking in low-income households. In: Blank, R. M., Barr, M. S. (Eds.), Insufficient Funds: Savings, Assets, Credit, and Banking Among Low-Income Households. Russell Sage Foundation, New York, Ch. 5, pp. 121–146.

Ouimet, P., Tate, G., 2020. Learning from coworkers: Peer effects on individual investment decisions. Journal of Finance 75 (1), 133–172.

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.

Sakong, J., Zentefis, A. K., 2022. Bank access across America, working paper, Federal Reserve Bank of Chicago.

Stein, L. C. D., Yannelis, C., 2020. Financial inclusion, human capital, and wealth accumulation: Evidence from the Freedman's Savings Bank. Review of Financial Studies 33 (11), 5333–5377.

Stock, J. H., Yogo, M., 2005. Testing for weak instruments in linear IV regression. In: Andrews, D. W. K., Stock, J. H. (Eds.), Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge University Press, Cambridge, Ch. 5, pp. 80–108.

U.S. Census Bureau, 2015–2019. American Community Survey demographic and housing five-year data, https://www.census.gov/data/developers/data-sets/acs-5year.html.

U.S. Department of Labor, 2009–2019. Form 5500 annual report, https://www.dol.gov/agencies/ebsa/about-ebsa/our-activities/public-disclosure/foia/form-5500-datasets.

Wang, Y., Zhai, M., Lynch, J. G., 2022. Cashing out retirement savings at job separation, working paper, University of British Columbia.

Table 1

Retirement account participation by income

	Percentile of usual income					
Year	0–20	20–40	40–60	60–80	80–100	All
Panel A. Administrative tax data						
2015	0.22	0.56	0.83	0.94	0.97	0.70
2016	0.21	0.55	0.82	0.94	0.97	0.70
2017	0.21	0.55	0.82	0.94	0.97	0.70
2018	0.21	0.55	0.81	0.93	0.97	0.69
2019	0.21	0.54	0.81	0.93	0.97	0.69
Panel B. Survey of Consumer Finances						
2016	0.33	0.62	0.83	0.90	0.95	0.72
2019	0.24	0.63	0.76	0.92	0.93	0.69

This table reports retirement account participation for households with a member aged 50 to 59. Retirement accounts include employer retirement plans (i.e., defined benefit and defined contribution plans) and IRAs.

Table 2

Bank account participation by income

	Percentile of usual income					
Year	0–20	20–40	40–60	60–80	80–100	All
Panel A. Administrative tax data						
2015	0.71	0.94	0.99	1.00	1.00	0.93
2016	0.70	0.94	0.99	1.00	1.00	0.92
2017	0.68	0.94	0.98	1.00	1.00	0.92
2018	0.67	0.93	0.98	0.99	1.00	0.92
2019	0.70	0.93	0.98	0.99	1.00	0.92
Panel B. Survey of Consumer Finances						
2016	0.76	0.93	0.98	0.99	1.00	0.93
2019	0.80	0.95	1.00	1.00	1.00	0.95

This table reports bank account participation for households with a member aged 50 to 59.

Table 3
Breakdown of retirement account participation

	Percentile of usual income					
Households with	0–20	20–40	40–60	60–80	80–100	All
Employer plan only	0.11	0.35	0.47	0.41	0.23	0.32
IRA only	0.05	0.06	0.06	0.05	0.05	0.05
Both	0.04	0.13	0.28	0.47	0.69	0.32
Total	0.21	0.54	0.81	0.93	0.97	0.69

This table reports a breakdown of retirement account participation for households with a member aged 50 to 59 in the 2019 administrative tax data.

Table 4
Access to an employer retirement plan

	Percentile of usual income					
Year	0–20	20–40	40–60	60–80	80–100	All
2015	0.42	0.81	0.93	0.97	0.96	0.82
2016	0.41	0.80	0.93	0.97	0.96	0.81
2017	0.39	0.80	0.93	0.97	0.96	0.81
2018	0.38	0.80	0.93	0.97	0.96	0.81
2019	0.38	0.79	0.93	0.96	0.96	0.80

This table reports access to an employer retirement plan (i.e., defined benefit and defined contribution plans) for households with a member aged 50 to 59 in the administrative tax data.

Table 5

Geography of retirement account participation in the lowest income quintile

Regressor	(1)	(2)	(3)	(4)	(5)	(6)
Log income	0.055	0.039	0.041	0.041	-0.006	-0.008
	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Single	-0.133	-0.123	-0.124	-0.124	-0.083	-0.076
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Regional price parities	0.001	0.003	0.012	0.009	0.012	0.012
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Average log income		0.012	0.012	0.012	0.008	0.008
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
SD of log income		-0.020	-0.017	-0.016	-0.011	-0.010
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Racial composition:						
Hispanic			-0.106	-0.089	-0.091	-0.095
			(0.004)	(0.004)	(0.002)	(0.003)
Black			-0.033	-0.021	-0.041	-0.044
			(0.003)	(0.003)	(0.003)	(0.003)
Asian			-0.056	-0.065	0.039	0.033
			(0.020)	(0.019)	(0.012)	(0.013)
Other nonwhite			-0.032	-0.027	-0.081	-0.078
			(0.008)	(0.008)	(0.006)	(0.006)
Economic connectedness				0.011		
				(0.001)		
Access to employer plan					0.406	0.396
					(0.001)	(0.001)
Bank account						0.039
						(0.001)
Constant	0.329	0.343	0.370	0.369	0.166	0.136
	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
R^2	0.03	0.04	0.04	0.04	0.26	0.26
Observations	7,189,565	7,189,541	7,189,541	7,173,593	7,189,541	7,189,541

This table reports regressions of retirement account participation on household and ZCTA-level characteristics. The household characteristics are log income, marital status, access to an employer retirement plan by either spouse from 2011 to 2019, and bank account participation. The ZCTA-level characteristics are regional price parities, average log income, standard deviation of log income, racial composition, and economic connectedness. All coefficients on the continuous variables, except for racial composition, are standardized. Robust standard errors, clustered by ZCTA, are reported in parentheses. The sample includes households in the lowest income quintile with a member aged 50 to 59 in the 2019 administrative tax data.

Table 6

Geography of bank account participation in the lowest income quintile

Regressor	(1)	(2)	(3)	(4)	(5)
Log income	0.064	0.062	0.062	0.061	0.062
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Single	-0.055	-0.053	-0.053	-0.053	-0.054
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Regional price parities	-0.003	0.003	0.002	0.004	0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Average log income		-0.005	-0.004	-0.004	-0.004
		(0.002)	(0.002)	(0.002)	(0.002)
SD of log income		-0.027	-0.028	-0.028	-0.027
		(0.002)	(0.003)	(0.003)	(0.003)
Racial composition:					
Hispanic			0.006	-0.001	0.014
			(0.004)	(0.004)	(0.004)
Black			0.032	0.027	0.039
			(0.004)	(0.004)	(0.004)
Asian			0.042	0.046	0.039
			(0.008)	(0.008)	(0.008)
Other nonwhite			-0.096	-0.098	-0.086
			(0.008)	(0.008)	(0.008)
Economic connectedness				-0.005	
				(0.001)	
Bank branch presence					0.012
					(0.003)
Bank branch density					0.005
					(0.001)
Constant	0.923	0.928	0.925	0.925	0.914
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
R^2	0.04	0.04	0.04	0.04	0.04
Observations	5,361,090	5,361,074	5,361,074	5,348,662	5,361,074

This table reports regressions of bank account participation on household and ZCTA-level characteristics. The household characteristics are log income and marital status. The ZCTA-level characteristics are regional price parities, average log income, standard deviation of log income, racial composition, economic connectedness, and bank branch presence and density. All coefficients on the continuous variables, except for racial composition, are standardized. Robust standard errors, clustered by ZCTA, are reported in parentheses. The sample includes households in the lowest income quintile with a member aged 50 to 59 in the 2019 administrative tax data, who have filed tax returns in the previous ten years.

Table 7

Instrumental variables regression for retirement account participation

	Percentile of usual income				
Regressor	0–20	20–40	40–60	60–80	80–100
Access to employer plan	0.32	0.36	0.34	0.31	0.24
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Additional years of access	0.01	0.03	0.04	0.03	0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Automatic enrollment	0.28	0.36	0.40	0.27	0.10
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Log income	0.00	0.04	0.04	0.04	0.02
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Single	-0.04	0.02	0.04	0.03	-0.06
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Employer size of					
Primary filer	-0.03	-0.04	-0.06	-0.05	-0.02
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Spouse	-0.06	-0.05	-0.06	-0.04	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.08	0.06	0.17	0.38	0.63
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Weak instrument test	2,439	8,453	7,090	4,425	8,165
Observations	5,840,537	3,537,129	1,707,164	791,418	718,509

The three endogenous regressors are an indicator variable for access to an employer retirement plan by either spouse from 2011 to 2019, additional years (beyond the first) that either spouse had access to an employer retirement plan without automatic enrollment, and an indicator variable for automatic enrollment in an employer retirement plan by either spouse. The intent-to-treat instrument is the counterfactual access to an employer retirement plan had the worker remained with the same employer from 2010 to 2019. Robust standard errors, clustered by ZCTA, are reported in parentheses. All specifications include indicator variables for the employer's two-digit NAICS code for the primary filer and spouse (if married), whose coefficients are not reported for brevity. The critical value for a test of weak instruments at the 5 percent significance level is 7.03 (Stock and Yogo, 2005, table 5.2). The sample includes all households with a member aged 50 to 59 in the 2019 administrative tax data, who did not have access to an employer retirement plan in 2010.

Table 8
State-sponsored retirement savings programs

State	Program	Effective year	Default contribution rate (%)	Employers with at least: Workers	Employers with at least: Years in business
California	CalSavers	2020	5	5	
Colorado	SecureSavings	2023	5	5	2
Connecticut	MyCTSavings	2022	3	5	
Illinois	Secure Choice	2018	5	5	2
Maine	MaineSaves	2023	5	5	2
Maryland	MarylandSaves	2022	5	1	2
New Jersey	Secure Choice	2022	3	25	2
New York	Secure Choice	2023	3	10	2
Oregon	OregonSaves	2017	5	5	
Virginia	VirginiaSaves	2023	5	25	2

This table reports features of the state-sponsored retirement savings programs as of December 2022.

Table 9
Predicted change in retirement account participation

	Percentile of usual income				
Participation rate	0–20	20–40	40–60	60–80	80–100
Actual	0.21	0.54	0.81	0.93	0.97
Counterfactual	0.40	0.70	0.88	0.96	0.98
Predicted change	0.19	0.16	0.07	0.02	0.01
Observations	7,214,614	7,214,482	7,214,243	7,214,417	7,214,423

In the counterfactual, workers have access to an employer retirement plan with automatic enrollment during all working years. The sample includes all households with a member aged 50 to 59 in the 2019 administrative tax data.

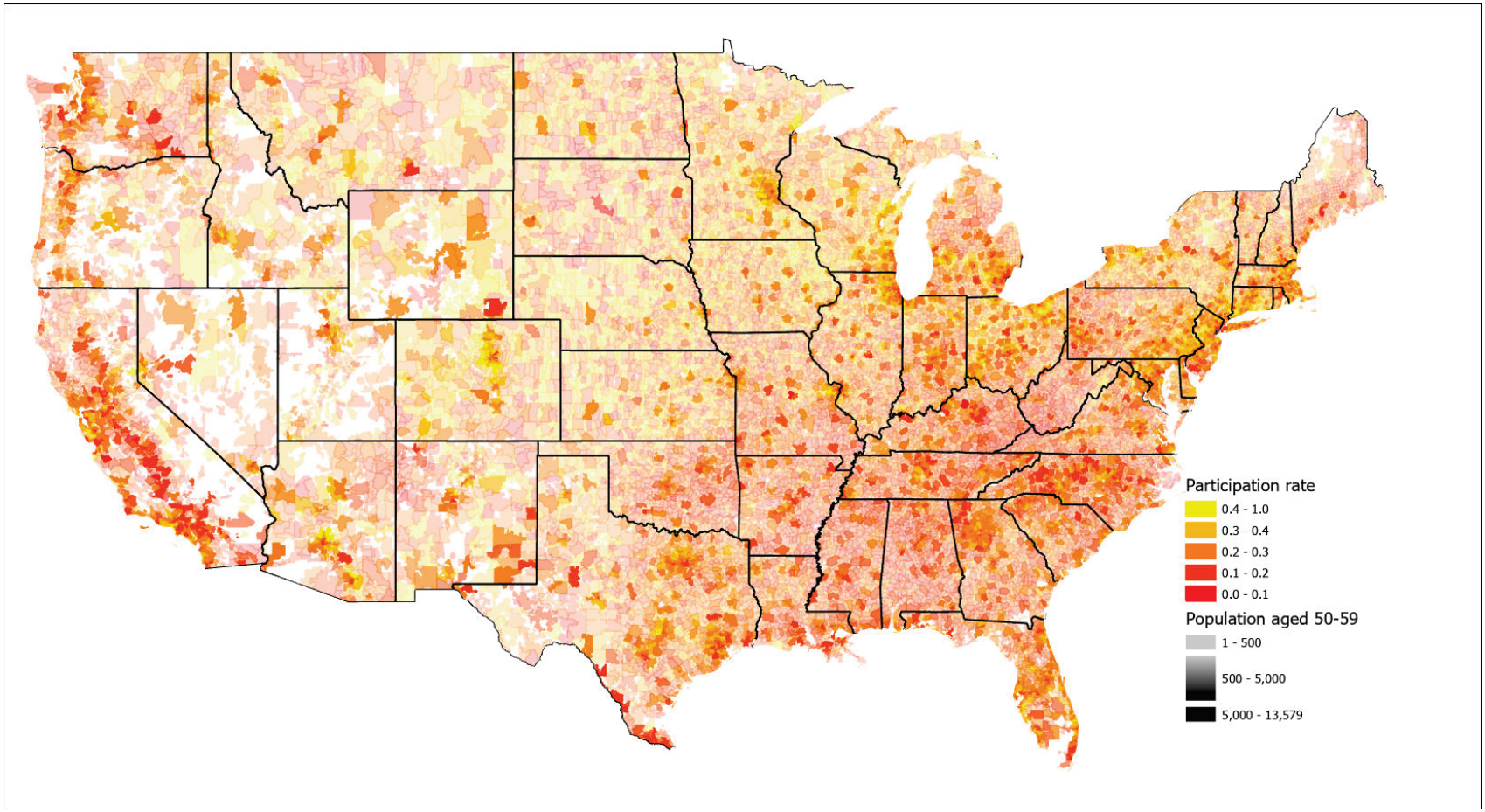

Fig. 1. Retirement account participation in the lowest income quintile. Retirement accounts include employer retirement plans (i.e., defined benefit and defined contribution plans) and IRAs. The sample includes all households in the lowest income quintile with a member aged 50 to 59 in the 2019 administrative tax data.

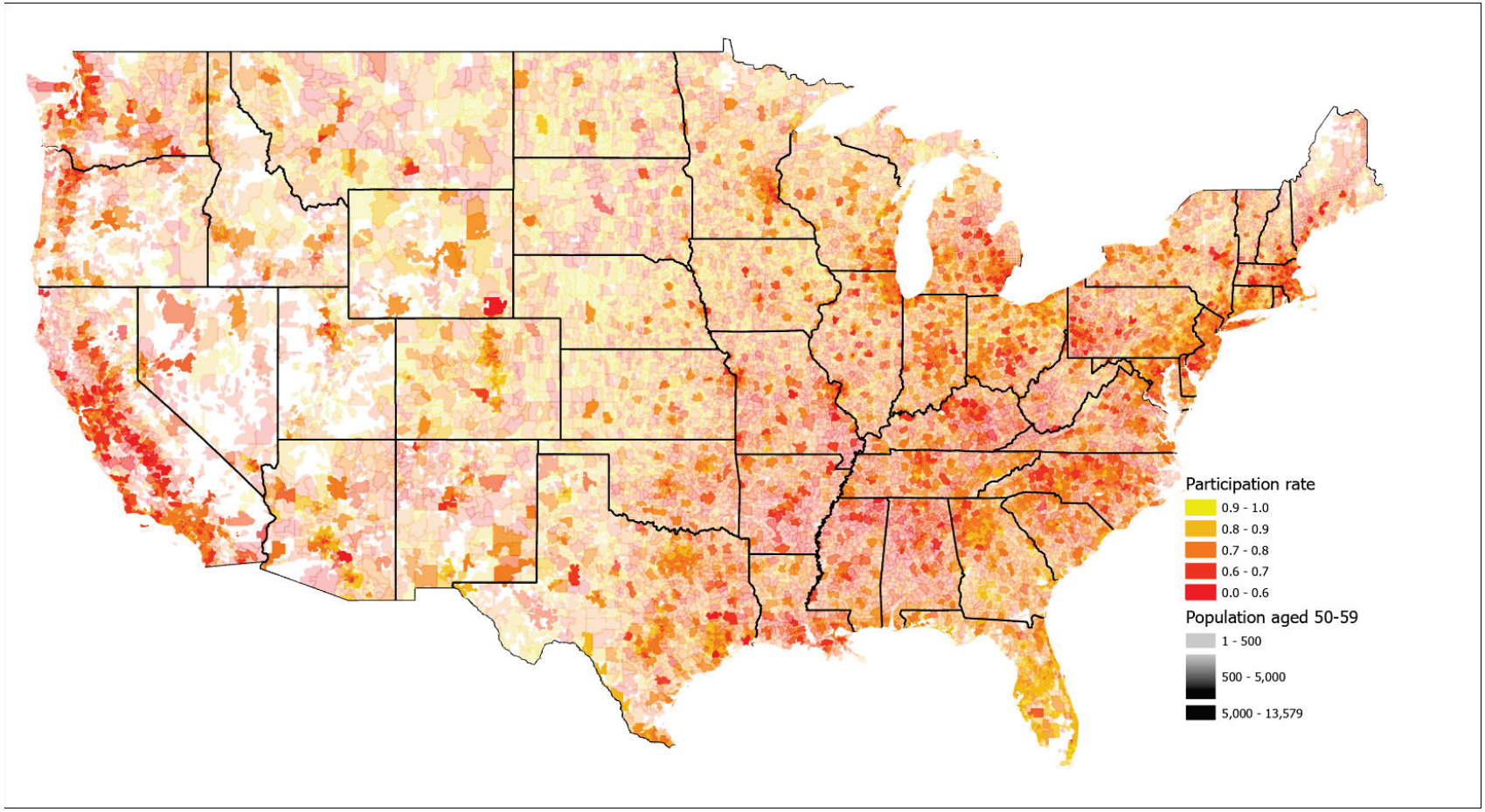

Fig. 2. Bank account participation in the lowest income quintile. The sample includes all households in the lowest income quintile with a member aged 50 to 59 in the 2019 administrative tax data.

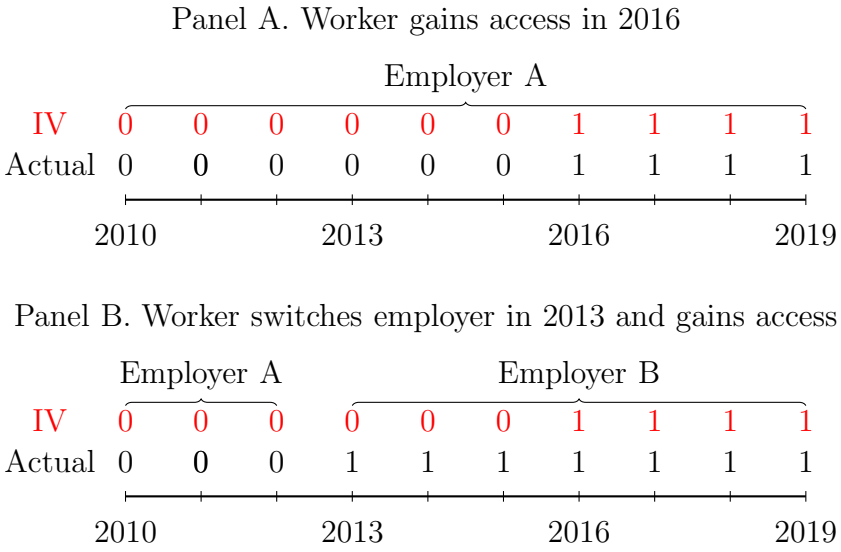

Fig. 3. Illustration of the instrument. In Panel A, the worker stays with employer A from 2010 to 2019 and gains access to an employer retirement plan in 2016. In Panel B, the worker switches from employer A to employer B and gains access to an employer retirement plan in 2013. The intent-to-treat instrument in red is the counterfactual access to an employer retirement plan had the worker remained with employer A from 2010 to 2019.

Appendix A. Administrative tax data

A.1. Population count

For each year from 2015 to 2019, we sample all individuals aged 50 to 59 on April 1 (i.e., census day), who have either a tax return or an information return with a ZIP Code within the U.S. states or Washington, DC. Table A.1 compares the population count in the administrative tax data with the resident population aged 50 to 59 in the census. The ratio of the population count in the administrative tax data to that in the census is slightly greater than one from 2015 to 2019.

Table A.1
Comparison of the population count

Year	Tax data	Census	Ratio
2015	44,221	43,985	1.01
2016	44,202	43,738	1.01
2017	43,920	43,278	1.01
2018	43,463	42,777	1.02
2019	43,332	42,355	1.02

The population counts are reported in thousands. The population count in the administrative tax data includes all individuals aged 50 to 59, who have either a tax return or an information return with a ZIP Code within the U.S. states or Washington, DC. The census population count is the resident population aged 50 to 59.

A potential reason for the slightly higher population count in the administrative tax data is the treatment of part-time U.S. residents and migrant workers. These individuals are in our sample if they receive information returns or file tax returns with a U.S. address, but they are not part of the census population if they do not reside in the United States on April 1. The difference in the population counts is sufficiently small to not be a concern for our purposes.

A.2. Measuring bank account participation

As we discussed in Section 2, we use a ten-year history to measure bank account participation. Thus, the definition of bank account participation in 2015 is having an account in the previous ten years, even if that account is closed as of 2015. As an alternative measure, we define bank account participation in 2015 as having an account as of the most recent tax filing in the previous ten years, which reduces the likelihood that the account is closed as of 2015. If bank account participation fluctuates over time, which may be more common among low-income households (Dlugosz et al., 2021), this alternative measure will be lower.

Table A.2 reports bank account participation under the alternative measure, which is indeed lower than that under the benchmark measure in Table 2. Under the alternative measure, the overall participation rate is 9 percentage points lower in 2019, and the participation rate is 14 percentage lower in the lowest income quinitle. However, we view these results as inconclusive, as the alternative measure could underestimate bank account participation if a tax filer alternates between checks and electronic funds transfer and chooses checks for the most recent tax filing.

Table A.2
Alternative measure of bank account participation by income

	Percentile of usual income					
Year	0–20	20–40	40–60	60–80	80–100	All
2015	0.56	0.78	0.86	0.92	0.95	0.82
2016	0.56	0.78	0.86	0.91	0.95	0.81
2017	0.55	0.78	0.86	0.92	0.95	0.81
2018	0.55	0.78	0.87	0.92	0.96	0.82
2019	0.56	0.80	0.88	0.93	0.96	0.83

This table reports an alternative measure of bank account participation for households with a member aged 50 to 59. We define bank account participation as having an account at the most recent tax filing in the previous ten years.

A.3. Aggregate datasets

Since we cannot disclose household-level data, we construct three aggregate datasets to facilitate future research. The first dataset (Data_incm) is aggregated by year and income quintile. The second dataset (Data_ZCTA) is aggregated by year and ZCTA. The third dataset (Data_ZCTA_incm) is aggregated by year, ZCTA, and income quintile.

For each of these datasets, we construct the following variables. We take three steps to avoid revealing information about specific households and to comply with data sharing requirements. First, we round usual household income to the nearest $100 after aggregation. We refer to Section 2 for the definition of usual household income. Second, we mask observations that would otherwise be derived from cells with fewer than 100 households, by aggregating with other cells. Data_incm is not subject to masking because the cells are sufficiently large. Third, we do not report the household count in each cell and instead estimate it using the census data.

- incm_usual: Average usual household income.
- Lincm_usual: Average log usual household income.

- SLincm_usual: Standard deviation of log usual household income.
- d_bank: Share of households that have a bank account.
- d_retire: Share of households that have a retirement account.
- d_emp_part: Share of households that have an employer retirement plan.
- d_emp_access: Share of households that have access to an employer retirement plan.
- obs: Census-derived household count in each cell, based on
  - hh2: Share of households with two people aged 50 to 59.
  - md_incm: Share of households in a given income quintile.

For Data_ZCTA, we define a cell as a {year, ZCTA} couplet and a small cell as one that has fewer than 100 households. We sort all small cells by the average of the variable that is being constructed. We group adjacent cells so that each group has between 100 and 300 households. By sorting before grouping, we maximize the chance that cells with similar average values are grouped together. We then calculate the weighted average of the variable within each group and assign it to all cells in that group. Finally, we discard the household count in each cell that was used for the weighted average. We repeat the masking procedure for all variables. About 28 percent of the 31,724 cells in 2019 are subject to masking.

For Data_ZCTA_incm, we apply the same masking procedure for all variables with three changes. First, we define a cell as a {year, ZCTA, income quintile} triplet. Second, we define a small cell as one that (a) has fewer than 20 households or (b) is nested in a {year, ZCTA} couplet with fewer than 100 households. Thus, we do not mask a cell if and only if it has at least 20 households and is part of a ZCTA with at least 100 households. Third, we group adjacent cells in two steps. We first group cells that are small according to definition (b) so that each group has between 100 and 300 households. We then group the remaining cells that are small according to definition (a) so that each group has between 20 and 60 households. About 31 percent of the 155,791 cells in 2019 are subject to masking.

We estimate household counts, based on the population aged 50 to 59 by ZCTA in the American Community Survey (U.S. Census Bureau, 2015–2019). Because these data are individual counts, we need to make an adjustment for households that have two people aged 50 to 59 to avoid double counting. For Data_ZCTA, we approximate the household count by ZCTA as Population/(1+hh2). For Data_ZCTA_incm, we approximate the household count by ZCTA and income quintile as Population×md_incm/(1+hh2).

Appendix B. Other data

B.1. Department of Labor Form 5500

We use the Forms 5500 and 5500-SF Annual Reports (U.S. Department of Labor, 2009–2019). We focus on the sample of defined contribution plans with active participants at either the beginning of the plan year (e.g., box 6a(1) on the 2019 form) or the end of the plan year (e.g., box 6a(2) on the 2019 form). For each employer, we construct an indicator variable for automatic enrollment if the pension feature code (e.g., box 8a of the 2019 form) is 2S on any of its active plans. To catch cases where the pension feature code is incomplete, we do a textual analysis of the plan description in the actual filing to search for "automatic enrollment", "auto enrollment", or "default enrollment".

As we discussed in Section 2, we measure access to an employer retirement plan if retirement plan (box 13) is checked on any Form W-2 issued by the employer. For this subset of employers, we construct an indicator variable for automatic enrollment by merging with the Form 5500 data by employer identification number (EIN) and tax year. If the EIN on the Form W-2 fails to merge with a Form 5500, we instead use the parent firm's EIN on Form 851 (Affiliations Schedule). Thus, we attribute the parent firm's retirement plan to their subsidiaries in cases where the subsidiary's EIN does not merge with a Form 5500.

B.2. Regional price parities

We use the regional price parities by state (Bureau of Economic Analysis, 2019) to control for geographic differences in the cost of living.

B.3. Racial composition

We construct the population shares by race at the ZCTA level, based on the American Community Survey (U.S. Census Bureau, 2015–2019). We group race into white, Hispanic, Black, Asian, or other nonwhite. Other nonwhite includes American Indian, Alaska Native, Native Hawaiian, other Pacific Islander, and multiple race.

B.4. Economic connectedness

We use the baseline measure of economic connectedness at the county level (ec_county) from Chetty et al. (2022). We use the county-level mesure because the ZCTA-level measure is not available for all ZCTAs in our sample. Their estimation sample consists of Facebook users aged 25 to 44 on May 28, 2022, who have a residential ZIP Code in the United States, have

at least 100 U.S.-based Facebook friends, and were active on the Facebook platform at least once in the previous 30 days.

B.5. Bank branch access

We count the number of bank branches by ZCTA, based on the Annual Survey of Branch Office Deposits (Federal Deposit Insurance Corporation, 2019). For each ZCTA, we construct bank branch density as the logarithm of the number of branches divided by its population from the American Community Survey. We winsorize the right tail at 10 branches per 1,000 residents (about 7 percent of observations) to reduce the impact of outliers. We set bank branch density to zero for ZCTAs without a bank branch.

Appendix C. First-stage regressions

Table C.1

First-stage regression for access to an employer retirement plan

	Percentile of usual income				
Regressor	0–20	20–40	40–60	60–80	80–100
Instrument for					
Access to employer plan	0.02	0.16	0.25	0.28	0.42
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Additional years of access	0.05	0.02	0.01	0.00	0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Automatic enrollment	0.23	0.06	0.01	0.00	0.02
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Log income	0.09	0.08	0.01	-0.01	-0.02
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Single	-0.02	0.12	0.06	-0.01	-0.11
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Employer size of					
Primary filer	0.00	0.04	0.04	0.04	0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Spouse	0.03	0.04	0.02	0.02	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.20	0.43	0.61	0.63	0.61
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	5,840,537	3,537,129	1,707,164	791,418	718,509

The endogenous regressor is an indicator variable for access to an employer retirement plan by either spouse from 2011 to 2019. The intent-to-treat instrument is the counterfactual access to an employer retirement plan had the worker remained with the same employer from 2010 to 2019. The coefficients for log income and employer size are standardized. Robust standard errors, clustered by ZCTA, are reported in parentheses. All specifications include indicator variables for the employer's two-digit NAICS code for the primary filer and spouse (if married), whose coefficients are not reported for brevity. The sample includes all households with a member aged 50 to 59 in the 2019 administrative tax data, who did not have access to an employer retirement plan in 2010.

Table C.2
First-stage regression for additional years of access to an employer retirement plan

	Percentile of usual income				
Regressor	0–20	20–40	40–60	60–80	80–100
Instrument for					
Access to employer plan	-0.18	0.08	0.36	0.47	0.64
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Additional years of access	0.19	0.24	0.26	0.29	0.44
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Automatic enrollment	0.48	0.14	-0.25	-0.51	-0.40
	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)
Log income	0.12	0.36	0.07	-0.07	-0.11
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Single	-0.04	0.49	0.37	0.01	-0.32
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Employer size of					
Primary filer	-0.02	0.08	0.10	0.10	-0.05
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Spouse	0.03	0.10	0.06	0.01	-0.17
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Constant	0.22	0.75	1.56	1.70	1.35
	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)
Observations	5,840,537	3,537,129	1,707,164	791,418	718,509

The endogenous regressor is additional years (beyond the first) that either spouse had access to an employer retirement plan without automatic enrollment from 2011 to 2019. The intent-to-treat instrument is the counterfactual access to an employer retirement plan had the worker remained with the same employer from 2010 to 2019. The coefficients for log income and employer size are standardized. Robust standard errors, clustered by ZCTA, are reported in parentheses. All specifications include indicator variables for the employer's two-digit NAICS code for the primary filer and spouse (if married), whose coefficients are not reported for brevity. The sample includes all households with a member aged 50 to 59 in the 2019 administrative tax data, who did not have access to an employer retirement plan in 2010.

Table C.3

First-stage regression for automatic enrollment in an employer retirement plan

	Percentile of usual income				
Regressor	0–20	20–40	40–60	60–80	80–100
Instrument for					
Access to employer plan	-0.02	-0.01	-0.01	-0.01	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Additional years of access	0.01	0.01	0.01	0.01	0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Automatic enrollment	0.18	0.27	0.36	0.41	0.54
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Log income	0.01	0.03	0.01	0.00	-0.03
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Single	0.00	0.04	0.03	0.00	-0.04
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Employer size of					
Primary filer	0.00	0.02	0.04	0.04	0.02
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Spouse	0.01	0.01	0.02	0.02	0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.02	0.07	0.14	0.18	0.17
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	5,840,537	3,537,129	1,707,164	791,418	718,509

The endogenous regressor is an indicator variable for automatic enrollment in an employer retirement plan by either spouse from 2011 to 2019. The intent-to-treat instrument is the counterfactual access to an employer retirement plan had the worker remained with the same employer from 2010 to 2019. The coefficients for log income and employer size are standardized. Robust standard errors, clustered by ZCTA, are reported in parentheses. All specifications include indicator variables for the employer's two-digit NAICS code for the primary filer and spouse (if married), whose coefficients are not reported for brevity. The sample includes all households with a member aged 50 to 59 in the 2019 administrative tax data, who did not have access to an employer retirement plan in 2010.